CHAPTER 1

# An Introduction to Groups

While we have no intention of presenting a comprehensive treatment of group theory in this text, there are a number of definitions that will facilitate a rigorous description of vector spaces. Furthermore, the concepts from abstract algebra that we shall introduce will be of great use to us throughout the text.

## 1.1 DEFINITIONS

A **group** $(G, \bullet)$ is a nonempty set G together with a binary operation called **multiplication** (or a **product**) and denoted by $\bullet$ that obeys the following axioms:

(G1) $a, b \in G$ implies $a \bullet b \in G$ (closure);
(G2) $a, b, c \in G$ implies $(a \bullet b) \bullet c = a \bullet (b \bullet c)$ (associativity);
(G3) There exists $e \in G$ such that $a \bullet e = e \bullet a = a$ for all $a \in G$ (identity);
(G4) For each $a \in G$, there exists $a^{-1} \in G$ such that $a \bullet a^{-1} = a^{-1} \bullet a = e$ (inverse).

Furthermore, a group is said to be **abelian** if it also has the property that

(G5) $a \bullet b = b \bullet a$ for all $a, b \in G$ (commutativity).

In the case of abelian groups, the group multiplication operation is frequently denoted by + and called **addition**. We will generally simplify our notation by leaving out the group multiplication symbol and assuming that it is understood for the particular group under discussion.

The number of elements in a group G is called its **order** of and will be denoted by o(G). (The order of G is frequently denoted by |G| although we shall not use this notation.) If this number is finite, then we say that G is a **finite** group. Otherwise, G is said to be **infinite**.

While we have defined a group in the usual manner, it should be realized that there is a certain amount of redundancy in our definition. In particular, it is not necessary to require that a "right inverse" also be the "left inverse." To see this, suppose that for any $a \in G$, we have the **right inverse** defined by $aa^{-1} = e$. Then multiplying from the left by $a^{-1}$ yields $a^{-1}aa^{-1} = a^{-1}$. But $a^{-1} \in G$ so there exists an $(a^{-1})^{-1} \in G$ such that $(a^{-1})(a^{-1})^{-1} = e$. Multiplying our previous expression from the right by $(a^{-1})^{-1}$ results in $a^{-1}a = e$, and hence we see that $a^{-1}$ is also a **left inverse**. Of course, we could have started with a left inverse and shown that it is also a right inverse.

Similarly, we could have defined a **right identity** by $ae = a$ for all $a \in G$. We then observe that $a = ae = a(a^{-1}a) = (aa^{-1})a = ea$, and hence e is also a **left identity**.

It is easy to show that the identity element is unique. To see this, suppose that there exist $e, \hat{e} \in G$ such that for every $a \in G$ we have $ea = ae = \hat{e}a = a\hat{e} = a$. Since $ea = a$ for every $a \in G$, we have in particular that $e\hat{e} = \hat{e}$. On the other hand, since we also have $a\hat{e} = a$, it follows that $e\hat{e} = e$. Therefore $\hat{e} = e\hat{e} = e$ so that $e = \hat{e}$.

Before showing the uniqueness of the inverse, we first prove an important basic result. Suppose that $ax = ay$ for $a, x, y \in G$. Let $a^{-1}$ be a (not necessarily unique) inverse to a. Then $x = ex = (a^{-1}a)x = a^{-1}(ax) = a^{-1}(ay) = (a^{-1}a)y = ey = y$. In other words, the equation $ax = ay$ means that $x = y$. This is sometimes called the (left) **cancellation law**. As a special case, we see that $aa^{-1} = e = a\hat{a}^{-1}$ implies $a^{-1} = \hat{a}^{-1}$ so that the inverse is indeed unique as claimed. This also shows that

$$(a^{-1})^{-1} = a$$

since $(a^{-1})^{-1}(a^{-1}) = e$ and $aa^{-1} = e$.

Finally, another important result follows by noting that $(ab)(b^{-1}a^{-1}) = a((bb^{-1})a^{-1}) = a(ea^{-1}) = aa^{-1} = e$. Since the inverse is unique, we then see that

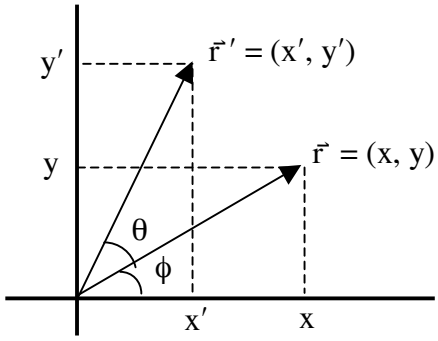
$$(ab)^{-1} = b^{-1}a^{-1} .$$

This clearly extends by induction to any finite product of group elements.

**Example 1.1** The set of integers $\mathbb{Z} = 0, \pm1, \pm2, \ldots$ forms an infinite abelian group where the group multiplication operation is just ordinary addition. It should be obvious that the (additive) identity element is 0, and the inverse of any number n is given by $-n$. However, it is easy to see that $\mathbb{Z}$ is not a group under the operation of ordinary multiplication. Indeed, while $\mathbb{Z}$ is both closed and associative under multiplication, and it also contains the (multiplicative) identity element 1, no element of $\mathbb{Z}$ (other than $\pm1$) has a multiplicative inverse in $\mathbb{Z}$ (for example, $2^{-1} = 1/2 \notin \mathbb{Z}$).

On the other hand, if we consider the set $\mathbb{Q}$ of all rational numbers, then $\mathbb{Q}$ forms a group under ordinary addition (with identity element 0 and inverse $-p/q \in \mathbb{Q}$ to any $p/q \in \mathbb{Q}$). Moreover, the *nonzero* elements of $\mathbb{Q}$ also form a group under ordinary multiplication (with identity element 1 and inverse $q/p \in \mathbb{Q}$ to any $p/q \in \mathbb{Q}$). //

**Example 1.2** A more complicated (but quite useful) example is given by the set of all rotations in the xy-plane. (This example uses some notation that we have not yet defined in this book, although most readers should have no difficulty following the discussion.) Consider the following figure that shows a vector $\vec{r} = (x, y)$ making an angle $\phi$ with the x-axis, and a vector $\vec{r}\,' = (x', y')$ making an angle $\theta + \phi$ with the x-axis:




We assume $r = |\vec{r}| = |\vec{r}\,'|$ so that the vector $\vec{r}\,'$ results from a counterclockwise rotation by an angle $\theta$ with respect to the vector $\vec{r}$. From the figure, we see that $\vec{r}\,'$ has components $x'$ and $y'$ given by

$$x' = r\cos(\theta+\phi) = r\cos\theta\cos\phi - r\sin\theta\sin\phi = x\cos\theta - y\sin\theta$$
$$y' = r\sin(\theta+\phi) = r\sin\theta\cos\phi + r\cos\theta\sin\phi = x\sin\theta + y\cos\theta.$$

Let $R(\alpha)$ denote a counterclockwise rotation by an angle $\alpha$. It should be clear that $R(0)$ is just the identity rotation (i.e., no rotation at all), and that the inverse is given by $R(\alpha)^{-1} = R(-\alpha)$. With these definitions, it is easy to see

that the set of all rotations in the plane forms an infinite (actually, continuous) abelian group. A convenient way of describing these rotations is with the matrix

$$R(\alpha) = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}.$$

(Such a matrix is said form a **representation** of the rotation group.) We then see that $\vec{r}' = R(\theta)\,\vec{r}$, which in matrix notation is just

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Using this notation, it is easy to see that R(0) is the identity since

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

and also that $R(\theta)^{-1} = R(-\theta)$ because

$$R(\theta)R(-\theta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = R(-\theta)R(\theta).$$

We remark that while the rotation group in two dimensions is abelian, the rotation group in three dimensions is not. For example, let $R_z(\theta)$ denote a rotation about the z-axis (in the "right-handed sense"). Then, applied to any vector $\hat{x}$ lying along the x-axis, we see that

$$R_y(90^\circ)R_z(45^\circ)\hat{x} \neq R_z(45^\circ)R_y(90^\circ)\hat{x}$$

since in the second case, the result lies along the z-axis, while in the first case it does not. //

While we will return shortly to discuss subgroups in more detail, it will be of use to define them now. If G is a group, then a subset $H \subset G$ is said to be a subgroup of G if the elements of H form a group under the same group multiplication rule as G. For example, the set $\mathbb{Z}$ of integers is a subgroup of the group $\mathbb{Q}$ of all rational numbers under ordinary addition. Furthermore, it is easy to show that a nonempty subset H of a group G is a subgroup of G if and only if a, $b \in H$ implies that $ab \in H$, and $a \in H$ implies that $a^{-1} \in H$ (see Exercise 1.1.9).

## Exercises

1. Decide which of the following sets G forms a group under the indicated operation. If G does not form a group, give the reason.
   (a) G = {all integers} under ordinary subtraction.
   (b) G = {all nonzero rational numbers} under ordinary division.
   (c) G = $\{a_0, a_1, \ldots, a_6\}$ where

$$a_i a_j = \begin{cases} a_{i+j} & \text{if } i + j < 7 \\ a_{i+j-7} & \text{if } i + j \geq 7 \end{cases}.$$

   (d) G = $\{2^m 3^n : m, n \in \mathbb{Z}\}$ under ordinary multiplication.

2. Let F denote the set of all mappings from $\mathbb{R}$ into $\mathbb{R}$. For any f, g ∈ F we define (f + g)(x) = f(x) + g(x) for each x ∈ $\mathbb{R}$ so that f + g ∈ F. Show that this defines a group.

3. Show that the collection of all subsets of a set S, with the operation of taking symmetric differences (see Exercise 0.1.2) as the group multiplication operation, forms a group. [*Hint*: Show that the identity element is ∅, and the inverse of any A ⊂ S is A itself.]

4. Prove that any group of order n ≤ 4 must be abelian.

5. Given two groups A and B, we can form the Cartesian product A × B = {(a, b): a ∈ A and b ∈ B} of these groups considered as sets. Prove that A × B can be made into a group with respect to the operation defined by $(a_1, b_1)(a_2, b_2) = (a_1a_2, b_1b_2)$ for all $a_1, a_2 \in A$ and $b_1, b_2 \in B$. This group is called the **direct product** of A and B.

6. Prove that {(x, x): x ∈ G} is a subgroup of G × G (see the previous problem). This is called the **diagonal subgroup** of G × G.

7. Let G = $\{g_1, \ldots, g_n\}$ be a group, and let h ∈ G be arbitrary but fixed. Define the set hG = $\{hg_1, \ldots, hg_n\} = \{g_{h_1}, \ldots, g_{h_n}\}$. Show that hG = G, and conclude that the ordered set $(h_1, \ldots, h_n)$ is a permutation of the ordered set (1, . . . , n). (This simple but very useful result is frequently referred to as the **rearrangement lemma**.)

8. Let H be a subgroup of a group G.

(a) If e is the identity element in G and f is the identity element in H, show that f = e.
(b) If a ∈ H, show that the inverse element $a^{-1}$ is the same in H as the $a^{-1}$ is in G.

9. Let H be a nonempty subset of a group G. Prove that H is a subgroup of G if and only if a, b ∈ H implies ab ∈ H and a ∈ H implies $a^{-1} \in H$.

10. Let $\mathcal{H}$ be a collection of subgroups of a group G. Show that the intersection of all H ∈ $\mathcal{H}$ is a subgroup of G.

11. Let G be a group. An element a ∈ G is said to be **conjugate** to an element b ∈ G if there exists g ∈ G such that $b = gag^{-1}$. Show that this defines an equivalence relation on G. (Mutually conjugate elements of G are said to form a (conjugate) **class**.)

12. Let X be a (nonempty) subset of a group G, and let $\{H_i: i \varepsilon I\}$ be the collection of all subgroups of G that contain X. Then $\cap H_i$ is called the **subgroup** of G **generated by the set X** and denoted ⟨X⟩. Prove that ⟨X⟩ consists of *all* finite products $a_1^{n_1} a_2^{n_2} \cdots a_r^{n_r}$ where $a_i \in X$ and $n_i \in \mathbb{Z}$. [*Hint* : Show that the set H of *all* such products is a subgroup of G that contains X and is contained in every subgroup containing X. Thus H < ⟨X⟩ < H.]

## 1.2 PERMUTATION GROUPS

Let G be any group and suppose a ∈ G. As a matter of notational convenience, we define $a^0 = e$, $a^1 = a$, $a^2 = aa$, . . . , $a^k = aa^{k-1}$, as well as $a^{-2} = (a^{-1})^2$, $a^{-3} = (a^{-1})^3$, . . . (where $a^{-1}$ is the usual inverse element to a). It is then easy to see that for any m, n ∈ ℤ we have $a^m a^n = a^{m+n}$ and $(a^m)^n = a^{mn}$. From now on we will assume the reader understands that this is what is meant when we write an element of any group to a power.

Now consider three objects (□ Δ, O) where the parentheses mean that the given order is relevant. We define this to be the **canonical** (or standard) order on the set {□ Δ, O}. Given any other ordered triple, for example (O, □ Δ), we define a **permutation** f of the set S = {□ Δ, O} by

$$f = \begin{pmatrix} \square & \Delta & O \\ O & \square & \Delta \end{pmatrix}$$

where the first line is the set of objects in their canonical order and the second line is the given order. In other words, a **permutation** on a set S is a bijection from S onto itself. Note that simply giving an arbitrary order to a collection of objects does not in itself define a permutation. It is necessary that some canonical order also be specified as a point of reference.

This notation, where the top row *defines* the canonical order, is referred to as **two-line** notation. However, it is very important to realize that as long as the same pairing of objects is maintained between the top and bottom rows, we may rearrange these pairs any way we please. For example, the above permutation f can equally well be written as

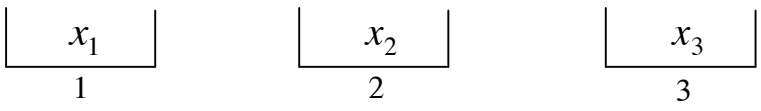
$$f = \begin{pmatrix} \text{O} & \square & \Delta \\ \Delta & \text{O} & \square \end{pmatrix}.$$

It is also common to use a simplified **one-line** notation. In this case, the canonical order must be understood. For example, in the first case above we would write simply f = (O, □, Δ) where the canonical order is understood to be (□, Δ, O).

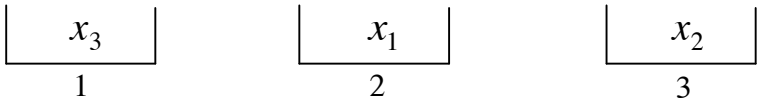
While we have now given a precise definition of the term "permutation," there are other ways of describing permutations that are very useful in practice. Two of these are given in the next (rather long) example, which will then be generalized to form one of the most useful groups in linear algebra.

**Example 1.3** Suppose we have three boxes that each contain a single object. Now, given three distinct objects and three boxes, any one of the three objects could go into the first box, then either of the two remaining objects could go into the second box, and finally only the remaining object can go into the third and last box. In other words, there are 3! = 6 possible placements of the three distinct objects in the three boxes such that each box receives a single object. Let us see how permutations can be used to describe the distribution of distinct objects among boxes. We give two common, intuitive interpretations.

Imagine three boxes labelled 1, 2, 3 that contain objects $x_1$, $x_2$, $x_3$ respectively, as shown below:




We now redistribute these objects among the boxes as follows:



One way to describe the transition from the first distribution to the second is by the permutation

$$\tilde{f} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$$

which is to be interpreted as a rule for redistributing objects by saying "take the object in box i (a number in the upper row) and place it in box $\tilde{f}$ (i) (the number in the lower row directly below it)." In this example, this means that we take the object in box i = 1 and place it in box $\tilde{f}$ (1) = 2, the object in box i = 2 goes into box $\tilde{f}$ (2) = 3, and the object in box i = 3 goes into box $\tilde{f}$ (3) = 1. This rule yields the second distribution from the first.

(Note also that in terms of our original definition of a permutation, we can interpret $\tilde{f}$ as a reordering of boxes in space. In other words, we can equally well describe the above redistribution in effect by leaving the objects fixed in space and rearranging the boxes underneath them. It is easy to see that if we leave the objects in the order ($x_1$, $x_2$, $x_3$) and label the boxes underneath them in the order (2, 3, 1), then we obtain the same pairing of objects and boxes.)

Another approach to describing this transition is by using permutations on the set of objects. For example, if we let

$$f = \begin{pmatrix} x_1 & x_2 & x_3 \\ x_3 & x_1 & x_2 \end{pmatrix}$$

then the second distribution (the lower row) is obtained from the first distribution (the upper row) by interpreting f as "replace object $x_1$ (wherever it is) by object $x_3$ , replace object $x_2$ (wherever it is) by object $x_1$ , and replace object $x_3$ (wherever it is) by object $x_2$." An equivalent way to describe this permutation is by the mapping f defined by

$$f(x_1) = x_3$$
$$f(x_2) = x_1$$
$$f(x_3) = x_2$$

which we can write in the simple one-line notation

$$f = (x_3, x_1, x_2).$$

Let us denote the set of objects {$x_1$, $x_2$, $x_3$} by S. Since there are only six possible distinct arrangements of S within the three boxes, there can be only six such permutations of S. We wish to make this set of permutations into a group. In particular, we will then have a group (denoted by $S_3$) of permuta-

tions defined on the set S. This group is called the **symmetric group** (or the **permutation group**) of **degree** 3. Since $S_3$ contains 3! = 6 elements, its order is 6.

We define the group multiplication as the composition of our permutations. For example, consider the permutation in $S_3$ defined by either

$$\tilde{g} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$$

or

$$g = \begin{pmatrix} x_1 & x_2 & x_3 \\ x_2 & x_1 & x_3 \end{pmatrix}$$

which in our one-line notation is simply $g = (x_2, x_1, x_3)$. Composing this with the above permutation $f = (x_3, x_1, x_2)$ we have, for example,

$$(fg)(x_1) = f(g(x_1)) = f(x_2) = x_1$$

and it is easy to see that the complete expression is given by $fg = (x_1, x_3, x_2)$. Note however, that

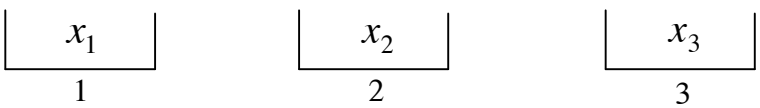
$$gf = (x_3, x_2, x_1) \neq fg$$

so that $S_3$ is a nonabelian group. This composition of mappings also shows us how to multiply our permutations. Indeed, if we write out the equation $fg = (x_1, x_3, x_2)$ in terms of our two-line notation, we obtain

$$fg = \begin{pmatrix} x_1 & x_2 & x_3 \\ x_3 & x_1 & x_2 \end{pmatrix} \begin{pmatrix} x_1 & x_2 & x_3 \\ x_2 & x_1 & x_3 \end{pmatrix} = \begin{pmatrix} x_1 & x_2 & x_3 \\ x_1 & x_3 & x_2 \end{pmatrix}.$$

Reading the product from right to left, we first see that $x_1$ is replaced by $x_2$, and then this $x_2$ is replaced by $x_1$, and the net result is that $x_1$ is replaced by $x_1$. Next we see that $x_2$ is first replaced by $x_1$, and then this $x_1$ is replaced by $x_3$ with the net result of replacing $x_2$ by $x_3$. Finally, $x_3$ is replaced by $x_3$, and then this $x_3$ is replaced by $x_2$, resulting in the replacement of $x_3$ by $x_2$. Therefore we see that combining the product from right to left results in exactly the same permutation as shown on the right hand side.

Now let us see how to combine the alternative descriptions in terms of $\tilde{f}$ and $\tilde{g}$. We know that $\tilde{f}$ takes the initial distribution



to the redistributed form ("contents of box 1 to box 2, contents of box 2 to box 3, and contents of box 3 to box 1")

| $x_3$ | $x_1$ | $x_2$ |
|---|---|---|
| 1 | 2 | 3 |

and $\tilde{g}$ takes the initial distribution to the redistributed form

| $x_2$ | $x_1$ | $x_3$ |
|---|---|---|
| 1 | 2 | 3 |

Applying $\tilde{f}$ to this last distribution we obtain (just take the contents of box 1 to box 2 etc.)

| $x_3$ | $x_2$ | $x_1$ |
|---|---|---|
| 1 | 2 | 3 |

With respect to the initial distribution, this composition of permutations is just the permutation

$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$$

In other words, simply following each permutation in sequence results in

$$\tilde{f}\tilde{g} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}.$$

Again reading the product from right to left, we see that the object in box 1 goes into box 2, and then the object in box 2 goes into box 3, with the net result that the object in box 1 goes into box 3. Next, the object in box 2 goes into box 1, and then the object in box 1 goes into box 2, resulting in the object in box 2 going into box 2. Finally, the object in box 3 goes into box 3, and then the object in box 3 goes into box 1, resulting in the object in box 3 going into box 1. Therefore, reading this type of product from right to left also results in the correct combination of permutations.

We now observe that $f^2(x_1) = f(x_3) = x_2$, and in general

$$f^2 = (x_2, x_3, x_1)$$

and

$$f^3 = (x_1, x_2, x_3)$$

which shows that $f^3 = ff^2 = e$, and hence $f^{-1} = f^2$. Similarly, we leave it to the reader to show that $g^2 = e$, and hence $g^{-1} = g$.

Since $S_3$ contains six elements and we have already constructed the six distinct mappings $\{e, f, g, f^2, fg, gf\}$, it must be true that any combination of mappings may be reduced to one of these six. To see this, all we really need to calculate is $(f^{-1}g)(x_1) = f^{-1}(x_2) = x_3$, and in general,

$$f^{-1}g = (x_3, x_2, x_1) = gf$$

so that $f^{-1}g = gf$. For example, we have $f(gf) = f(f^{-1}g) = (ff^{-1})g = g$. Other combinations are proved in a similar manner. ///

We now generalize this example to the case of an arbitrary (but finite) number of elements. Let S be a set containing a finite number n of elements. Then the set $S_n$ of all one-to-one mappings of S onto itself is called the **permutation group** of **degree** n. It should be clear that $S_n$ is of order n!. If $f \in S_n$, then f has the effect of taking $x_i \rightarrow f(x_i)$ which we may write as

$$f = \begin{pmatrix} x_1 & x_2 & \cdots & x_n \\ x_{i_1} & x_{i_2} & \cdots & x_{i_n} \end{pmatrix}$$

where $(i_1, \ldots, i_n)$ is some permutation of $(1, \ldots, n)$. To simplify our notation, let us write this mapping as

$$\begin{pmatrix} 1 & 2 & \cdots & n \\ i_1 & i_2 & \cdots & i_n \end{pmatrix}$$

where the top row stands for $(x_1, x_2, \ldots, x_n)$ and the bottom row represents $(x_{i_1}, x_{i_2}, \ldots, x_{i_n})$ which is just $(x_1, \ldots, x_n)$ in some permuted order. This should not be confused with the interpretation (which we will no longer use) of permutations as "the object in box 1 goes into box $i_1$" etc.

The identity element in $S_n$ is

$$\begin{pmatrix} 1 & 2 & \cdots & n \\ 1 & 2 & \cdots & n \end{pmatrix}$$

and the inverse to any given permutation is just the permutation that restores the original order. For instance, the inverse to the permutation f defined in Example 1.3 is the permutation $f^{-1} = f^2$ given in this notation by

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}.$$

In general, we will denote elements of $S_n$ by Greek letters such as $\theta$, $\phi$ and $\sigma$, so that we have expressions such as $\theta x$, $\theta^2 x = \theta\theta x$, and so forth. In other words, if $\theta \in S_3$ is just the mapping f in the previous example, then we would have $\theta 1 = 3$, $\theta 2 = 1$ and $\theta 3 = 2$.

Now let S be any set of n elements, and consider any element $\theta \in S_n$. Given any $x, y \in S$, we say that x is **equivalent** to y if $y = \theta^i x$ for some $i \in \mathbb{Z}$, and we write this as $x \approx_\theta y$. Since $x = \theta^0 x = ex = x$, we see that $x \approx_\theta x$. Next, note that if $x \approx_\theta y$, then $x = \theta^i y$ so that $y = \theta^{-i} x$, and hence $y \approx_\theta x$. In addition, if $x \approx_\theta y$ and $y \approx_\theta z$, then $x = \theta^i y = \theta^i \theta^j z = \theta^{i+j} z$, and hence $x \approx_\theta z$. We have therefore defined an equivalence relation on S as described in Section 0.3. Furthermore, Theorem 0.2 shows that this equivalence relation induces a decomposition of S into disjoint subsets called the equivalence classes of S.

For each $x \in S$, the equivalence class of x is the set $[x] = \{\theta^i x: i \in \mathbb{Z}\}$ which is called the **orbit** of x under $\theta$. Since S is finite, sooner or later repeated applications of $\theta$ to x must give back x. In other words, for each $x \in S$ there exists some smallest positive integer m such that $\theta^m x = x$ (where the value of m need not be the same for every $x \in S$). Thus the orbit of x under $\theta$ will be the set $\{x, \theta x, \ldots, \theta^{m-1} x\}$. If we consider these elements as being in a particular order, we then obtain what is called a **cycle** of $\theta$, and we write this as $(x, \theta x, \ldots, \theta^{m-1} x)$. In words, this means "x is replaced by $\theta x$, $\theta x$ is replaced by $\theta^2 x$, . . . , and $\theta^{m-1} x$ is replaced by x." It should be clear that a knowledge of all the cycles of $\theta$ is the same as knowing $\theta$, because we would then know the result of applying $\theta$ to any $x \in S$. (While the cycle notation is the same as the one-line notation for a permutation, the context should always make it clear which is meant.)

**Example 1.4** Let $S = \{x_1, \ldots, x_6\}$ which we denote by $\{1, \ldots, 6\}$ for simplicity. We consider the element $\theta \in S_6$ given by

$$\theta = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 1 & 3 & 5 & 6 & 4 \end{pmatrix}.$$

Now observe that $\theta 1 = 2$ and $\theta^2 1 = \theta 2 = 1$, so the orbit of 1 is the set $\{1, 2\}$ and the corresponding cycle is (1, 2). Since this cycle is the equivalence class of 1 and equivalence classes are disjoint, we see that it must also be the equivalence class of 2. Continuing, the orbit of 3 is just $\{3\}$, and for 4 we have $\theta 4 = 5$, $\theta^2 4 = 6$, and $\theta^3 4 = 4$ so that the orbit of 4 is $\{4, 5, 6\}$. Thus the cycles

of θ are (1, 2), (3) and (4, 5, 6). Notice that these cycles are disjoint ordered equivalence classes of S under the mapping $\theta \in S_6$. //

We can carry this idea one step further as follows. Consider a cycle of the form $(i_1, \ldots, i_m)$ which we now interpret as that permutation which replaces $i_1$ by $i_2$, $i_2$ by $i_3$, ..., $i_{m-1}$ by $i_m$, and $i_m$ by $i_1$. For example, using the set $S = \{1, \ldots, 6\}$, the cycle (2, 6, 3) is to be interpreted as the permutation

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 6 & 2 & 4 & 5 & 3 \end{pmatrix}.$$

Since we already know how to multiply permutations, we now have a way to multiply cycles. Thus, using this same S and, for example, the cycles (1, 5) and (2, 6, 3), we have

$$\begin{aligned}(1, 5)(2, 6, 3) &= \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 5 & 2 & 3 & 4 & 1 & 6 \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1 & 6 & 2 & 4 & 5 & 3 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 5 & 6 & 2 & 4 & 1 & 3 \end{pmatrix}.\end{aligned}$$

Note that while we have defined our multiplication as proceeding from right to left, in this case we would have obtained the same result by multiplying the cycles in either order. In fact, it should not be hard to convince yourself that this will always be the case when *disjoint* cycles are multiplied together. In other words, disjoint cycles commute. This is because each cycle only acts on a specific subset of elements that are not acted on by any other (disjoint) cycle.

As another example, let us now find the cycles of the permutation

$$\theta = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 5 & 6 & 2 & 4 & 1 & 3 \end{pmatrix}.$$

We have $\theta 1 = 5$ and $\theta^2 1 = 1$ so that the orbit of 1 is {1, 5}. Also, $\theta 2 = 6$, $\theta^2 2 = 3$, and $\theta^3 2 = 2$ so the orbit of 2 is {2, 6, 3}. Therefore θ has the cycles (1, 5) and (2, 6, 3) (and of course, also (4)). But now notice that θ is just the product of these cycles (which contain no elements in common) taken in any order. A little thought as we just mentioned shows that this is not unexpected, as we prove in our first theorem.

**Theorem 1.1** Every permutation can be expressed as the product of disjoint cycles.

*Proof* Consider any permutation $\theta \in S_n$ on a set S, and assume that $\theta$ has k cycles where each cycle is of the form $(x, \theta x, \theta^2 x, \ldots, \theta^{m_i - 1}x)$ for some i with $1 \leq i \leq k$. (Note that since each $x \in S$ must be in some cycle, and since the cycles are disjoint, we must have $\sum_{i=1}^{k} m_i = n$ where n is the number of elements in S.) When these cycles are multiplied together, we see that each of the corresponding permutations affects only those elements contained in the orbit (i.e., cycle) it represents. Hence, by multiplying together all of the cycles, each element of S will be accounted for with the same result as $\theta$.

Another way to see this is to consider the effect of $\theta$ on any $x \in S$. The resulting element $\theta x$ is exactly the same as the image of x under the product of all the (disjoint) cycles of $\theta$ since only the cycle containing x will have any effect on it. Since both $\theta$ and the product of its cycles have the same effect on any $x \in S$, it must be true that $\theta$ equals the product of its cycles. ∎

At this point, there is no substitute for simply working out an example for yourself. Thus, the reader should pick some permutation, find its cycles, and then multiply them together. In so doing, the proof of Theorem 1.1 should become quite obvious (or see the exercises at the end of this section).

Suppose that $S = \{1, 2, \ldots, m\}$ and consider the product of the 2-cycles $(1, m), (1, m-1), \ldots, (1, 3), (1, 2)$. Expressing these in terms of their corresponding permutations, we have (note the order of factors since we are multiplying from right to left, and these cycles are not disjoint)

$$\begin{pmatrix} 1 & 2 & 3 & 4 & \cdots & m \\ m & 2 & 3 & 4 & \cdots & 1 \end{pmatrix} \cdots \begin{pmatrix} 1 & 2 & 3 & 4 & \cdots & m \\ 3 & 2 & 1 & 4 & \cdots & m \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 & \cdots & m \\ 2 & 1 & 3 & 4 & \cdots & m \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 2 & 3 & 4 & \cdots & m \\ 2 & 3 & 4 & 5 & \cdots & 1 \end{pmatrix}.$$

But this last permutation is just the m-cycle $(1, 2, \ldots, m)$. A similar calculation shows that in fact any m-cycle of the form $(a_1, a_2, \ldots, a_m)$ may be written as the product $(a_1, a_m) \cdots (a_1, a_3)(a_1, a_2)$. (We remark that the multiplication of 2-cycles is one place where multiplying from left to right would be more natural.)

**Example 1.5** Consider the permutation

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 5 & 2 & 6 & 1 & 4 \end{pmatrix}$$

and its cycle (1, 3, 2, 5). We claim that this cycle may be written as the product (1, 5)(1, 2)(1, 3). There are actually two equivalent ways of seeing this. First, we could write out all of the complete permutations as

$$(1, 3, 2, 5) = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 5 & 2 & 4 & 1 & 6 \end{pmatrix} \qquad (1, 3) = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 2 & 1 & 4 & 5 & 6 \end{pmatrix}$$

$$(1, 2) = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 1 & 3 & 4 & 5 & 6 \end{pmatrix} \qquad (1, 5) = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 5 & 2 & 3 & 4 & 1 & 6 \end{pmatrix}$$

It is then easy to see that (1, 3, 2, 5) = (1, 5)(1, 2)(1, 3).

On the other hand, we could also leave out those elements in each permutation that are not affected by any of the cycles, and simply write

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

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

Again, we obtain (1, 3, 2, 5) = (1, 5)(1, 2)(1, 3). At this point you should be sufficiently familiar with the cycle notation to be able to multiply cycles without reverting to two-line notation. //

All of this discussion has shown that any m-cycle may be written as a product of 2-cycles, which are usually called **transpositions**. However we could also write, for example, (1, 2, . . . , m) = (m, m − 1) · · · (m, 2)(m, 1) so that this decomposition is by no means unique.

With all of this background, it is now easy to prove an important result in the description of permutations.

**Theorem 1.2** Every permutation can be written as the product of transpositions.

*Proof* Theorem 1.1 showed that every permutation can be written as the product of disjoint cycles, while we just showed that any cycle can be written (in a non-unique manner) as the product of transpositions. ∎

In view of this theorem, we say that a permutation is **even** (**odd**) if it can be written as the product of an even (odd) number of transpositions. Of

course, since the decomposition of cycles into transpositions is not unique, we must be sure that such a designation is unambiguous. This is the intent of our next theorem.

**Theorem 1.3** If a permutation can be represented by an even (odd) number of transpositions in one decomposition, then any other decomposition must also be an even (odd) number of transpositions.

*Proof* Define the polynomial p in n real variables by

$$\begin{aligned} p(x_1,\ldots,x_n) &= \prod\nolimits_{i<j}(x_i - x_j) \\ &= (x_1 - x_2)(x_1 - x_3)\cdots(x_2 - x_3)(x_2 - x_4)\cdots(x_{n-1} - x_n) \end{aligned}$$

and let $\sigma \in S_n$ be any transposition. By $\sigma p$ we mean

$$\sigma p(x_1, \ldots, x_n) = p(x_{\sigma 1}, \ldots, x_{\sigma n}) \ .$$

We claim that $\sigma p = -p$. To see this in detail, let $\sigma$ be the transposition $(x_a, x_b)$. We assume without loss of generality that $x_a < x_b$ , and write out all of those terms in $p(x_1, \ldots, x_n)$ that contain either $x_a$ or $x_b$ (or both). Thus, those terms containing $x_a$ are

$$\underbrace{\{(x_1 - x_a)(x_2 - x_a)\cdots(x_{a-1} - x_a)\}}_{a-1 \text{ terms}}$$

$$\times \underbrace{\{(x_a - x_{a+1})(x_a - x_{a+2})\cdots(x_a - x_{b-1})\}}_{b-a-1 \text{ terms}}$$

$$\times \underbrace{\{(x_a - x_b)\}}_{1 \text{ term}} \times \underbrace{\{(x_a - x_{b+1})\cdots(x_a - x_n)\}}_{n-b \text{ terms}}$$

while those containing $x_b$ are

$$\underbrace{\{(x_1 - x_b)(x_2 - x_b)\cdots(x_{a-1} - x_b)\}}_{a-1 \text{ terms}} \times \underbrace{\{(x_a - x_b)\}}_{\substack{\text{already} \\ \text{counted}}}$$

$$\times \underbrace{\{(x_{a+1} - x_b)(x_{a+2} - x_b)\cdots(x_{b-1} - x_b)\}}_{b-a-1 \text{ terms}}$$

$$\times \underbrace{\{(x_b - x_{b+1}) \cdots (x_b - x_n)\}}_{n-b \text{ terms}}$$

Since all of these terms are multiplied together, we see that if $x_a$ and $x_b$ are interchanged (but *not* $x_{a+1}$ and $x_{b+1}$ etc.), there will be no net effect on the polynomial $p(x_1, \ldots, x_n)$ except for the unpaired term $(x_a - x_b)$ which results in a single change of sign. This shows that $\sigma p = -p$ as claimed.

Now, for any other $\theta \in S_n$, Theorem 1.2 shows that $\theta = \prod_i \sigma_i$ where each $\sigma_i$ is a transposition. Thus, if

$$\theta = \prod_{i=1}^{k} \sigma_i$$

we see that

$$\theta p = \left( \prod_{i=1}^{k} \sigma_i \right) p = (-1)^k p.$$

Similarly, if

$$\theta = \prod_{i=1}^{m} \sigma_i$$

we have $\theta p = (-1)^m p$. Therefore, if $\theta$ is represented by k transpositions and by m transpositions, we must have $(-1)^k p = (-1)^m p$, and hence k and m must both be even or both be odd. ∎

This result allows us to make the unambiguous definition of the sign of a permutation as follows. We define the sign of a permutation $\theta$, $\operatorname{sgn} \theta$, by

$$\operatorname{sgn} \theta = \begin{cases} +1 \text{ if } \theta \text{ is even} \\ -1 \text{ if } \theta \text{ is odd} \end{cases}.$$

Our next theorem will be of great benefit to us when we come to discuss the theory of determinants in Chapter 4.

**Theorem 1.4** For any two permutations $\theta, \phi \in S_n$ we have

$$\operatorname{sgn}(\theta\phi) = (\operatorname{sgn} \theta)(\operatorname{sgn} \phi) .$$

*Proof* By Theorem 1.2, we may write $\theta$ as a product of k transpositions and $\phi$ as a product of m transpositions. Therefore it follows from Theorem 1.3 that $\operatorname{sgn} \theta = (-1)^k$ and $\operatorname{sgn} \phi = (-1)^m$. But then

$$\operatorname{sgn}(\theta\phi) = (-1)^{k+m} = (-1)^k(-1)^m = (\operatorname{sgn}\theta)(\operatorname{sgn}\phi) \ . \ \blacksquare$$

As the final topic in our treatment of permutations, we take a look at the inverse of a given transposition. For any given transposition $(a_1, a_2)$, it should be obvious that $(a_1, a_2)^2$ is just the identity transposition. This may be formally shown by noting that

$$(a_1, a_2)^2 = (a_1, a_2)(a_1, a_2) = \begin{pmatrix} a_1 & a_2 \\ a_2 & a_1 \end{pmatrix}\begin{pmatrix} a_1 & a_2 \\ a_2 & a_1 \end{pmatrix} = \begin{pmatrix} a_1 & a_2 \\ a_1 & a_2 \end{pmatrix} = e.$$

Since the identity element in any group is unique, this means that for any transposition $\sigma$ we have $\sigma^{-1} = \sigma$. In view of this result, one might rightfully expect that the sign of an inverse permutation is the same as the sign of the permutation itself.

**Theorem 1.5** For any $\theta \in S_n$ we have $\operatorname{sgn}\theta^{-1} = \operatorname{sgn}\theta$.

*Proof* By Theorem 1.2, we write $\theta = \sigma_1\sigma_2 \cdots \sigma_m$ where each $\sigma_i$ is a transposition. Then, using the fact that $\theta$ is just a product of elements in the group $S_2$, we see that

$$\theta^{-1} = (\sigma_1 \sigma_2 \cdots \sigma_m)^{-1} = \sigma_m^{-1} \cdots \sigma_2^{-1} \sigma_1^{-1} = \sigma_m \cdots \sigma_2 \sigma_1$$

and hence $\operatorname{sgn}\theta^{-1} = (-1)^m = \operatorname{sgn}\theta$. $\blacksquare$

## Exercises

1. Consider the following permutations

$$\theta = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 4 & 3 & 2 \end{pmatrix} \qquad \phi = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 4 & 2 \end{pmatrix}$$

and compute each of the following:

(a) $\theta\phi$ (b) $\phi\theta$ (c) $\theta^{-1}$ (d) $\phi^{-1}$
(e) $\theta^{-1}\phi^{-1}$ (f) $\phi^{-1}\theta^{-1}$ (g) $(\theta\phi)^{-1}$ (h) $(\phi\theta)^{-1}$

2. Referring to Example 1.3, evaluate $gfgf^3gf$. How is $\tilde{f}$ related to f?

3. Find all of the orbits and cycles of the following permutations:

(*a*) $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 2 & 3 & 4 & 5 & 1 & 6 & 7 & 9 & 8 \end{pmatrix}$.

(*b*) $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 6 & 5 & 4 & 3 & 1 & 2 \end{pmatrix}$.

4. Express each of the following as the product of disjoint cycles:
   (a) (1, 2, 3)(4, 5)(1, 6, 7, 8, 9)(1, 5)
   (b) (1, 2)(1, 2, 3)(1, 2)

5. Determine which of the following products of cycles is an even permutation:
   (a) (1, 2, 3)(1, 2)
   (b) (1, 2, 3, 4, 5)(1, 2, 3)(4, 5)
   (c) (1, 2)(1, 3)(1, 4)(2, 5)

6. Show that the set $A_n \subset S_n$ consisting of *even* permutations forms a group. Show that $S_n$ consists of n!/2 even permutations and n!/2 odd permutations.

7. Compute $\theta^{-1}\phi\theta$ for each of the following:
   (a) $\theta = (1, 3, 5)(1, 2)$ $\quad \phi = (1, 5, 7, 9)$.
   (b) $\theta = (5, 7, 9)$ $\quad \phi = (1, 2, 3)$.

8. Show that permutations with the same cycle structure belong to the same class (see Exercise 1.1.8). In other words, if $\theta, \phi \in S_n$, show that $\theta\phi\theta^{-1}$ has the same cycle structure as $\phi$. [*Hint*: Using two-line notation, show that $\theta\phi\theta^{-1}$ may be evaluated by simply applying $\theta$ to the top and bottom rows of $\phi$ separately.]

9. Show that $S_n$ is non-abelian if $n \geq 3$.

## 1.3 HOMOMORPHISMS OF GROUPS

We now turn our attention to a discussion of mappings from one group to another. These results will be absolutely fundamental to everything else that follows, and it is essential that the reader thoroughly understand the concepts to be presented in this section.

Let $\phi$: G $\rightarrow$ G′ be a mapping from a group G to a group G′. If for every x, y $\in$ G we have

$$\phi(xy) = \phi(x)\,\phi(y)$$

then $\phi$ is said to be a **homomorphism**, and the groups G and G′ are said to be **homomorphic**. In other words, a homomorphism preserves group multiplication, but is not in general either surjective or injective. It should also be noted that the product xy is an element of G while the product $\phi(x)\,\phi(y)$ is an element of G′.

**Example 1.6** Let G be the (abelian) group of all real numbers under addition, and let G′ be the group of nonzero real numbers under multiplication. If we define $\phi$: G $\rightarrow$ G′ by $\phi(x) = 2^x$, then

$$\phi(x + y) = 2^{x+y} = 2^x 2^y = \phi(x)\,\phi(y)$$

so that $\phi$ is indeed a homomorphism. //

**Example 1.7** Let G be the group of all real (or complex) numbers under ordinary addition. For any real (or complex) number a, we define the mapping $\phi$ of G onto itself by $\phi(x) = ax$. This $\phi$ is clearly a homomorphism since

$$\phi(x + y) = a(x + y) = ax + ay = \phi(x) + \phi(y) \ .$$

However, if b is any other nonzero real (or complex) number, then we leave it to the reader to easily show that the ("nonhomogeneous") mapping $\psi(x) = ax + b$ is not a homomorphism. //

Let e be the identity element of G, and let e′ be the identity element of G′. If $\phi$: G $\rightarrow$ G′ is a homomorphism, then $\phi(x)e' = \phi(x) = \phi(xe) = \phi(x)\phi(e)$, and we have the important result

$$\phi(e) = e'.$$

Using this result, we then see that $e' = \phi(e) = \phi(xx^{-1}) = \phi(x)\,\phi(x^{-1})$, and hence the uniqueness of the inverse tells us that

$$\phi(x^{-1}) = \phi(x)^{-1} .$$

It is very important to note that in general $\phi(x)^{-1} \neq \phi^{-1}(x)$ since if $x \in G$ we have $\phi(x)^{-1} \in G'$ while if $x \in G'$, then $\phi^{-1}(x) \in G$. Using these results, it should now be easy for the reader to show that $\phi(G)$ forms a subgroup of $G'$ (see Exercise 1.3.1).

In general, there may be many elements $x \in G$ that map into the same element $x' \in G'$ under $\phi$. It is of particular interest to see what happens if more than one element of G (besides e) maps into $e'$. If $k \in G$ is such that $\phi(k) = e'$, then for any $x \in G$ we have $\phi(xk) = \phi(x)\,\phi(k) = \phi(x)e' = \phi(x)$. Therefore if $xk \neq x$ we see that $\phi$ could not possibly be a one-to-one mapping. To help us get a hold on when a homomorphism is one-to-one, we define the **kernel** of $\phi$ to be the set

$$\text{Ker } \phi = \{x \in G: \phi(x) = e'\} .$$

It is also easy to see that Ker $\phi$ is a subgroup of G (see Exercise 1.3.1).

If a homomorphism $\phi: G \to G'$ is one-to-one (i.e., injective), we say that $\phi$ is an **isomorphism**. If, in addition, $\phi$ is also onto (i.e., surjective), then we say that G and $G'$ are **isomorphic**. In other words, G and $G'$ are isomorphic if $\phi$ is a bijective homomorphism. (We point out that many authors use the word "isomorphism" to implicitly mean that $\phi$ is a bijection.) In particular, an isomorphism of a group onto itself is called an **automorphism**.

From the definition, it appears that there is a relationship between the kernel of a homomorphism and whether or not it is an isomorphism. We now proceed to show that this is indeed the case. By way of notation, if H is a subset of a group G, then by Hg we mean the set $Hg = \{hg \in G: h \in H\}$. Recall also that if $\phi: G \to G'$ and $x' \in G'$ then, by an inverse image of $x'$, we mean any element $x \in G$ such that $\phi(x) = x'$.

**Theorem 1.6** Let $\phi$ be a homomorphism of a group G onto a group $G'$, and let $K_\phi$ be the kernel of $\phi$. Then given any $x' \in G'$, the set of all inverse images of $x'$ is given by $K_\phi x$ where $x \in G$ is any particular inverse image of $x'$.

*Proof* Consider any $k \in K_\phi$. Then by definition of homomorphism, we must have

$$\phi(kx) = \phi(k)\,\phi(x) = e'x' = x' .$$

In other words, if x is any inverse image of $x'$, then so is any $kx \in K_\phi x$. We must be sure that there is no other element $y \in G$, $y \notin K_\phi x$ with the property that $\phi(y) = x'$.

To see that this is true, suppose $\phi(y) = x' = \phi(x)$. Then $\phi(y) = \phi(x)$ implies that

$$e' = \phi(y)\phi(x)^{-1} = \phi(y)\phi(x^{-1}) = \phi(yx^{-1}) .$$

But this means that $yx^{-1} \in K_\phi$, and hence $yx^{-1} = k$ for some $k \in K_\phi$. Therefore $y = kx \in K_\phi x$ and must have already been taken into account. ∎

**Corollary** A homomorphism $\phi$ mapping a group G to a group G′ is an isomorphism if and only if $\text{Ker } \phi = \{e\}$.

*Proof* Note that if $\phi(G) \neq G'$, then we may apply Theorem 1.6 to G and $\phi(G)$. In other words, it is trivial that $\phi$ always maps G *onto* $\phi(G)$. Now, if $\phi$ is an isomorphism, then it is one-to-one by definition, so that there can be no element of G other than e that maps into e′. Conversely, if $\text{Ker } \phi = \{e\}$ then Theorem 1.6 shows that any $x' \in \phi(G) \subset G'$ has exactly one inverse image. ∎

Of course, if $\phi$ is surjective, then $\phi(G)$ is just equal to G′. In other words, we may think of isomorphic groups as being essentially identical to each other.

**Example 1.8** Let G be any group, and let $g \in G$ be fixed. We define the mapping $\phi: G \to G$ by $\phi(x) = gxg^{-1}$, and we claim that $\phi$ is an automorphism. To see this, first note that $\phi$ is indeed a homomorphism since for any $x, y \in G$ we have

$$\begin{aligned}\phi(xy) &= g(xy)g^{-1} = g(xey)g^{-1} = g(xg^{-1}gy)g^{-1} = (gxg^{-1})(gyg^{-1}) \\ &= \phi(x)\phi(y).\end{aligned}$$

To see that $\phi$ is surjective, simply note that for any $y \in G$ we may define $x = g^{-1}yg$ so that $\phi(x) = y$. Next, we observe that if $\phi(x) = gxg^{-1} = e$, then right-multiplying by g and left multiplying by $g^{-1}$ yields

$$x = (g^{-1}g)x(g^{-1}g) = g^{-1}eg = e$$

and hence $\text{Ker } \phi = \{e\}$. From the corollary to Theorem 1.6, we now see that $\phi$ must be an isomorphism. ∎

**Exercises**

1. Let $\phi: G \to G'$ be a homomorphism.
   (a) Show that $\phi(G)$ is a subgroup of $G'$.
   (b) Show that Ker $\phi$ is a subgroup of G.

2. Show that the composition $\phi \circ \psi: A \to C$ is a homomorphism if both $\phi: B \to C$ and $\psi: A \to B$ are.

3. Determine which of the following mappings $\phi: G \to G'$ are homomorphisms, and for those that are, determine their kernel:
   (a) $G = G'$ = the group of nonzero real numbers under multiplication, and $\phi(x) = x^2$ for all $x \in G$.
   (b) Repeat part (a) but with $\phi(x) = 2^x$.
   (c) $G = G'$ = the group of all real numbers under addition, and $\phi(x) = 1 + x$ for all $x \in G$.
   (d) Repeat part (c), but with $\phi(x) = kx$ for any (fixed) number k.

4. Show that an isomorphism $\phi$ defines an equivalence relation on the set of all groups.

5. If G is abelian and $G'$ is isomorphic to G, prove that $G'$ is also abelian.

6. Let $\mathbb{R}^+$ denote the set of all real numbers $> 0$, and define the mapping $\phi: \mathbb{R}^+ \to \mathbb{R}$ by $\phi(x) = \log_{10} x$ for each $x \in \mathbb{R}^+$. Let $\mathbb{R}^+$ be a group with respect to multiplication, and let $\mathbb{R}$ be a group with respect to addition. Show that $\phi$ is an isomorphism.

7. Let A and B be groups (with their own group operations). Show that $A \times B$ is isomorphic to $B \times A$ (see Exercise 1.1.5).

8. (a) (**Cayley's theorem**) Prove that every group G of order n is isomorphic to a subgroup of $S_n$ for some S. [*Hint*: By the rearrangement lemma (Exercise 1.1.7), we know that $hG = G$ for any $h \in G$. If $G = \{g_1, \ldots, g_n\}$, define the mapping $\psi: G \to S_n$ by

$$\psi(a) = \begin{pmatrix} g_1 & \cdots & g_n \\ ag_1 & \cdots & ag_n \end{pmatrix}$$

   for every $a \in G$. Using the techniques of Exercise 1.2.8, show that $\psi$ is a homomorphism, i.e., $\psi(ab) = \psi(a)\psi(b)$.]

(b) Explain why this result shows that there can only be a finite number of non-isomorphic groups of order n.

## 1.4 RINGS AND FIELDS

Before starting our discussion of vector spaces, let us first define precisely what is meant by a field. We shall see that this is simply a generalization of those essential properties of the real and complex numbers that we have been using all along. For the sake of completeness and future reference, we will do this in a somewhat roundabout manner.

A nonempty set R together with two operations denoted by + and • is said to be an **associative ring** if it obeys the following axioms for all a, b, c ∈ R:

(R1) a + b ∈ R;
(R2) a + b = b + a;
(R3) (a + b) + c = a + (b + c);
(R4) There exists an element 0 ∈ R such that a + 0 = a;
(R5) There exists and element −a ∈ R such that a + (−a) = 0;
(R6) a•b ∈ R;
(R7) (a•b)•c = a•(b•c);
(R8) a•(b + c) = a•b + a•c and (a + b)•c = a•c + b•c.

Since every ring that we will ever discuss obeys (R7), we henceforth drop the adjective "associative" when discussing rings. It should also be noticed that (R1) – (R5) simply require that R be an abelian group under the operation + which we call addition. In addition to these axioms, if there exists an element 1 ∈ R such that a•1 = 1•a = a for every a ∈ R, then R is said to be a **ring with unit element**. Furthermore, if for every a, b ∈ R we have a•b = b•a, then R is called a **commutative ring**. As usual, we shall generally leave out the multiplication sign when dealing with rings.

**Example 1.9** The set $\mathbb{Z}$ of all real integers under the usual operations of addition and multiplication is a commutative ring with unit element. However, the set of even integers under addition and multiplication is a commutative ring with no unit element. Note also that the set $\mathbb{Z}^+$ of positive integers is not a ring since there are no additive inverse (i.e., negative) elements in this set. //

Note that while the elements of a ring form an additive abelian group, we have not required that each element have a multiplicative inverse. However, if the nonzero elements of a ring R happen to form a group under multiplication, we say that R is a **division ring**. In this case we denote the unit element of R by 1, and we let $a^{-1}$ denote the inverse of any element a ∈ R. The reason that

only the nonzero elements are considered is that 0 has no inverse $0^{-1}$ such that $0 \bullet 0^{-1} = 1$. Finally, a **field** is defined to be a commutative division ring. We will generally denote an arbitrary field by the symbol $\mathcal{F}$.

**Example 1.10** It should be clear that the real numbers $\mathbb{R}$ form a field with the usual operations of addition and multiplication. However, the set $\mathbb{Z}$ of integers does not form a field because for any $n \in \mathbb{Z}$ with $n \neq 0$, we have $n^{-1} = 1/n \notin \mathbb{Z}$ (except for $n = \pm 1$).

It is also true that the complex numbers $\mathbb{C}$ form a field, but this is slightly more difficult to prove. To do so, let us denote a complex number $a + ib$ by the ordered pair $(a, b) \in \mathbb{R} \times \mathbb{R}$. Referring to Section 0.6 for motivation, we define addition and multiplication on these pairs by

$$\begin{aligned}(a, b) + (c, d) &= (a + c, b + d)\\ (a, b)(c, d) &= (ac - bd, ad + bc)\end{aligned}$$

for all $a, b, c, d \in \mathbb{R}$. We claim that the set $\mathbb{C}$ consisting of all such ordered pairs is a field. Some of the details will be left to the reader to fill in, but we will show the important points here. The additive identity element is clearly $(0, 0)$, the negative of any $(a, b) \in \mathbb{C}$ is $(-a, -b)$, and the multiplicative identity is $(1, 0)$. Multiplication is commutative since

$$(a, b)(c, d) = (ac - bd, ad + bc) = (ca - db, cb + da) = (c, d)(a, b)\ .$$

To prove associativity, we have

$$\begin{aligned}(a, b)[(c, d)(e, f)] &= (a, b)(ce - df, cf + de)\\ &= (ace - adf - bcf - bde, acf + ade + bce - bdf)\\ &= (ac - bd, ad + bc)(e, f)\\ &= [(a, b)(c, d)](e, f).\end{aligned}$$

Finally, we show that every $(a, b) \neq (0, 0)$ has an inverse in $\mathbb{C}$. Since $a$ and $b$ can not both be 0, we have $a^2 + b^2 > 0$. We leave it to the reader to show that

$$(a, b)\left(\frac{a}{a^2 + b^2}, \frac{-b}{a^2 + b^2}\right) = (1, 0).$$

(In the notation of Chapter 0, we see that this is just the statement that $zz^* = |z|^2$ implies $z^{-1} = z^*/|z|^2$.)

We will be using the fields $\mathbb{R}$ and $\mathbb{C}$ almost exclusively in this text. //

Since we will be using fields (and therefore rings) as our usual number system, let us use the defining relations to prove formally that a ring behaves in the manner we are accustomed to and expect.

**Theorem 1.7** Let R be a ring with unit element. Then for all a, b $\in$ R we have

(a) $a0 = 0a = 0$.
(b) $a(-b) = (-a)b = -(ab)$.
(c) $(-a)(-b) = ab$.
(d) $(-1)a = -a$.
(e) $(-1)(-1) = 1$.

*Proof* (a) $a0 = a(0 + 0)$ (by (R4)) $= a0 + a0$ (by (R8)). But R is an additive abelian group, so that canceling a0 from both sides of this equation says that $a0 = 0$. Similarly, we see that $0a = (0 + 0)a = 0a + 0a$ implies $0a = 0$.

(b) $ab + a(-b) = a(b + (-b))$ (by (R8)) $= a0$ (by (R5)) $= 0$ (by (a)). Therefore the group property of R shows that $a(-b) = -(ab)$. It is clear that we also have $(-a)b = -(ab)$.

(c) $(-a)(-b) = -(a(-b))$ (by (b)) $= -(-(ab))$ (by (b) again). But $-(-(ab))$ is the unique inverse to $-(ab)$, and since we also have $ab + (-(ab)) = 0$, it follows that $-(-(ab)) = ab$.

(d) $a + (-1)a = 1a + (-1)a = (1 + (-1))a = 0a = 0$ so that $(-1)a = -a$.

(e) This follows from (d) using $a = -1$ since $(-1)(-1) = -(-1) = 1$. ∎

Note the fact that R contains a unit element was actually only required for parts (d) and (e) of this theorem.

## Exercises

1. Let R be a ring. Prove that $a^2 - b^2 = (a + b)(a - b)$ and that $(a + b)^2 = a^2 + 2ab + b^2$ for all a, b $\in$ R if and only if R is commutative (where by terms of the form $a^2$ we mean aa).

2. Let F denote the set of all mappings from $\mathbb{R}$ into $\mathbb{R}$. For any f, g $\in$ F, we define $(f + g)(x) = f(x) + g(x)$ and $(fg)(x) = f(x)g(x)$ for each x $\in \mathbb{R}$. In other words, f + g and fg are in F. Show that this defines a ring of functions.

3. In the previous problem, show that if we replace the product fg by the composition $f \circ g$ then this does not define a ring.

4. Show that the set $\mathbb{Q}$ of all rational numbers forms a field.

5. Consider the set $\mathbb{Z}[\sqrt{2}] = \{a + b\sqrt{2}: a, b \in \mathbb{Z}\}$. We define addition and multiplication in $\mathbb{Z}[\sqrt{2}]$ by

$$(a_1 + b_1\sqrt{2}) + (a_2 + b_2\sqrt{2}) = (a_1 + a_2) + (b_1 + b_2)\sqrt{2}$$

and

$$(a_1 + b_1\sqrt{2})(a_2 + b_2\sqrt{2}) = (a_1a_2 + 2b_1b_2) + (a_1b_2 + b_1a_2)\sqrt{2} \ .$$

Show that the set $\mathbb{Z}[\sqrt{2}]$ with these operations forms a ring. Does it form a field?

6. Repeat the previous problem with $\mathbb{Q}$ instead of $\mathbb{Z}$.

## 1.5 MORE ON GROUPS AND RINGS

In this section we lay the foundation for future work in our chapter on polynomials. If the reader has not had much experience with abstract algebra, this section may prove somewhat long and difficult on a first reading. Because of this, the student should feel free to skim this section now, and return to it only when it becomes necessary in later chapters.

We have seen that fields offer a distinct advantage over rings in that elements of the field can be divided (since the field contains the multiplicative inverse of each nonzero element). It will be of interest to know how certain rings can be "enlarged" to form a field. Rather than treat this problem directly, we choose to introduce some additional terminology that will be of use in discussing further properties of polynomials.

In view of the fact that a ring has both addition and multiplication defined on it, we make the following definition. Let R and R′ be rings. A mapping $\phi$: $R \to R'$ is said to be a **ring homomorphism** if

$$\phi(a + b) = \phi(a) + \phi(b)$$

and

$$\phi(ab) = \phi(a)\,\phi(b)$$

for all $a, b \in R$. We see that

$$\phi(a) = \phi(0 + a) = \phi(0) + \phi(a)$$

and therefore $\phi(0) = 0$. Then we also have

$$0 = \phi(a - a) = \phi(a) + \phi(-a)$$

so that adding $-\phi(a)$ to both sides yields

$$\phi(-a) = -\phi(a) \ .$$

While these last two results are the exact analogues of what we found for groups, not all of our results can be carried over directly. In particular, it must be remembered that every element of a group had an inverse in the group, while no such requirement is made on the multiplication in an arbitrary ring (recall that a ring in which the nonzero elements form a multiplicative group is called a division ring).

If R is a commutative ring, a nonzero element $a \in R$ is said to be a **zero divisor** if there exists a nonzero element $b \in R$ such that $ab = 0$. We then say that a commutative ring is an **integral domain** if it contains no zero divisors. For example, the ring of integers is an integral domain.

**Example 1.11** Consider the set $\mathbb{Z}$ of all integers, and let $n \in \mathbb{Z}^+$ be fixed. A notation that is frequently used in algebra is to write $a|b$ to mean "a divides b," (i.e., in this case, that b is an integral multiple of a) and $c \nmid d$ to mean "c does not divide d." We define a relationship between the integers a and b by writing

$$a \equiv b (\text{mod } n)$$

if $n|(a - b)$. This relation is called **congruence modulo n**, and we read it as "a is congruent to b modulo n." We leave it as an exercise for the reader to show that this defines an equivalence relation on the set of integers (see Exercise 1.5.2). For example, it should be clear that $5 = 2 (\text{mod } 3)$, $23 = 5 (\text{mod } 6)$ and $21 = -9 (\text{mod } 10)$.

Now suppose we define a ring R to be the set of integers mod 6 (this ring is usually denoted by $\mathbb{Z}_6$). Then the elements of R are the equivalence classes of the integers, and we denote the equivalence class of an integer n by [n]. Then the elements of R are [0], [1], [2], [3], [4] and [5]. For example, from the previous paragraph we see that [5] = [23] because $6|(23 - 5)$.

We define addition in R by [a] + [b] = [a + b], and thus [0] is the zero element of R. Defining multiplication in R by [a][b] = [ab], we see, for example,

that [2][5] = [4]. However, note that [2][3] = [0] even though [2] ≠ [0] and [3] ≠ [0], and thus R is not an integral domain. We will have much more to say about this ring in Section 6.6. //

It should now be clear that arbitrary rings can have a number of properties that we generally find rather unpleasant. Another type of pathology that is worth pointing out is the following. Let D be an integral domain. We say that D is of **finite characteristic** if there exists some integer $m > 0$ and some nonzero $a \in D$ such that $ma = 0$. Then the smallest positive integer p such that $pa = 0$ for some nonzero $a \in D$ is called the **characteristic** of D.

If D is an integral domain of characteristic p, then there exists a nonzero element $a \in D$ such that $pa = 0$. Then for any $x \in D$ we also have

$$0 = (pa)x = (a + \cdots + a)x = ax + \cdots + ax = a(x + \cdots + x) = a(px) .$$

But D has no zero divisors, and hence we must have $px = 0$ for every $x \in D$. If D has a unit element , then an equivalent requirement is to say that if D is of characteristic p, then $1 + \cdots + 1 = 0$, where there are p terms in the sum. Furthermore, any such sum consisting of less than p terms is nonzero.

Obviously, the most important types of integral domain for our purposes are those of characteristic 0. In other words, to say that D is of characteristic 0 means that if m is an integer and $a \in D$ is nonzero, then $ma = 0$ if and only if $m = 0$. The reason that we even bother to mention this is because most of the theory of matrices and determinants that we shall develop is valid over an arbitrary field $\mathcal{F}$. For example, we shall obtain results such as $\det A = -\det A$ which implies that $2 \det A = 0$. However, if $\mathcal{F}$ happens to be of characteristic 2, then we can not conclude from this that $\det A = 0$. In this book, we will always assume that our fields are of characteristic 0 (except in Section 6.6).

Returning to our general discussion, let 1 and $1'$ be the multiplicative identities of the rings R and $R'$ respectively, and consider any ring homomorphism $\phi: R \to R'$. Then

$$\phi(a) = \phi(1a) = \phi(1)\,\phi(a)$$

but this does *not* in general imply that $\phi(1) = 1'$. However, if $R'$ is an integral domain and $\phi(a) \neq 0$, then we have

$$0 = \phi(a) - \phi(1)\,\phi(a) = \phi(a)[1' - \phi(1)]$$

and hence $\phi(1) = 1'$ (note that we do not distinguish in our notation between 0 and $0'$).

As was the case with groups, we define the kernel of $\phi$ to be the set

$$\text{Ker } \phi = \{a \in R: \phi(a) = 0\} .$$

If $a, b \in \text{Ker } \phi$ then

$$\phi(a + b) = \phi(a) + \phi(b) = 0 + 0 = 0$$

so that $a + b \in \text{Ker } \phi$ also. Furthermore, if $a \in \text{Ker } \phi$ then

$$\phi(-a) = -\phi(a) = 0$$

so that the (additive) inverse of a is also in Ker $\phi$. Thus Ker $\phi$ forms a subgroup of R under addition.

As we also did with groups, we say that a ring homomorphism of R into $R'$ is a (**ring**) **isomorphism** if it is an injective (i.e., one-to-one) mapping. If there exists a bijective ring homomorphism of R onto $R'$, then we say that R and $R'$ are isomorphic. Theorem 1.6 also carries over directly to the present case, and we then have that a ring homomorphism is an isomorphism if and only if Ker $\phi = \{0\}$.

Now note that another very important property of Ker $\phi$ comes from the observation that if $a \in \text{Ker } \phi$ and $r \in R$, then

$$\phi(ar) = \phi(a)\,\phi(r) = 0\phi(r) = 0 .$$

Similarly $\phi(ra) = 0$, and therefore both ar and ra are in Ker $\phi$. We take this property as the prototype of a new object defined as follows.

A nonempty subset I of a ring R is said to be a (two-sided) **ideal** of R if I is a subgroup of R under addition, and if $ar \in I$ and $ra \in I$ for *all* $a \in I$ and *all* $r \in R$. It is important to realize that the element r can be any element of R, not just an element of I.

Now let R be a *commutative* ring with unit element, and let $a \in R$ be arbitrary. We denote by (a) the set of all multiples of a by elements of R. (While this is a somewhat confusing notation, it nevertheless conforms to standard usage.) In other words,

$$(a) = \{ra: r \in R\} .$$

We claim that (a) is actually an ideal of R. Indeed, if $r, s \in R$, then $ra, sa \in (a)$ and therefore $ra + sa = (r + s)a \in (a)$. Next, we have $0 = 0a \in (a)$, and finally, the negative (i.e., additive inverse) of $ra \in (a)$ is $(-r)a$ which is also in (a). This shows that (a) is a subgroup of R under addition. Lastly, for any $ra \in (a)$ and any $s \in R$, we see that $(ra)s = s(ra) = (sr)a \in (a)$. We have thus shown that (a) is an ideal. In general, any ideal of the form (a) is called a **principal** ideal,

and the element $a \in R$ is called a **generator** of (a). A principal ideal (a) is thus the smallest ideal of R that contains a.

**Example 1.12** We show that any field $\mathcal{F}$ has no ideals other than (0) and $\mathcal{F}$. Since the ideal (0) is quite trivial, let I be an ideal and assume that $I \neq (0)$. If $a \in I$, $a \neq 0$, then $a \in \mathcal{F}$ implies that $a^{-1} \in \mathcal{F}$ so that $1 = aa^{-1} \in I$ by the definition of ideal. But now, for any $r \in \mathcal{F}$ we have $r = 1r \in I$ (again by the definition of ideal), and hence $I = \mathcal{F}$. ∥

The converse of this example is given in the next theorem. Recall that a field is a commutative division ring, and hence a commutative ring R with unit element 1 is a field if every nonzero $a \in R$ has an inverse $b \in R$ with $ab = 1$.

**Theorem 1.8** If R is a commutative ring with unit element whose only ideals are (0) and R, then R is a field.

*Proof* Part of this was proved in the above discussion, but for the sake of completeness we repeat it here. Let $a \in R$ be nonzero, and consider the set

$$Ra = \{ra: r \in R\} .$$

We shall first show that this set is an ideal of R. To see this, suppose $x, y \in Ra$. Then there exist $r_1, r_2 \in R$ such that $x = r_1a$ and $y = r_2a$. But then (using the definition of a ring) we see that

$$x + y = r_1a + r_2a = (r_1 + r_2)a \in Ra .$$

Next we note that

$$-x = -r_1a = (-r_1)a \in Ra$$

and therefore Ra is a subgroup of R under addition. Now, given any $r \in R$ we have

$$rx = r(r_1a) = (rr_1)a \in Ra$$

and since R is commutative, it also follows that $xr \in Ra$. This shows that Ra is an ideal of R.

By hypothesis, we see that Ra must equal either (0) or R. Since R is a ring with unit element, we have $0 \neq a = 1a \in Ra$ and hence $Ra \neq (0)$. This means that we must have $Ra = R$ so that every element of R is a multiple of a. In particular, since $1 \in R$, there must exist an element $b \in R$ with the property that $ba = 1$. In other words, $b = a^{-1}$ and thus R is a field. ∎

Now let H be a subgroup of a group G, and let $a \in G$ be arbitrary. Then the set

$$Ha = \{ha: h \in H\}$$

is called a **right coset** of H in G. Let $a, b \in G$ be arbitrary, and suppose that the cosets Ha and Hb have an element in common. This means that $h_1a = h_2b$ for some $h_1, h_2 \in H$. But then using the fact that H is a subgroup, we see that

$$a = h_1^{-1}h_1a = h_1^{-1}h_2b \in Hb \ .$$

Since this means that $a = hb$ for some $h = h_1^{-1}h_2 \in H$, we see (using the rearrangement lemma of Exercise 1.1.7) that this implies

$$Ha = Hhb = Hb$$

and therefore if any two right cosets have an element in common, then they must in fact be identical. It is easy to see that the set of all right cosets of H in G defines a partition of G and hence an equivalence relation that decomposes G into disjoint subsets (see Exercise 1.5.15).

Recall that o(G) denotes the order of G (i.e., the number of elements in the group G). We claim that if H is a subgroup of G, then o(H) = o(Ha) for any $a \in G$. Indeed, to prove this we show that there is a bijection of H to Ha. Define the mapping $\alpha: H \to Ha$ by $\alpha(h) = ha$. This is clearly a surjective mapping since Ha consists precisely of elements of the form ha for $h \in H$. To see that it is also injective, suppose that for some $h_1, h_2 \in H$ we have $\alpha(h_1) = \alpha(h_2)$ or, equivalently, $h_1a = h_2a$. Multiplying from the right by $a^{-1}$ then implies that $h_1 = h_2$, thus showing that $\alpha$ is one-to-one.

In the particular case of finite groups, the previous paragraph shows that any two right cosets of H in G must have the same number o(H) of elements. We also showed above that any two distinct right cosets have no elements in common. It then follows that any $a \in G$ is in the unique right coset Ha, and therefore the set of all right cosets of H in G must contain every element of G. This means that if there are k distinct right cosets of H in G, then we must have ko(H) = o(G) (i.e., $o(H) | o(G)$), and hence we have proved **Lagrange's theorem**:

**Theorem 1.9** If G is a finite group and H is a subgroup of G, then o(H) is a divisor of o(G).

The number o(G)/o(H) will be denoted by $i_G(H)$, and is usually called the **index** of H in G. (This is frequently denoted by [G : H].) The index of H in G is thus the number of distinct right cosets of H in G.

While we have restricted our discussion to right cosets, it is clear that everything could be repeated using **left cosets** defined in the obvious way. It should also be clear that for a general subgroup H of a group G, we need not have Ha = aH for any $a \in G$. However, if N is a subgroup of G such that for every $n \in N$ and $g \in G$ we have $gng^{-1} \in N$, then we say that N is a **normal** subgroup of G. An equivalent way of phrasing this is to say that N is a normal subgroup of G if and only if $gNg^{-1} \subset N$ for all $g \in G$ (where by $gNg^{-1}$ we mean the set of all $gng^{-1}$ with $n \in N$).

**Theorem 1.10** A subgroup N of G is normal if and only if $gNg^{-1} = N$ for every $g \in G$.

*Proof* If $gNg^{-1} = N$ for every $g \in G$, then clearly $gNg^{-1} \subset N$ so that N is normal. Conversely, suppose that N is normal in G. Then, for each $g \in G$ we have $gNg^{-1} \subset N$, and hence

$$g^{-1}Ng = g^{-1}N(g^{-1})^{-1} \subset N .$$

Using this result, we see that

$$N = (gg^{-1})N(gg^{-1}) = g(g^{-1}Ng)g^{-1} \subset gNg^{-1}$$

and therefore $N = gNg^{-1}$ (This also follows from Example 1.8). ∎

The reader should be careful to note that this theorem does not say that $gng^{-1} = n$ for every $n \in N$ and $g \in G$. This will in general not be true. The usefulness of this theorem is that it allows us to prove the following result.

**Theorem 1.11** A subgroup N of G is normal if and only if every left coset of N in G is also a right coset of N in G.

*Proof* If N is normal, then $gNg^{-1} = N$ for every $g \in G$, and hence gN = Ng. Conversely, suppose that every left coset gN is also a right coset. We show that in fact this right coset must be Ng. Since N is a subgroup it must contain the identity element e, and therefore $g = ge \in gN$ so that g must also be in whatever right coset it is that is identical to gN. But we also have eg = g so that g is in the right coset Ng. Then, since any two right cosets with an element in common must be identical, it follows that gN = Ng. Thus, we see that $gNg^{-1} = Ngg^{-1} = N$ so that N is normal. ∎

If G is a group and A, B are subsets of G, we define the set

$$AB = \{ab \in G: a \in A, b \in B\} .$$

In particular, if H is a subgroup of G, then $HH \subset H$ since H is closed under the group multiplication operation. But we also have $H = He \subset HH$ (since $e \in H$), and hence $HH = H$.

Now let N be a normal subgroup of G. By Theorem 1.11 we then see that

$$(Na)(Nb) = N(aN)b = N(Na)b = NNab = Nab .$$

In other words, the product of right cosets of a normal subgroup is again a right coset. This closure property suggests that there may be a way to construct a group out of the cosets Na where a is any element of G. We now show that there is indeed a way to construct such a group. Our method is used frequently throughout mathematics, and entails forming what is called a **quotient structure**.

Let G/N denote the collection of all right cosets of N in G. In other words, an element of G/N is a right coset of N in G. We use the product of subsets as defined above to define a product on G/N.

**Theorem 1.12** Let N be a normal subgroup of a group G. Then G/N is a group.

*Proof* We show that the product in G/N obeys properties (G1) – (G4) in the definition of a group.

(1) If $A, B \in G/N$, then $A = Na$ and $B = Nb$ for some $a, b \in G$, and hence (since $ab \in G$)

$$AB = NaNb = Nab \in G/N .$$

(2) If $A, B, C \in G/N$, then $A = Na$, $B = Nb$ and $C = Nc$ for some $a, b, c \in G$ and hence

$$\begin{aligned}(AB)C = (NaNb)Nc = (Nab)Nc = N(abN)c = N(Nab)c = N(ab)c \\ = Na(bc) = Na(Nbc) = Na(NbNc) = A(BC).\end{aligned}$$

(3) If $A = Na \in G/N$, then

$$AN = NaNe = Nae = Na = A$$

and similarly

$$NA = NeNa = Nea = Na = A .$$

Thus $N = Ne \in G/N$ serves as the identity element in G/N.

(4) If $Na \in G/N$, then $Na^{-1}$ is also in G/N, and we have

$$NaNa^{-1} = Naa^{-1} = Ne$$

as well as

$$Na^{-1}Na = Na^{-1}a = Ne \ .$$

Therefore $Na^{-1} \in G/N$ is the inverse to any element $Na \in G/N$. ∎

**Corollary** If N is a normal subgroup of a finite group G, then o(G/N) = o(G)/o(N).

*Proof* By construction, G/N consists of all the right cosets of N in G, and since this number is just the definition of $i_G(N)$, we see that o(G/N) = o(G)/o(N). ∎

The group defined in Theorem 1.12 is called the **quotient group** (or **factor group**) of G by N.

Let us now apply this quotient structure formalism to rings. Since any subgroup of an abelian group is automatically normal, and since any ring R is an abelian group under addition, any ideal I of R is therefore a normal subgroup of R (under addition). It is clear that we can now form the quotient group R/I where the elements of R/I are the cosets of I in R (since R is abelian, there is no need to distinguish between right and left cosets). We write these cosets as I + r (or r + I) for each $r \in R$. In the next theorem we show that R/I can in fact be made into a ring which is called the **quotient ring** of R by I.

**Theorem 1.13** Let I be an ideal of a ring R. For any $I + a, I + b \in R/I$ we define

$$(I + a) + (I + b) = I + (a + b)$$

and

$$(I + a)(I + b) = I + ab \ .$$

Then, with these operations, R/I forms a ring.

*Proof* From the proof of Theorem 1.12, it is obvious that R/I forms a group under addition if we use the composition rule (I + a) + (I + b) = I + (a + b) for all $a, b \in R$. We now turn our attention to the multiplication rule on R/I, and we begin by showing that this rule is well-defined. In other words, we must show that if I + a = I + a′ and I + b = I + b′, then I + ab = I + a′b′. From I + a = I + a′, we have a = x + a′ for some $x \in I$, and similarly b = y + b′ for some $y \in I$. Then

$$ab = (x + a')(y + b') = xy + xb' + a'y + a'b' .$$

But I is an ideal so that xy, xb′, and a′y are all elements of I, and hence z = xy + xb′ + a′y ∈ I. Therefore, ab = z + a′b′ so that

$$ab + I = a'b' + z + I = a'b' + I$$

as desired.

To show that R/I is a ring, we must verify that the properties (R1) – (R8) given in Section 1.4 hold in R/I. This is straightforward to do, and we give one example, leaving the rest to the reader (Exercise 1.5.5). To prove the first part of (R8), suppose a, b, c ∈ R. Then I + a, I + b, I + c ∈ R/I and hence

$$\begin{aligned}(I + a)[(I + b) + (I + c)] &= (I + a)[I + (b + c)] \\ &= I + a(b + c) \\ &= I + (ab + ac) \\ &= (I + ab) + (I + ac) \\ &= (I + a)(I + b) + (I + a)(I + c).\end{aligned}$$

**Example 1.13** Recall that the set $\mathbb{Z}$ of all integers forms a commutative ring with unit element (see Example 1.9). If we choose any n ∈ $\mathbb{Z}$, then n generates a principal ideal (n) that consists of all numbers of the form na for each a ∈ $\mathbb{Z}$. For example, the number 2 generates the principal ideal (2) that is nothing more than the ring of all even integers. The quotient ring $\mathbb{Z}$/(2) is then the set of all cosets of (2). Each of these cosets is either the set of even integers, or the even integers plus some odd integer. ∥

We have now finished essentially all of the mathematical formalism necessary to undertake a rigorous study of linear algebra. In the next chapter we begin our treatment of the subject matter proper of this text.

## Exercises

1. Let ϕ be a homomorphism of a group G into a group G′, and let $K_\phi$ be the kernel of ϕ. Prove that $K_\phi$ is a normal subgroup of G.

2. This exercise refers to the relation "congruence modulo n" defined in Example 1.11. Throughout this exercise, let n ∈ $\mathbb{Z}^+$ be arbitrary but fixed.
   (a) Show that this relation defines an equivalence relation.
   (b) Using Theorem 0.8 to divide a by n, show that the congruence relation has exactly n distinct equivalence sets.

(c) If $a \equiv b \pmod{n}$ and $c \equiv d \pmod{n}$, show that $a + c \equiv (b + d) \pmod{n}$ and $ac \equiv bd \pmod{n}$.
(d) Show $\mathbb{Z}/(n)$ is isomorphic to the integers mod n.

3. Let $\phi: R \to R'$ be a ring isomorphism. Show that $R'$ is commutative if R is.

4. Let $\phi: R \to R'$ be a ring isomorphism. Show that $R'$ is an integral domain if R is.

5. Finish the proof that R/I forms a ring in Theorem 1.13.

6. Prove that an integral domain is a field if and only if every nonzero element has a multiplicative inverse.

7. Show that the kernel of a ring homomorphism is an ideal.

8. Determine all the subgroups of the permutation group $S_3$. Which of these is normal?

9. Let $\mathcal{N}$ be a collection of normal subgroups of a group G. Show that the intersection of all $N \in \mathcal{N}$ is a normal subgroup of G.

10. Prove or disprove the following statement: If $\phi: R \to R'$ is a ring homomorphism, then the image of $\phi$ is an ideal of $R'$.

11. Let $\phi$ be a homomorphism of a group G onto a group $G'$, and let K be the kernel of $\phi$. By Exercise 5.1, we know that K is a normal subgroup of G, and hence we may form the quotient group G/K. Prove that G/K is isomorphic to $G'$. [*Hint*: Since any element in $X \in G/K$ is of the form Kg where $g \in G$, define the mapping $\psi: G/K \to G'$ by $\psi(X) = \psi(Kg) = \phi(g)$. To show that $\psi$ is an isomorphism, first show that $\psi$ is well-defined, that is, $X = Kg = Kg'$ implies $\phi(g) = \phi(g')$. Next, show that $\psi$ is a homomorphism, i.e., that $\psi(XY) = \psi(X)\,\psi(Y)$. Now show that $\psi$ is surjective (use the fact that $\phi$ is surjective). Finally, show that Ker $\psi = \{0\}$ (you will need the additional fact that the identity in G/K is K = Ke).]

12. Show that a field $\mathcal{F}$ can have no zero divisors.

13. Let H be a subgroup of a group G. Show that the set of all right cosets of H in G decomposes G into disjoint subsets.

14. The center of a group G is the set $Z = \{z \in G: zg = gz \text{ for all } g \in G\}$. Show that Z is a normal subgroup of G.

15. Show that the set {0, 1} with the usual addition and multiplication operations, but subject to $1 + 1 = 0$, forms a field of characteristic 2. (This is an example of a finite field.)

16. Let G be a group and let $G_1$, $G_2$ be subgroups of G with $G_1 \cap G_2 = e$. Show that $G_1$ and $G_2$ commute if and only if they are normal subgroups.