9

Permutations and Group Actions

This chapter contains an introduction to some aspects of group theory that are directly related to combinatorial problems. The first part of the chapter gives the basic definitions of group theory and derives some fundamental properties of *symmetric groups*. We apply this material to give combinatorial derivations of the basic properties of determinants. The second part of the chapter discusses *group actions*, which have many applications to algebra and combinatorics. In particular, group actions can be used to solve counting problems in which symmetry must be taken into account. For example, how many ways can we color a 5×5 chessboard with seven colors, if all rotations and reflections of a given colored board are considered the same? The theory of group actions provides systematic methods for solving problems like this one.

9.1 Definition and Examples of Groups

9.1. Definition: Groups. A *group* consists of a set G and a binary operation $\star : G \times G \to G$ subject to the following axioms:

$$\begin{array}{ll} \forall x, y, z \in G, x \star (y \star z) = (x \star y) \star z & \text{(associativity);} \\ \exists e \in G, \forall x \in G, x \star e = x = e \star x & \text{(identity);} \\ \forall x \in G, \exists y \in G, x \star y = e = y \star x & \text{(inverses).} \end{array}$$

The requirement that $\star$ map $G \times G$ *into* G is often stated explicitly as the following axiom:

$$\forall x, y \in G, x \star y \in G \qquad \text{(closure).}$$

A group G is called *abelian* or *commutative* iff G satisfies the additional axiom

$$\forall x, y \in G, x \star y = y \star x \qquad \text{(commutativity).}$$

9.2. Example: Additive Groups. The set $\mathbb{Z}$ of all integers, with addition as the operation, is a commutative group. The identity element is $e = 0$ and the (additive) inverse of $x \in \mathbb{Z}$ is $-x \in \mathbb{Z}$. Similarly, $\mathbb{Q}$ and $\mathbb{R}$ and $\mathbb{C}$ are all commutative groups under addition. $\mathbb{N}^+$ is not a group under addition because there is no identity element *in the set* $\mathbb{N}^+$. $\mathbb{N}$ is not a group under addition because $1 \in \mathbb{N}$ has no additive inverse *in the set* $\mathbb{N}$. The three-element set $S = \{-1, 0, 1\}$ is not a group under addition because closure fails ($1 + 1 = 2 \notin S$).

9.3. Example: Multiplicative Groups. The set $\mathbb{Q}^+$ of strictly positive rational numbers is a commutative group under multiplication. The identity element is $e = 1$ and the inverse of $a/b \in \mathbb{Q}^+$ is b/a. Similarly, $\mathbb{R}^+$ is a group under multiplication. The set $\mathbb{Q}$ is not a group under multiplication because 0 has no inverse. On the other hand, $\mathbb{Q} \sim \{0\}$, $\mathbb{R} \sim \{0\}$, and $\mathbb{C} \sim \{0\}$ are groups under multiplication. So is the two-element set $\{-1, 1\} \subseteq \mathbb{Q}$, and the four-element set $\{1, i, -1, -i\} \subseteq \mathbb{C}$.

9.4. Example: Symmetric Groups. Let X be any set, and let $\mathrm{Sym}(X)$ be the set of all bijections $f : X \to X$. For $f, g \in \mathrm{Sym}(X)$, define $f \circ g$ to be the composite function that sends $x \in X$ to $f(g(x))$. Then $f \circ g \in \mathrm{Sym}(X)$ since the composition of bijections is a bijection, so the axiom of closure holds. Given $f, g, h \in \mathrm{Sym}(X)$, note that both of the functions $(f \circ g) \circ h : X \to X$ and $f \circ (g \circ h) : X \to X$ send $x \in X$ to $f(g(h(x)))$. So these functions are equal, proving the axiom of associativity. Next, take e to be the bijection $\mathrm{id}_X : X \to X$, which is defined by $\mathrm{id}_X(x) = x$ for all $x \in X$. One immediately checks that $f \circ \mathrm{id}_X = f = \mathrm{id}_X \circ f$ for all $f \in \mathrm{Sym}(X)$, so the identity axiom holds. Finally, given a bijection $f \in \mathrm{Sym}(X)$, there exists an inverse function $f^{-1} : X \to X$ that is also a bijection, and which satisfies $f \circ f^{-1} = \mathrm{id}_X = f^{-1} \circ f$. So the axiom of inverses holds. This completes the verification that $(\mathrm{Sym}(X), \circ)$ is a group. This group is called the *symmetric group on* X, and elements of $\mathrm{Sym}(X)$ are called *permutations of* X. Symmetric groups play a central role in group theory and are closely related to group actions. In the special case when $X = \{1, 2, \ldots, n\}$, we write S_n to denote the group $\mathrm{Sym}(X)$.

Most of the groups $\mathrm{Sym}(X)$ are *not* commutative. For instance, consider $f, g \in S_3$ given by

$$f(1) = 2,\ f(2) = 1,\ f(3) = 3; \quad g(1) = 3,\ g(2) = 2,\ g(3) = 1.$$

We see that $(f \circ g)(1) = f(g(1)) = 3$, whereas $(g \circ f)(1) = g(f(1)) = 2$. So $f \circ g \neq g \circ f$, and the axiom of commutativity fails.

9.5. Example: Integers modulo n. Let n be a fixed positive integer. Consider the set $\mathbb{Z}_n = \underline{\mathbf{n}} = \{0, 1, 2, \ldots, n-1\}$. We define a binary operation on $\mathbb{Z}_n$ by setting, for all $x, y \in \mathbb{Z}_n$,

$$x \oplus y = \begin{cases} x + y & \text{if } x + y < n; \\ x + y - n & \text{if } x + y \geq n. \end{cases}$$

Closure follows from this definition, once we note that $0 \leq x + y \leq 2n - 2$ for $x, y \in \mathbb{Z}_n$. The identity element is 0. The inverse of 0 is 0, while for $x > 0$ in $\mathbb{Z}_n$, the inverse of x is $n - x \in \mathbb{Z}_n$. To verify associativity, one may prove the relations

$$(x \oplus y) \oplus z = \left\{ \begin{array}{ll} x + y + z & \text{if } x + y + z < n; \\ x + y + z - n & \text{if } n \leq x + y + z < 2n; \\ x + y + z - 2n & \text{if } 2n \leq x + y + z < 3n; \end{array} \right\} = x \oplus (y \oplus z), \tag{9.1}$$

which can be established by a tedious case analysis. Commutativity of $\oplus$ follows from the definition and the commutativity of ordinary integer addition. We conclude that $(\mathbb{Z}_n, \oplus)$ is a commutative group containing n elements. In particular, for every positive integer n, there exists a group of cardinality n.

9.6. Definition: Multiplication Tables. If $(G, \star)$ is a group and $G = \{x_1, \ldots, x_n\}$ is finite, then a *multiplication table* for G is an $n \times n$ table, with rows and columns labeled by the x_i's, such that the element in row i and column j is $x_i \star x_j$. When the operation is written additively, we refer to this table as the *addition table* for G. It is customary, but not mandatory, to take x_1 to be the identity element of G.

9.7. Example. The multiplication tables for $(\{1, i, -1, -i\}, \times)$ and for $(\mathbb{Z}_4, \oplus)$ are shown here:

$\times$	1	i	-1	$-i$
1	1	i	-1	$-i$
i	i	-1	$-i$	1
-1	-1	$-i$	1	i
$-i$	$-i$	1	i	-1

$\oplus$	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

The reader may notice a relationship between the two tables: each row within the table is

obtained from the preceding one by a cyclic shift one step to the left. (Using terminology to be discussed later, this happens because each of the two groups under consideration is cyclic of size four.)

One can define a group operation by specifying its multiplication table. For example, here is the table for another group of size four (which turns out not to be cyclic):

$\star$	a	b	c	d
a	a	b	c	d
b	b	a	d	c
c	c	d	a	b
d	d	c	b	a

The identity and inverse axioms can be checked from inspection of the table (here a is the identity, and every element is equal to its own inverse). There is no quick way to verify associativity by visual inspection of the table, but this axiom can be checked exhaustively using the table entries.

All of the groups in this example are commutative. This can be read off from the multiplication tables by noting the symmetry about the main diagonal line (the entry $x_i \star x_j$ in row i and column j always equals the entry $x_j \star x_i$ in row j and column i).

9.2 Basic Properties of Groups

We now collect some facts about groups that follow from the defining axioms.

First, the identity element e in a group $(G, \star)$ is *unique*. For, suppose $e' \in G$ also satisfies the identity axiom. On one hand, $e \star e' = e$ since e' is an identity element. On the other hand, $e \star e' = e'$ since e is an identity element. So $e = e'$. (The very statement of the inverse axiom makes implicit use of the uniqueness of the identity.) We use the symbol e_G to denote the identity element of an abstract group G. When the operation is addition or multiplication, we write 0_G or 1_G instead, dropping the G if it is understood from context.

Similarly, the inverse of an element x in a group G is unique. For suppose $y, y' \in G$ both satisfy the condition in the inverse axiom. Then

$$y = y \star e = y \star (x \star y') = (y \star x) \star y' = e \star y' = y'.$$

We denote the unique inverse of x in G by the symbol x^{-1}. When the operation is written additively, the symbol $-x$ is used.

A product such as $x \star y$ is often written xy, except in the additive case. The associativity axiom can be used to show that any parenthesization of a product $x_1 x_2 \cdots x_n$ will give the same answer (see 2.148), so it is permissible to omit parentheses in products like these.

9.8. Theorem: Cancellation Laws and Inverse Rules. Suppose a, x, y are elements in a group G. (a) $ax = ay$ implies $x = y$ (left cancellation); (b) $xa = ya$ implies $x = y$ (right cancellation); (c) $(x^{-1})^{-1} = x$; (d) $(xy)^{-1} = y^{-1}x^{-1}$ (inverse rule for products).

Proof. Starting from $ax = ay$, multiply both sides on the left by a^{-1} to get $a^{-1}(ax) = a^{-1}(ay)$. Then the associativity axiom gives $(a^{-1}a)x = (a^{-1}a)y$; the inverse axiom gives $ex = ey$; and the identity axiom gives $x = y$. Right cancellation is proved similarly. Next, note that

$$(x^{-1})^{-1}x^{-1} = e = xx^{-1}$$

by the definition of the inverse of x and of x^{-1}; right cancellation of x^{-1} yields $(x^{-1})^{-1} = x$. Similarly, routine calculations using the group axioms show that

$$(xy)^{-1}(xy) = e = (y^{-1}x^{-1})(xy),$$

so right cancellation of xy gives the inverse rule for products. □

9.9. Definition: Exponent Notation. Let G be a group written multiplicatively. Given $x \in G$, recursively define $x^0 = 1 = e_G$ and $x^{n+1} = x^n \star x$ for all $n \geq 0$. To define negative powers of x, set $x^{-n} = (x^{-1})^n$ for all $n > 0$.

Informally, for positive n, x^n is the product of n copies of x. For negative n, x^n is the product of $|n|$ copies of the inverse of x. Note in particular that $x^1 = x$ and x^{-1} is the inverse of x (in accordance with the conventions introduced before this definition). When G is written additively, we write nx instead of x^n; this denotes the sum of n copies of x for $n > 0$, or the sum of $|n|$ copies of $-x$ for $n < 0$.

9.10. Theorem: Laws of Exponents. Suppose G is a group written multiplicatively, $x \in G$, and $m, n \in \mathbb{Z}$. Then $x^{m+n} = x^m x^n$ and $x^{mn} = (x^n)^m$. If $x, y \in G$ satisfy $xy = yx$, then $(xy)^n = x^n y^n$. In additive notation, these results read: $(m+n)x = mx + nx$; $(mn)x = m(nx)$; and $n(x+y) = nx + ny$ when $x + y = y + x$.

The idea of the proof is to use induction to establish the results for $m, n \geq 0$, and then use case analyses to handle the situations where m or n or both is negative. We leave the details as an exercise for the reader.

9.3 Notation for Permutations

Permutations and symmetric groups arise frequently in the theory of groups and group actions. So we will now develop some notation for describing permutations.

9.11. Definition: Two-Line Form of a Function. Let X be a finite set, and let $x_1, \ldots, x_n$ be a list of all the distinct elements of X in some fixed order. The *two-line form* of a function $f : X \to X$ relative to this ordering is the array

$$f = \begin{pmatrix} x_1 & x_2 & \cdots & x_n \\ f(x_1) & f(x_2) & \cdots & f(x_n) \end{pmatrix}.$$

If $X = \{1, 2, \ldots, n\}$, we usually display the elements of X on the top line in the order $1, 2, \ldots, n$.

9.12. Example. The notation $f = \begin{pmatrix} a & b & c & d & e \\ b & c & e & a & b \end{pmatrix}$ defines a function on the set $X = \{a, b, c, d, e\}$ such that $f(a) = b$, $f(b) = c$, $f(c) = e$, $f(d) = a$, and $f(e) = b$. This function is not a permutation, since b occurs twice in the bottom row and d never occurs.

The notation $g = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 5 & 1 & 3 \end{pmatrix}$ defines an element of S_5 such that $g(1) = 2$, $g(2) = 4$, $g(3) = 5$, $g(4) = 1$, and $g(5) = 3$. Observe that the *inverse* of g sends 2 to 1, 4 to 2, and so on. So, we obtain one possible two-line form of g^{-1} by interchanging the lines in the two-line form of g:

$$g^{-1} = \begin{pmatrix} 2 & 4 & 5 & 1 & 3 \\ 1 & 2 & 3 & 4 & 5 \end{pmatrix}.$$

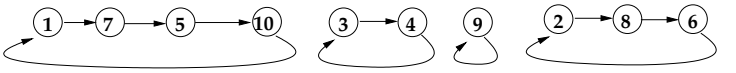

FIGURE 9.1
Digraph associated to the permutation h.

It is customary to write the numbers in the top line in increasing order. This can be accomplished by sorting the columns of the previous array:

$$g^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 1 & 5 & 2 & 3 \end{pmatrix}.$$

Recall that the group operation in $\mathrm{Sym}(X)$ is composition. We can compute the composition of two functions written in two-line form by tracing the effect of the composite function on each element. For instance,

$$\begin{pmatrix} a & b & c & d \\ b & d & a & c \end{pmatrix} \circ \begin{pmatrix} a & b & c & d \\ a & c & d & b \end{pmatrix} = \begin{pmatrix} a & b & c & d \\ b & a & c & d \end{pmatrix},$$

because the left side maps a to a and then to b; b maps to c and then to a; and so on.

If the ordering of X is fixed and known from context, we may omit the top line of the two-line form. This leads to one-line notation for a function defined on X.

9.13. Definition: One-Line Form of a Function. Let $X = \{x_1 < x_2 < \cdots < x_n\}$ be a finite totally ordered set. The *one-line form* of a function $f : X \to X$ is the array $[f(x_1)\ f(x_2) \cdots f(x_n)]$. We use square brackets to avoid a conflict with the cycle notation to be introduced below. Sometimes we omit the brackets, identifying f with the word $f(x_1)f(x_2)\cdots f(x_n)$.

9.14. Example. The functions f and g in the preceding example are given in one-line form by writing $f = [b\ c\ e\ a\ b]$ and $g = [2\ 4\ 5\ 1\ 3]$. Note that the one-line form of an element of $\mathrm{Sym}(X)$ is literally a permutation of the elements of X, as defined in §1.4. This explains why elements of this group are called permutations.

9.15. Cycle Notation for Permutations. Assume X is a finite set. Recall from §3.6 that any function $f : X \to X$ can be represented by a digraph with vertex set X and directed edges $\{(i, f(i)) : i \in X\}$. A digraph on X arises from a function in this way iff every vertex in X has outdegree 1. In 3.45 we proved that the digraph of a permutation is a disjoint union of directed cycles. For example, Figure 9.1 displays the digraph of the permutation

$$h = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 7 & 8 & 4 & 3 & 10 & 2 & 5 & 6 & 9 & 1 \end{pmatrix}.$$

We can describe a directed cycle in a digraph by traversing the edges in the cycle and listing the elements we encounter in the order of visitation, enclosing the whole list in parentheses. For example, the cycle containing 1 in Figure 9.1 can be described by writing $(1, 7, 5, 10)$. The cycle containing 9 is denoted by (9). To describe the entire digraph of a permutation, we write down all the cycles in the digraph, one after the other. For example, h can be written in cycle notation as

$$h = (1, 7, 5, 10)(2, 8, 6)(3, 4)(9).$$

This cycle notation is not unique. We are free to begin our description of each cycle at any vertex in the cycle, and we are free to rearrange the order of the cycles. Furthermore, by convention it is permissible to omit some or all cycles of length 1. For example, some other cycle notations for h are

$$h = (5,10,1,7)(3,4)(9)(6,2,8) = (2,8,6)(4,3)(7,5,10,1).$$

To compute the inverse of a permutation written in cycle notation, we reverse the orientation of each cycle. For example,

$$h^{-1} = (10,5,7,1)(6,8,2)(4,3)(9).$$

9.16. Example. Using cycle notation, we can list the six elements of S_3 as follows:

$$S_3 = \{(1)(2)(3),\ (1,2),\ (1,3),\ (2,3),\ (1,2,3),\ (1,3,2)\}.$$

9.17. Example. To compose permutations written in cycle notation, we must see how the composite function acts on each element. For instance, consider the product $(3,5)(1,2,4) \circ (3,5,2,1)$ in S_5. This composite function sends 1 to 3 and then 3 to 5, so 1 maps to 5. Next, 2 maps first to 1 and then to 2, so 2 maps to 2. Continuing similarly, we find that

$$(3,5)(1,2,4) \circ (3,5,2,1) = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 2 & 3 & 1 & 4 \end{pmatrix} = (1,5,4)(2)(3).$$

With enough practice, one can proceed immediately to the cycle form of the answer without writing down the two-line form or other scratch work.

9.18. Definition: k-cycles. For $k > 1$, a permutation $f \in \mathrm{Sym}(X)$ whose digraph consists of one cycle of length k and all other cycles of length 1 is called a *k-cycle*.

9.19. Remark. We can view the cycle notation for a permutation f as a way of factorizing f *in the group* S_n into a product of cycles. For example,

$$(1,7,5,10)(2,8,6)(3,4)(9) = (1,7,5,10) \circ (2,8,6) \circ (3,4) \circ (9).$$

Here we have expressed the single permutation on the left side as a product of four other permutations in S_{10}. The stated equality may be verified by checking that both sides have the same effect on each $k \in \{1,2,\ldots,10\}$.

9.20. Definition: $\mathrm{cyc}(f)$ **and** $\mathrm{type}(f)$**.** Given a permutation $f \in \mathrm{Sym}(X)$, let $\mathrm{cyc}(f)$ be the number of components (cycles) in the digraph for f. Let $\mathrm{type}(f)$ be the list of sizes of these components, including repetitions and written in weakly decreasing order.

Note that $\mathrm{type}(f)$ is an integer partition of $n = |X|$.

9.21. Example. The permutation h in Figure 9.1 has $\mathrm{cyc}(h) = 4$ and $\mathrm{type}(h) = (4,3,2,1)$. The identity element of S_n, namely $\mathrm{id} = (1)(2)\cdots(n)$, has $\mathrm{cyc}(\mathrm{id}) = n$ and $\mathrm{type}(\mathrm{id}) = (1,\ldots,1)$. Table 9.1 displays the 24 elements of S_4 in cycle notation, collecting together all permutations with the same type and counting the number of permutations of each type. In 9.134, we will give a general formula for the number of permutations of n objects having a given type.

TABLE 9.1
Elements of S_4.

Type	Permutations of this type	Count
$(1,1,1,1)$	$(1)(2)(3)(4)$	1
$(2,1,1)$	$(1,2),\ (1,3),\ (1,4),\ (2,3),\ (2,4),\ (3,4)$	6
$(2,2)$	$(1,2)(3,4),\ (1,3)(2,4),\ (1,4)(2,3)$	3
$(3,1)$	$(1,2,3),\ (1,2,4),\ (1,3,4),\ (1,3,2),$ $(1,4,2),\ (1,4,3),\ (2,3,4),\ (2,4,3)$	8
(4)	$(1,2,3,4),\ (1,2,4,3),\ (1,3,2,4),$ $(1,3,4,2),\ (1,4,2,3),\ (1,4,3,2)$	6

9.4 Inversions and Sign

In this section, we use inversions of permutations to define the *sign* function $\text{sgn} : S_n \to \{+1, -1\}$. We then study factorizations of permutations into products of transpositions to derive facts about the sgn function. Let us begin by recalling the definition of inversions (§6.2).

9.22. Definition: Inversions and Sign of a Permutation. Let $w = w_1 w_2 \cdots w_n \in S_n$ be a permutation written in one-line form. An *inversion* of w is a pair of indices $i < j$ such that $w_i > w_j$. The number of inversions of w is denoted $\text{inv}(w)$. Furthermore, the *sign* of w is defined to be $\text{sgn}(w) = (-1)^{\text{inv}(w)}$.

9.23. Example. Given $w = 42531$, we have $\text{inv}(w) = 7$ and $\text{sgn}(w) = -1$. The seven inversions of w are $(1,2)$, $(1,4)$, $(1,5)$, $(2,5)$, $(3,4)$, $(3,5)$, and $(4,5)$. For instance, $(1,4)$ is an inversion because $w_1 = 4 > 3 = w_4$. The following table displays $\text{inv}(f)$ and $\text{sgn}(f)$ for all $f \in S_3$:

$f \in S_3$	$\text{inv}(f)$	$\text{sgn}(f)$
123	0	$+1$
132	1	-1
213	1	-1
231	2	$+1$
312	2	$+1$
321	3	-1

We want to understand how the group operation in S_n (composition of permutations) is related to inversions and sign. For this purpose, we introduce the concept of a *transposition*.

9.24. Definition: Transpositions. A *transposition* in S_n is a permutation f of the form (i,j), for some $i, j \leq n$. Note that $f(i) = j$, $f(j) = i$, and $f(k) = k$ for all $k \neq i, j$. A *basic transposition* in S_n is a transposition $(i, i+1)$, for some $i < n$.

The following lemmas illuminate the connection between basic transpositions and the process of sorting the one-line form of a permutation into increasing order.

9.25. Lemma: Basic Transpositions and Sorting. Let $w = w_1 \cdots w_i w_{i+1} \cdots w_n \in S_n$ be a permutation in one-line form. For each $i < n$,

$$w \circ (i, i+1) = w_1 \cdots w_{i+1} w_i \cdots w_n.$$

So *right-multiplication by the basic transposition* $(i, i+1)$ *interchanges the elements in positions* i *and* $i+1$ *of* w.

Proof. Let us evaluate the function $f = w \circ (i, i+1)$ at each $k \leq n$. When $k = i$, $f(i) = w(i+1)$. When $k = i+1$, $f(i+1) = w(i)$. When $k \neq i$ and $k \neq i+1$, $f(k) = k$. So the one-line form of f is $w_1 \cdots w_{i+1} w_i \cdots w_n$, as desired. □

9.26. Lemma: Basic Transpositions and Inversions. Let $w = w_1 \cdots w_n \in S_n$ be a permutation in one-line form, and let $i < n$. Then

$$\mathrm{inv}(w \circ (i, i+1)) = \begin{cases} \mathrm{inv}(w) + 1 & \text{if } w_i < w_{i+1}; \\ \mathrm{inv}(w) - 1 & \text{if } w_i > w_{i+1}. \end{cases}$$

Consequently, in all cases, we have

$$\mathrm{sgn}(w \circ (i, i+1)) = -\mathrm{sgn}(w).$$

Proof. We use the result of the previous lemma to compare the inversions of w and $w' = w \circ (i, i+1)$. Let $j < k$ be two indices between 1 and n, and consider various cases. First, if $j \neq i, i+1$ and $k \neq i, i+1$, then (j, k) is an inversion of w iff (j, k) is an inversion of w', since $w_j = w'_j$ and $w_k = w'_k$. Second, if $j = i$ and $k > i+1$, then (i, k) is an inversion of w iff $(i+1, k)$ is an inversion of w', since $w_i = w'_{i+1}$ and $w_k = w'_k$. Similar results hold in the cases $(j = i+1 < k)$, $(j < k = i)$, and $(j < i,\ k = i+1)$. The critical case is when $j = i$ and $k = i+1$. If $w_i < w_{i+1}$, then (j, k) is an inversion of w' but not of w. If $w_i > w_{i+1}$, then (j, k) is an inversion of w but not of w'. This establishes the first formula in the lemma. The remaining formula follows since $(-1)^{+1} = (-1)^{-1} = -1$. □

The proof of the next lemma is left as an exercise.

9.27. Lemma. For all $n \geq 1$, the identity permutation $\mathrm{id} = 1, 2, \ldots, n$ is the unique element of S_n satisfying $\mathrm{inv}(\mathrm{id}) = 0$. We have $\mathrm{sgn}(\mathrm{id}) = +1$.

If $f = (i, i+1)$ is a basic transposition, then the ordered pair $(i, i+1)$ is the only inversion of f, so $\mathrm{inv}(f) = 1$ and $\mathrm{sgn}(f) = -1$. More generally, we now show that any transposition has sign -1.

9.28. Lemma. If $f = (i, j)$ is any transposition, then $\mathrm{sgn}(f) = -1$.

Proof. Since $(i, j) = (j, i)$, we may assume that $i < j$. Let us write f in two-line form:

$$f = \begin{pmatrix} 1 & \cdots & i & \cdots & j & \cdots & n \\ 1 & \cdots & j & \cdots & i & \cdots & n \end{pmatrix}.$$

We can find the inversions of f by inspecting the two-line form. The inversions are: all (i, k) with $i < k \leq j$; and all (k, j) with $i < k < j$. There are $j - i$ inversions of the first type and $j - i - 1$ inversions of the second type, hence $2(j - i) - 1$ inversions total. Since this number is odd, we conclude that $\mathrm{sgn}(f) = -1$. □

9.29. Theorem: Inversions and Sorting. Let $w = w_1 w_2 \cdots w_n \in S_n$ be a permutation in one-line form. The number $\mathrm{inv}(w)$ is the minimum number of steps required to sort the word w into increasing order by repeatedly interchanging two adjacent elements. Furthermore, w can be factored in S_n into the product of $\mathrm{inv}(w)$ basic transpositions.

Proof. Given $w \in S_n$, it is certainly possible to sort w into increasing order in finitely many steps by repeatedly swapping adjacent elements. For instance, we can move 1 to the far left position in at most $n-1$ moves, then move 2 to its proper position in at most $n-2$ moves, and so on. Let m be the *minimum* number of moves of this kind that are needed to sort w. By 9.25, we can accomplish each sorting move by starting with w and repeatedly multiplying on the right by suitable basic transpositions. Each such multiplication either increases or decreases the inversion count by 1, according to 9.26. At the end, we have transformed w into the identity permutation. Combining these observations, we see that $0 = \mathrm{inv}(\mathrm{id}) \geq \mathrm{inv}(w) - m$, so that $m \geq \mathrm{inv}(w)$. On the other hand, consider the following particular sequence of sorting moves starting from w. If the current permutation w^* is not the identity, there exists a smallest index i with $w_i^* > w_{i+1}^*$. Apply the basic transposition $(i, i+1)$, which reduces $\mathrm{inv}(w^*)$ by 1, and continue. This sorting method will end in exactly $\mathrm{inv}(w)$ steps, since id is the unique permutation with zero inversions. This proves it is possible to sort w in $\mathrm{inv}(w)$ steps, so that $m \leq \mathrm{inv}(w)$.

To prove the last part of the theorem, recall that the sorting process just described can be implemented by right-multiplying by suitable basic transpositions. We therefore have an equation in S_n of the form

$$w \circ (i_1, i_1+1) \circ (i_2, i_2+1) \circ \cdots \circ (i_m, i_m+1) = \mathrm{id}\,.$$

Solving for w, and using the fact that $(i,j)^{-1} = (j,i) = (i,j)$, we get

$$w = (i_m, i_m+1) \circ \cdots \circ (i_2, i_2+1) \circ (i_1, i_1+1),$$

which expresses w as a product of m basic transpositions. □

9.30. Example. Let us apply the sorting algorithm in the preceding proof to write $w = 42531$ as a product of $\mathrm{inv}(w) = 7$ basic transpositions. Since $4 > 2$, we first multiply w on the right by $(1,2)$ to obtain

$$w \circ (1,2) = 24531.$$

(Observe that $\mathrm{inv}(24531) = 6 = \mathrm{inv}(w) - 1$.) Next, since $5 > 3$, we multiply on the right by $(3,4)$ to get

$$w \circ (1,2) \circ (3,4) = 24351.$$

The computation continues as follows:

$$\begin{aligned}
w \circ (1,2) \circ (3,4) \circ (2,3) &= 23451;\\
w \circ (1,2) \circ (3,4) \circ (2,3) \circ (4,5) &= 23415;\\
w \circ (1,2) \circ (3,4) \circ (2,3) \circ (4,5) \circ (3,4) &= 23145;\\
w \circ (1,2) \circ (3,4) \circ (2,3) \circ (4,5) \circ (3,4) \circ (2,3) &= 21345;\\
w \circ (1,2) \circ (3,4) \circ (2,3) \circ (4,5) \circ (3,4) \circ (2,3) \circ (1,2) &= 12345 = \mathrm{id}\,.
\end{aligned}$$

We now solve for w, which has the effect of reversing the order of the basic transpositions we used to reach the identity:

$$w = (1,2) \circ (2,3) \circ (3,4) \circ (4,5) \circ (2,3) \circ (3,4) \circ (1,2).$$

It is also possible to find such a factorization by starting with the identity word and "unsorting" to reach w. Here it will not be necessary to reverse the order of the transpositions

at the end. We illustrate this idea with the following computation:

$$\begin{aligned}
\mathrm{id} &= 12345;\\
\mathrm{id}\circ(3,4) &= 12435;\\
\mathrm{id}\circ(3,4)\circ(2,3) &= 14235;\\
\mathrm{id}\circ(3,4)\circ(2,3)\circ(1,2) &= 41235;\\
\mathrm{id}\circ(3,4)\circ(2,3)\circ(1,2)\circ(2,3) &= 42135;\\
\mathrm{id}\circ(3,4)\circ(2,3)\circ(1,2)\circ(2,3)\circ(4,5) &= 42153;\\
\mathrm{id}\circ(3,4)\circ(2,3)\circ(1,2)\circ(2,3)\circ(4,5)\circ(3,4) &= 42513;\\
\mathrm{id}\circ(3,4)\circ(2,3)\circ(1,2)\circ(2,3)\circ(4,5)\circ(3,4)\circ(4,5) &= 42531 = w.
\end{aligned}$$

So $w = (3,4)\circ(2,3)\circ(1,2)\circ(2,3)\circ(4,5)\circ(3,4)\circ(4,5)$. Observe that this is a different factorization of w from the one obtained earlier, although both involve seven basic transpositions. This shows that *factorizations of permutations into products of basic transpositions are not unique*. It is also possible to find factorizations involving more than seven factors, by interchanging two entries that are already in the correct order during the sorting of w into id. So the number of factors in such factorizations is not unique either; but we will see shortly that the *parity* of the number of factors (odd or even) *is* uniquely determined by w. In fact, the parity is odd when $\mathrm{sgn}(w) = -1$ and even when $\mathrm{sgn}(w) = +1$.

We now have enough machinery to prove the fundamental properties of sgn.

9.31. Theorem: Properties of Sign. (a) For all $f, g \in S_n$, $\mathrm{sgn}(f\circ g) = \mathrm{sgn}(f)\cdot\mathrm{sgn}(g)$. (b) For all $f \in S_n$, $\mathrm{sgn}(f^{-1}) = \mathrm{sgn}(f)$.

Proof. (a) If $g = \mathrm{id}$, then the result is true since $f\circ g = f$ and $\mathrm{sgn}(g) = 1$ in this case. If $t = (i, i+1)$ is a basic transposition, then 9.26 shows that $\mathrm{sgn}(f\circ t) = -\mathrm{sgn}(f)$. Given a non-identity permutation g, use 9.29 to write g as a nonempty product of basic transpositions, say $g = t_1\circ t_2\circ\cdots t_k$. Then, for every $f \in S_n$, iteration of 9.26 gives

$$\begin{aligned}
\mathrm{sgn}(f\circ g) &= \mathrm{sgn}(ft_1\cdots t_{k-1}t_k) = -\mathrm{sgn}(ft_1\cdots t_{k-1})\\
&= (-1)^2\,\mathrm{sgn}(ft_1\cdots t_{k-2}) = \cdots = (-1)^k\,\mathrm{sgn}(f).
\end{aligned}$$

In particular, this equation is true when $f = \mathrm{id}$; in that case, we obtain $\mathrm{sgn}(g) = (-1)^k$. Using this fact in the preceding equation produces $\mathrm{sgn}(f\circ g) = \mathrm{sgn}(g)\,\mathrm{sgn}(f) = \mathrm{sgn}(f)\,\mathrm{sgn}(g)$ for all $f \in S_n$.

(b) By part (a), $\mathrm{sgn}(f)\cdot\mathrm{sgn}(f^{-1}) = \mathrm{sgn}(f\circ f^{-1}) = \mathrm{sgn}(\mathrm{id}) = +1$. If $\mathrm{sgn}(f) = +1$, it follows that $\mathrm{sgn}(f^{-1}) = +1$. If instead $\mathrm{sgn}(f) = -1$, then it follows that $\mathrm{sgn}(f^{-1}) = -1$. □

Iteration of 9.31 shows that

$$\mathrm{sgn}(f_1\circ\cdots\circ f_k) = \prod_{i=1}^{k}\mathrm{sgn}(f_i). \tag{9.2}$$

9.32. Theorem: Factorizations into Transpositions. Let $f = t_1\circ t_2\circ\cdots\circ t_k$ be *any* factorization of $f \in S_n$ into a product of transpositions (not necessarily basic ones). Then $\mathrm{sgn}(f) = (-1)^k$. In particular, the parity of k (odd or even) is uniquely determined by f.

Proof. By 9.28, $\mathrm{sgn}(t_i) = -1$ for all i. The conclusion now follows by setting $f_i = t_i$ in (9.2). □

9.33. Theorem: Sign of a k-cycle. The sign of any k-cycle $(i_1, i_2, \ldots, i_k)$ is $(-1)^{k-1}$.

Proof. The result is already known for $k = 1$ and $k = 2$. For $k > 2$, one may check that the given k-cycle can be written as the following product of $k - 1$ transpositions:

$$(i_1, i_2, \ldots, i_k) = (i_1, i_2) \circ (i_2, i_3) \circ (i_3, i_4) \circ \cdots (i_{k-1}, i_k).$$

So the result follows from 9.32. □

We can now show that the sign of a permutation f is completely determined by type(f).

9.34. Theorem: Cycle Type and Sign. Suppose $f \in S_n$ has type$(f) = \mu$. Then

$$\operatorname{sgn}(f) = \prod_{i=1}^{\ell(\mu)} (-1)^{\mu_i - 1} = (-1)^{n - \ell(\mu)} = (-1)^{n - \operatorname{cyc}(f)}.$$

Proof. Let the cycle decomposition of f be $f = C_1 \circ \cdots \circ C_{\ell(\mu)}$, where C_i is a μ_i-cycle. The result follows from the relations $\operatorname{sgn}(f) = \prod_{i=1}^{\ell(\mu)} \operatorname{sgn}(C_i)$ and $\operatorname{sgn}(C_i) = (-1)^{\mu_i - 1}$. □

9.35. Example. The permutation $f = (4, 6, 2, 8)(3, 9, 1)(5, 10, 7)$ in S_{10} has $\operatorname{sgn}(f) = (-1)^{10-3} = -1$.

9.5 Determinants

In the next three sections, we interrupt our exposition of group theory to give an application of the preceding material to determinants. We will see that the combinatorial properties of permutations underlie many commonly used facts about determinants.

9.36. Definition: Matrix Rings. For every commutative ring R and positive integer n, let $M_n(R)$ be the set of $n \times n$ matrices with entries in R. Formally, an element of $M_n(R)$ is a function $A : \{1, 2, \ldots, n\} \times \{1, 2, \ldots, n\} \to R$. $A(i, j)$ is called the *i, j-entry of A*. We often display A as a square array in which $A(i, j)$ appears in row i and column j. For $A, B \in M_n(R)$ and $c \in R$, define $A + B$, AB, and cA by setting

$$\begin{aligned} (A + B)(i, j) &= A(i, j) + B(i, j); \\ (AB)(i, j) &= \sum_{k=1}^{n} A(i, k) B(k, j); \\ (cA)(i, j) &= c(A(i, j)) \qquad (1 \leq i, j \leq n). \end{aligned}$$

Routine verifications (see 2.151) show that $M_n(R)$ with these operations is a ring, whose multiplicative identity element I_n is given by $I_n(i, j) = 1_R$ if $i = j$, and $I_n(i, j) = 0_R$ if $i \neq j$. One also checks that $M_n(R)$ is non-commutative if $n > 1$ and $R \neq \{0\}$.

9.37. Definition: Determinants. For a matrix $A \in M_n(R)$, the *determinant* of A is

$$\det(A) = \sum_{w \in S_n} \operatorname{sgn}(w) \prod_{i=1}^{n} A(i, w(i)) \in R.$$

9.38. Example. When $n = 1$, $\det(A) = A(1, 1)$. When $n = 2$, the possible permutations w

(in one-line form) are $w = 12$ with $\operatorname{sgn}(w) = +1$, and $w = 21$ with $\operatorname{sgn}(w) = -1$. Therefore, the definition gives

$$\det(A) = \det \begin{bmatrix} A(1,1) & A(1,2) \\ A(2,1) & A(2,2) \end{bmatrix} = +A(1,1)A(2,2) - A(1,2)A(2,1).$$

When $n = 3$, the definition and the table in 9.23 lead to the formula

$$\begin{aligned} \det(A) &= \det \begin{bmatrix} A(1,1) & A(1,2) & A(1,3) \\ A(2,1) & A(2,2) & A(2,3) \\ A(3,1) & A(3,2) & A(3,3) \end{bmatrix} \\ &= +A(1,1)A(2,2)A(3,3) - A(1,1)A(2,3)A(3,2) - A(1,2)A(2,1)A(3,3) \\ &\quad +A(1,2)A(2,3)A(3,1) + A(1,3)A(2,1)A(3,2) - A(1,3)A(2,2)A(3,1). \end{aligned}$$

In general, we see that $\det(A)$ is a sum of $n!$ signed terms. A given term arises by choosing one factor $A(i, w(i))$ from each row of A; since w is a permutation, each of the chosen factors must come from a different column of A. The term in question is the product of the n chosen factors, times $\operatorname{sgn}(w)$. Since $\operatorname{sgn}(w) = (-1)^{\operatorname{inv}(w)}$, the sign attached to this term depends on the parity of the number of basic transpositions needed to sort the column indices $w(1), w(2), \ldots, w(n)$ into increasing order.

The next result shows that we can replace $A(i, w(i))$ by $A(w(i), i)$ in the defining formula for $\det(A)$. This corresponds to interchanging the roles of rows and columns in the description above.

9.39. Definition: Transpose of a Matrix. Given $A \in M_n(R)$, the *transpose* of A is the matrix $A^t \in M_n(R)$ such that $A^t(i,j) = A(j,i)$ for all $i, j \leq n$.

9.40. Theorem: Determinant of a Transpose. For all $A \in M_n(R)$, $\det(A^t) = \det(A)$.

Proof. By definition,

$$\det(A^t) = \sum_{w \in S_n} \operatorname{sgn}(w) \prod_{k=1}^{n} A^t(k, w(k)) = \sum_{w \in S_n} \operatorname{sgn}(w) \prod_{k=1}^{n} A(w(k), k).$$

For a fixed $w \in S_n$, we make a change of variables in the product indexed by w by letting $j = w(k)$, so $k = w^{-1}(j)$. Since w is a permutation and R is commutative, we have

$$\prod_{k=1}^{n} A(w(k), k) = \prod_{j=1}^{n} A(j, w^{-1}(j))$$

because the second product contains the same factors as the first product in a different order (cf. 2.149). We now calculate

$$\det(A^t) = \sum_{w \in S_n} \operatorname{sgn}(w) \prod_{j=1}^{n} A(j, w^{-1}(j)) = \sum_{w \in S_n} \operatorname{sgn}(w^{-1}) \prod_{j=1}^{n} A(j, w^{-1}(j)).$$

Now consider the change of variable $v = w^{-1}$. As w ranges over S_n, so does v, since $w \mapsto w^{-1}$ is a bijection on S_n. Furthermore, we can reorder the terms of the sum since addition in R is commutative (see 2.149). We conclude that

$$\det(A^t) = \sum_{v \in S_n} \operatorname{sgn}(v) \prod_{j=1}^{n} A(j, v(j)) = \det(A). \qquad \square$$

Next we derive a formula for the determinant of an upper-triangular matrix.

9.41. Theorem: Determinant of Triangular and Diagonal Matrices. Suppose $A \in M_n(R)$ satisfies $A(i,j) = 0$ whenever $i > j$. Then $\det(A) = \prod_{i=1}^{n} A(i,i)$. Consequently, if A is either upper-triangular, lower-triangular, or diagonal, then $\det(A)$ is the product of the diagonal entries of A.

Proof. By definition, $\det(A) = \sum_{w \in S_n} \operatorname{sgn}(w) \prod_{i=1}^{n} A(i, w(i))$. In order for a given summand to be nonzero, we must have $i \leq w(i)$ for all $i \leq n$. Since w is a permutation, we successively deduce that $w(n) = n$, $w(n-1) = n-1$, ..., $w(1) = 1$. Thus, the only possibly nonzero summand comes from $w = \mathrm{id}$. Since $\operatorname{sgn}(\mathrm{id}) = +1$ and $\mathrm{id}(i) = i$ for all i, the stated formula for $\det(A)$ follows when A is upper-triangular. The result for lower-triangular A follows by considering A^t. Since diagonal matrices are upper-triangular, the proof is complete. □

9.42. Corollary: Determinant of Identity Matrix. For all $n \in \mathbb{N}^+$, $\det(I_n) = 1_R$.

9.6 Multilinearity and Laplace Expansions

This section continues our development of the properties of determinants.

9.43. Definition: R-Linear Maps. Let R be a commutative ring and $n \in \mathbb{N}^+$. A map $T : R^n \to R$ is called *R-linear* iff $T(v+z) = T(v) + T(z)$ and $T(cv) = cT(v)$ for all $v, z \in R^n$ and all $c \in R$.

9.44. Example. Suppose $b_1, \ldots, b_n \in R$ are fixed constants, and $T : R^n \to R$ is defined by

$$T(v_1, \ldots, v_n) = b_1 v_1 + b_2 v_2 + \cdots + b_n v_n.$$

It is routine to check that the map T is R-linear. Conversely, one can show that every R-linear map from R^n to R must be of this form.

9.45. Theorem: Multilinearity of Determinants. Let $A \in M_n(R)$, and let $k \leq n$ be a fixed row index. For every row vector $v \in R^n$, let $A[v]$ denote the matrix A with row k replaced by v. Then the map $T : R^n \to R$ given by $T(v) = \det(A[v])$ is R-linear. A similar result holds for the columns of A.

Proof. By 9.44, it suffices to show that there exist constants $b_1, \ldots, b_n \in R$ such that for all $v = (v_1, v_2, \ldots, v_n) \in R^n$,

$$T(v) = b_1 v_1 + b_2 v_2 + \cdots + b_n v_n. \tag{9.3}$$

To establish this, consider the defining formula for $\det(A[v])$:

$$T(v) = \det(A[v]) = \sum_{w \in S_n} \operatorname{sgn}(w) \prod_{i=1}^{n} A[v](i, w(i)) = \sum_{w \in S_n} \operatorname{sgn}(w) \left[\prod_{\substack{i=1 \\ i \neq k}}^{n} A(i, w(i)) \right] v_{w(k)}.$$

The terms in brackets depend only on the fixed matrix A, not on v. So (9.3) holds with

$$b_j = \sum_{\substack{w \in S_n \\ w(k) = j}} \operatorname{sgn}(w) \prod_{\substack{i=1 \\ i \neq k}}^{n} A(i, w(i)) \qquad (1 \leq j \leq n). \tag{9.4}$$

To obtain the multilinearity result for the columns of A, apply the result just proved to A^t. □

We sometimes use the following notation when invoking the multilinearity of determinants. For $A \in M_n(R)$, let $A_1, A_2, \ldots, A_n$ denote the n rows of A; thus each A_i lies in R^n. We write $\det(A) = \det(A_1, \ldots, A_n)$, viewing the determinant as a function of n arguments (row vectors). The previous result says that if we fix any $n-1$ of these arguments and let the other one vary, the resulting map $v \mapsto \det(A_1, \ldots, v, \ldots, A_n)$ (for $v \in R^n$) is R-linear.

9.46. Theorem: Alternating Property of Determinants. If $A \in M_n(R)$ has two equal rows or two equal columns, then $\det(A) = 0$.

Proof. Recall $\det(A)$ is a sum of $n!$ signed terms of the form $T(w) = \operatorname{sgn}(w) \prod_{i=1}^{n} A(i, w(i))$, where w ranges over S_n. Suppose rows r and s of A are equal, so $A(r,k) = A(s,k)$ for all k. We will define a involution I on S_n with no fixed points such that $T(I(w)) = -T(w)$ for all $w \in S_n$. It will follow that the $n!$ terms cancel in pairs, so that $\det(A) = 0$. Define $I(w) = w \circ (r,s)$ for $w \in S_n$; evidently $I \circ I = \mathrm{id}_{S_n}$ and I has no fixed points. On one hand, $\operatorname{sgn}(I(w)) = \operatorname{sgn}(w) \cdot \operatorname{sgn}((r,s)) = -\operatorname{sgn}(w)$ by 9.31. On the other hand,

$$\begin{aligned}\prod_{i=1}^{n} A(i,[w\circ(r,s)](i)) &= A(r,w(s))A(s,w(r)) \prod_{i\neq r,s} A(i,w(i)) \\ &= A(r,w(r))A(s,w(s)) \prod_{i\neq r,s} A(i,w(i)) = \prod_{i=1}^{n} A(i,w(i)).\end{aligned}$$

Combining these facts, we see that $T(I(w)) = -T(w)$, as desired. If A has two equal columns, then A^t has two equal rows, so $\det(A) = \det(A^t) = 0$. □

9.47. Theorem: Effect of Elementary Row Operations on Determinants. Let $A \in M_n(R)$, let j, k be distinct indices, and let $c \in R$.
(a) If B is obtained from A by multiplying row j by c, then $\det(B) = c\det(A)$.
(b) If B is obtained from A by interchanging rows j and k, then $\det(B) = -\det(A)$.
(c) If B is obtained from A by adding c times row j to row k, then $\det(B) = \det(A)$.
Analogous results hold for elementary column operations.

Proof. Part (a) is a special case of the multilinearity of determinants (see 9.45). Part (b) is a consequence of multilinearity and the alternating property. Specifically, define $T : R^n \times R^n \to R$ by letting $T(v,w) = \det(A_1, \ldots, v, \ldots, w, \ldots, A_n)$ (where the v and w occur in positions j and k). Since det is multilinear and alternating, we get

$$0 = T(v+w, v+w) = T(v,v)+T(w,v)+T(v,w)+T(w,w) = T(w,v)+T(v,w) \quad (v,w \in R^n).$$

Thus, $T(w,v) = -T(v,w)$ for all v, w, which translates to statement (b) after taking $v = A_j$ and $w = A_k$. Part (c) follows for similar reasons, since

$$T(v, cv+w) = cT(v,v) + T(v,w) = T(v,w).$$ □

9.48. Theorem: Laplace Expansions of Determinants. For $A \in M_n(R)$ and $i, j \leq n$, let $A[i|j]$ be the matrix in $M_{n-1}(R)$ obtained by deleting row i and column j of A. For $1 \leq k \leq n$, we have

$$\begin{aligned}\det(A) &= \sum_{i=1}^{n} (-1)^{i+k} A(i,k) \det(A[i|k]) \quad \text{(expansion along column } k) \\ &= \sum_{j=1}^{n} (-1)^{j+k} A(k,j) \det(A[k|j]) \quad \text{(expansion along row } k).\end{aligned}$$

Proof. Let us first prove the Laplace expansion formula along row $k = n$. By the proof of multilinearity (see equations (9.3) and (9.4)), we know that

$$\det(A) = b_1 A(n,1) + b_2 A(n,2) + \cdots + b_n A(n,n)$$

where

$$b_j = \sum_{\substack{w \in S_n \\ w(n)=j}} \operatorname{sgn}(w) \prod_{i=1}^{n-1} A(i, w(i)) \quad (1 \leq j \leq n).$$

Comparing to the desired formula, we need only show that $b_j = (-1)^{j+n} \det(A[n|j])$ for all j.

Fix an index j. Let $S_{n,j} = \{w \in S_n : w(n) = j\}$. We define a bijection $f : S_{n,j} \to S_{n-1}$ as follows. Every $w \in S_{n,j}$ can be written in one-line form as $w = w_1 w_2 \cdots w_{n-1} w_n$ where $w_n = j$. Define $f(w) = w'_1 w'_2 \cdots w'_{n-1}$ where $w'_t = w_t$ if $w_t < j$, and $w'_t = w_t - 1$ if $w_t > j$. In other words, we drop the j at the end of w and decrement all letters larger than j. The inverse map increments all letters $\geq j$ and then adds a j at the end. Observe that the deletion of j decreases $\operatorname{inv}(w)$ by $n - j$ (the number of letters to the left of j that are greater than j), and the decrementing operation has no further effect on the inversion count. So, $\operatorname{inv}(f(w)) = \operatorname{inv}(w) - (n - j)$ and $\operatorname{sgn}(f(w)) = (-1)^{j+n} \operatorname{sgn}(w)$. We also note that for $w' = f(w)$, we have $A(i, w(i)) = A[n|j](i, w'(i))$ for all $i < n$, since all columns in A after column j get shifted one column left when column j is deleted. Now use the bijection f to change the summation variable in the formula for b_j. Writing $w' = f(w)$, we obtain

$$\begin{aligned} b_j &= \sum_{w \in S_{n,j}} \operatorname{sgn}(w) \prod_{i=1}^{n-1} A(i, w(i)) \\ &= \sum_{w' \in S_{n-1}} (-1)^{j+n} \operatorname{sgn}(w') \prod_{i=1}^{n-1} A[n|j](i, w'(i)) = (-1)^{j+n} \det(A[n|j]). \end{aligned}$$

The Laplace expansion along an arbitrary row k follows from the special case $k = n$. Given k, let B be the matrix obtained from A by successively interchanging row k with row $k+1$, $k+2$, …, n. These $n - k$ row interchanges multiply the determinant by $(-1)^{n-k}$. It is evident that $B(n, j) = A(k, j)$ and $B[n|j] = A[k|j]$ for all j. So

$$\begin{aligned} \det(A) &= (-1)^{n-k} \det(B) = (-1)^{k-n} \sum_{j=1}^{n} (-1)^{j+n} B(n,j) \det(B[n|j]) \\ &= \sum_{j=1}^{n} (-1)^{j+k} A(k,j) \det(A[k|j]). \end{aligned}$$

Finally, to derive the Laplace expansion along column k, pass to transposes:

$$\begin{aligned} \det(A) &= \det(A^t) = \sum_{j=1}^{n} (-1)^{j+k} A^t(k,j) \det(A^t[k|j]) \\ &= \sum_{j=1}^{n} (-1)^{j+k} A(j,k) \det(A[j|k]). \quad \square \end{aligned}$$

We now use Laplace expansions to derive the classical formula for the inverse of a matrix.

9.49. Definition: Classical Adjoint of a Matrix. Given $A \in M_n(R)$, let $\operatorname{adj} A \in M_n(R)$ be the matrix with i,j-entry $(-1)^{i+j} \det(A[j|i])$ for $i, j \leq n$.

The next result explains why we wrote $A[j|i]$ instead of $A[i|j]$ in the preceding definition.

9.50. Theorem: Adjoint Formula. For all $A \in M_n(R)$, we have

$$A(\operatorname{adj} A) = (\det(A))I_n = (\operatorname{adj} A)A.$$

Proof. For $1 \leq i \leq n$, the i,i-entry of the product $A(\operatorname{adj} A)$ is

$$\sum_{k=1}^{n} A(i,k)[\operatorname{adj} A](k,i) = \sum_{k=1}^{n} (-1)^{i+k} A(i,k) \det(A[i|k]) = \det(A),$$

by Laplace expansion along row i of A. Now suppose $i \neq j$. The i,j-entry of $A(\operatorname{adj} A)$ is

$$\sum_{k=1}^{n} A(i,k)[\operatorname{adj} A](k,j) = \sum_{k=1}^{n} (-1)^{j+k} A(i,k) \det(A[j|k]).$$

Let C be the matrix obtained from A by replacing row j of A by row i of A. Then $C(j,k) = A(i,k)$ and $C[j|k] = A[j|k]$ for all k. So the preceding expression is the Laplace expansion for $\det(C)$ along row j. On the other hand, $\det(C) = 0$ because C has two equal rows. So $[A(\operatorname{adj} A)](i,j) = 0$. We have proved that $A(\operatorname{adj} A)$ is a diagonal matrix with all diagonal entries equal to $\det(A)$, as desired. The analogous result for $(\operatorname{adj} A)A$ is proved similarly, using column expansions. □

9.51. Corollary: Formula for the Inverse of a Matrix. If $A \in M_n(R)$ and $\det(A)$ is an invertible element of R, then the matrix A is invertible in $M_n(R)$ with inverse

$$A^{-1} = \frac{1}{\det(A)} \operatorname{adj} A.$$

9.52. Remark. Conversely, if A is invertible in $M_n(R)$, then $\det(A)$ is an invertible element of R. The proof uses the following *product formula for determinants*:

$$\det(AB) = \det(A)\det(B) = \det(B)\det(A) \qquad (A, B \in M_n(R)).$$

Taking $B = A^{-1}$, the left side becomes $\det(I_n) = 1_R$, so $\det(B)$ is a two-sided inverse of $\det(A)$ in R. We will deduce the product formula as a consequence of the Cauchy-Binet formula, which is proved in the next section.

9.7 Cauchy-Binet Formula

This section discusses the Cauchy-Binet formula, which expresses the determinant of a product of rectangular matrices as the sum of a product of determinants of suitable submatrices. The proof of this formula is a nice application of the properties of inversions and determinants.

To state the Cauchy-Binet formula, we need the following notation. Given a $c \times d$ matrix M, write M_i for the ith row of M and M^j for the jth column of M. Given indices $j_1, \ldots, j_c \in \{1, 2, \ldots, d\}$, let $(M^{j_1}, \ldots, M^{j_c})$ denote the $c \times c$ matrix whose columns are $M^{j_1}, \ldots, M^{j_c}$ in this order. Similarly, $(M_{i_1}, \ldots, M_{i_d})$ is the matrix whose rows are $M_{i_1}, \ldots, M_{i_d}$ in this order.

9.53. Theorem: Cauchy-Binet Formula. Suppose $m \leq n$, A is an $m \times n$ matrix, and B is an $n \times m$ matrix. Let J be the set of all lists $j = (j_1, j_2, \ldots, j_m)$ such that $1 \leq j_1 < j_2 < \cdots < j_m \leq n$. Then

$$\det(AB) = \sum_{j \in J} \det(A^{j_1}, A^{j_2}, \ldots, A^{j_m}) \det(B_{j_1}, B_{j_2}, \ldots, B_{j_m}).$$

Proof. Note that all matrices appearing in the formula are $m \times m$, so all the determinants are defined. We begin by using the definitions of matrix products and determinants (§9.5) to write

$$\det(AB) = \sum_{w \in S_m} \operatorname{sgn}(w) \prod_{i=1}^{m} (AB)(i, w(i)) = \sum_{w \in S_m} \operatorname{sgn}(w) \prod_{i=1}^{m} \sum_{k_i=1}^{n} A(i, k_i) B(k_i, w(i)).$$

The generalized distributive law (§2.1) changes the product of sums into a sum of products:

$$\det(AB) = \sum_{w \in S_m} \sum_{k_1=1}^{n} \cdots \sum_{k_m=1}^{n} \operatorname{sgn}(w) \prod_{i=1}^{m} A(i, k_i) \prod_{i=1}^{m} B(k_i, w(i)).$$

Let K be the set of all lists $k = (k_1, \ldots, k_m)$ with every $k_i \in \{1, 2, \ldots, n\}$, and let K' be the set of lists in K whose entries k_i are distinct. We can combine the m separate sums over the k_i's into a single sum over lists $k \in K$. We can also reorder the summation to get

$$\det(AB) = \sum_{k \in K} \sum_{w \in S_m} \operatorname{sgn}(w) \prod_{i=1}^{m} A(i, k_i) \prod_{i=1}^{m} B(k_i, w(i)).$$

Next, factor out quantities that do not depend on w:

$$\det(AB) = \sum_{k \in K} \prod_{i=1}^{m} A(i, k_i) \left[\sum_{w \in S_m} \operatorname{sgn}(w) \prod_{i=1}^{m} B(k_i, w(i)) \right].$$

The term in brackets is the defining formula for $\det(B_{k_1}, \ldots, B_{k_m})$. If any two entries in $(k_1, \ldots, k_m)$ are equal, this matrix has two equal rows, so its determinant is zero. Discarding these terms, we are reduced to summing over lists $k \in K'$. So now we have

$$\det(AB) = \sum_{k \in K'} \prod_{i=1}^{m} A(i, k_i) \det(B_{k_1}, \ldots, B_{k_m}).$$

To continue, observe that for every list $k \in K'$ there exists a unique sorted list $j \in J$ with $j = \operatorname{sort}(k)$. Grouping summands gives

$$\det(AB) = \sum_{j \in J} \sum_{\substack{k \in K' \\ \operatorname{sort}(k) = j}} \prod_{i=1}^{m} A(i, k_i) \det(B_{k_1}, \ldots, B_{k_m}).$$

Given that $\operatorname{sort}(k) = j$, we can change the matrix $(B_{k_1}, \ldots, B_{k_m})$ into the matrix $(B_{j_1}, \ldots, B_{j_m})$ by repeatedly switching adjacent rows. Each such switch flips the sign of the determinant, and the number of row switches required is readily seen to be $\operatorname{inv}(k_1 k_2 \cdots k_m)$. (To see this, adapt the proof of 9.29 to the case where the objects being sorted are $\{j_1 < j_2 < \cdots < j_m\}$ instead of $\{1 < 2 < \cdots < m\}$; cf. 9.179.) Letting $\operatorname{sgn}(k) = (-1)^{\operatorname{inv}(k)}$, we can therefore write

$$\det(AB) = \sum_{j \in J} \sum_{\substack{k \in K' \\ \operatorname{sort}(k) = j}} \operatorname{sgn}(k) \prod_{i=1}^{m} A(i, k_i) \det(B_{j_1}, \ldots, B_{j_m}).$$

The determinant in this formula depends only on j, not on k, so it can be brought out of the inner summation:

$$\det(AB) = \sum_{j \in J} \det(B_{j_1}, \ldots, B_{j_m}) \sum_{\substack{k \in K' \\ \mathrm{sort}(k)=j}} \mathrm{sgn}(k) \prod_{i=1}^{m} A(i, k_i).$$

To finish, note that every $k \in K'$ that sorts to j can be written as $(k_1, \ldots, k_m) = (j_{v(1)}, \ldots, j_{v(m)})$ for a uniquely determined permutation $v \in S_m$. Since j is an increasing sequence, it follows that $\mathrm{inv}(k) = \mathrm{inv}(v)$ and $\mathrm{sgn}(k) = \mathrm{sgn}(v)$. Changing variables in the inner summation, we get

$$\det(AB) = \sum_{j \in J} \det(B_{j_1}, \ldots, B_{j_m}) \left[\sum_{v \in S_m} \mathrm{sgn}(v) \prod_{i=1}^{m} A(i, j_{v(i)}) \right].$$

The term in brackets is none other than $\det(A^{j_1}, \ldots, A^{j_m})$, so the proof is complete. □

9.54. Theorem: Product Formula for Determinants. If A and B are $m \times m$ matrices, then $\det(AB) = \det(A)\det(B)$.

Proof. Take $n = m$ in the Cauchy-Binet formula. The index set J consists of the single list $(1, 2, \ldots, m)$, and the summand corresponding to this list reduces to $\det(A)\det(B)$. □

Other examples of combinatorial proofs of determinant formulas appear in §11.14 and §12.9.

9.8 Subgroups

Suppose $(G, \star)$ is a group, and H is a subset of G. One might hope that $(H, \star')$ is also a group, where $\star'$ is the restriction of $\star$ to $H \times H$. This is not true in general, but it will be true if H is a *subgroup*.

9.55. Definition: Subgroups. Let $(G, \star)$ be a group and let H be a subset of G. H is called a *subgroup* of G, written $H \leq G$, iff the following three "closure conditions" are satisfied:

$e_G \in H$	(closure under identity);
$\forall a, b \in H,\ a \star b \in H$	(closure under the operation);
$\forall a \in H,\ a^{-1} \in H$	(closure under inverses).

A subgroup H is called *normal in* G, written $H \trianglelefteq G$, iff

$$\forall g \in G, \forall h \in H, ghg^{-1} \in H \qquad \text{(closure under conjugation)}.$$

Let us verify that $(H, \star')$ is indeed a group when $H \leq G$. Since H is closed under the operation, $\star'$ does map $H \times H$ *into* H (not just into G), so the closure axiom holds for $(H, \star')$. Since H is a subset of G, associativity holds in H because it is known to hold in G. The identity e of G lies in H by assumption. Since $e \star h = h = h \star e$ holds for all $h \in G$, the relation $e \star' h = h = h \star' e$ certainly holds for all $h \in H \subseteq G$. Finally, every element x of H has an inverse y (relative to $\star$) that lies *in* H, by assumption. Now y is still an inverse of x relative to $\star'$, so the proof is complete. One also sees that H is commutative if G is commutative, but the converse statement is not always true. Usually, we use the same symbol $\star$ (instead of $\star'$) to denote the operation in the subgroup H.

9.56. Example. We have the following chain of subgroups of the additive group $(\mathbb{C}, +)$:

$$\{0\} \leq \mathbb{Z} \leq \mathbb{Q} \leq \mathbb{R} \leq \mathbb{C}.$$

Similarly, $\{-1, 1\}$ and $\mathbb{Q}^+$ are both subgroups of $(\mathbb{Q} \sim \{0\}, \times)$. The set $\{0, 3, 6, 9\}$ is a subgroup of $(\mathbb{Z}_{12}, \oplus)$; one can prove closure under addition and inverses by a finite case analysis, or by inspection of the relevant portion of the addition table for $\mathbb{Z}_{12}$.

9.57. Example. The sets $H = \{(1)(2)(3), (1,2,3), (1,3,2)\}$ and $K = \{(1)(2)(3), (1,3)\}$ are subgroups of S_3, as one readily verifies. Moreover, H is normal in S_3, but K is not. The set $J = \{(1)(2)(3), (1,3), (2,3), (1,3)\}$ is not a subgroup of S_3, since closure under the operation fails: $(1,3) \circ (2,3) = (1,3,2) \notin J$. Here is a four-element normal subgroup of S_4:

$$V = \{(1)(2)(3)(4),\ (1,2)(3,4),\ (1,3)(2,4),\ (1,4)(2,3)\}.$$

Each element of V is its own inverse, and one confirms closure of V under the operation by checking all possible products. To prove the normality of V in S_4, it is helpful to use 9.131 below.

9.58. Example. The set of *even integers* is a subgroup of $(\mathbb{Z}, +)$. For, the identity element zero is even; the sum of two even integers is again even; and x even implies $-x$ is even. More generally, let k be any fixed integer, and let $H = \{kn : n \in \mathbb{Z}\}$ consist of all integer multiples of k. A routine verification shows that H is a subgroup of $(\mathbb{Z}, +)$. We write $H = k\mathbb{Z}$ for brevity. The next theorem shows that we have found *all* the subgroups of the additive group $\mathbb{Z}$.

9.59. Theorem: Subgroups of $\mathbb{Z}$. Every subgroup H of $(\mathbb{Z}, +)$ has the form $k\mathbb{Z}$ for a unique integer $k \geq 0$.

Proof. We have noted that all the subsets $k\mathbb{Z}$ are indeed subgroups. Given an arbitrary subgroup H, consider two cases. If $H = \{0\}$, then $H = 0\mathbb{Z}$. Otherwise, H contains at least one nonzero integer m. If m is negative, then $-m \in H$ since H is closed under inverses. So, H contains strictly positive integers. Take k to be the least positive integer in H. We claim that $H = k\mathbb{Z}$. Let us prove that $kn \in H$ for all $n \in \mathbb{Z}$, so that $k\mathbb{Z} \subseteq H$. For $n \geq 0$, we argue by induction on n. When $n = 0$, we must prove $k0 = 0 \in H$, which holds since H contains the identity of $\mathbb{Z}$. When $n = 1$, we must prove $k1 = k \in H$, which is true by choice of k. Assume $n \geq 1$ and $kn \in H$. Then $k(n+1) = kn + k \in H$ since $kn \in H$, $k \in H$, and H is closed under addition. Finally, for negative n, write $n = -m$ and note that $kn = -(km) \in H$ since $km \in H$ and H is closed under inverses.

The key step is to prove the reverse inclusion $H \subseteq k\mathbb{Z}$. Fix $z \in H$. Dividing z by k, we obtain $z = kq + r$ for some integers q, r with $0 \leq r < k$. By what we proved in the last paragraph, $k(-q) \in H$. So, $r = z - kq = z + k(-q) \in H$ since H is closed under addition. Now, since k is the *least* positive integer in H, we cannot have $0 < r < k$. The only possibility left is $r = 0$, so $z = kq \in k\mathbb{Z}$, as desired.

Finally, to prove uniqueness, suppose $k\mathbb{Z} = m\mathbb{Z}$ for $k, m \geq 0$. Note $k = 0$ iff $m = 0$, so assume $k, m > 0$. Since $k \in k\mathbb{Z} = m\mathbb{Z}$, k is a multiple of m. Similarly, m is a multiple of k. As both k and m are positive, this forces $k = m$, completing the proof. □

How can we find subgroups of a given group G? As we see next, each element $x \in G$ gives rise to a subgroup of G in a natural way.

9.60. Definition: Cyclic Subgroups and Cyclic Groups. Let G be a group written multiplicatively, and let $x \in G$. The *cyclic subgroup of G generated by x* is $\langle x \rangle = \{x^n : n \in \mathbb{Z}\}$. One sees, using the laws of exponents, that this subset of G really is a subgroup. G is called a *cyclic* group iff there exists $x \in G$ with $G = \langle x \rangle$. When G is written additively, we have $\langle x \rangle = \{nx : n \in \mathbb{Z}\}$.

9.61. Example. The group $(\mathbb{Z}, +)$ is cyclic, since $\mathbb{Z} = \langle 1 \rangle = \langle -1 \rangle$. The subgroups $k\mathbb{Z} = \langle k \rangle$ considered above are cyclic subgroups of $\mathbb{Z}$. Our last theorem implies that *every subgroup of* $\mathbb{Z}$ *is cyclic*. The groups $(\mathbb{Z}_n, \oplus)$ are also cyclic; each of these groups is generated by 1. The group $(\{a, b, c, d\}, \star)$ discussed at the end of 9.7 is *not* cyclic. To prove this, we compute all the cyclic subgroups of this group:

$$\langle a \rangle = \{a\}, \quad \langle b \rangle = \{a, b\}, \quad \langle c \rangle = \{a, c\}, \quad \langle d \rangle = \{a, d\}.$$

None of the cyclic subgroups equals the whole group, so the group is not cyclic. For a bigger example of a non-cyclic group, consider $(\mathbb{Q}, +)$. Any nonzero cyclic subgroup has the form $\langle a/b \rangle$ for some positive rational number a/b. One may check that $a/2b$ does not lie in this subgroup, so $\mathbb{Q} \neq \langle a/b \rangle$. Noncommutative groups furnish additional examples of non-cyclic groups, as the next result shows.

9.62. Theorem: Cyclic Groups are Commutative.

Proof. Let $G = \langle x \rangle$ be cyclic. Given $y, z \in G$, we can write $y = x^n$ and $z = x^m$ for some $n, m \in \mathbb{Z}$. Since integer addition is commutative, the laws of exponents give

$$yz = x^n x^m = x^{n+m} = x^{m+n} = x^m x^n = zy. \qquad \square$$

By adapting the argument in 9.59, one can show that *every subgroup of a cyclic group is cyclic*; we leave this as an exercise for the reader.

9.63. Example. The cyclic group $\mathbb{Z}_6$ has the following cyclic subgroups (which are *all* the subgroups of this group):

$$\langle 0 \rangle = \{0\}; \quad \langle 1 \rangle = \{0, 1, 2, 3, 4, 5\} = \langle 5 \rangle; \quad \langle 2 \rangle = \{0, 2, 4\} = \langle 4 \rangle; \quad \langle 3 \rangle = \{0, 3\}.$$

In the group S_4, we have

$$\langle (1,3,4,2) \rangle = \{(1,3,4,2),\ (1,4)(3,2),\ (1,2,4,3),\ (1)(2)(3)(4)\}.$$

9.9 Automorphism Groups of Graphs

This section uses graphs to construct examples of subgroups of symmetric groups. These subgroups will be used later when we discuss applications of group theory to counting problems.

9.64. Definition: Automorphism Group of a Graph. Let K be a simple graph with vertex set X and edge set E. A *graph automorphism* of K is a bijection $f : X \to X$ such that, for all $u \neq v$ in X, $\{u, v\} \in E$ iff $\{f(u), f(v)\} \in E$. Let $\mathrm{Aut}(K)$ denote the set of all graph automorphisms of K. Analogous definitions are made for directed simple graphs; here, the requirement on f is that $(u, v) \in E$ iff $(f(u), f(v)) \in E$ for all $u, v \in X$.

One verifies immediately from the definition that $\mathrm{Aut}(K) \leq (\mathrm{Sym}(X), \circ)$. Thus, automorphism groups of graphs are subgroups of symmetric groups.

9.65. Example. Consider the graphs shown in Figure 9.2. The undirected cycle C_5 has exactly ten automorphisms. They are given in one-line form in the following list:

$$\begin{array}{llllll} [1\ 2\ 3\ 4\ 5], & [2\ 3\ 4\ 5\ 1], & [3\ 4\ 5\ 1\ 2], & [4\ 5\ 1\ 2\ 3], & [5\ 1\ 2\ 3\ 4], \\ [5\ 4\ 3\ 2\ 1], & [4\ 3\ 2\ 1\ 5], & [3\ 2\ 1\ 5\ 4], & [2\ 1\ 5\ 4\ 3], & [1\ 5\ 4\ 3\ 2]. \end{array}$$

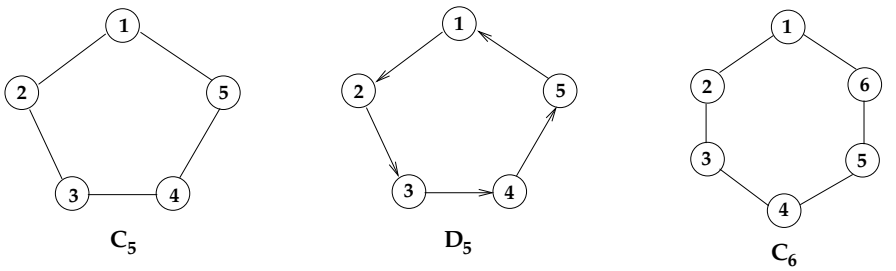

FIGURE 9.2
Graphs used to illustrate automorphism groups.

The same automorphisms, written in cycle notation, look like this:

$$(1)(2)(3)(4)(5), \quad (1,2,3,4,5), \quad (1,3,5,2,4), \quad (1,4,2,5,3), \quad (1,5,4,3,2),$$
$$(1,5)(2,4)(3), \quad (1,4)(2,3)(5), \quad (1,3)(4,5)(2), \quad (1,2)(3,5)(4), \quad (2,5)(3,4)(1).$$

Geometrically, we can think of C_5 as a *necklace* with five beads. The first five automorphisms on each list arise by rotating the necklace through various angles (rotation by zero is the identity map). The next five automorphisms arise by reflecting the necklace in five possible axes of symmetry.

Now consider the automorphism group of the *directed* cycle D_5. Every automorphism of the directed graph D_5 is automatically an automorphism of the associated undirected graph C_5, so $\mathrm{Aut}(D_5) \leq \mathrm{Aut}(C_5)$. However, not every automorphism of C_5 is an automorphism of D_5. In this example, the five "rotations" preserve the direction of the edges, hence are automorphisms of D_5. But the five "reflections" reverse the direction of the edges, so these are not elements of $\mathrm{Aut}(D_5)$. We can write $\mathrm{Aut}(D_5) = \langle (1,2,3,4,5) \rangle$, so that this automorphism group is cyclic of size 5.

The 6-cycle C_6 can be analyzed in a similar way. The automorphism group consists of six "rotations" and six "reflections," which are given in cycle form below:

$$(1)(2)(3)(4)(5)(6), \quad (1,2,3,4,5,6), \quad (1,3,5)(2,4,6), \quad (1,4)(2,5)(3,6),$$
$$(1,5,3)(2,6,4), \quad (1,6,5,4,3,2), \quad (2,6)(3,5)(1)(4), \quad (1,2)(3,6)(4,5),$$
$$(1,3)(4,6)(2)(5), \quad (2,3)(5,6)(1,4), \quad (1,5)(2,4)(3)(6), \quad (1,6)(2,5)(3,4).$$

The observations in the previous example generalize as follows.

9.66. Theorem: Automorphism Group of a Cycle. For $n \geq 3$, let C_n be the graph with vertex set $X = \{1, 2, \ldots, n\}$ and edge set $E = \{\{i, i+1\} : 1 \leq i < n\} \cup \{\{1, n\}\}$. Then $\mathrm{Aut}(C_n)$ is a subgroup of S_n of size $2n$. The elements of this group (in one-line form) are the n permutations

$$[i,\ i+1,\ i+2,\ \ldots\ n,\ 1,\ 2,\ \ldots\ i-1] \qquad (1 \leq i \leq n) \tag{9.5}$$

together with the reversals of these n words.

Proof. It is routine to check that all of the displayed permutations do preserve the edges of C_n, hence are automorphisms of this graph. We must show that these are the *only* automorphisms of C_n. Let g be any automorphism of C_n, and put $i = g(1)$. Now, since 1 and 2 are adjacent in C_n, $g(1)$ and $g(2)$ must also be adjacent in C_n. There are two cases: $g(2) = i + 1$ or $g(2) = i - 1$ (reading values mod n). Suppose the first case occurs. Since 2 is adjacent to 3, we must have $g(3) = i + 2$ or $g(3) = i$. But $i = g(1)$ and g is injective, so it

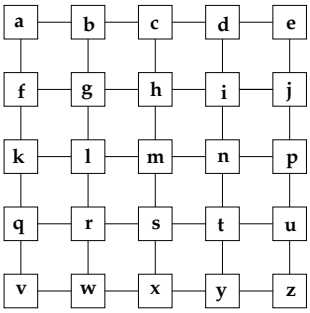

FIGURE 9.3
Graph representing a 5×5 chessboard.

must be that $g(3) = i + 2$. Continuing around the cycle in this way, we see that g must be one of the permutations displayed in (9.5). Similarly, in the case where $g(2) = i - 1$, we see that $g(3) = i - 2$, etc., and g must be the reversal of one of the permutations in (9.5). □

The reasoning used in the preceding proof can be adapted to determine the automorphism groups of more complicated graphs.

9.67. Example. Consider the graph B displayed in Figure 9.3, which models a 5×5 chessboard. What are the automorphisms of B? We note that B has four vertices of degree 2: a, e, v, and z. An automorphism ϕ of B must restrict to give a permutation of these four vertices, since automorphisms preserve degree. Suppose, for example, that $\phi(a) = v$. What can $\phi(b)$ be in this situation? Evidently, $\phi(b)$ must be q or w. In the former case, $\phi(c) = k$ is forced by degree considerations, whereas $\phi(c) = x$ is forced in the latter case. Continuing around the "edge" of the graph, we see that the action of ϕ on all of the "border" vertices is completely determined by where a and b go. A tedious but routine argument then shows that the images of the remaining vertices are also forced. Since a can map to one of the four corners, and then b can map to one of the two neighbors of $\phi(a)$, there are at most $4 \times 2 = 8$ automorphisms of B. Here are the eight possibilities in cycle form:

$$\begin{aligned}
r_0 &= (a)(b)(c)\cdots(x)(y)(z) = \mathrm{id};\\
r_1 &= (a,e,z,v)(b,j,y,q)(c,p,x,k)(d,u,w,f)(g,i,t,r)(h,n,s,l)(m);\\
r_2 &= (a,z)(b,y)(c,x)(d,w)(e,v)(j,q)(p,k)(u,f)(g,t)(h,s)(i,r)(n,l)(m);\\
r_3 &= (a,v,z,e)(b,q,y,j)(c,k,x,p)(d,f,w,u)(g,r,t,i)(h,l,s,n)(m);
\end{aligned}$$

$$\begin{aligned}
s_{y=0} &= (a,v)(b,w)(c,x)(d,y)(e,z)(f,q)(g,r)(h,s)(i,t)(j,u)(k)(l)(m)(n)(p);\\
s_{x=0} &= (a,e)(b,d)(f,j)(g,i)(k,p)(l,n)(q,u)(r,t)(v,z)(w,y)(c)(h)(m)(s)(x);\\
s_{y=x} &= (a,z)(b,u)(c,p)(d,j)(f,y)(g,t)(h,n)(k,x)(l,s)(q,w)(e)(i)(m)(r)(v);\\
s_{y=-x} &= (b,f)(c,k)(d,q)(e,v)(h,l)(i,r)(j,w)(n,s)(p,x)(u,y)(a)(g)(m)(t)(z).
\end{aligned}$$

One may check that all of these maps really are automorphisms of B, so $|\operatorname{Aut}(B)| = 8$. The reader will perceive that this graph has essentially the same symmetries as C_4: four rotations and four reflections. (The subscripts of the reflections indicate the axis of reflection, taking m to be located at the origin.) By directing the edges in a suitable way, we could produce a graph with only four automorphisms (the rotations). These graphs and groups will play a crucial role in solving the chessboard-coloring problem mentioned in the introduction to this chapter.

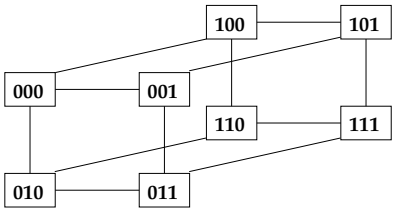

FIGURE 9.4
The cube graph.

9.68. Example. As a final illustration of the calculation of an automorphism group, consider the graph C shown in Figure 9.4, which models a three-dimensional cube. We have taken the vertex set of C to be $\{0,1\}^3$, the set of binary words of length 3. Which bijections $f : \{0,1\}^3 \to \{0,1\}^3$ might be automorphisms of C? First, $f(000)$ can be any of the eight vertices. Next, the three neighbors of 000 (namely 001, 010, and 100) can be mapped bijectively onto the three neighbors of $f(000)$ in any of $3! = 6$ ways. The images of the remaining four vertices are now uniquely determined, as one may check. By the product rule, there are at most $8 \times 6 = 48$ automorphisms of C. A routine but tedious verification shows that all of these potential automorphisms really are automorphisms, so $|\operatorname{Aut}(C)| = 48$. The geometrically inclined reader may like to visualize these automorphisms as arising from suitable rotations and reflections in three-dimensional space. Here are the six automorphisms of C that send 000 to 110:

$$f_1 = \begin{pmatrix} 000 & 001 & 010 & 011 & 100 & 101 & 110 & 111 \\ 110 & 100 & 111 & 101 & 010 & 000 & 011 & 001 \end{pmatrix},$$

$$f_2 = \begin{pmatrix} 000 & 001 & 010 & 011 & 100 & 101 & 110 & 111 \\ 110 & 100 & 010 & 000 & 111 & 101 & 011 & 001 \end{pmatrix},$$

$$f_3 = \begin{pmatrix} 000 & 001 & 010 & 011 & 100 & 101 & 110 & 111 \\ 110 & 111 & 100 & 101 & 010 & 011 & 000 & 001 \end{pmatrix},$$

$$f_4 = \begin{pmatrix} 000 & 001 & 010 & 011 & 100 & 101 & 110 & 111 \\ 110 & 111 & 010 & 011 & 100 & 101 & 000 & 001 \end{pmatrix},$$

$$f_5 = \begin{pmatrix} 000 & 001 & 010 & 011 & 100 & 101 & 110 & 111 \\ 110 & 010 & 100 & 000 & 111 & 011 & 101 & 001 \end{pmatrix},$$

$$f_6 = \begin{pmatrix} 000 & 001 & 010 & 011 & 100 & 101 & 110 & 111 \\ 110 & 010 & 111 & 011 & 100 & 000 & 101 & 001 \end{pmatrix}.$$

9.10 Group Homomorphisms

9.69. Definition: Group Homomorphisms. Let $(G, \star)$ and $(H, \bullet)$ be groups. A function $f : G \to H$ is called a *group homomorphism* iff

$$f(x \star y) = f(x) \bullet f(y) \qquad \text{for all } x, y \in G.$$

A *group isomorphism* is a bijective group homomorphism.

9.70. Example. Define $f : \mathbb{R} \to \mathbb{R}^+$ by $f(x) = e^x$. This function is a group homomorphism from $(\mathbb{R}, +)$ to $(\mathbb{R}^+, \times)$, since $f(x+y) = e^{x+y} = e^x \times e^y = f(x) \times f(y)$ for all $x, y \in \mathbb{R}$. In fact, f is a group isomorphism since $g : \mathbb{R}^+ \to \mathbb{R}$ given by $g(x) = \ln x$ is a two-sided inverse for f.

9.71. Example. Define $h : \mathbb{C} \to \mathbb{R}$ by $h(x+iy) = x$ for all $x, y \in \mathbb{R}$. One checks that h is a group homomorphism from $(\mathbb{C}, +)$ to $(\mathbb{R}, +)$ that is surjective but not injective. Next, define $r : \mathbb{C} \sim \{0\} \to \mathbb{R} \sim \{0\}$ by setting $r(x+iy) = |x+iy| = \sqrt{x^2+y^2}$. Given nonzero $w = x+iy$ and $z = u+iv$, we calculate

$$\begin{aligned} r(wz) &= r((xu-yv)+i(yu+xv)) = \sqrt{(xu-yv)^2+(yu+xv)^2} \\ &= \sqrt{(x^2+y^2)(u^2+v^2)} = r(w)r(z). \end{aligned}$$

So r is a homomorphism of multiplicative groups.

9.72. Example. For any group G, the identity map $\mathrm{id}_G : G \to G$ is a group isomorphism. More generally, if $H \leq G$, then the inclusion map $j : H \to G$ given by $j(h) = h$ for $h \in H$ is a group homomorphism. If $f : G \to K$ and $g : K \to P$ are group homomorphisms, then $g \circ f : G \to P$ is a group homomorphism, since

$$(g \circ f)(xy) = g(f(xy)) = g(f(x)f(y)) = g(f(x))g(f(y)) = (g \circ f)(x)(g \circ f)(y) \qquad (x, y \in G).$$

Moreover, $g \circ f$ is an isomorphism if f and g are isomorphisms, since the composition of bijections is a bijection. If $f : G \to K$ is an isomorphism, then $f^{-1} : K \to G$ is also an isomorphism. For suppose $u, v \in K$. Write $x = f^{-1}(u)$ and $y = f^{-1}(v)$, so $u = f(x)$ and $v = f(y)$. Since f is a group homomorphism, it follows that $uv = f(xy)$. Applying f^{-1} to this relation, we get $f^{-1}(uv) = xy = f^{-1}(u)f^{-1}(v)$.

9.73. Definition: Automorphism Groups. Let $(G, \star)$ be a group. An *automorphism* of G is a group isomorphism $f : G \to G$. Let $\mathrm{Aut}(G)$ denote the set of all such automorphisms.

The remarks in the preceding example (with $K = P = G$) show that $\mathrm{Aut}(G)$ *is a subgroup of* $(\mathrm{Sym}(G), \circ)$.

9.74. Example: Inner Automorphisms. Let G be any group, and fix an element $g \in G$. Define a map $C_g : G \to G$ (called *conjugation by* g) by setting $C_g(x) = gxg^{-1}$. This map is a group homomorphism, since $C_g(xy) = g(xy)g^{-1} = (gxg^{-1})(gyg^{-1}) = C_g(x)C_g(y)$ for all $x, y \in G$. Furthermore, C_g is a group isomorphism, since a calculation shows that $C_{g^{-1}}$ is the two-sided inverse of C_g. It follows that $C_g \in \mathrm{Aut}(G)$ for every $g \in G$. We call automorphisms of the form C_g *inner automorphisms of* G. It is possible for different group elements to induce the same inner automorphism of G. For example, if G is commutative, then $C_g(x) = gxg^{-1} = gg^{-1}x = x$ for all $g, x \in G$, so that all of the conjugation maps C_g reduce to id_G.

9.75. Theorem: Properties of Group Homomorphisms. Let $f : G \to H$ be a group homomorphism. For all $n \in \mathbb{Z}$ and all $x \in G$, $f(x^n) = f(x)^n$. In particular, $f(e_G) = e_H$ and $f(x^{-1}) = f(x)^{-1}$. We say that f *preserves powers, identities, and inverses.*

Proof. First we prove the result for all $n \geq 0$ by induction on n. When $n = 0$, we must prove that $f(e_G) = e_H$. Note that $e_G e_G = e_G$. Applying f to both sides of this equation gives

$$f(e_G)f(e_G) = f(e_G e_G) = f(e_G) = f(e_G)e_H.$$

By left cancellation of $f(e_G)$ in H, we conclude that $f(e_G) = e_H$. For the induction step,

assume $n \geq 0$ and $f(x^n) = f(x)^n$; we will prove $f(x^{n+1}) = f(x)^{n+1}$. Using the definition of exponent notation, we calculate

$$f(x^{n+1}) = f(x^n x) = f(x^n)f(x) = f(x)^n f(x) = f(x)^{n+1}.$$

Next, let us prove the result when $n = -1$. Given $x \in G$, apply f to the equation $xx^{-1} = e_G$ to obtain

$$f(x)f(x^{-1}) = f(xx^{-1}) = f(e_G) = e_H = f(x)f(x)^{-1}.$$

Left cancellation of $f(x)$ gives $f(x^{-1}) = f(x)^{-1}$. Finally, consider an arbitrary negative integer $n = -m$, where $m > 0$. We have

$$f(x^n) = f((x^m)^{-1}) = f(x^m)^{-1} = (f(x)^m)^{-1} = f(x)^{-m} = f(x)^n. \qquad \square$$

We can use group homomorphisms to construct more examples of subgroups.

9.76. Definition: Kernel and Image of a Homomorphism. Let $f : G \to H$ be a group homomorphism. The *kernel* of f, denoted $\ker(f)$, is the set of all $x \in G$ such that $f(x) = e_H$. The *image* of f, denoted $\text{img}(f)$, is the set of all $y \in H$ such that $y = f(z)$ for some $z \in G$.

The reader may check that $\ker(f) \trianglelefteq G$ and $\text{img}(f) \leq H$.

9.77. Example. Consider the homomorphisms h and r from 9.71, given by $h(x + iy) = x$ and $r(z) = |z|$ for $x, y \in \mathbb{R}$ and nonzero $z \in \mathbb{C}$. The kernel of h is the set of pure imaginary numbers $\{iy : y \in \mathbb{R}\}$. The kernel of r is the unit circle $\{z \in \mathbb{C} : |z| = 1\}$. The image of h is all of $\mathbb{R}$, while the image of r is $\mathbb{R}^+$.

9.78. Example: Even Permutations. By 9.31, the function $\text{sgn} : S_n \to \{+1, -1\}$ is a group homomorphism. The kernel of this homomorphism, which is denoted A_n, consists of all $f \in S_n$ such that $\text{sgn}(f) = +1$. Such f are called *even permutations*. A_n is called the *alternating group on n letters*. We will see later (9.121) that $|A_n| = |S_n|/2 = n!/2$ for all $n \geq 2$.

9.79. Example: Analysis of Cyclic Subgroups. Let G be any group, written multiplicatively, and fix an element $x \in G$. Define $f : \mathbb{Z} \to G$ by setting $f(n) = x^n$ for all $n \in \mathbb{Z}$. By the laws of exponents, f is a group homomorphism. The image of f is precisely $\langle x \rangle$, the cyclic subgroup of G generated by x. The kernel of f is some subgroup of $\mathbb{Z}$, which by 9.59 has the form $m\mathbb{Z}$ for some integer $m \geq 0$. Consider the case where $m = 0$. Then $x^i = e_G$ iff $f(i) = e_G$ iff $i \in \ker(f)$ iff $i = 0$, so x^0 is the only power of x that equals the identity of G. We say that x has *infinite order* in this case. We remark that $i \neq j$ implies $x^i \neq x^j$, since

$$x^i = x^j \Rightarrow x^{i-j} = e_G \Rightarrow i - j = 0 \Rightarrow i = j.$$

In other words, all integer powers of x are distinct elements of G. This means that $f : \mathbb{Z} \to G$ is injective. So f induces a group isomorphism $f' : \mathbb{Z} \to \langle x \rangle$.

Now consider the case where $m > 0$. Then $x^i = e_G$ iff $f(i) = e_G$ iff $i \in \ker(f)$ iff i is a multiple of m. We say that x *has order* m in this case; thus, the order of x is the least positive exponent i such that $x^i = e_G$. We claim that the cyclic group $\langle x \rangle$ consists of the m *distinct* elements $x^0, x^1, x^2, \ldots, x^{m-1}$. For, given an arbitrary element $x^n \in \langle x \rangle$, we can divide n by m to get $n = mq + r$ for some r with $0 \leq r < m$. Then $x^n = x^{mq+r} = (x^m)^q x^r = e_G^q x^r = x^r$, so x^n is equal to one of the elements in our list. Furthermore, the listed elements are distinct. For suppose $0 \leq i \leq j < m$ and $x^i = x^j$. Then $x^{j-i} = e_G$, forcing m to divide $j - i$. But $0 \leq j - i < m$, so the only possibility is $j - i = 0$, hence $i = j$. Consider the function $g : \mathbb{Z}_m \to \langle x \rangle$ given by $g(i) = x^i$ for $0 \leq i < m$. This function is a well-defined bijection

by the preceding remarks. Furthermore, g is a group homomorphism. To check this, let $i, j \in \mathbb{Z}_m$. If $i + j < m$, then

$$g(i \oplus j) = g(i + j) = x^{i+j} = x^i x^j = g(i)g(j).$$

If $i + j \geq m$, then

$$g(i \oplus j) = g(i + j - m) = x^{i+j-m} = x^i x^j (x^m)^{-1} = x^i x^j = g(i)g(j).$$

So g is an isomorphism from $(\mathbb{Z}_m, \oplus)$ to the cyclic subgroup generated by x.

We have just shown that *every cyclic group $\langle x \rangle$ is isomorphic to one of the additive groups $\mathbb{Z}$ or $\mathbb{Z}_m$ for some $m > 0$.* The first case occurs when x has infinite order, and the second case occurs when x has order m.

9.11 Group Actions

The fundamental tool needed to solve counting problems involving symmetry is the notion of a *group action*.

9.80. Definition: Group Actions. Suppose G is a group and X is a set. An *action* of G on X is a function $* : G \times X \to X$ satisfying the following axioms.

1. For all $g \in G$ and all $x \in X$, $g * x \in X$ (closure).
2. For all $x \in X$, $e_G * x = x$ (identity).
3. For all $g, h \in G$ and all $x \in X$, $g * (h * x) = (gh) * x$ (associativity).

The pair $(X, *)$ is called a *G-set*.

9.81. Example. For any set X, the group $G = (\mathrm{Sym}(X), \circ)$ acts on X via the rule $g * x = g(x)$ for $g \in G$ and $x \in X$. Axiom 1 holds because each $g \in G$ is a function from X to X, hence $g * x = g(x) \in X$ for all $x \in X$. Axiom 2 holds since $e_G * x = \mathrm{id}_X * x = \mathrm{id}_X(x) = x$ for all $x \in X$. Axiom 3 holds because

$$g * (h * x) = g(h(x)) = (g \circ h)(x) = (gh) * x \qquad (g, h \in \mathrm{Sym}(X), x \in X).$$

9.82. Example. Let G be any group, written multiplicatively, and let X be the set G. Define $* : G \times X \to X$ by $g * x = gx$ for all $g, x \in G$. We say that "G acts on itself by left multiplication." In this example, the action axioms reduce to the corresponding group axioms for G.

We can define another action of G on $X = G$ by letting $g \bullet x = xg^{-1}$ for all $g, x \in G$. The first two axioms for an axiom are immediately verified; the third axiom follows from the calculation

$$g \bullet (h \bullet x) = g \bullet (xh^{-1}) = (xh^{-1})g^{-1} = x(h^{-1}g^{-1}) = x(gh)^{-1} = (gh) \bullet x \quad (g, h, x \in G).$$

We say that "G acts on itself by inverted right multiplication." One can check that the rule $g \cdot x = xg$ (for $g, x \in G$) does *not* define a group action for non-commutative groups G, because axiom 3 fails. (But see the discussion of right group actions below.)

9.83. Example. Let the group G act on the set $X = G$ as follows:

$$g * x = gxg^{-1} \qquad (g \in G, x \in X).$$

We say that "G acts on itself by conjugation." The reader should verify that the axioms for an action are satisfied.

9.84. Example. Suppose we are given a group action $* : G \times X \to X$. Let H be any subgroup of G. By restricting the action function to $H \times X$, we obtain an action of H on X, as one immediately verifies. Combining this construction with previous examples, we obtain quite a few additional instances of group actions. For example, any subgroup H of a group G acts on G by left multiplication, and by inverted right multiplication, and by conjugation. Any subgroup H of $\mathrm{Sym}(X)$ acts on X via $f \star x = f(x)$ for $f \in H$ and $x \in X$. In particular, the automorphism group $\mathrm{Aut}(G)$ of a group G is a subgroup of $\mathrm{Sym}(G)$, so $\mathrm{Aut}(G)$ acts on G via $f \star x = f(x)$ for $f \in \mathrm{Aut}(G)$ and $x \in G$. Similarly, if K is a graph with vertex set X, then $\mathrm{Aut}(K)$ is a subgroup of $\mathrm{Sym}(X)$, and therefore $\mathrm{Aut}(K)$ acts on X via $f \star x = f(x)$ for $f \in \mathrm{Aut}(K)$ and $x \in X$.

9.85. Example. Suppose $(X, *)$ is a G-set. Let $\mathcal{P}(X)$ be the power set of X, which consists of all subsets of X. It is routine to check that $\mathcal{P}(X)$ is a G-set under the action

$$g \bullet S = \{g * s : s \in S\} \qquad (g \in G, S \in \mathcal{P}(X)).$$

9.86. Example. Consider a polynomial ring $R = F[x_1, x_2, \ldots, x_n]$, where F is a field (see §7.16). The symmetric group S_n acts on $\{1, 2, \ldots, n\}$ via $f * i = f(i)$ for $f \in S_n$ and $1 \leq i \leq n$. We can transfer this to an action of S_n on $\{x_1, \ldots, x_n\}$ by defining

$$f * x_i = x_{f(i)} \qquad (f \in S_n, 1 \leq i \leq n).$$

Using the universal mapping property of polynomial rings (see 7.102), each bijection $(x_i \mapsto x_{f(i)} : 1 \leq i \leq n)$ extends to a ring isomorphism $\overline{f}$ sending $p = p(x_1, \ldots, x_n) \in R$ to $\overline{f}(p) = p(x_{f(1)}, \ldots, x_{f(n)})$. One may check that the rule $f * p = \overline{f}(p)$ (for $f \in S_n$ and $p \in R$) defines an action of S_n on R. In particular, $g * (h * p) = (g \circ h) * p$ follows by the uniqueness part of the universal mapping property, since both sides are the image of p under the unique ring homomorphism sending x_i to $x_{g(h(i))}$ for all i.

9.87. Example. By imitating ideas in the previous example, we can define certain group actions on vector spaces. Suppose V is a vector space over a field F and let $X = (x_1, \ldots, x_n)$ be an ordered basis of V. For $f \in S_n$, the map $x_i \mapsto x_{f(i)}$ on basis vectors extends by linearity to a unique linear map $T_f : V \to V$, given explicitly by

$$T_f(a_1 x_1 + \cdots + a_n x_n) = a_1 x_{f(1)} + \cdots + a_n x_{f(n)} \qquad (a_i \in F).$$

One may check that $f * v = T_f(v)$ (for $f \in S_n$ and $v \in V$) defines an action of the group S_n on the set V.

9.88. Example. Suppose G is a group, $(X, *)$ is a G-set, and W and Z are any sets. Recall that ${}^W X$ is the set of all functions $F : W \to X$. This set of functions can be turned into a G-set by defining

$$(g \bullet F)(w) = g * (F(w)) \qquad (g \in G, F \in {}^W X, w \in W).$$

We leave the verification of the action axioms as an exercise.

Now consider the set ${}^X Z$ of all functions $F : X \to Z$. We claim this set of functions becomes a G-set if we define

$$(g \bullet F)(x) = F(g^{-1} * x) \qquad (g \in G, F \in {}^X Z, x \in X).$$

Let us carefully prove this claim. First, given $g \in G$ and $F \in {}^X Z$, the map $g \bullet F$ is a well-defined function from X to Z because $g^{-1} * x \in X$ and F maps X into Z. So, $g \bullet F \in {}^X Z$, verifying closure. Second, letting e be the identity of G and letting $F \in {}^X Z$, we have

$$(e \bullet F)(x) = F(e^{-1} * x) = F(e * x) = F(x) \qquad (x \in X),$$

so that we have the equality of functions $e \bullet F = F$. (Recall that two functions are equal iff they have the same domain X, have the same codomain, and take the same values at each $x \in X$.) Third, we verify the associativity axiom for $\bullet$. Fix $g, h \in G$ and $F \in {}^{X}Z$. The two functions $g \bullet (h \bullet F)$ and $(gh) \bullet F$ both have domain X and codomain Z. Fix $x \in X$. On one hand,

$$[(gh) \bullet F](x) = F((gh)^{-1} * x) = F((h^{-1}g^{-1}) * x).$$

On the other hand, using the definition of $\bullet$ twice,

$$[g \bullet (h \bullet F)](x) = [h \bullet F](g^{-1} * x) = F(h^{-1} * (g^{-1} * x)).$$

Since $*$ is known to be an action, we see that $g \bullet (h \bullet F)$ and $(gh) \bullet F$ take the same value at x. So the third action axiom is proved. The reader may check that this axiom would fail, in general, if we omitted the inverse in the definition of $\bullet$.

9.89. Example. Let n be a fixed integer, let Y be a set, and let

$$U = \{(y_1, \ldots, y_n) : y_i \in Y\}$$

be the set of all sequences of n elements of Y. The group S_n acts on U via the rule

$$f \cdot (y_1, y_2, \ldots, y_n) = (y_{f^{-1}(1)}, y_{f^{-1}(2)}, \ldots, y_{f^{-1}(n)}) \qquad (f \in S_n; y_1, \ldots, y_n \in Y).$$

The inverses in this formula are *essential*. To see why, we observe that the action here is actually a special case of the previous example. For, a sequence in U is officially defined to be a function $y : \{1, 2, \ldots, n\} \to Y$ where $y(i) = y_i$. Using this function notation for sequences, we have (for $f \in S_n$)

$$(f \cdot y)(i) = y(f^{-1}(i)) = (f \bullet y)(i) \qquad (1 \leq i \leq n),$$

in agreement with the previous example. One should also note that acting by f moves the object z originally in *position* i to *position* $f(i)$ in the new sequence. This is true because $(f \cdot y)(f(i)) = y(f^{-1}(f(i))) = y(i) = z$.

The reader may now be disturbed by the *lack* of inverses in the formula $f * x_i = x_{f(i)}$ from 9.87. However, there is no contradiction since the x_i's in the latter example are fixed basis elements in a vector space V, not the entries in a sequence. Indeed, recall that the action on V is given by $f * v = T_f(v)$ where T_f is the linear extension of the map $x_i \mapsto x_{f(i)}$. Writing $v = \sum_i a_i x_i$, the *coordinates* of v relative to this basis are the entries in the sequence $(a_1, a_2, \ldots, a_n)$. Applying f to v gives

$$\sum_i a_i x_{f(i)} = \sum_j a_{f^{-1}(j)} x_j,$$

where we changed variables by letting $j = f(i)$, $i = f^{-1}(j)$. We now see that the coordinates of $f * v$ relative to the ordered basis $(x_1, \ldots, x_n)$ are $(a_{f^{-1}(1)}, \ldots, a_{f^{-1}(n)})$. For example,

$$(1,2,3) * (a_1x_1 + a_2x_2 + a_3x_3) = (a_1x_2 + a_2x_3 + a_3x_1) = (a_3x_1 + a_1x_2 + a_2x_3),$$

or equivalently, in coordinate notation,

$$(1,2,3) * (a_1, a_2, a_3) = (a_3, a_1, a_2).$$

To summarize, when f acts directly on the *objects* x_i, no inverse is needed; but when f permutes the *positions* in a list, one must apply f^{-1} to each subscript.

9.90. Remark: Right Actions. A *right action* of a group G on a set X is a map $* : X \times G \to X$ such that $x * e = x$ and $x * (gh) = (x * g) * h$ for all $x \in X$ and all $g, h \in G$. For example, $x * g = xg$ (with no inverse) defines a right action of a group G on the set $X = G$. Similarly, we get a *right* action of S_n on the set of sequences in the previous example by writing

$$(y_1, \ldots, y_n) * f = (y_{f(1)}, \ldots, y_{f(n)}).$$

Group actions (as defined at the beginning of this section) are sometimes called *left* actions to avoid confusion with right actions. We shall mostly consider left group actions in the sequel, but right actions are occasionally more convenient to use (cf. 9.109).

9.12 Permutation Representations

Group actions are closely related to symmetric groups. To understand the precise nature of this relationship, we need the following definition.

9.91. Definition: Permutation Representations. A *permutation representation* of a group G on a set X is a group homomorphism $\phi : G \to \mathrm{Sym}(X)$.

This definition seems quite different from the definition of a group action given in the last section. But we will see in this section that group actions and permutation representations are essentially the same thing. Both viewpoints turn out to be pertinent in the application of group actions to problems in combinatorics and algebra.

We first show that any group action of G on X gives rise to a permutation representation of G on X in a canonical way. The key idea appears in the next definition.

9.92. Definition: Left Multiplication Maps. Let $* : G \times X \to X$ be an action of the group G on the set X. For each $g \in G$, *left multiplication by* G (relative to this action) is the function $L_g : X \to X$ defined by

$$L_g(x) = g * x \qquad (x \in X).$$

Note that L_g does take values in X, by the closure axiom for group actions.

9.93. Theorem: Properties of Left Multiplication Maps. Let $(X, *)$ be a G-set. (a) $L_e = \mathrm{id}_X$. (b) For all $g, h \in G$, $L_{gh} = L_g \circ L_h$. (c) For all $g \in G$, $L_g \in \mathrm{Sym}(X)$, and $L_g^{-1} = L_{g^{-1}}$.

Proof. All functions appearing here have domain X and codomain X. So it suffices to check that the relevant functions take the same value at each $x \in X$. For (a), $L_e(x) = e * x = x = \mathrm{id}_X(x)$ by the identity axiom for group actions. For (b), $L_{gh}(x) = (gh) * x = g * (h * x) = L_g(L_h(x)) = (L_g \circ L_h)(x)$ by the associativity axiom for group actions. Finally, using (a) and (b) with $h = g^{-1}$ shows that $\mathrm{id}_X = L_g \circ L_{g^{-1}}$. Similarly, $\mathrm{id}_X = L_{g^{-1}} \circ L_g$. This means that $L_{g^{-1}}$ is the two-sided inverse of L_g; in particular, both of these maps must be bijections. □

Using the theorem, we can pass from a group action $*$ to a permutation representation ϕ as follows. Define $\phi : G \to \mathrm{Sym}(X)$ by setting $\phi(g) = L_g \in \mathrm{Sym}(X)$ for all $g \in G$. By part (b) of the theorem,

$$\phi(gh) = L_{gh} = L_g \circ L_h = \phi(g) \circ \phi(h) \qquad (g, h \in G),$$

and so ϕ is a group homomorphism.

9.94. Example: Cayley's Theorem. We have seen that any group G acts on the set $X = G$ by left multiplication. The preceding construction produces a group homomorphism $\phi : G \to \mathrm{Sym}(G)$, such that $\phi(g) = L_g = (x \mapsto gx : x \in G)$. We claim that ϕ is injective in this situation. For, suppose $g, h \in G$ and $L_g = L_h$. Applying these two functions to e (the identity of G) gives $g = ge = L_g(e) = L_h(e) = he = h$. It follows that G is isomorphic (via ϕ) to the image of ϕ, which is a subgroup of the symmetric group $\mathrm{Sym}(G)$. We have just proved *Cayley's Theorem*, which says that *any group is isomorphic to a subgroup of some symmetric group*. If G has n elements, one can check that $\mathrm{Sym}(G)$ is isomorphic to S_n. So every n-element group is isomorphic to a subgroup of the specific symmetric group S_n.

9.95. Example. Recall that, for any set X, $\mathrm{Sym}(X)$ acts on X via $f * x = f(x)$ for $f \in \mathrm{Sym}(X)$ and $x \in X$. What is the associated permutation representation $\phi : \mathrm{Sym}(X) \to \mathrm{Sym}(X)$? First note that for $f \in \mathrm{Sym}(X)$, left multiplication by f is the map $L_f : X \to X$ such that $L_f(x) = f * x = f(x)$. In other words, $L_f = f$, so that $\phi(f) = L_f = f$. This means that ϕ is the identity homomorphism. More generally, whenever a subgroup H of $\mathrm{Sym}(X)$ acts on X in the canonical way, the corresponding permutation representation is the inclusion map of H into $\mathrm{Sym}(X)$.

So far, we have seen that every group action of G on X has an associated permutation representation. We can reverse this process by starting with an arbitrary permutation representation $\phi : G \to \mathrm{Sym}(X)$ and building a group action, as follows. Given ϕ, define $* : G \times X \to X$ by setting $g * x = \phi(g)(x)$ for all $g \in G$ and $x \in X$. Note that $\phi(g)$ is a function with domain X, so the expression $\phi(g)(x)$ denotes a well-defined element of X. In particular, $*$ satisfies the closure axiom in 9.80. Since group homomorphisms preserve identities, $\phi(e) = \mathrm{id}_X$, and so $e * x = \phi(e)(x) = \mathrm{id}_X(x) = x$ for all $x \in X$. So the identity axiom holds. Finally, using the fact that ϕ is a group homomorphism, we calculate

$$\begin{aligned} (gh) * x &= \phi(gh)(x) = (\phi(g) \circ \phi(h))(x) \\ &= \phi(g)(\phi(h)(x)) = g * (h * x). \end{aligned}$$

So the associativity axiom holds, completing the proof that $*$ is a group action.

The following theorem is the formal enunciation of our earlier claim that group actions and permutation representations are "essentially the same concept."

9.96. Theorem: Equivalence of Group Actions and Permutation Representations. Fix a group G and a set X. Let A be the set of all group actions of G on X, and let P be the set of all permutation representations of G on X. There are mutually inverse bijections $F : A \to P$ and $H : P \to A$, given by

$$F(*) = \phi : G \to \mathrm{Sym}(X) \text{ where } \phi(g) = L_g = (x \mapsto g * x : x \in X);$$

$$H(\phi) = * : G \times X \to X \text{ where } g * x = \phi(g)(x) \qquad (g \in G, x \in X).$$

Proof. The discussion preceding the theorem has shown that F does map the set A into the stated codomain P, and that H does map the set P into the stated codomain A. We need only verify that $F \circ H = \mathrm{id}_P$ and $H \circ F = \mathrm{id}_A$.

To show $F \circ H = \mathrm{id}_P$, fix $\phi \in P$, and write $* = H(\phi)$ and $\psi = F(*)$. We must confirm that $\psi = \phi : G \to \mathrm{Sym}(X)$. To do this, fix $g \in G$, and ask whether the two functions $\psi(g), \phi(g) : X \to X$ are equal. For each $x \in X$,

$$\psi(g)(x) = L_g(x) = g * x = \phi(g)(x).$$

So $\psi(g) = \phi(g)$ for all g, hence $\psi = \phi$ as desired.

To show $H \circ F = \mathrm{id}_A$, fix $* \in A$, and write $\phi = F(*)$, $\bullet = H(\phi)$. We must confirm that $\bullet = *$. For this, fix $g \in G$ and $x \in X$. Now compute

$$g \bullet x = \phi(g)(x) = L_g(x) = g * x. \qquad \square$$

9.97. Example. We can use permutation representations to generate new constructions of group actions. For instance, suppose $(X, *)$ is a G-set with associated permutation representation $\phi : G \to \mathrm{Sym}(X)$. Now suppose we are given a group homomorphism $u : K \to G$. Composing with ϕ gives a homomorphism $\phi \circ u : K \to \mathrm{Sym}(X)$. This is a permutation representation of K on X, which means that X can be made into a K-set in a canonical way. Specifically, by applying the map H from the theorem, we see that the K-action on X is given by

$$k \bullet x = u(k) * x \qquad (k \in K, x \in X).$$

9.13 Stable Subsets and Orbits

One way to gain information about a group is to study its subgroups. The analogous concept for G-sets appears in the next definition.

9.98. Definition: G-Stable Subsets. Let $(X, *)$ be a G-set. A subset Y of X is called a *G-stable subset* iff $g * y \in Y$ for all $g \in G$ and all $y \in Y$.

When Y is a G-stable subset, the restriction of $*$ to $G \times Y$ maps into the codomain Y, by definition. Since the identity axiom and associativity axiom still hold for the restricted action, we see that Y is a G-set.

Recall that every element of a group generates a cyclic subgroup. Similarly, we can pass from an element of a G-set to a G-stable subset as follows.

9.99. Definition: Orbits. Suppose $(X, *)$ is a G-set, and $x \in X$. The *G-orbit* of x is the set

$$Gx = G * x = \{g * x : g \in G\} \subseteq X.$$

Every orbit is a G-stable subset: for, given $h \in G$ and $g * x \in Gx$, the associativity axiom gives $h * (g * x) = (hg) * x \in Gx$. Furthermore, by the identity axiom, $x = e * x \in Gx$ for each $x \in X$.

9.100. Example. Let S_5 act on the set $X = \{1, 2, 3, 4, 5\}$ via $f * x = f(x)$ for $f \in S_5$ and $x \in X$. For each $i \in X$, the orbit $S_5 * i = \{f(i) : f \in S_5\}$ is all of X. The reason is that for any given j in X, we can find an $f \in S_5$ such that $f(i) = j$; for instance, take $f = (i, j)$. On the other hand, consider the subgroup $H = \langle (1, 3)(2, 4, 5) \rangle$ of S_5. If we let H act on X via $f * x = f(x)$ for $f \in H$ and $x \in X$, we get different orbits. One may check directly that

$$H * 1 = H * 3 = \{1, 3\}, \qquad H * 2 = H * 4 = H * 5 = \{2, 4, 5\}.$$

Note that the H-orbits are precisely the connected components of the digraph representing the generator $(1, 3)(2, 4, 5)$ of H. One can verify that this holds in general whenever a cyclic subgroup of S_n acts on $\{1, 2, \ldots, n\}$.

Now consider the action of A_5 on X. As in the case of S_5, we have $A_5 * i = X$ for all $i \in X$, but for a different reason. Given $j \in X$, we must now find an *even* permutation sending i to j. We can use the identity permutation if $i = j$. Otherwise, choose two distinct elements k, l that are different from i and j, and use the permutation $(i, j)(k, l)$.

9.101. Example. Let S_4 act on the set X of all 4-tuples of integers by permuting positions:

$$f * (x_1, x_2, x_3, x_4) = (x_{f^{-1}(1)}, x_{f^{-1}(2)}, x_{f^{-1}(3)}, x_{f^{-1}(4)}) \qquad (f \in S_4, x_i \in \mathbb{Z}).$$

The S_4-orbit of a sequence $x = (x_1, x_2, x_3, x_4)$ consists of all possible sequences obtainable from x by permuting the entries. For example,

$$S_4 * (5, 1, 5, 1) = \{(1, 1, 5, 5), (1, 5, 1, 5), (1, 5, 5, 1), (5, 1, 1, 5), (5, 1, 5, 1), (5, 5, 1, 1)\}.$$

As another example, $S_4 * (3, 3, 3, 3) = \{(3, 3, 3, 3)\}$ and $S_4 * (1, 3, 5, 7)$ is the set of all 24 permutations of this list. Now consider the cyclic subgroup $H = \langle (1, 2, 3, 4) \rangle$ of S_4. Restricting the action turns X into an H-set. When computing orbits relative to the H-action, we are only allowed to cyclically shift the elements in each 4-tuple. So, for instance,

$$\begin{aligned} H * (5, 1, 5, 1) &= \{(5, 1, 5, 1), (1, 5, 1, 5)\}; \\ H * (1, 3, 5, 7) &= \{(1, 3, 5, 7), (3, 5, 7, 1), (5, 7, 1, 3), (7, 1, 3, 5)\}. \end{aligned}$$

As before, the orbit of a given $x \in X$ depends heavily on which group is acting on X.

9.102. Example. Let a group G act on itself by left multiplication: $g * x = gx$ for $g, x \in G$. For every $x \in G$, the orbit Gx is all of G. For, given any $y \in G$, we have $(yx^{-1}) * x = y$. In the next section, we will study what happens when a subgroup H acts on G by left (or right) multiplication.

9.103. Example: Conjugacy Classes. Let G be a group. We have seen that G acts on itself by conjugation: $g * x = gxg^{-1}$ for $g, x \in G$. The orbit of $x \in G$ under this action is the set

$$G * x = \{gxg^{-1} : g \in G\}.$$

This set is called the *conjugacy class of x in G*. For example, when $G = S_3$, the conjugacy classes are

$$\begin{aligned} G * \mathrm{id} &= \{\mathrm{id}\}; \\ G * (1, 2) = G * (1, 3) = G * (2, 3) &= \{(1, 2), (1, 3), (2, 3)\}; \\ G * (1, 2, 3) = G * (1, 3, 2) &= \{(1, 2, 3), (1, 3, 2)\}. \end{aligned}$$

One can confirm this with the aid of the identities

$$f \circ (i, j) \circ f^{-1} = (f(i), f(j)); \qquad f \circ (i, j, k) \circ f^{-1} = (f(i), f(j), f(k)) \qquad (f \in S_3).$$

(The generalization of this example to any S_n is discussed in §9.16.) We observe in passing that $G * x = \{x\}$ iff $gxg^{-1} = x$ for all $g \in G$ iff $gx = xg$ for all $g \in G$ iff x commutes with every element of G. In particular, for G commutative, every conjugacy class consists of a single element.

9.104. Example. Let $B = (X, E)$ be the graph representing a 5×5 chessboard shown in Figure 9.3. Let the graph automorphism group $G = \mathrm{Aut}(B)$ act on X via $f * x = f(x)$ for $f \in G$ and $x \in X$. We explicitly determined the elements of G in 9.67. We can use this calculation to find all the distinct G-orbits. They are:

$$\begin{aligned} Ga &= \{a, e, z, v\} = Ge = Gz = Gv; \\ Gb &= \{b, d, j, u, y, w, q, f\} = Gd = Gj = \cdots; \\ Gc &= \{c, p, x, k\}; \\ Gg &= \{g, i, t, r\}; \\ Gh &= \{h, n, s, l\}; \\ Gm &= \{m\}. \end{aligned}$$

The reader may have noticed in these examples that distinct orbits of the G-action on X are always pairwise disjoint. We now prove that this always happens.

9.105. Theorem: Orbit Decomposition of a G-set. Let $(X, *)$ be a G-set. Every element $x \in X$ belongs to exactly one G-orbit, namely Gx. In other words, the distinct G-orbits of the action $*$ form a set partition of X.

Proof. Define a relation on X by setting, for $x, y \in X$, $x \sim y$ iff $y = g * x$ for some $g \in G$. This relation is reflexive on X: given $x \in G$, we have $x = e * x$, so $x \sim x$. This relation is symmetric: given $x, y \in X$ with $x \sim y$, we know $y = g * x$ for some $g \in G$. A routine calculation shows that $x = g^{-1} * y$, so $y \sim x$. This relation is transitive: given $x, y, z \in X$ with $x \sim y$ and $y \sim z$, we know $y = g * x$ and $z = h * y$ for some $g, h \in G$. So $z = h * (g * x) = (hg) * x$, and $x \sim z$. Thus we have an equivalence relation on X. Recall from the proof of 2.55 that the equivalence classes of any equivalence relation on X form a set partition of X. In this situation, the equivalence classes are precisely the G-orbits, since the equivalence class of x is

$$\{y \in X : x \sim y\} = \{y \in X : y = g * x \text{ for some } g \in G\} = Gx. \qquad \square$$

9.106. Corollary. Every group G is the disjoint union of its conjugacy classes.

Everything we have said can be adapted to give results on right actions. In particular, if G acts on X on the right, then X is partitioned into a disjoint union of the right G-orbits

$$xG = \{x * g : g \in G\} \qquad (x \in X).$$

9.14 Cosets

The idea of a *coset* plays a central role in group theory. Cosets arise as the orbits of a certain group action.

9.107. Definition: Right Cosets. Let G be a group, and let H be any subgroup of G. Let H act on G by left multiplication: $h * x = hx$ for $h \in H$ and $x \in G$. The orbit of x under this action, namely

$$Hx = \{hx : h \in H\}$$

is called the *right coset of x relative to H*.

By the general theory of group actions, we know that G is the disjoint union of its right cosets.

9.108. Example. Let $G = S_3$ and $H = \{\mathrm{id}, (1,2)\}$. The right cosets of H in G are

$$\begin{aligned} H\,\mathrm{id} = H(1,2) &= \{\mathrm{id}, (1,2)\} = H; \\ H(1,3) = H(1,3,2) &= \{(1,3), (1,3,2)\}; \\ H(2,3) = H(1,2,3) &= \{(2,3), (1,2,3)\}. \end{aligned}$$

For the subgroup $K = \{\mathrm{id}, (1,2,3), (1,3,2)\}$, the right cosets are

$$K\,\mathrm{id} = K \text{ and } K(1,2) = \{(1,2), (2,3), (1,3)\}.$$

Note that the subgroup itself is always a right coset, but the other right cosets are not subgroups (they do not contain the identity of G).

By letting H act on the right, we obtain the notion of a *left* coset, which will be used frequently in the sequel.

9.109. Definition: Left Cosets. Let G be a group, and let H be any subgroup of G. Let H act on G by right multiplication: $x * h = xh$ for $h \in H$ and $x \in G$. The orbit of x under this action, namely

$$xH = \{xh : h \in H\}$$

is called the *left coset of x relative to H*.

By 9.105, G is the disjoint union of its left cosets.

9.110. Example. Let $G = S_3$ and $H = \{\mathrm{id}, (1,2)\}$ as above. The left cosets of H in G are

$$\begin{aligned} \mathrm{id}\, H = (1,2)H &= \{\mathrm{id}, (1,2)\} = H; \\ (1,3)H = (1,2,3)H &= \{(1,3), (1,2,3)\}; \\ (2,3)H = (1,3,2)H &= \{(2,3), (1,3,2)\}. \end{aligned}$$

Observe that $xH \neq Hx$ except when $x \in H$. This shows that left cosets and right cosets do not coincide in general. On the other hand, for the subgroup $K = \{\mathrm{id}, (1,2,3), (1,3,2)\}$, the left cosets are K and $(1,2)K = \{(1,2), (1,3), (2,3)\}$. One checks that $xK = Kx$ for all $x \in S_3$, so that left cosets and right cosets do coincide for some subgroups.

Although $x = y$ certainly implies $xH = yH$, one must remember that the converse is almost always false. The next result gives criteria for deciding when two cosets xH and yH are equal; it is used constantly in arguments involving cosets.

9.111. Coset Equality Theorem. Let H be a subgroup of G. For all $x, y \in G$, the following conditions are logically equivalent:

(a) $xH = yH$.	(a′) $yH = xH$.
(b) $x \in yH$.	(b′) $y \in xH$.
(c) There exists $h \in H$ with $x = yh$.	(c′) There exists $h' \in H$ with $y = xh'$.
(d) $y^{-1}x \in H$.	(d′) $x^{-1}y \in H$.

Proof. We first prove (a)⇒(b)⇒(c)⇒(d)⇒(a). If $xH = yH$, then $x = xe \in xH = yH$, so $x \in yH$. If $x \in yH$, then $x = yh$ for some $h \in H$ by definition of yH. If $x = yh$ for some $h \in H$, then multiplying by y^{-1} on the left gives $y^{-1}x \in H$. Finally, assume that $y^{-1}x \in H$. Then $y(y^{-1}x) = x$ lies in the orbit yH. We also have $x = xe \in xH$. As orbits are either disjoint or equal, we must have $xH = yH$.

Interchanging x and y in the last paragraph proves the equivalence of (a′), (b′), (c′), and (d′). Since (a) and (a′) are visibly equivalent, the proof is complete. □

9.112. Remark. The equivalence of (a) and (d) in the last theorem is used quite frequently. Note too that the subgroup H is a coset (namely eH), and $xH = H$ iff $xH = eH$ iff $e^{-1}x \in H$ iff $x \in H$. Finally, one can prove an analogous theorem for right cosets. The key difference is that $Hx = Hy$ iff $xy^{-1} \in H$ iff $yx^{-1} \in H$ (so that inverses occur on the right for right cosets).

We can use cosets to construct more examples of G-sets.

9.113. The G-set G/H. Let G be a group, and let H be *any* subgroup of G. Let G/H be the set of all distinct left cosets of H in G. Every element of G/H is a subset of G of the

form $xH = \{xh : h \in H\}$ for some $x \in G$ (which is usually not unique). So, G/H is a subset of the power set $\mathcal{P}(G)$. Let the group G act on the set $X = G/H$ by left multiplication:

$$g * S = \{gs : s \in S\} \qquad (g \in G, S \in X).$$

Note that this action is the restriction of the action from 9.85 to $G \times X$. To see that the action makes sense, we must check that X is a *G-stable subset* of $\mathcal{P}(G)$. Let xH be an element of X and let $g \in G$; then

$$g * (xH) = \{g(xh) : h \in H\} = \{(gx)h : h \in H\} = (gx)H \in X.$$

Let $[G : H] = |G/H|$ (which may be infinite); this cardinal number is called the *index of H in G*. Lagrange's Theorem (below) will show that $|G/H| = |G|/|H|$ when G is finite.

9.114. Remark. Using the coset equality theorems, one can show that $xH \mapsto Hx^{-1}$ gives a well-defined bijection between G/H and the set of right cosets of H in G. So, we would obtain the same number $[G : H]$ if we had used right cosets in the definition of G/H. It is more convenient to use left cosets here, so that G can act on G/H on the left.

9.115. Example. If $G = S_3$ and $H = \{\mathrm{id}, (1,2)\}$, then

$$G/H = \{\{\mathrm{id}, (1,2)\}, \{(1,3), (1,2,3)\}, \{(2,3), (1,3,2)\}\} = \{\mathrm{id}\, H, (1,3)H, (2,3)H\}.$$

We have $[G : H] = |G/H| = 3$. Note that $|G|/|H| = 6/2 = 3 = |G/H|$. This is a special case of Lagrange's Theorem, proved below.

To prepare for Lagrange's Theorem, we first show that every left coset of H in G has the same cardinality as G.

9.116. Coset Size Theorem. Let H be a subgroup of G. For all $x \in G$, $|xH| = |H|$.

Proof. We have seen that the left multiplication $L_x : G \to G$, given by $g \mapsto xg$ for $g \in G$, is a bijection (with inverse $L_{x^{-1}}$). Restricting the domain of L_x to H gives an injective map $L'_x : H \to G$. The image of this map is $\{xh : h \in H\} = xH$. So, restricting the codomain gives a bijection from H to xH. Thus, the sets H and xH have the same cardinality. □

9.117. Lagrange's Theorem. Let H be any subgroup of a finite group G. Then

$$[G : H] \cdot |H| = |G|.$$

So $|H|$ and $[G : H]$ are divisors of $|G|$, and $|G/H| = [G : H] = |G|/|H|$.

Proof. We know that G is the disjoint union of its distinct left cosets: $G = \bigcup_{S \in G/H} S$. By the previous theorem, $|S| = |H|$ for every $S \in G/H$. So, by the sum rule,

$$|G| = \sum_{S \in G/H} |S| = \sum_{S \in G/H} |H| = |G/H| \cdot |G| = [G : H] \cdot |H|. \qquad \square$$

9.118. Remark. The equality of cardinal numbers $|H| \cdot [G : H] = |G|$ holds even when G is infinite, with the same proof.

9.119. Theorem: Order of Group Elements. If G is a finite group of size n and $x \in G$, then the order of x is a divisor of n, and $x^n = e_G$.

Proof. Consider the subgroup $H = \langle x \rangle$ generated by x. The order d of x is $|H|$, which divides $|G| = n$ by Lagrange's theorem. Writing $n = cd$, we see that $x^n = (x^d)^c = e^c = e$. □

The next result gives an interpretation for cosets xK in the case where K is the kernel of a group homomorphism.

9.120. Theorem: Cosets of the Kernel of a Homomorphism. Let $f : G \to L$ be a group homomorphism with kernel K. For every $x \in G$,

$$xK = \{y \in G : f(y) = f(x)\} = Kx.$$

If G is finite and I is the image of f, it follows that $|G| = |K| \cdot |I|$.

Proof. Fix $x \in G$, and set $S = \{y \in G : f(y) = f(x)\}$. We will prove that $xK = S$. First suppose $y \in xK$, so $y = xk$ for some $k \in K$. Applying f, we find that $f(y) = f(xk) = f(x)f(k) = f(x)e_L = f(x)$, so $y \in S$. Next suppose $y \in S$, so $f(y) = f(x)$. Note that $f(x^{-1}y) = f(x)^{-1}f(y) = e$, so $x^{-1}y \in \ker(f) = K$. So $y = x(x^{-1}y) \in xK$. The proof that $S = Kx$ is analogous. To obtain the formula for $|G|$, note that G is the disjoint union

$$G = \bigcup_{z \in I} \{y \in G : f(y) = z\}.$$

Every $z \in I$ has the form $z = f(x)$ for some $x \in G$. So, by what we have just proved, each set appearing in the union is a coset of K, which has the same cardinality as K. So the sum rule gives $|G| = \sum_{z \in I} |K| = |K| \cdot |I|$. □

9.121. Corollary: Size of A_n. For $n > 1$, $|A_n| = n!/2$.

Proof. We know that $\text{sgn} : S_n \to \{1, -1\}$ is a surjective group homomorphism with kernel A_n. So $n! = |S_n| = |A_n| \cdot 2$. □

9.15 The Size of an Orbit

In 9.105, we saw that every G-set X breaks up into a disjoint union of orbits. This result suggests two combinatorial questions. First, given $x \in X$, what is the size of the orbit Gx? Second, how many orbits are there? We answer the first question here; the second question will be solved in §9.18.

The key to computing the orbit size $|Gx|$ is to relate the G-set Gx to one of the special G-sets G/H defined in 9.113. For this purpose, we need to associate a subgroup H of G to the given orbit Gx.

9.122. Definition: Stabilizers. Let $(X, *)$ be a G-set. For each $x \in X$, the *stabilizer of x in G* is

$$\text{Stab}(x) = \{g \in G : g * x = x\}.$$

Sometimes the notation G_x is used to denote $\text{Stab}(x)$.

The following calculations show that $\text{Stab}(x)$ is a *subgroup* of G for each $x \in X$: $e*x = x$; $g * x = x$ implies $x = g^{-1} * x$ for $g \in G$; $g * x = x = h * x$ implies $(gh) * x = g * (h * x) = x$ for $g, h \in G$.

9.123. Example. Let S_n act on $X = \{1, 2, \ldots, n\}$ via $f * x = f(x)$ for $f \in S_n$ and $x \in X$. The stabilizer of a point $i \in X$ consists of all permutations of X for which i is a fixed point. In particular, $\text{Stab}(n)$ consist of all bijections $f : X \to X$ with $f(n) = n$. Restricting the domain to $\{1, 2, \ldots, n-1\}$ defines a group isomorphism between $\text{Stab}(n)$ and S_{n-1}.

9.124. Example. Let a group G act on itself by left multiplication. Right cancellation of x shows that $gx = x$ iff $g = e$. Therefore, $\text{Stab}(x) = \{e\}$ for all $x \in G$. At the other extreme, we can let G act on any set X by declaring $g * x = x$ for all $g \in G$ and all $x \in X$. Relative to this action, $\text{Stab}(x) = G$ for all $x \in X$.

9.125. Example: Centralizers. Let G act on itself by conjugation: $g * x = gxg^{-1}$ for all $g, x \in G$. For a given $x \in G$, $g \in \text{Stab}(x)$ iff $gxg^{-1} = x$ iff $gx = xg$ iff g commutes with x. This stabilizer subgroup is often denoted $C_G(x)$ and called the *centralizer of x in G*. The intersection $\bigcap_{x \in G} C_G(x)$ consists of all $g \in G$ that commute with *every* $x \in G$. This is a subgroup called the *center* of G and denoted $Z(G)$.

9.126. Example: Normalizers. Let G be a group, and let X be the set of all subgroups of G. G acts on X by conjugation: $g * H = gHg^{-1} = \{ghg^{-1} : h \in H\}$. (Note that $g * H$ *is* a subgroup, since it is the image of a subgroup under the inner automorphism "conjugation by g"; cf. 9.74.) For this action, $g \in \text{Stab}(H)$ iff $gHg^{-1} = H$. This stabilizer subgroup is denoted $N_G(H)$ and called the *normalizer of H in G*. One may check that $N_G(H)$ always contains H.

9.127. Example. Let S_4 act on 4-tuples of integers by permuting the positions. Then $\text{Stab}((5,1,5,1)) = \{\text{id}, (1,3), (2,4), (1,3)(2,4)\}$; $\text{Stab}((2,2,2,2)) = S_4$; $\text{Stab}((1,2,3,4)) = \{\text{id}\}$; and $\text{Stab}((2,5,2,2))$ is a subgroup of S_4 isomorphic to $\text{Sym}(\{1,3,4\})$, which is in turn isomorphic to S_3.

The following fundamental theorem calculates the size of an orbit of a group action.

9.128. Theorem: Size of an Orbit. Let $(X, *)$ be a G-set. For each $x \in X$, there is a bijection $f : G/\text{Stab}(x) \to Gx$ given by $f(g\,\text{Stab}(x)) = g * x$ for all $g \in G$. So, when G is finite, *the size of the orbit of x is the index of the stabilizer of x, which is a divisor of $|G|$:*

$$|Gx| = [G : \text{Stab}(x)] = |G|/|\,\text{Stab}(x)|.$$

Proof. Write $H = \text{Stab}(x)$ for convenience. We first check that the function $f : G/H \to Gx$ is well defined. Assume $g, k \in G$ satisfy $gH = kH$; we must check that $g * x = k * x$. Now, $gH = kH$ means $k^{-1}g \in H = \text{Stab}(x)$, and hence $(k^{-1}g) * x = x$. Acting on both sides by k and simplifying, we obtain $g * x = k * x$. Second, is f one-to-one? Fix $g, k \in G$ with $f(gH) = f(kH)$; we must prove $gH = kH$. Now, $f(gH) = f(kH)$ means $g * x = k * x$. Acting on both sides by k^{-1}, we find that $(k^{-1}g) * x = x$, so $k^{-1}g \in H$, so $gH = kH$. Third, is f surjective? Given $y \in Gx$, the definition of Gx says that $y = g * x$ for some $g \in G$, so $y = f(gH)$. In summary, f is a well-defined bijection. □

9.129. Remark. One can prove a stronger version of the theorem, analogous to the "fundamental homomorphism theorem for groups," by introducing the following definition. Given two G-sets $(X, *)$ and $(Y, \bullet)$, a *G-map* is a function $p : X \to Y$ such that $p(g * x) = g \bullet p(x)$ for all $g \in G$ and all $x \in X$. A *G-isomorphism* is a bijective G-map. The theorem gives us a bijection p from the G-set Gx to the G-set $G/\text{Stab}(x)$ such that $p(g_0 * x) = g_0\,\text{Stab}(x)$. This bijection is in fact a G-isomorphism, because

$$p(g * (g_0 * x)) = p((gg_0) * x) = (gg_0)\,\text{Stab}(x) = g \bullet (g_0\,\text{Stab}(x)) = g \bullet p(g_0 * x).$$

Since every G-set is a disjoint union of orbits, this result shows that the special G-sets of the form G/H are the "building blocks" from which all G-sets are constructed.

Applying 9.128 to some of the preceding examples gives the following corollary.

9.130. Corollary: Counting Conjugates of Group Elements and Subgroups. The size of the conjugacy class of x in a finite group G is $[G : \text{Stab}(x)] = [G : C_G(x)] = |G|/|C_G(x)|$. If H is a subgroup of G, the number of distinct conjugates of H (subgroups of the form gHg^{-1}) is $[G : \text{Stab}(H)] = [G : N_G(H)] = |G|/|N_G(H)|$.

9.16 Conjugacy Classes in S_n

The conjugacy classes in the symmetric groups S_n can be described explicitly. We shall prove that the conjugacy class of $f \in S_n$ consists of all $g \in S_n$ with the same cycle type as f (see 9.20). The proof employs the following computational result.

9.131. Theorem: Conjugation in S_n. For $f, g \in S_n$, the permutation gfg^{-1} can be obtained by applying g to each entry in the disjoint cycle decomposition of f. In other words, if

$$f = (i_1, i_2, i_3, \ldots)(j_1, j_2, \ldots)(k_1, k_2, \ldots)\cdots,$$

then

$$gfg^{-1} = (g(i_1), g(i_2), g(i_3), \ldots)(g(j_1), g(j_2), \ldots)(g(k_1), g(k_2), \ldots)\cdots.$$

In particular, $\text{type}(gfg^{-1}) = \text{type}(f)$.

Proof. First assume f is a k-cycle, say $f = (i_1, i_2, \ldots, i_k)$. We prove that the functions gfg^{-1} and $h = (g(i_1), g(i_2), \ldots, g(i_k))$ are equal by showing that both have the same effect on every $x \in \{1, 2, \ldots, n\}$. We consider various cases. First, if $x = g(i_s)$ for some $s < k$, then $gfg^{-1}(x) = gfg^{-1}(g(i_s)) = g(f(i_s)) = g(i_{s+1}) = h(x)$. Second, if $x = g(i_k)$, then $gfg^{-1}(x) = g(f(i_k)) = g(i_1) = h(x)$. Finally, if x does not equal any $g(i_s)$, then $g^{-1}(x)$ does not equal any i_s. So f fixes $g^{-1}(x)$, and $gfg^{-1}(x) = g(g^{-1}(x)) = x = h(x)$.

In the general case, write $f = C_1 \circ C_2 \circ \cdots \circ C_t$ where each C_i is a cycle. Since conjugation by g is a homomorphism,

$$gfg^{-1} = (gC_1g^{-1}) \circ (gC_2g^{-1}) \circ \cdots \circ (gC_tg^{-1}).$$

By the previous paragraph, we can compute gC_ig^{-1} by applying g to each element of C_i. This completes the proof. □

9.132. Theorem: Conjugacy Classes of S_n. The conjugacy class of $f \in S_n$ consists of all $h \in S_n$ with $\text{type}(h) = \text{type}(f)$. The number of conjugacy classes is $p(n)$, the number of integer partitions of n.

Proof. Fix $f \in S_n$; let $T = \{gfg^{-1} : g \in S_n\}$ be the conjugacy class of f, and let $U = \{h \in S_n : \text{type}(h) = \text{type}(f)\}$. Using 9.131, we see that $T \subseteq U$. For the reverse inclusion, let $h \in S_n$ have the same cycle type of f. We give an algorithm for finding a $g \in S_n$ such that $h = gfg^{-1}$. Write down any complete cycle decomposition of f (including 1-cycles), writing longer cycles before shorter cycles. Immediately below this, write down a complete cycle decomposition of h. Now erase all the parentheses and regard the resulting array as the two-line form of a permutation g. Then 9.131 shows that $gfg^{-1} = h$. For example, suppose

$$\begin{aligned} f &= (1,7,3)(2,8,9)(4,5)(6) \\ h &= (4,9,2)(6,3,5)(1,8)(7) \end{aligned}$$

Then

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

The g constructed here is not unique; we could obtain different g's satisfying $gfg^{-1} = h$ by starting with a different complete cycle decomposition for f or h.

The last statement of the theorem follows since the possible cycle types of permutations of n objects are exactly the integer partitions of n (weakly decreasing sequences of positive integers that sum to n). □

We now apply 9.130 to determine the sizes of the conjugacy classes of S_n.

9.133. Definition: z_μ. Let μ be an integer partition of n consisting of a_1 ones, a_2 twos, etc. Define

$$z_\mu = 1^{a_1}2^{a_2}\cdots n^{a_n}a_1!a_2!\cdots a_n!.$$

For example, for $\mu = (3,3,2,2,2,2,1,1,1,1,1)$, we have $a_1 = 5$, $a_2 = 4$, $a_3 = 2$, and $z_\mu = 1^5 2^4 3^2 5!4!2! = 829,440$.

9.134. Theorem: Size of Conjugacy Classes of S_n. For each $\mu \in \mathrm{Par}(n)$, the number of permutations $f \in S_n$ with $\mathrm{type}(f) = \mu$ is $n!/z_\mu$.

Proof. Fix a particular $f \in S_n$ with $\mathrm{type}(f) = \mu$. By 9.130 and the fact that $|S_n| = n!$, it is enough to show that $|C_{S_n}(f)| = z_\mu$. The argument is most readily understood by consideration of a specific example. Let $\mu = (3,3,2,2,2,2,1,1,1,1,1)$ as above, and take

$$f = (1,2,3)(4,5,6)(7,8)(9,10)(11,12)(13,14)(15)(16)(17)(18)(19).$$

A permutation $g \in S_n$ lies in $C_{S_n}(f)$ iff $gfg^{-1} = f$ iff applying g to each symbol in the cycle decomposition above produces another cycle decomposition of f. So we are reduced to counting the number of ways of writing down a complete cycle decomposition of f such that longer cycles come before shorter cycles. Note that we have freedom to rearrange the order of all cycles of a given length, and we also have freedom to cyclically permute the entries in any given cycle of f. For example, we could permute the five 1-cycles of f in any of 5! ways; we could replace $(4,5,6)$ by one of the three cyclic shifts $(4,5,6)$ or $(5,6,4)$ or $(6,4,5)$; and so on. For this particular f, the product rule gives $2!4!5!3^2 2^4 1^5 = z_\mu$ different possible complete cycle decompositions. The argument for the general case is similar: the term $a_i!$ in z_μ accounts for permuting the a_i cycles of length i, while the term i^{a_i} accounts for the i possible cyclic shifts of each of the a_i cycles of length i. Multiplying these contributions gives z_μ, as desired. □

9.17 Applications of the Orbit Size Formula

When a finite group G acts on a finite set X, 9.128 asserts that the size of the orbit Gx is $|G|/|\mathrm{Stab}(x)|$, which is a divisor of $|G|$. We now use this fact to establish several famous theorems from algebra, number theory, and combinatorics.

9.135. Fermat's Little Theorem. For every integer $a > 0$ and every prime p, $a^p \equiv a \pmod{p}$.

Proof. Let $Y = \{1,2,\ldots,a\}$, and let $X = Y^p$ be the set of all p-tuples $(y_1,\ldots,y_p)$ of elements of Y. By the product rule, $|X| = a^p$. We know that S_p acts on X by permuting positions (see 9.89). Let $H = \langle(1,2,\ldots,p)\rangle$, which is a cyclic subgroup of S_p of size p. Restricting the action to H, we see that H acts on X by cyclically shifting positions. The only divisors of the prime p are 1 and p, so all orbits of X under the H-action have size 1 or p. Since X is the disjoint union of the orbits, $|X|$ is congruent modulo p to the number of orbits of size 1. But one sees immediately that $w = (y_1,\ldots,y_p)$ is in an orbit of size 1 iff all cyclic shifts of w are equal to w iff $y_1 = \cdots = y_p \in Y$. So there are precisely a orbits of size 1, as desired. □

9.136. Cauchy's Theorem. Suppose G is a finite group and p is a prime divisor of $|G|$. Then there exists an element $x \in G$ of order p.

Proof. As in the previous proof, the group $H = \langle (1, 2, \ldots, p) \rangle$ acts on the set G^p by cyclically permuting positions. Let X consist of all p-tuples $(g_1, \ldots, g_p) \in G^p$ such that $g_1 g_2 \cdots g_p = e$. We can build a typical element of X by choosing $g_1, \ldots, g_{p-1}$ arbitrarily from G; then we are forced to choose $g_p = (g_1 \cdots g_{p-1})^{-1}$ to achieve the condition $g_1 g_2 \cdots g_{p-1} g_p = e$. The product rule therefore gives $|X| = |G|^{p-1}$, which is a multiple of p.

We next claim that X is an *H-stable subset* of G^p. This means that for all $i \leq p$, $g_1 g_2 \cdots g_p = e$ implies $g_i g_{i+1} \cdots g_p g_1 \cdots g_{i-1} = e$. To prove this, multiply the equation $g_1 g_2 \cdots g_p = e$ by $(g_1 g_2 \cdots g_{i-1})^{-1}$ on the left and by $(g_1 g_2 \cdots g_{i-1})$ on the right. We now know that X is an H-set, so it is a union of orbits of size 1 and size p. Since $|X|$ is a multiple of p, the number of orbits of size 1 must be a multiple of p as well. Now, $(e, e, \ldots, e)$ is one orbit of size 1; so there must exist at least $p - 1 > 0$ additional orbits of size 1. By definition of the H-action, such an orbit looks like $(x, x, \ldots, x)$ where $x \neq e$. By definition of X, we must have $x^p = e$. Since p is prime, we have proved the existence of an element x of order p (in fact, we know there are at least $p - 1$ such elements). □

9.137. Lucas' Congruence for Binomial Coefficients. Suppose p is prime and $0 \leq k \leq n$ are integers. Let n and k have base-p expansions $n = \sum_{i \geq 0} n_i p^i$, $k = \sum_{i \geq 0} k_i p^i$, where $0 \leq n_i, k_i < p$ (see 5.5). Then

$$\binom{n}{k} \equiv \prod_{i \geq 0} \binom{n_i}{k_i} \pmod{p}, \tag{9.6}$$

where we set $\binom{0}{0} = 1$ and $\binom{a}{b} = 0$ whenever $b > a$.

Proof. Step 1: For all $j \geq 0$, $m \geq 0$, and p prime, we show that

$$\binom{m+p}{j} \equiv \binom{m}{j} + \binom{m}{j-p} \pmod{p}. \tag{9.7}$$

To prove this identity, let $X = \{1, 2, \ldots, m+p\}$, and let Y be the set of all j-element subsets of X. We know that $|Y| = \binom{m+p}{j}$. Consider the subgroup $G = \langle (1, 2, \ldots, p) \rangle$ of $\mathrm{Sym}(X)$, which is cyclic of size p. G acts on Y via $g \star S = \{g(s) : s \in S\}$ for $g \in G$ and $S \in Y$.

Y is a disjoint union of orbits under this action. Since every orbit has size 1 or p, $|Y|$ is congruent modulo p to the number M of orbits of size 1. We will show that $M = \binom{m}{j} + \binom{m}{j-p}$. The orbits of size 1 correspond to the j-element subsets S of X such that $g \star S = S$ for all $g \in G$. It is equivalent to require that $f \star S = S$ for the generator $f = (1, 2, \ldots, p)$ of G. Suppose S satisfies this condition, and consider two cases. Case 1: $S \cap \{1, 2, \ldots, p\} = \emptyset$. Since $f(x) = x$ for $x > p$, we have $f \star S = S$ for all such subsets S. Since S can be an arbitrary subset of the m-element set $\{p+1, \ldots, m+p\}$, there are $\binom{m}{j}$ subsets of this form. Case 2: $S \cap \{1, 2, \ldots, p\} \neq \emptyset$. Say $i \in S$ where $1 \leq i \leq p$. Applying f repeatedly and noting that $f \star S = S$, we see that $\{1, 2, \ldots, p\} \subseteq S$. The remaining $j - p$ elements of S can be chosen arbitrarily from the m-element set $\{p+1, \ldots, m+p\}$. So there are $\binom{m}{j-p}$ subsets of this form. Combining the two cases, we see that $M = \binom{m}{j} + \binom{m}{j-p}$.

Step 2: Assume p is prime, $a, c \geq 0$, and $0 \leq b, d < p$; we show that $\binom{ap+b}{cp+d} \equiv \binom{a}{c}\binom{b}{d} \pmod{p}$. This will follow from step 1 and the identity $\binom{n+1}{k} = \binom{n}{k} + \binom{n}{k-1}$ (see 2.25). We argue by induction on a. The base step is $a = 0$. If $a = 0$ and $c > 0$, both sides of the congruence are zero; if $a = 0 = c$, then both sides of the congruence are $\binom{b}{d}$. Assuming that the result holds for a given a (and all b, c, d), the following computation shows that it holds

for $a+1$:

$$\binom{(a+1)p+b}{cp+d} = \binom{(ap+b)+p}{cp+d} \equiv \binom{ap+b}{cp+d} + \binom{ap+b}{(c-1)p+d} \equiv \binom{a}{c}\binom{b}{d} + \binom{a}{c-1}\binom{b}{d}$$
$$= \left[\binom{a}{c} + \binom{a}{c-1}\right]\binom{b}{d} = \binom{a+1}{c}\binom{b}{d} \pmod{p}.$$

Step 3: We prove Lucas' congruence (9.6) by induction on n. If $k > n$, then $k_i > n_i$ for some i, so that both sides of the congruence are zero. From now on, assume $k \leq n$. The result holds in the base cases $0 \leq n < p$, since $n = n_0$, $k = k_0$, and all higher digits of the base p expansion are zero. For the induction step, note that $n = ap + n_0$, $k = cp + k_0$, where $a = \sum_{i \geq 0} n_{i+1} p^i$ and $c = \sum_{i \geq 0} k_{i+1} p^i$ in base p. (We obtain a and c from n and k, respectively, by chopping off the final base p digits n_0 and k_0.) By step 2 and induction, we have

$$\binom{n}{p} \equiv \binom{a}{c}\binom{n_0}{k_0} \equiv \binom{n_0}{k_0}\prod_{i \geq 1}\binom{n_i}{k_i} = \prod_{i \geq 0}\binom{n_i}{k_i} \pmod{p}. \qquad \square$$

9.138. Corollary. Given $a, b, p \in \mathbb{N}^+$ with p prime and p not dividing b,

$$p \text{ does not divide } \binom{p^a b}{p^a}.$$

Proof. Write $b = \sum_{i \geq 0} b_i p^i$ in base p. The base-p expansions of $p^a b$ and p^a are $p^a b = \cdots b_3 b_2 b_1 b_0 00 \cdots 0$ and $p^a = 1000 \cdots 0$, respectively, where each expansion ends in a zeroes. Since $b_0 \neq 0$ by hypothesis, Lucas' congruence gives

$$\binom{p^a b}{p^a} \equiv \binom{b_0}{1} = b_0 \not\equiv 0 \pmod{p}. \qquad \square$$

This corollary can also be proved directly, by writing out the fraction defining $\binom{p^a b}{p^a}$ and counting powers of p in numerator and denominator. We leave this as an exercise for the reader.

9.139. Sylow's First Theorem. Let G be a finite group of size $p^a b$, where p is prime, $a > 0$, and p does not divide b. There exists a subgroup H of G of size p^a.

Proof. Let X be the collection of all subsets of G of size p^a. We know from 1.42 that $|X| = \binom{p^a b}{p^a}$. By 9.138, p does not divide $|X|$. Now, G acts on X by left multiplication: $g * S = \{gs : s \in S\}$ for $g \in G$ and $S \in X$. (The set $g * S$ still has size p^a, since left multiplication by g is injective.) Not every orbit of X has size divisible by p, since $|X|$ itself is not divisible by p. Choose $T \in X$ such that $|GT| \not\equiv 0 \pmod{p}$. Let $H = \text{Stab}(T) = \{g \in G : g * T = T\}$, which is a subgroup of G. The size of the orbit of T is $|G|/|H| = p^a b/|H|$. This integer is not divisible by p, forcing $|H|$ to be a multiple of p^a. So $|H| \geq p^a$. To obtain the reverse inequality, let t_0 be any fixed element of T. Given any $h \in H$, $h * T = T$ implies $ht_0 \in T$. So the right coset $Ht_0 = \{ht_0 : h \in H\}$ is contained in T. We conclude that $|H| = |Ht_0| \leq |T| = p^a$. Thus H is a subgroup of size p^a (and T is in fact one of the right cosets of H). $\square$

9.18 The Number of Orbits

The following theorem, which is traditionally known as "Burnside's Lemma," allows us to count the number of orbits in a given G-set.

9.140. Orbit-Counting Theorem. Let a finite group G act on a finite set X. For each $g \in G$, let $\mathrm{Fix}(g) = \{x \in X : gx = x\}$ be the set of "fixed points" of g, and let N be the number of distinct orbits. Then

$$N = \frac{1}{|G|} \sum_{g \in G} |\operatorname{Fix}(g)|.$$

So *the number of orbits is the average number of fixed points of elements of G.*

Proof. Define $f : X \to \mathbb{R}$ by setting $f(x) = 1/|Gx|$ for each $x \in X$. We will compute $\sum_{x \in X} f(x)$ in two ways. Let $\{O_1, \ldots, O_N\}$ be the distinct orbits of the G-action. On one hand, grouping summands based on which orbit they are in, we get

$$\sum_{x \in X} f(x) = \sum_{i=1}^{N} \sum_{x \in O_i} f(x) = \sum_{i=1}^{N} \sum_{x \in O_i} \frac{1}{|O_i|} = \sum_{i=1}^{N} 1 = N.$$

On the other hand, 9.128 says that $|Gx| = |G|/|\operatorname{Stab}(x)|$. Therefore

$$\begin{aligned} \sum_{x \in X} f(x) &= \sum_{x \in X} \frac{|\operatorname{Stab}(x)|}{|G|} = \frac{1}{|G|} \sum_{x \in X} \sum_{g \in G} \chi(gx = x) \\ &= \frac{1}{|G|} \sum_{g \in G} \sum_{x \in X} \chi(gx = x) = \frac{1}{|G|} \sum_{g \in G} |\operatorname{Fix}(g)|. \quad \square \end{aligned}$$

We are finally ready to solve the counting problems involving symmetry that were mentioned in the introduction to this chapter. The strategy is to introduce a set of objects X on which a certain group of symmetries acts. Each orbit of the group action consists of a set of objects in X that get identified with one another when symmetries are taken into account. So the solution to the counting problem is the number of orbits, which may be calculated by the formula of the previous theorem.

9.141. Example: Counting Necklaces. How many ways can we build a five-bead circular necklace if there are seven available types of gemstones (repeats allowed) and all rotations of a given necklace are considered equivalent? We can model the set of necklaces (before accounting for symmetries) by the set of words $X = \{(y_1, y_2, y_3, y_4, y_5) : 1 \leq y_i \leq 7\}$. Now let $G = \langle (1,2,3,4,5) \rangle$ act on X by cyclically permuting positions (see 9.89). Every orbit of G consists of a set of necklaces that get identified with one another when symmetry is taken into account. To count the orbits, let us compute $|\operatorname{Fix}(g)|$ for each $g \in G$. First, $\mathrm{id} = (1)(2)(3)(4)(5)$ fixes every object in X, so $|\operatorname{Fix}(\mathrm{id})| = |X| = 7^5$ by the product rule. Second, the generator $g = (1,2,3,4,5)$ fixes $(y_1, y_2, y_3, y_4, y_5)$ iff

$$(y_1, y_2, y_3, y_4, y_5) = (y_5, y_1, y_2, y_3, y_4).$$

Comparing coordinates, this holds iff $y_1 = y_2 = y_3 = y_4 = y_5$ iff all the y_i's are equal to one another. So $|\operatorname{Fix}((1,2,3,4,5))| = 7$ since there are seven choices for y_1. Next, what is $|\operatorname{Fix}(g^2)|$? We have $g^2 = (1,3,5,2,4)$, so that g^2 fixes $(y_1, y_2, y_3, y_4, y_5)$ iff

$$(y_1, y_2, y_3, y_4, y_5) = (y_4, y_5, y_1, y_2, y_3),$$

which holds iff $y_1 = y_3 = y_5 = y_2 = y_4$. So $|\operatorname{Fix}(g^2)| = 7$. Similarly, $|\operatorname{Fix}(g^3)| = |\operatorname{Fix}(g^4)| = 7$, so the answer is

$$\frac{7^5 + 7 + 7 + 7 + 7}{5} = 3367.$$

Now suppose we are counting six-bead necklaces, identifying all rotations of a given necklace. Here, the group of symmetries is

$$G = \{\text{id},\ (1,2,3,4,5,6),\ (1,3,5)(2,4,6),\ (1,4)(2,5)(3,6),\ (1,5,3)(2,6,4),\ (1,6,5,4,3,2)\}.$$

As before, id has 7^6 fixed points, and each of the two six-cycles has 7 fixed points. What about $\text{Fix}((1,3,5)(2,4,6))$? We have

$$(1,3,5)(2,4,6) * (y_1,y_2,y_3,y_4,y_5,y_6) = (y_5,y_6,y_1,y_2,y_3,y_4),$$

and this equals $(y_1,\ldots,y_6)$ iff $y_1 = y_3 = y_5$ and $y_2 = y_4 = y_6$. Here there are 7 choices for y_1, 7 choices for y_2, and the remaining y_i's are then forced. So $|\text{Fix}((1,3,5)(2,4,6))| = 7^2$. Likewise, $|\text{Fix}((1,5,3)(2,6,4))| = 7^2$. Similarly, we find that $(y_1,\ldots,y_6)$ is fixed by $(1,4)(2,5)(3,6)$ iff $y_1 = y_4$ and $y_2 = y_5$ and $y_3 = y_6$, so that there are 7^3 such fixed points. In each case, $\text{Fix}(f)$ turned out to be $7^{\text{cyc}(f)}$ where $\text{cyc}(f)$ is the number of cycles in the complete cycle decomposition of f (including 1-cycles). The number of necklaces is

$$\frac{7^6 + 7 + 7^2 + 7^3 + 7^2 + 7}{6} = 19,684.$$

Now consider the question of counting five-bead necklaces using q types of beads, where rotations and reflections of a given necklace are considered equivalent. For this problem, the group of symmetries to use is the automorphism group of the cycle graph C_5 (see 9.65). In addition to the five powers of $(1,2,3,4,5)$, this group contains the following five permutations corresponding to reflections of the necklace:

$$(1,5)(2,4)(3),\ (1,4)(2,3)(5),\ (1,3)(4,5)(2),\ (1,2)(3,5)(4),\ (2,5)(3,4)(1).$$

The reader may check that each of the five new permutations has $q^3 = q^{\text{cyc}(f)}$ fixed points. For example, a necklace $(y_1,\ldots,y_5)$ is fixed by $(1,5)(2,4)(3)$ iff $y_1 = y_5$ (q choices) and $y_2 = y_4$ (q choices) and y_3 is arbitrary (q choices). So, the number of necklaces is

$$\frac{q^5 + 5q^3 + 4q^1}{10}.$$

The following general example can be used to solve many counting problems involving symmetry.

9.142. Example: Counting Colorings under Symmetries. Suppose V is a finite set of objects, C is a finite set of q colors, and $G \subseteq \text{Sym}(V)$ is a group of symmetries of the objects V. (For example, if V is the vertex set of a graph, we could take G to be the automorphism group of the graph.) G acts on V via $g \cdot x = g(x)$ for $g \in G$ and $x \in V$. Now let $X = {}^VC$ be the set of all functions $f : V \to C$. We think of a function f as a *coloring* of V such that x receives color $f(x)$ for all $x \in V$. As we saw in 9.88, G acts on X via $g * f = f \circ g^{-1}$ for $g \in G$ and $f \in X$. Informally, if f assigns color c to object x, then $g * f$ assigns color c to object $g(x)$. The G-orbits consist of colorings that get identified when we take into account the symmetries in G. So the number of colorings "up to symmetry" is $\frac{1}{|G|}\sum_{g\in G} |\text{Fix}(g)|$. In the previous example, we observed that $|\text{Fix}(g)| = q^{\text{cyc}(g)}$. To see why this holds in general, fix $g \in G$ and write a complete cycle decomposition $g = C_1C_2\cdots C_k$, so $k = \text{cyc}(g)$. Let V_i be the elements appearing in cycle C_i, so V is the disjoint union of the sets V_i. Consider C_1, for example. Say $C_1 = (x_1,x_2,\ldots,x_s)$, so that $V_1 = \{x_1,\ldots,x_s\}$. Suppose $f \in X$ is fixed by g, so $f = g * f$. Then

$$f(x_2) = (g * f)(x_2) = f(g^{-1}(x_2)) = f(x_1).$$

Similarly, $f(x_3) = f(x_2)$, and in general $f(x_{j+1}) = f(x_j)$ for all $j < s$. It follows that f is constant on V_1. Similarly, f is constant on every V_i in the sense that f assigns the same color to every $x \in V_i$. This argument is reversible, so $\mathrm{Fix}(g)$ consists precisely of the colorings $f \in {}^V C$ that are constant on each V_i. To build such an f, choose a common color for all the vertices in V_i (for $1 \leq i \leq k$). By the product rule, $|\mathrm{Fix}(g)| = q^k = q^{\mathrm{cyc}(g)}$ as claimed. Therefore, the answer to the counting problem is

$$\frac{1}{|G|} \sum_{g \in G} q^{\mathrm{cyc}(g)}. \tag{9.8}$$

9.143. Example: Counting Chessboards. We now answer the question posed at the beginning of this chapter: how many ways can we color a 5×5 chessboard with seven colors, if all rotations and reflections of a given colored board are considered the same? We apply the method of the preceding example. Let $B = (V, E)$ be the graph that models the chessboard (Figure 9.3). Let $C = \{1, 2, \ldots, 7\}$ be the set of colors, and let $X = {}^V C$ be the set of colorings before accounting for symmetry. The symmetry group $G = \mathrm{Aut}(B)$ was computed in 9.67. By inspecting the cycle decompositions for the eight elements $g \in G$, the answer follows from (9.8):

$$\frac{7^{25} + 7^7 + 7^{13} + 7^7 + 4 \cdot 7^{15}}{8} = 167,633,579,843,887,699,759.$$

9.19 Pólya's Formula

Consider the following variation of the chessboard coloring example: how many ways can we color a 5×5 chessboard so that 10 squares are red, 12 are blue, and 3 are green, if all rotations and reflections of a colored board are equivalent? We can answer questions like this with the aid of "Pólya's Formula," which extends Burnside's Lemma to *weighted* sets.

Let a finite group G act on a finite set X. Let $\{O_1, \ldots, O_N\}$ be the orbits of this action. Suppose each $x \in X$ has a weight $\mathrm{wt}(x)$ in some polynomial ring R, and suppose that the weights are *G-invariant:* $\mathrm{wt}(g * x) = \mathrm{wt}(x)$ for all $g \in G$ and all $x \in X$. This condition implies that every object in a given G-orbit has the same weight. So we can assign a well-defined weight to each orbit by letting $\mathrm{wt}(O_i) = \mathrm{wt}(x_i)$ for any $x_i \in O_i$. The next result lets us compute the generating function for the set of weighted orbits.

9.144. Orbit-Counting Theorem for Weighted Sets. With the above notation,

$$\sum_{i=1}^{N} \mathrm{wt}(O_i) = \frac{1}{|G|} \sum_{g \in G} \sum_{x \in \mathrm{Fix}(g)} \mathrm{wt}(x).$$

So, *the weighted sum of the orbits is the average over G of the weighted fixed point sets of elements of G.*

Proof. We adapt the proof of the original orbit-counting theorem to include weights. Define $f : X \to \mathbb{R}$ by setting $f(x) = \mathrm{wt}(x)/|Gx|$ for each $x \in X$. On one hand,

$$\sum_{x \in X} f(x) = \sum_{i=1}^{N} \sum_{x \in O_i} f(x) = \sum_{i=1}^{N} \sum_{x \in O_i} \frac{\mathrm{wt}(x)}{|O_i|} = \sum_{i=1}^{N} \sum_{x \in O_i} \frac{\mathrm{wt}(O_i)}{|O_i|} = \sum_{i=1}^{N} \mathrm{wt}(O_i).$$

On the other hand, using $|Gx| = |G|/|\operatorname{Stab}(x)|$, we get

$$\begin{aligned}\sum_{x\in X} f(x) &= \sum_{x\in X} \frac{|\operatorname{Stab}(x)|\operatorname{wt}(x)}{|G|} = \frac{1}{|G|}\sum_{x\in X}\sum_{g\in G}\chi(gx=x)\operatorname{wt}(x)\\ &= \frac{1}{|G|}\sum_{g\in G}\sum_{x\in X}\chi(gx=x)\operatorname{wt}(x) = \frac{1}{|G|}\sum_{g\in G}\sum_{x\in \operatorname{Fix}(g)}\operatorname{wt}(x). \quad \square\end{aligned}$$

We now extend the setup of 9.142 to count weighted colorings. We are given finite sets V and $C = \{1,\ldots,q\}$, a subgroup G of $\operatorname{Sym}(V)$, and the set of colorings $X = {}^VC$. G acts on X by permuting the domain: $g * f = f \circ g^{-1}$ for $g \in G$ and $f \in X$. We define a *weight* for a given coloring by setting

$$\operatorname{wt}(f) = \prod_{x\in V} z_{f(x)} \in \mathbb{R}[z_1, z_2, \ldots, z_q].$$

Note that $\operatorname{wt}(f) = z_1^{e_1}\cdots z_q^{e_q}$ iff f colors e_i of the objects in V with color i. We see that

$$\operatorname{wt}(g*f) = \prod_{x\in V} z_{f(g^{-1}(x))} = \prod_{v\in V} z_{f(v)} = \operatorname{wt}(f) \qquad (g\in G, f\in X)$$

by making the change of variable $v = g^{-1}(x)$. So the weighted orbit-counting theorem is applicable. In the unweighted case (see 9.142), we found that $|\operatorname{Fix}(g)| = q^{\operatorname{cyc}(g)}$ by arguing that $f \in \operatorname{Fix}(g)$ must be constant on each connected component $V_1,\ldots,V_k$ of the digraph of the permutation g. To take weights into account, let us construct such an f using the product rule for weighted sets. Suppose the components $V_1, V_2, \ldots, V_k$ have sizes $n_1 \geq n_2 \geq \cdots \geq n_k$ (so that $\operatorname{type}(g) = (n_1, n_2, \ldots, n_k)$). First choose a common color for the n_1 vertices in V_1. The generating function for this choice is $z_1^{n_1} + z_2^{n_1} + \cdots + z_q^{n_1}$; the term $z_i^{n_1}$ arises by coloring all n_1 vertices in V_1 with color i. Second, choose a common color for the n_2 vertices in V_2. The generating function for this choice is $z_1^{n_2} + \cdots + z_q^{n_2}$. Continuing similarly, we arrive at the formula

$$\sum_{x\in\operatorname{Fix}(g)} \operatorname{wt}(x) = \prod_{i=1}^{k}(z_1^{n_i} + z_2^{n_i} + \cdots + z_q^{n_i}).$$

We can abbreviate this formula by introducing the power-sum polynomials (which are studied in more detail in Chapter 10). For each integer $k \geq 1$, set $p_k(z_1,\ldots,z_q) = z_1^k + z_2^k + \cdots + z_q^k$. For each integer partition $\mu = (\mu_1, \mu_2, \ldots, \mu_k)$, set $p_\mu(z_1,\ldots,z_q) = \prod_{i=1}^k p_{\mu_i}(z_1,\ldots,z_q)$. Then the weighted orbit-counting formula assumes the following form.

9.145. Pólya's Formula. With the above notation, the generating function for weighted colorings with q colors relative to the symmetry group G is

$$\sum_{i=1}^{N}\operatorname{wt}(O_i) = \frac{1}{|G|}\sum_{g\in G} p_{\operatorname{type}(g)}(z_1, z_2, \ldots, z_q) \in \mathbb{R}[z_1,\ldots,z_q].$$

The coefficient of $z_1^{e_1}\cdots z_q^{e_q}$ in this polynomial is the number of colorings (taking the symmetries in G into account) in which color i is used e_i times.

9.146. Example. The generating function for five-bead necklaces using q types of beads (identifying all rotations and reflections of a given necklace) is

$$(p_{(1,1,1,1,1)} + 4p_{(5)} + 5p_{(2,2,1)})/10,$$

where all power-sum polynomials are evaluated at $(z_1,\ldots,z_q)$.

9.147. Example. Let us use Pólya's formula to count 5×5 chessboards with 10 red squares, 12 blue squares, and 3 green squares. We may as well take $q = 3$ here. Consulting the cycle decompositions in 9.67 again, we find that the group G has one element of type $(1^{25}) = (1, 1, \cdots, 1)$, two elements of type $(4^6, 1)$, one element of type $(2^{12}, 1)$, and four elements of type $(2^{10}, 1^5)$. Therefore, $\sum_{i=1}^{N} \mathrm{wt}(O_i)$ is given by

$$\frac{p_{(1^{25})}(z_1, z_2, z_3) + 2p_{(4^6,1)}(z_1, z_2, z_3) + p_{(2^{12},1)}(z_1, z_2, z_3) + 4p_{(2^{10},1^5)}(z_1, z_2, z_3)}{8}.$$

Using a computer algebra system, we can compute this polynomial and extract the coefficient of $z_1^{10} z_2^{12} z_3^3$. The final answer is $185,937,878$.

Summary

Table 9.2 summarizes some definitions from group theory used in this chapter. Table 9.3 contains definitions pertinent to the theory of group actions.

- *Examples of Groups.* (i) additive commutative groups: $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, and $\mathbb{Z}_n = \{0, 1, \ldots, n-1\}$ (under addition modulo n); (ii) multiplicative commutative groups: invertible elements in $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, and $\mathbb{Z}_n$; (iii) non-commutative groups: invertible matrices in $M_n(R)$, the group $\mathrm{Sym}(X)$ of bijections on X under composition, dihedral groups (automorphism groups of cycle graphs); (iv) constructions of groups: product groups (see 9.153), subgroups, quotient groups (see 9.205), cyclic subgroup generated by a group element, automorphism group of a graph, automorphism group of a group.

- *Basic Properties of Groups.* The identity of a group is unique, as is the inverse of each group element. In a group, there are left and right cancellation laws: $(ax = ay) \Rightarrow (x = y)$ and $(xa = ya) \Rightarrow (x = y)$; inverse rules: $(x^{-1})^{-1} = x$ and $(x_1 \cdots x_n)^{-1} = x_n^{-1} \cdots x_1^{-1}$; and the laws of exponents: $x^{m+n} = x^m x^n$; $(x^m)^n = x^{mn}$; and, when $xy = yx$, $(xy)^n = x^n y^n$.

- *Notation for Permutations.* A bijection $f \in S_n$ can be described in two-line form $\begin{pmatrix} 1 & 2 & \cdots & n \\ f(1) & f(2) & \cdots & f(n) \end{pmatrix}$, in one-line form $[f(1), f(2), \ldots, f(n)]$, or in cycle notation. The cycle notation is obtained by listing the elements going around each directed cycle in the digraph of f, enclosing each cycle in parentheses, and optionally omitting cycles of length 1. The cycle notation for f is not unique.

- *Sorting, Inversions, and Sign.* A permutation $w = w_1 w_2 \cdots w_n \in S_n$ can be sorted to the identity permutation $\mathrm{id} = 12 \cdots n$ by applying $\mathrm{inv}(w)$ basic transpositions to switch adjacent elements that are out of order. It follows that w can be written as the composition of $\mathrm{inv}(w)$ basic transpositions. Any factorization of w into a product of transpositions must involve an even number of terms when $\mathrm{sgn}(w) = +1$, or an odd number when $\mathrm{sgn}(w) = -1$. Sign is a group homomorphism: $\mathrm{sgn}(f \circ g) = \mathrm{sgn}(f) \cdot \mathrm{sgn}(g)$ for $f, g \in S_n$. The sign of a k-cycle is $(-1)^{k-1}$. For all $f \in S_n$, $\mathrm{sgn}(f) = (-1)^{n - \mathrm{cyc}(f)}$.

- *Properties of Determinants.* The determinant of a matrix $A \in M_n(R)$ is an R-multilinear, alternating function of the rows (resp. columns) of A such that $\det(I_n) = 1_R$. This means that $\det(A)$ is an R-linear function of any given row when the other rows are fixed, and the determinant is zero if A has two equal rows; similarly for columns.

TABLE 9.2
Definitions in group theory.

Concept	Definition
group axioms for $(G, \star)$	$\begin{cases} \forall x, y \in G, x \star y \in G \text{ (closure)} \\ \forall x, y, z \in G, x \star (y \star z) = (x \star y) \star z \text{ (associativity)} \\ \exists e \in G, \forall x \in G, x \star e = x = e \star x \text{ (identity)} \\ \forall x \in G, \exists y \in G, x \star y = e = y \star x \text{ (inverses)} \end{cases}$
commutative group	group G with $xy = yx$ for all $x, y \in G$
H is a subgroup of G	$\begin{cases} e_G \in H \text{ (closure under identity)} \\ \forall a, b \in H, ab \in H \text{ (closure under operation)} \\ \forall a \in H, a^{-1} \in H \text{ (closure under inverses)} \end{cases}$
H is *normal* in G ($H \trianglelefteq G$)	$\forall g \in G, \forall h \in H, ghg^{-1} \in H$ (closure under conjugation)
exponent notation in $(G, \cdot)$	$x^0 = 1_G$, $x^{n+1} = x^n \cdot x$, $x^{-n} = (x^{-1})^n$ $(n \geq 0)$
multiple notation in $(G, +)$	$0x = 0_G$, $(n+1)x = nx + x$, $(-n)x = n(-x)$ $(n \geq 0)$
k-cycle	$f \in \mathrm{Sym}(X)$ of the form $(i_1, i_2, \cdots, i_k)$ (cycle notation)
transposition	a 2-cycle (i, j)
basic transposition	a 2-cycle $(i, i+1)$ in S_n
$\mathrm{cyc}(f)$	number of components in digraph of $f \in \mathrm{Sym}(X)$
$\mathrm{type}(f)$	list of cycle lengths of $f \in \mathrm{Sym}(X)$ in decreasing order
$\mathrm{inv}(w_1 \cdots w_n)$	number of $i < j$ with $w_i > w_j$
$\mathrm{sgn}(w)$ for $w \in S_n$	$(-1)^{\mathrm{inv}(w)}$
determinant of $A \in M_n(R)$	$\det(A) = \sum_{w \in S_n} \mathrm{sgn}(w) \prod_{i=1}^n A(i, w(i))$
classical adjoint $\mathrm{adj}(A)$	$\mathrm{adj}(A)_{i,j} = (-1)^{i+j} \det(A[j\|i])$
cyclic subgroup $\langle x \rangle$	$\{x^n : n \in \mathbb{Z}\}$ or $\{nx : n \in \mathbb{Z}\}$ (additive notation)
cyclic group	group G such that $G = \langle x \rangle$ for some $x \in G$
order of $x \in G$	least $n > 0$ with $x^n = e_G$, or ∞ if no such n
graph automorphism of K	bijection on vertex set of K preserving edges of K
group homomorphism	map $f : G \to H$ with $f(xy) = f(x)f(y)$ for all $x, y \in G$
kernel of hom. $f : G \to H$	$\ker(f) = \{x \in G : f(x) = e_H\}$
image of hom. $f : G \to H$	$\mathrm{img}(f) = \{y \in H : y = f(x) \text{ for some } x \in G\}$
group isomorphism	bijective group homomorphism
group automorphism	group isomorphism from G to itself
inner automorphism C_g	the automorphism $x \mapsto gxg^{-1}$ $(g, x \in G)$

We have $\det(A^t) = \det(A)$. For triangular or diagonal A, $\det(A) = \prod_{i=1}^n A(i,i)$. The Laplace expansion for $\det(A)$ along row k (resp. column k) is

$$\det(A) = \sum_{j=1}^{n} (-1)^{j+k} A(k,j) \det(A[k|j]) = \sum_{i=1}^{n} (-1)^{i+k} A(i,k) \det(A[i|k]).$$

We have $A(\mathrm{adj}\, A) = (\det(A)) I_n = (\mathrm{adj}\, A)A$, so that $A^{-1} = (\det(A))^{-1} \mathrm{adj}(A)$ when $\det(A)$ is invertible in R. If $m \leq n$, A is $m \times n$, and B is $n \times m$, the *Cauchy-Binet formula* says

$$\det(AB) = \sum_{1 \leq j_1 < j_2 < \cdots < j_m \leq n} \det(A^{j_1}, \ldots, A^{j_m}) \det(B_{j_1}, \ldots, B_{j_m}),$$

where A^j is the jth column of A, and B_j is the jth row of B. In particular, $\det(AB) = \det(A)\det(B)$ for $A, B \in M_n(R)$.

- *Properties of Cyclic Groups.* Every cyclic group is commutative and isomorphic to $\mathbb{Z}$

TABLE 9.3
Definitions in the theory of group actions.

Concept	Definition
action axioms for G-set X	$\begin{cases} \forall g \in G, \forall x \in X, g * x \in X \text{ (closure)} \\ \forall x \in X, e_G * x = x \text{ (identity)} \\ \forall g, h \in G, \forall x \in X, g * (h * x) = (gh) * x \text{ (assoc.)} \end{cases}$
perm. representation of G on X	group homomorphism $R : G \to \mathrm{Sym}(X)$
G-stable subset Y of X	$\forall g \in G, \forall y \in Y, g * y \in Y$ (closure under action)
orbit of x in G-set X	$Gx = G * x = \{g * x : g \in G\}$
stabilizer of x rel. to G-set X	$\mathrm{Stab}(x) = \{g \in G : g * x = x\} \leq G$
fixed points of g in G-set X	$\mathrm{Fix}(g) = \{x \in X : g * x = x\}$
conjugacy class of x in G	$\{gxg^{-1} : g \in G\}$
centralizer of x in G	$C_G(x) = \{g \in G : gx = xg\} \leq G$
center of G	$Z(G) = \{g \in G : gx = xg \text{ for all } x \in G\} \trianglelefteq G$
normalizer of H in G	$N_G(H) = \{g \in G : gHg^{-1} = H\} \leq G$
left coset of H	$xH = \{xh : h \in H\}$
right coset of H	$Hx = \{hx : h \in H\}$
set of left cosets G/H	for $H \leq G$, $G/H = \{xH : x \in G\}$
index $[G : H]$	$[G : H] = \lvert G/H \rvert$ =number of left cosets of H in G

or $\mathbb{Z}_n$ for some $n \geq 1$. More precisely, if $G = \langle x \rangle$ is infinite, then $f : \mathbb{Z} \to G$ given by $f(i) = x^i$ for $i \in \mathbb{Z}$ is a group isomorphism. If $G = \langle x \rangle$ has size n, then $g : \mathbb{Z}_n \to G$ given by $g(i) = x^i$ for $i \in \mathbb{Z}_n$ is a group isomorphism; moreover, $x^m = e$ iff n divides m. Every subgroup of the additive group $\mathbb{Z}$ has the form $k\mathbb{Z}$ for a unique $k \geq 0$. Every subgroup of a cyclic group is cyclic.

- *Properties of Group Homomorphisms.* If $f : G \to H$ is a group homomorphism, then $\ker(f) \trianglelefteq G$ and $\mathrm{img}(f) \leq H$. Moreover, $f(x^n) = f(x)^n$ for all $x \in G$ and $n \in \mathbb{Z}$. The composition of group homomorphisms (resp. isomorphisms) is a group homomorphism (resp. isomorphism), and the inverse of a group isomorphism is a group isomorphism.

- *Main Results on Group Actions.* Actions $*$ of a group G on a set X correspond bijectively to permutation representations $R : G \to \mathrm{Sym}(X)$, via the formula $R(g)(x) = g * x$ for $g \in G$ and $x \in X$. Every G-set X is the disjoint union of orbits; more precisely, each $x \in X$ lies in a unique orbit Gx. The size of the orbit Gx is the index (number of cosets) of the stabilizer $\mathrm{Stab}(x)$ in G, which (for finite G) is a divisor of $|G|$. The number of orbits is the average number of fixed points of elements of G (for G finite); this extends to weighted sets where the weight is constant on each orbit.

- *Examples of Group Actions.* A subgroup H of a group G acts on G by left multiplication ($h * x = hx$), and by inverted right multiplication ($h * x = xh^{-1}$), and by conjugation ($h * x = hxh^{-1}$). The orbits of x under these respective actions are the right coset Hx, the left coset xH, and (when $H = G$) the conjugacy class of x in G. Similarly, G (or its subgroups) act on the set of all subsets of G by left multiplication, and G acts by conjugation on the set of subgroups of G. The set of subsets of a fixed size k are also G-sets under these actions. Centralizers of elements and normalizers of subgroups are stabilizers under suitable actions, hence subgroups of G. Any subgroup G of $\mathrm{Sym}(X)$ acts on X by $g * x = g(x)$ for $g \in G$ and $x \in X$. For any set X, S_n (or its subgroups) acts on X^n via $f \cdot (x_1, \ldots, x_n) = (x_{f^{-1}(1)}, \ldots, x_{f^{-1}(n)})$. For any subgroup H of G, G acts on G/H via $g * (xH) = (gx)H$ for $g, x \in G$.

- *Facts about Cosets.* Given a subgroup H of a group G, G is the disjoint union of its left (resp. right) cosets, which all have the same cardinality as H. This implies Lagrange's theorem: $|G| = |H| \cdot [G : H]$, so that (for finite G), the order and index of any subgroup of G are both divisors of $|G|$. To test equality of left cosets, one may check any of the following equivalent conditions: $xH = yH$; $x \in yH$; $x = yh$ for some $h \in H$; $y^{-1}x \in H$; $x^{-1}y \in H$. Similarly, $Hx = Hy$ iff $xy^{-1} \in H$ iff $yx^{-1} \in H$. Left and right cosets coincide (i.e., $xH = Hx$ for all $x \in G$) iff H is normal in G iff all conjugates xHx^{-1} equal H iff H is a union of conjugacy classes of G. Given a group homomorphism $f : G \to L$ with kernel K, $Kx = xK = \{y \in G : f(y) = f(x)\}$ for all $x \in G$.

- *Conjugacy Classes.* Every group G is the disjoint union of its conjugacy classes, where the conjugacy class of x is $\{gxg^{-1} : g \in G\}$. Conjugacy classes need not all have the same size. The size of the conjugacy class of x is the index $[G : C_G(x)]$, where $C_G(x)$ is the subgroup $\{y \in G : xy = yx\}$; this index is a divisor of $|G|$ for G finite. For $x \in G$, the conjugacy class of x has size 1 iff x is in the center $Z(G)$. This can be used to show that groups G of size p^n (where p is prime and $n \geq 1$) have $|Z(G)| > 1$. Each conjugacy class of S_n consists of those $f \in S_n$ with a given cycle type $\mu \in \mathrm{Par}(n)$. This follows from the fact that the cycle notation for gfg^{-1} is the cycle notation for f with each value x replaced by $g(x)$. The size of the conjugacy class indexed by μ is $n!/z_\mu$.

- *Cayley's Theorem on Permutation Representations.* Every group G is isomorphic to a subgroup of $\mathrm{Sym}(G)$, via the homomorphism sending $g \in G$ to the left multiplication $L_g = (x \mapsto gx : x \in G)$. Every n-element group is isomorphic to a subgroup of S_n.

- *Theorems Provable by Group Actions.* (i) Fermat's Little Theorem: $a^p \equiv a \pmod{p}$ for $a \in \mathbb{N}^+$ and p prime. (ii) Cauchy's Theorem: If G is a group and p is a prime divisor of $|G|$, then there exists $x \in G$ of order p. (iii) Lucas' Congruence: For $0 \leq k \leq n$ and prime p, $\binom{n}{k} \equiv \prod_{i \geq 0} \binom{n_i}{k_i} \pmod{p}$, where the n_i and k_i are the base-p digits of n and k. (iv) Sylow's First Theorem: If G is a group and $|G|$ has prime factorization $|G| = p_1^{n_1} \cdots p_k^{n_k}$, then G has a subgroup of size $p_i^{n_i}$ for $1 \leq i \leq k$.

- *Counting Colorings under Symmetries.* Given a finite set V, a group of symmetries $G \leq \mathrm{Sym}(V)$, and a set C of q colors, the number of colorings $f : V \to C$ taking symmetries into account is $|G|^{-1} \sum_{g \in G} q^{\mathrm{cyc}(g)}$. If the colors are weighted using $z_1, \ldots, z_q$, the generating function for weighted colorings is given by Pólya's formula

$$\frac{1}{|G|} \sum_{g \in G} p_{\mathrm{type}(g)}(z_1, \ldots, z_q),$$

  where p_μ is a power-sum symmetric polynomial. The coefficient of $z_1^{e_1} \cdots z_q^{e_q}$ gives the number of colorings (taking the symmetries in G into account) where color i is used e_i times.

Exercises

9.148. Let X be a set with more than one element. Define $a \star b = b$ for all $a, b \in X$. (a) Prove that $(X, \star)$ satisfies the closure axiom and associativity axiom in 9.1. (b) Does there exist $e \in X$ such that $e \star x = x$ for all $x \in X$? If so, is this e unique? (c) Does there exist $e \in X$ such that $x \star e = x$ for all $x \in X$? If so, is this e unique? (d) Is $(X, \star)$ a group?

9.149. Let G be the set of odd integers. For all $x, y \in G$, define $x \star y = x + y + 5$. Prove that $(G, \star)$ is a commutative group.

9.150. Let G be the set of real numbers unequal to 1. For each $a, b \in G$, define $a \star b = a + b - ab$. Prove that $(G, \star)$ is a commutative group.

9.151. Assume $(G, \star)$ is a group such that $x \star x = e$ for *all* $x \in G$, where e is the identity element of G. Prove that G is commutative.

9.152. Let $(G, \star)$ be a group. Define $\bullet : G \times G \to G$ by setting $a \bullet b = b \star a$ for all $a, b \in G$. Prove that $(G, \bullet)$ is a group.

9.153. Product Groups. Let $(G, \star)$ and $(H, \bullet)$ be groups. (a) Show that $G \times H$ becomes a group under the operation $(g_1, h_1) * (g_2, h_2) = (g_1 \star g_2, h_1 \bullet h_2)$ for $g_1, g_2 \in G$, $h_1, h_2 \in H$. (b) Show $G \times H$ is commutative iff G and H are commutative.

9.154. Prove the associative axiom for $(\mathbb{Z}_n, \oplus)$ by verifying (9.1).

9.155. Suppose G is a set, $\star : G \times G \to G$ is associative, and there exists $e \in G$ such that for all $x \in G$, $e \star x = x$ and there is $y \in G$ with $y \star x = e$. Prove $(G, \star)$ is a group.

9.156. For x, y in a group G, define the *commutator* $[x, y] = xyx^{-1}y^{-1}$, and let $C_x(y) = xyx^{-1}$ (conjugation by x). Verify that the following identities hold for all $x, y, z \in G$: (a) $[x, y]^{-1} = [y, x]$; (b) $[x, yz] = [x, y]C_y([x, z])$; (c) $[x, yz][y, zx][z, xy] = e_G$; (d) $[[x, y], C_y(z)][[y, z], C_z(x)][[z, x], C_x(y)] = e_G$.

9.157. Give complete proofs of the three laws of exponents in 9.10.

9.158. Let G be a group. For each $g \in G$, define a function $R_g : G \to G$ by setting $R_g(x) = xg$ for each $x \in G$. R_g is called "right multiplication by g." (a) Prove that R_g is one-to-one and onto. (b) Prove that $R_e = \mathrm{id}_G$ (where e is the identity of G) and $R_g \circ R_h = R_{hg}$ for all $g, h \in G$. (c) Point out why R_g is an element of $\mathrm{Sym}(G)$. Give two answers, one based on (a) and one based on (b). (d) Define $\phi : G \to \mathrm{Sym}(G)$ by setting $\phi(g) = R_g$ for $g \in G$. Prove that ϕ is one-to-one. (e) Prove that for all $g, h \in G$, $L_g \circ R_h = R_h \circ L_g$ (where L_g is left multiplication by g).

9.159. Let G be a group. (a) Prove that for all $a, b \in G$, there exists a unique $x \in G$ with $ax = b$. (b) Prove that in the multiplication table for a group G, every group element appears exactly once in each row and column.

9.160. A certain group $(G, \star)$ has a multiplication table that has been partly filled in below:

$\star$	1	2	3	4
1	4			
2		1		
3			1	
4				

Use properties of groups to fill in the rest of the table.

9.161. Let $f, g \in S_8$ be given in one-line form by $f = [3, 2, 7, 5, 1, 4, 8, 6]$ and $g = [4, 5, 1, 3, 2, 6, 8, 7]$. (a) Write f and g in cycle notation. (b) Compute $f \circ g$, $g \circ f$, $g \circ g$, and f^{-1}, giving final answers in one-line form.

9.162. Let $h = [4, 1, 3, 6, 5, 2]$ in one-line form. Compute $\mathrm{inv}(h)$ and $\mathrm{sgn}(h)$. Write h as a product of $\mathrm{inv}(h)$ basic transpositions.

9.163. Let $f = (1,3,6)(2,8)(4)(5,7)$ and $g = (5,4,3,2,1)(7,8)$. (a) Compute fg, gf, fgf^{-1}, and gfg^{-1}, giving all answers in cycle notation. (b) Compute $\operatorname{sgn}(f)$ and $\operatorname{sgn}(g)$ without counting inversions. (c) Find an $h \in S_8$ such that $hfh^{-1} = (1,2,3)(4,5)(6)(7,8)$; give the answer in two-line form.

9.164. Suppose that $f \in S_n$ has cycle type $\mu = (\mu_1, \ldots, \mu_k)$. What is the order of f?

9.165. The *support* of a bijection $f \in \operatorname{Sym}(X)$ is the set $\operatorname{supp}(f) = \{x \in X : f(x) \neq x\}$. Two permutations $f, g \in \operatorname{Sym}(X)$ are called *disjoint* iff $\operatorname{supp}(f) \cap \operatorname{supp}(g) = \emptyset$. (a) Prove that for all $x \in X$ and $f \in \operatorname{Sym}(X)$, $x \in \operatorname{supp}(f)$ implies $f(x) \in \operatorname{supp}(f)$. (b) Prove that *disjoint permutations commute*, i.e., for all disjoint $f, g \in \operatorname{Sym}(X)$, $f \circ g = g \circ f$. (c) Suppose $f \in \operatorname{Sym}(X)$ is given in cycle notation by $f = C_1 C_2 \cdots C_k$, where the C_i are cycles involving pairwise disjoint subsets of X. Show that the C_i's commute with one another, and prove carefully that $f = C_1 \circ C_2 \circ \cdots \circ C_k$ (cf. 9.19).

9.166. Prove 9.27.

9.167. (a) Verify the formula $(i_1, i_2, \ldots, i_k) = (i_1, i_2) \circ (i_2, i_3) \circ (i_3, i_4) \circ \cdots (i_{k-1}, i_k)$ used in the proof of 9.33. (b) Prove that every transposition has sign -1 by finding an explicit formula for (i, j) as a product of an odd number of basic transpositions (which have sign -1 by 9.26 with $w = \mathrm{id}$).

9.168. Given $f \in S_n$, what is the relationship between the one-line forms of f and $f \circ (i, j)$? What about f and $(i, j) \circ f$?

9.169. Let $f \in S_n$ and $h = (i, i+1) \circ f$. (a) Prove an analogue of 9.26 relating $\operatorname{inv}(f)$ to $\operatorname{inv}(h)$ and $\operatorname{sgn}(f)$ to $\operatorname{sgn}(h)$. (b) Use (a) to give another proof of the formula $\operatorname{sgn}(f \circ g) = \operatorname{sgn}(f)\operatorname{sgn}(g)$ that proceeds by induction on $\operatorname{inv}(f)$.

9.170. Prove that for $n \geq 3$, every $f \in A_n$ can be written as a product of 3-cycles.

9.171. Suppose an $n \times n$ matrix A is given in block form as $A = \begin{bmatrix} B & 0 \\ C & D \end{bmatrix}$, where B is $k \times k$, C is $(n-k) \times k$, D is $(n-k) \times (n-k)$, and 0 denotes a $k \times (n-k)$ block of zeroes. Prove that $\det(A) = \det(B)\det(D)$.

9.172. Algorithmic Complexity of Determinant Evaluation. Let $A \in M_n(F)$ where F is a field. (a) How many additions and multiplications in F are needed to compute $\det(A)$ directly from 9.37? (b) How many additions and multiplications in F are needed to compute $\det(A)$ recursively, using 9.48? (c) Explain how to use 9.41 and 9.47 to compute $\det(A)$ efficiently (using about cn^3 field operations for some constant c).

9.173. Permanents. The *permanent* of an $n \times n$ matrix $A \in M_n(R)$ is defined as $\operatorname{per}(A) = \sum_{w \in S_n} \prod_{i=1}^{n} A(i, w(i))$. Prove the following facts about permanents: (a) $\operatorname{per}(A^t) = \operatorname{per}(A)$; (b) if A is diagonal, then $\operatorname{per}(A) = \prod_{i=1}^{n} A(i,i)$; (c) $\operatorname{per}(I_n) = 1_R$; (d) $\operatorname{per}(A)$ is an R-multilinear function of the rows and columns of A (cf. 9.45); (e) if B is obtained from A by permuting the rows in any fashion, then $\operatorname{per}(B) = \operatorname{per}(A)$.

9.174. State and prove analogues for permanents of the Laplace expansions in 9.48.

9.175. Verify the characterization of R-linear maps stated in 9.44.

9.176. Complete the proof of 9.50 by showing that $(\operatorname{adj} A)A = \det(A) I_n$.

9.177. Cramer's Rule. Let $A \in M_n(R)$ where $\det(A)$ is invertible in R, let b be a given $n \times 1$ vector, and let $x = [x_1 \ \cdots \ x_n]^t$. Show that the unique solution of the linear system $Ax = b$ is given by $x_i = \det(A_i)/\det(A)$, where A_i is the matrix obtained from A by replacing the ith column by b.

9.178. Verify the Cauchy-Binet formula for the matrices

$$A = \begin{bmatrix} 2 & 1 & 0 & 3 \\ 1 & -1 & 1 & 2 \\ 4 & 0 & 2 & 1 \end{bmatrix}, \qquad B = \begin{bmatrix} -1 & 1 & -1 \\ 0 & 2 & 5 \\ 1 & 1 & 4 \\ -2 & 0 & -1 \end{bmatrix}.$$

9.179. Consider a function $w : \{1, 2, \ldots, k\} \to \{1, 2, \ldots, n\}$, which we regard as a word $w = w_1 w_2 \cdots w_k$. Show that there exist basic transpositions $t_1, \ldots, t_m \in S_k$ such that $w \circ (t_1 t_2 \cdots t_m)$ is a weakly increasing word, and the minimum possible value of m is $\mathrm{inv}(w) = \sum_{i<j} \chi(w_i > w_j)$.

9.180. Let A and B be $n \times n$ matrices. Prove that $\det(AB) = \det(A)\det(B)$ by imitating (and simplifying) the proof of the Cauchy-Binet formula 9.53.

9.181. Verify all the assertions in 9.57.

9.182. Let x be an element of a group G, written multiplicatively. Use the laws of exponents to verify that $\langle x \rangle = \{x^n : n \in \mathbb{Z}\}$ is a subgroup of G, as stated in 9.60.

9.183. Subgroup Generated by a Set. Let S be a nonempty subset of a group G. Let $\langle S \rangle$ be the set of elements of G of the form $x_1 x_2 \cdots x_n$, where $n \in \mathbb{N}^+$ and, for $1 \leq i \leq n$, either $x_i \in S$ or $x_i^{-1} \in S$. Prove that $\langle S \rangle \leq G$, and for all T with $S \subseteq T \leq G$, $\langle S \rangle \leq T$.

9.184. Prove that every subgroup of a cyclic group is cyclic.

9.185. For subsets S and T of a multiplicative group G, define $ST = \{st : s \in S, t \in T\}$. (a) Show that if $S \trianglelefteq G$ and $T \leq G$, then $ST = TS$ and $ST \leq G$. Give an example to show ST may not be normal in G. (b) Show that if $S \trianglelefteq G$ and $T \trianglelefteq G$, then $ST \trianglelefteq G$. (c) Give an example of a group G and subgroups S and T such that ST is not a subgroup of G.

9.186. Let S and T be finite subgroups of a group G. Prove that $|S| \cdot |T| = |ST| \cdot |S \cap T|$.

9.187. Assume that G is a group and $H \leq G$. Let $H^{-1} = \{h^{-1} : h \in H\}$. (a) Show that $HH = H^{-1} = HH^{-1} = H$. (b) Prove that $H \trianglelefteq G$ iff $gHg^{-1} = H$ for all $g \in G$.

9.188. Show that a subgroup H of a group G is normal in G iff H is a union of conjugacy classes of G.

9.189. Find all the subgroups of S_4. Which subgroups are normal? Confirm that Sylow's theorem 9.139 is true for this group.

9.190. Find all *normal* subgroups of S_5, and prove that you have found them all (Lagrange's theorem and 9.188 can be helpful here).

9.191. Suppose H is a *finite*, nonempty subset of a group G such that $xy \in H$ for all $x, y \in H$. Prove that $H \leq G$. Give an example to show this result may not be true if H is not finite.

9.192. Given any simple graph or digraph K with vertex set X, show that $\mathrm{Aut}(K)$ is a subgroup of $\mathrm{Sym}(X)$.

9.193. Determine the automorphism groups of the following graphs and digraphs: (a) the path graph P_n (see 3.124); (b) the complete graph K_n (see 3.124); (c) the empty graph on $\{1, 2, \ldots, n\}$ with no edges; (d) the directed cycle with vertex set $\{1, 2, \ldots, n\}$ and edges $(n, 1)$ and $(i, i+1)$ for $i < n$; (e) the simple graph with vertex set $\{\pm 1, \pm 2, \ldots, \pm n\}$ and edge set $\{\{i, -i\} : 1 \leq i \leq n\}$.

9.194. Let K be the Petersen graph (defined in 3.215). (a) Given two paths $P = (y_0, y_1, y_2, y_3)$ and $Q = (z_0, z_1, z_2, z_3)$ in K, prove that there exists a unique automorphism of K that maps y_i to z_i for $0 \leq i \leq 3$. (b) Prove that K has exactly $5! = 120$ automorphisms. (c) Is $\mathrm{Aut}(K)$ isomorphic to S_5?

9.195. Let Q_k be the simple graph with vertex set $V = \{0,1\}^k$ and edge set $E = \{(v, w) \in V : v, w \text{ differ in exactly one position}\}$. Q_k is called a *k-dimensional hypercube.* (a) Compute $|V(Q_k)|$, $|E(Q_k)|$, and $\deg(Q_k)$. (b) Show that Q_k has exactly $\binom{k}{i}2^{k-i}$ induced subgraphs isomorphic to Q_i. (c) Find all the automorphisms of Q_k. How many are there?

9.196. (a) Construct an *undirected* graph whose automorphism group has size three. What is the minimum number of vertices in such a graph? (b) For each $n \geq 1$, construct an undirected graph whose automorphism group is cyclic of size n.

9.197. Let G be a simple graph with connected components $C_1, \ldots, C_k$. What is the relation between $|\mathrm{Aut}(G)|$ and $(|\mathrm{Aut}(C_i)| : 1 \leq i \leq k)$?

9.198. Let $f : G \to H$ be a group homomorphism. (a) Show that if $K \leq G$, then $f[K] = \{f(x) : x \in K\}$ is a subgroup of H. If $K \trianglelefteq G$, must $f[K]$ be normal in H? (b) Show that if $L \leq H$, then $f^{-1}[L] = \{x \in G : f(x) \in L\}$ is a subgroup of G. If $L \trianglelefteq H$, must $f^{-1}[L]$ be normal in G? (c) Deduce from (a) and (b) that the kernel and image of a group homomorphism are subgroups.

9.199. Show that the group of nonzero complex numbers under multiplication is isomorphic to the product of the subgroups $\mathbb{R}^+$ and $\{z \in \mathbb{C} : |z| = 1\}$.

9.200. Give examples of four non-isomorphic groups of size 12.

9.201. Suppose G is a commutative group with subgroups H and K, such that $G = HK$ and $H \cap K = \{e_G\}$. (a) Prove that the map $(h, k) \mapsto hk$ is a group isomorphism from $H \times K$ onto G. (b) Does any analogous result hold if G is not commutative? What if H and K are normal in G?

9.202. (a) Let G be a group and $x \in G$. Show there exists a unique group homomorphism $f : \mathbb{Z} \to G$ with $f(1) = x$. (b) Use (a) to determine the group $\mathrm{Aut}(\mathbb{Z})$.

9.203. (a) Suppose G is a group, $x \in G$, and $x^n = e_G$ for some $n \geq 2$. Show there exists a unique group homomorphism $f : \mathbb{Z}_n \to G$ with $f(1) = x$. (b) Use (a) to prove that $\mathrm{Aut}(\mathbb{Z}_n)$ is isomorphic to the group $\mathbb{Z}_n^*$ of invertible elements of $\mathbb{Z}_n$ under multiplication modulo n.

9.204. Properties of Order. Let G be a group and $x \in G$. (a) Prove x and x^{-1} have the same order. (b) Show that if x has infinite order, then so does x^i for all nonzero integers i. (c) Suppose x has finite order n. Show that the order of x^k is $n/\gcd(k, n)$ for all $k \in \mathbb{Z}$. (d) Show that if $f : G \to H$ is a group isomorphism, then x and $f(x)$ have the same order. What can be said if f is only a group homomorphism?

9.205. Quotient Groups. (a) Suppose H is a *normal* subgroup of G. Show that the set G/H of left cosets of H in G becomes a group of size $[G : H]$ if we define $(xH) \star (yH) = (xy)H$ for all $x, y \in G$. (One must first show that this operation is *well-defined*: i.e., for all $x_1, x_2, y_1, y_2 \in G$, $x_1H = x_2H$ and $y_1H = y_2H$ imply $x_1y_1H = x_2y_2H$. For this, use the coset equality theorem.) (b) With the notation in (a), define $\pi : G \to G/H$ by $\pi(x) = xH$ for $x \in G$. Show that π is a surjective group homomorphism with kernel H. (c) Let $H = \{\mathrm{id}, (1,2)\} \leq S_3$. Find $x_1, x_2, y_1, y_2 \in S_3$ with $x_1H = x_2H$ and $y_1H = y_2H$, but $x_1y_1H \neq x_2y_2H$. This shows that normality of H is needed for the product in (a) to be well defined.

9.206. Let H be a normal subgroup of a group G. (a) Prove that G/H is commutative if G is commutative. (b) Prove that G/H is cyclic if G is cyclic. (c) Does the converse of (a) or (b) hold? Explain.

9.207. Fundamental Homomorphism Theorem for Groups. Suppose G and H are groups and $f : G \to H$ is a group homomorphism. Let $K = \{x \in G : f(x) = e_H\}$ be the kernel of f, and let $I = \{y \in H : \exists x \in G, y = f(x)\}$ be the image of f. Show that $K \trianglelefteq G$, $I \leq H$ and there exists a unique group isomorphism $\overline{f} : G/K \to I$ given by $\overline{f}(xK) = f(x)$ for $x \in G$.

9.208. Universal Mapping Property for Quotient Groups. Let G be a group with normal subgroup N, let $\pi : G \to G/N$ be the homomorphism $\pi(x) = xN$ for $x \in G$, and let H be any group. (a) Show that if $h : G/N \to H$ is a group homomorphism, then $h \circ \pi$ is a group homomorphism from G to H sending each $n \in N$ to e_H. (b) Conversely, given any group homomorphism $f : G \to H$ such that $f(n) = e_H$ for all $n \in N$, show that there exists a unique group homomorphism $h : G/N \to H$ such that $f = h \circ \pi$. (c) Conclude that the map $h \mapsto h \circ \pi$ is a *bijection* from the set of all group homomorphisms from G/N to H to the set of all group homomorphisms from G to H that map everything in N to e_H.

9.209. Diamond Isomorphism Theorem for Groups. Suppose G is a group, $S \trianglelefteq G$, and $T \leq G$. Show (cf. 9.185) $TS = ST \leq G$, $S \trianglelefteq TS$, $(S \cap T) \trianglelefteq T$, and there is a well-defined group isomorphism $f : T/(S \cap T) \to (TS)/S$ given by $f(x(S \cap T)) = xS$ for all $x \in T$. Use this to give another solution of 9.186 in the case where S is *normal* in G.

9.210. Double-Quotient Isomorphism Theorem for Groups. Assume $A \leq B \leq C$ are groups with A and B both normal in C. Show that $A \trianglelefteq B$, $B/A \trianglelefteq C/A$, and $(C/A)/(B/A)$ is isomorphic to C/B via the map $(xA)B/A \mapsto xB$ for $x \in C$.

9.211. Correspondence Theorem for Quotient Groups. Let H be a normal subgroup of a group G. Let X be the set of subgroups of G containing H, and let Y be the set of subgroups of G/H. Show that the map $L \mapsto L/H = \{xH : x \in L\}$ is an inclusion-preserving bijection of X onto Y with inverse $M \mapsto \{x \in G : xH \in M\}$. If L maps to M under this correspondence, show that $[G : L] = [G/H : M]$, that $[L : H] = |M|$, that $L \trianglelefteq G$ iff $M \trianglelefteq G/H$, and that G/L is isomorphic to $(G/H)/M$ whenever $L \trianglelefteq G$.

9.212. Let G be a non-commutative group. Show that the rule $g \cdot x = xg$ (for $g, x \in G$) does not define a left action of G on the set G.

9.213. Let G act on itself by conjugation: $g * x = gxg^{-1}$ for $g, x \in G$. Verify that the axioms for a left group action are satisfied.

9.214. Let $(X, *)$ be a G-set and $f : K \to G$ a group homomorphism. Verify the K-set axioms for the action $k \bullet x = f(k) * x$ ($k \in K$, $x \in X$).

9.215. Suppose $* : G \times X \to X$ is a group action. (a) Show that $\mathcal{P}(X)$ is a G-set via the action $g \bullet S = \{g * s : s \in S\}$ for $g \in G$ and $S \in \mathcal{P}(X)$. (b) For fixed k, show that the set of all k-element subsets of X is a G-stable subset of $\mathcal{P}(X)$.

9.216. Verify the action axioms for the action of S_n on V in 9.87.

9.217. Suppose $(X, *)$ is a G-set and W is a set. Show that the set of functions $F : W \to X$ is a G-set via the action $(g \bullet F)(w) = g * (F(w))$ for all $g \in G$, $F \in {}^W X$, and $w \in W$.

9.218. Let a subgroup H of a group G act on G via $h * x = xh^{-1}$ for $h \in H$ and $x \in G$. Show that the orbit $H * x$ is the left coset xH, for $x \in G$.

9.219. (a) Suppose $f : X \to Y$ is a bijection. Show that the map $T : \mathrm{Sym}(X) \to \mathrm{Sym}(Y)$ given by $T(g) = f \circ g \circ f^{-1}$ for $g \in \mathrm{Sym}(X)$ is a group isomorphism. (b) Use (a) and Cayley's theorem to conclude that every n-element group is isomorphic to a subgroup of S_n.

9.220. Let a group G act on a set X. Show that $\bigcap_{x \in X} \mathrm{Stab}(x)$ is a *normal* subgroup of G. Give an example to show that a stabilizer subgroup $\mathrm{Stab}(x)$ may not be normal in G.

9.221. Let G act on itself by conjugation. (a) By considering the associated permutation representation and using the fundamental homomorphism theorem 9.207, deduce that $G/Z(G)$ is isomorphic to the subgroup of inner automorphisms in $\mathrm{Aut}(G)$. (b) Show that the subgroup of inner automorphisms is *normal* in $\mathrm{Aut}(G)$.

9.222. Let $(\mathbb{R}, +)$ act on $\mathbb{R}^2$ (viewed as column vectors) by the rule

$$\theta * \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \qquad (\theta, x, y \in \mathbb{R}).$$

Verify that this is an action, and describe the orbit and stabilizer of each point in $\mathbb{R}^2$.

9.223. Let $f \in S_n$, and let $\langle f \rangle$ act on $\{1, 2, \ldots, n\}$ via $g \cdot x = g(x)$ for all $g \in \langle f \rangle$ and $x \in \{1, 2, \ldots, n\}$. Prove that the orbits of this action are the connected components of the digraph of f.

9.224. Suppose X is a G-set and $x, y \in X$. Without appealing to equivalence relations, give a direct proof that $Gx \cap Gy \neq \emptyset$ implies $Gx = Gy$.

9.225. Let $*$ be a right action of a group G on a set X. (a) Prove that X is the disjoint union of orbits $x * G$. (b) Prove that $|x * G| = [G : \mathrm{Stab}(x)]$, where $\mathrm{Stab}(x) = \{g \in G : x * g = x\}$.

9.226. State and prove a version of the coset equality theorem 9.111 for right cosets.

9.227. Let G be a group with subgroup H. Prove that the map $T(xH) = Hx^{-1}$ for $x \in G$ is a well-defined bijection from the set of left cosets of H in G onto the set of right cosets of H in G.

9.228. Let X be a G-set. For $x \in X$ and $g \in G$, prove that $g\,\mathrm{Stab}(x) = \{h \in G : h * x = g * x\}$. (This shows that each left coset of the stabilizer of x consists of those group elements sending x to a particular element in its orbit Gx. Compare to 9.120.)

9.229. Let G be a group with subgroup H. Prove the following facts about the normalizer of H in G (see 9.126). (a) $N_G(H)$ contains H; (b) $H \trianglelefteq N_G(H)$; (c) for any $L \leq G$ such that $H \trianglelefteq L$, $L \leq N_G(H)$; (d) $H \trianglelefteq G$ iff $N_G(H) = G$.

9.230. Let X be a G-set. Prove: for $g \in G$ and $x \in X$, $\mathrm{Stab}(gx) = g\,\mathrm{Stab}(x)g^{-1}$.

9.231. Let H and K be subgroups of a group G. Prove that the G-sets G/H and G/K are isomorphic (as defined in 9.129) iff H and K are conjugate subgroups of G (i.e., $K = gHg^{-1}$ for some $g \in G$).

9.232. Calculate z_μ for every $\mu \in \mathrm{Par}(6)$.

9.233. Explicitly write down all elements in the centralizer of $g = (2,4,7)(1,6)(3,8)(5) \in S_8$. How large is this centralizer? How large is the conjugacy class of g?

9.234. Suppose $f = (2,4,7)(8,10,15)(1,9)(11,12)(17,20)(18,19)$ and $g = (7,8,9)(1,4,5)(11,20)(2,6)(3,18)(13,19)$. How many $h \in S_{20}$ satisfy $h \circ f = g \circ h$?

9.235. Find all integer partitions μ of n for which $z_\mu = n!$. Use your answer to calculate $Z(S_n)$ for all $n \geq 1$.

9.236. Prove that for all $n \geq 1$, $p(n) = \frac{1}{n!} \sum_{f \in S_n} z_{\text{type}(f)}$.

9.237. Conjugacy Classes of A_n. For $f \in A_n$, write $[f]_{A_n}$ (resp. $[f]_{S_n}$) to denote the conjugacy class of f in A_n (resp. S_n). (a) Show that $[f]_{A_n} \subseteq [f]_{S_n}$ for all $f \in A_n$. (b) Prove: for all $f \in A_n$, if there exists $g \in S_n \sim A_n$ with $fg = gf$, then $[f]_{A_n} = [f]_{S_n}$; but if no such g exists, then $[f]_{S_n}$ is the disjoint union of $[f]_{A_n}$ and $[(1,2) \circ f \circ (1,2)]_{A_n}$, and the latter two conjugacy classes are equal in size. (c) What are the conjugacy classes of A_5? How large are they? Use this to prove that A_5 is *simple*, i.e., the only normal subgroups of A_5 are $\{\text{id}\}$ and A_5.

9.238. Suppose G is a finite group and p is a prime divisor of $|G|$. Show that the number of elements in G of order p is congruent to $-1 \pmod{p}$.

9.239. (a) Compute $\binom{8936}{5833} \bmod 7$. (b) Compute $\binom{843}{212} \bmod 10$.

9.240. Prove 9.138 without using Lucas' congruence, by counting powers of p in the numerator and denominator of $\binom{p^a b}{p^a} = (p^a b)\downarrow_{p^a} / (p^a)!$.

9.241. Class Equation. Let G be a finite group with center $Z(G)$ (see 9.125), and let $x_1, \ldots, x_k \in G$ be such that each conjugacy class of G of size greater than 1 contains exactly one x_i. Prove that $|G| = |Z(G)| + \sum_{i=1}^{k} [G : C_G(x_i)]$, where each term in the sum is a divisor of $|G|$ greater than 1.

9.242. A *p-group* is a finite group of size p^e for some $e \geq 1$. Prove that every p-group G has $|Z(G)| > 1$.

9.243. Wilson's Theorem. Use group actions to prove that if an integer $p > 1$ is prime, then $(p-1)! \equiv -1 \pmod{p}$. Is the converse true?

9.244. How many ways are there to color an $n \times n$ chessboard with q possible colors if: (a) no symmetries are allowed; (b) rotations of a given board are considered equivalent; (c) rotations and reflections of a given board are considered equivalent?

9.245. Consider an $m \times n$ chessboard where $m \neq n$. (a) Describe all symmetries of this board. (b) How many ways can we color such a board with q possible colors?

9.246. How many n-letter words can be made using a k-letter alphabet if we identify each word with its reversal?

9.247. Consider necklaces that can use q kinds of gemstones, where rotations and reflections of a given necklace are considered equivalent. How many such necklaces are there with: (a) eight stones; (b) nine stones; (c) n stones?

9.248. Taking rotational symmetries into account, how many ways can we color the vertices of a regular tetrahedron with 7 available colors?

9.249. Taking rotational symmetries into account, how many ways can we color the vertices of a cube with 8 available colors?

9.250. Taking rotational symmetries into account, how many ways can we color the faces of a cube with q available colors?

9.251. Taking rotational symmetries into account, how many ways can we color the edges of a cube with q available colors?

9.252. Taking all symmetries into account, how many ways are there to color the vertices of the cycle C_3 with three *distinct* colors chosen from a set of five colors?

9.253. Taking all symmetries into account, how many ways are there to color the vertices of the cycle C_6 so that three vertices are blue, two are red, and one is yellow?

9.254. Taking rotational symmetries into account, how many ways are there to color the vertices of a regular tetrahedron so that: (a) two are blue and two are red; (b) one is red, one is blue, one is green, and one is yellow?

9.255. Taking rotational symmetries into account, how many ways are there to color the vertices of a cube so that four are blue, two are red, and two are green?

9.256. Taking rotational symmetries into account, how many ways are there to color the faces of a cube so that: (a) three are red, two are blue, and one is green; (b) two are red, two are blue, one is green, and one is yellow?

9.257. Taking rotational symmetries into account, how many ways are there to color the edges of a cube so that four are red, four are blue, and four are yellow?

9.258. How many ways can we color a 4×4 chessboard with five colors (identifying rotations of a given board) if each color must be used at least once?

9.259. How many ways can we build an eight-stone necklace using five kinds of gems (identifying rotations and reflections of a given necklace) if each type of gem must be used at least once?

Notes

For a more detailed development of group theory, we recommend the excellent book by Rotman [119]. More information on groups, rings, and fields may be found in textbooks on abstract algebra such as [29, 70, 71]. Many facts about matrices and determinants, including the Cauchy-Binet formula, appear in the matrix theory text by Lancaster [82]. The proof of Cauchy's theorem given in 9.136 is due to McKay [91]. The proof of Lucas' congruence in 9.137 is due to Sagan [120]. The proof of Sylow's theorem given in 9.139 is usually attributed to Wielandt [137], although Miller [92] gave a proof in a similar spirit over 40 years earlier. Proofs of Fermat's little theorem and Wilson's theorem using group actions were given by Peterson [103].