VII. Advanced Group Theory, 306-369

DOI: 10.3792/euclid/9781429799980-7

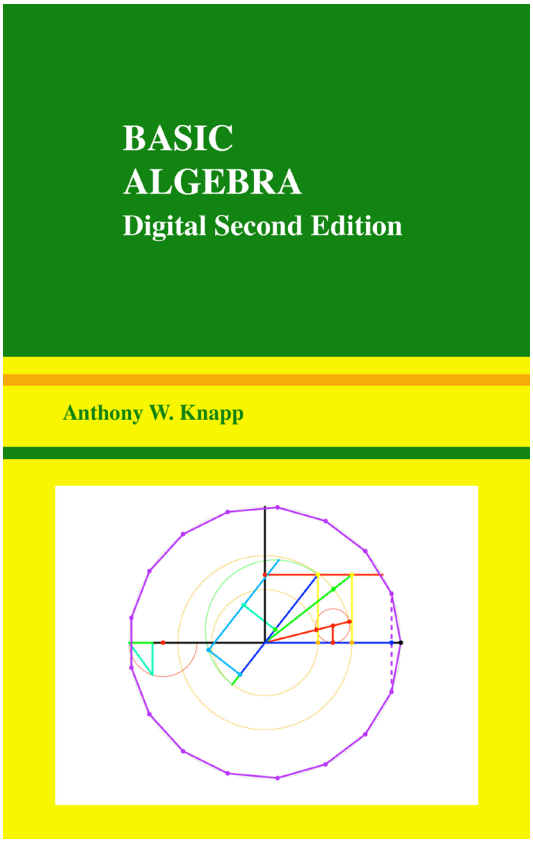
from

Basic Algebra

Digital Second Edition

Anthony W. Knapp

Full Book DOI: 10.3792/euclid/9781429799980
ISBN: 978-1-4297-9998-0



Anthony W. Knapp
81 Upper Sheep Pasture Road
East Setauket, N.Y. 11733–1729, U.S.A.
Email to: aknapp@math.stonybrook.edu
Homepage: www.math.stonybrook.edu/~aknapp

Title: Basic Algebra
Cover: Construction of a regular heptadecagon, the steps shown in color sequence; see page 505.

Mathematics Subject Classification (2010): 15–01, 20–01, 13–01, 12–01, 16–01, 08–01, 18A05, 68P30.

First Edition, ISBN-13 978-0-8176-3248-9

Published by Birkhäuser Boston

Digital Second Edition, not to be sold, no ISBN

Published by the Author

CHAPTER VII

Advanced Group Theory

Abstract. This chapter continues the development of group theory begun in Chapter IV, the main topics being the use of generators and relations, representation theory for finite groups, and group extensions. Representation theory uses linear algebra and inner-product spaces in an essential way, and a structure-theory theorem for finite groups is obtained as a consequence. Group extensions introduce the subject of cohomology of groups.

Sections 1–3 concern generators and relations. The context for generators and relations is that of a free group on the set of generators, and the relations indicate passage to a quotient of this free group by a normal subgroup. Section 1 constructs free groups in terms of words built from an alphabet and shows that free groups are characterized by a certain universal mapping property. This universal mapping property implies that any group may be defined by generators and relations. Computations with free groups are aided by the fact that two reduced words yield the same element of a free group if and only if the reduced words are identical. Section 2 obtains the Nielsen–Schreier Theorem that subgroups of free groups are free. Section 3 enlarges the construction of free groups to the notion of the free product of an arbitrary set of groups. Free product is what coproduct is for the category of groups; free groups themselves may be regarded as free products of copies of the integers.

Sections 4–5 introduce representation theory for finite groups and give an example of an important application whose statement lies outside representation theory. Section 4 contains various results giving an analysis of the space $C(G, \mathbb{C})$ of all complex-valued functions on a finite group G. In this analysis those functions that are constant on conjugacy classes are shown to be linear combinations of the characters of the irreducible representations. Section 5 proves Burnside's Theorem as an application of this theory—that any finite group of order $p^a q^b$ with p and q prime and with $a+b > 1$ has a nontrivial normal subgroup.

Section 6 introduces cohomology of groups in connection with group extensions. If N is to be a normal subgroup of G and Q is to be isomorphic to G/N, the first question is to parametrize the possibilities for G up to isomorphism. A second question is to parametrize the possibilities for G if G is to be a semidirect product of N and Q.

1. Free Groups

This section and the next two introduce some group-theoretic notions that in principle apply to all groups but in practice are used with countable groups, often countably infinite groups that are nonabelian. The material is especially useful in applications in topology, particularly in connection with fundamental groups and covering spaces. But the formal development here will be completely algebraic, not making use of any definitions or theorems from topology.

In the case of abelian groups, every abelian group G is a quotient of a suitable free abelian group, i.e., a suitable direct sum of copies of the additive group $\mathbb{Z}$ of integers.[1] Recall the discussion of Section IV.9: We introduce a copy $\mathbb{Z}_g$ of $\mathbb{Z}$ for each g in G, define $\widetilde{G} = \bigoplus_{g\in G} \mathbb{Z}_g$, let $i_g : \mathbb{Z}_g \to \widetilde{G}$ be the standard embedding, and let $\varphi_g : \mathbb{Z}_g \to G$ be the group homomorphism written additively as $\varphi_g(n) = ng$. The universal mapping property of direct sums that was stated as Proposition 4.17 produces a unique group homomorphism $\varphi : \widetilde{G} \to G$ such that $\varphi \circ i_g = \varphi_g$ for all g, and φ is the required homomorphism of a free abelian group onto G.

The goal in this section is to carry out an analogous construction for groups that are not necessarily abelian. The constructed groups, to be called "free groups," are to be rather concrete, and the family of all of them is to have the property that every group is the quotient of some member of the family.

If S is any set, we construct a "free group $F(S)$ on the set S." Let us speak of S as a set of "symbols" or as the members of an "alphabet," possibly infinite, with which we are working. If S is empty, the group $F(S)$ is taken to be the one-element trivial group, and we shall therefore now assume that S is not empty. If a is a symbol in S, we introduce a new symbol a^{-1} corresponding to it, and we let S^{-1} denote the set of all such symbols a^{-1} for $a \in S$. Define $S' = S \cup S^{-1}$. A **word** is a finite string of symbols from S', i.e., an ordered n-tuple for some n of members of S' with repetitions allowed. Words that are n-tuples are said to have **length** n. The empty word, with length 0, will be denoted by 1. Other words are usually written with the symbols juxtaposed and all commas omitted, as in $abca^{-1}cb^{-1}$. The set of words will be denoted by $W(S')$. We introduce a multiplication $W(S') \times W(S') \to W(S')$ by writing end-to-end the words that are to be multiplied: $(abca^{-1}, cb^{-1}) \mapsto abca^{-1}cb^{-1}$. The length of a product is the sum of the lengths of the factors. It is plain that this multiplication is associative and that 1 is a two-sided identity. It is not a group operation, however, since most elements of $W(S')$ do not have inverses: multiplication never decreases length, and thus the only way that 1 can be a product of two elements is as the product 11. To obtain a group from $W(S')$, we shall introduce an equivalence relation in $W(S')$.

Two words are said to be **equivalent** if one of the words can be obtained from the other by a finite succession of insertions and deletions of expressions aa^{-1} or $a^{-1}a$ within the word; here a is assumed to be an element of S. It will be convenient to refer to the pairs aa^{-1} and $a^{-1}a$ together; therefore when $b = a^{-1}$ is in S^{-1}, let us define $b^{-1} = (a^{-1})^{-1}$ to be a. Then two words are equivalent if one of the words can be obtained from the other by a finite succession of insertions and deletions of expressions of the form bb^{-1} with b in S'. This definition is

[1] Direct sum here is what coproduct, in the sense of Section IV.11, amounts to in the category of all abelian groups.

arranged so that "equivalent" is an equivalence relation. We write $x \sim y$ if x and y are words that are equivalent. The underlying set for the free group $F(S)$ will be taken to be the set of equivalence classes of members of $W(S')$.

Theorem 7.1. If S is a set and $W(S')$ is the corresponding set of words built from $S' = S \cup S^{-1}$, then the product operation defined on $W(S')$ descends in a well-defined fashion to the set $F(S)$ of equivalence classes of members of $W(S')$, and $F(S)$ thereby becomes a group. Define $\iota : S \to F(S)$ to be the composition of the inclusion into words of length one followed by passage to equivalence classes. Then the pair $(F(S), \iota)$ has the following universal mapping property: whenever G is a group and $\varphi : S \to G$ is a function, then there exists a unique group homomorphism $\widetilde{\varphi} : F(S) \to G$ such that $\varphi = \widetilde{\varphi} \circ \iota$.

REMARK. The group $F(S)$ is called the **free group** on S. Figure 7.1 illustrates its universal mapping property. The brief form in words of the property is that any function from S into a group G extends uniquely to a group homomorphism of $F(S)$ into G. This universal mapping property actually characterizes $F(S)$, as will be seen in Proposition 7.2.

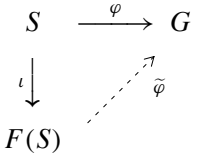


FIGURE 7.1. Universal mapping property of a free group.

PROOF. Let us denote equivalence classes by brackets. We want to define multiplication in $F(S)$ by $[w_1][w_2] = [w_1 w_2]$. To see that this formula makes sense in $F(S)$, let x_1, x_2, and y be words, and let b be in S'. Define $x = x_1 x_2$ and $x' = x_1 b b^{-1} x_2$, so that $x' \sim x$. Then it is evident that $x'y \sim xy$ and $yx' \sim yx$. Iteration of this kind of relationship shows that $w_1' \sim w_1$ and $w_2' \sim w_2$ implies $w_1' w_2' \sim w_1 w_2$, and hence multiplication of equivalence classes is well defined.

Since multiplication in $W(S')$ is associative, we have $[w_1]([w_2][w_3]) = [w_1][w_2 w_3] = [w_1(w_2 w_3)] = [(w_1 w_2)w_3] = [w_1 w_2][w_3] = ([w_1][w_2])[w_3]$. Thus multiplication is associative in $F(S)$. The class $[1]$ of the empty word 1 is a two-sided identity. If $b_1, \ldots, b_n$ are in S', then $b_n^{-1} \cdots b_2^{-1} b_1^{-1} b_1 b_2 \cdots b_n$ is equivalent to 1, and so is $b_1 b_2 \cdots b_n b_n^{-1} \cdots b_2^{-1} b_1^{-1}$. Consequently $[b_n^{-1} \cdots b_2^{-1} b_1^{-1}]$ is a two-sided inverse of $[b_1 b_2 \cdots b_n]$, and $F(S)$ is a group.

Now we address the universal mapping property, first proving the stated uniqueness of the homomorphism. Every member of $F(S)$ is the product of classes $[b]$ with b in S'. In turn, if b is of the form a^{-1} with a in S, then $[b] = [a]^{-1}$. Hence $F(S)$ is generated by all classes $[a]$ with a in S, i.e., by $\iota(S)$. Any homomorphism

of a group is determined by its values on the members of a generating set, and uniqueness therefore follows from the formula $\widetilde{\varphi}([a]) = \widetilde{\varphi}(\iota(a)) = \varphi(a)$.

For existence we begin by defining a function $\Phi : W(S') \to G$ such that

$$\begin{aligned} \Phi(a) &= \varphi(a) && \text{for } a \text{ in } S, \\ \Phi(a^{-1}) &= \varphi(a)^{-1} && \text{for } a^{-1} \text{ in } S^{-1}, \\ \Phi(w_1 w_2) &= \Phi(w_1)\Phi(w_2) && \text{for } w_1 \text{ and } w_2 \text{ in } W(S'). \end{aligned}$$

We use the formulas $\Phi(a) = \varphi(a)$ for a in S and $\Phi(a^{-1}) = \varphi(a)^{-1}$ for a^{-1} in S^{-1} as a definition of $\Phi(b)$ for b in S'. Any member of $W(S')$ can be written uniquely as $b_1 \cdots b_n$ with each b_j in S', and we set $\Phi(b_1 \cdots b_n) = \Phi(b_1) \cdots \Phi(b_n)$. (If $n = 0$, the understanding is that $\Phi(1) = 1$.) Then Φ has the required properties.

Let us show that $w' \sim w$ implies $\Phi(w') = \Phi(w)$. If $b_1, \ldots, b_n$ are in S' and b is in S', then the question is whether

$$\Phi(b_1 \cdots b_k b b^{-1} b_{k+1} \cdots b_n) \stackrel{?}{=} \Phi(b_1 \cdots b_k b_{k+1} \cdots b_n).$$

If g and g' denote the elements $\Phi(b_1) \cdots \Phi(b_k)$ and $\Phi(b_{k+1}) \cdots \Phi(b_n)$ of G, then the two sides of the queried formula are

$$g\Phi(b)\Phi(b^{-1})g' \qquad \text{and} \qquad gg'.$$

Thus the question is whether $\Phi(b)\Phi(b^{-1})$ always equals 1 in G. If $b = a$ is in S, this equals $\varphi(a)\varphi(a)^{-1} = 1$, while if $b = a^{-1}$ is in S^{-1}, it equals $\varphi(a)^{-1}\varphi(a) = 1$. We conclude that $w' \sim w$ implies $\Phi(w') = \Phi(w)$.

We may therefore define $\widetilde{\varphi}([w]) = \Phi(w)$ for $[w]$ in $F(S)$. Since $\widetilde{\varphi}([w][w']) = \widetilde{\varphi}([ww']) = \Phi(ww') = \Phi(w)\Phi(w') = \widetilde{\varphi}([w])\widetilde{\varphi}([w'])$, $\widetilde{\varphi}$ is a homomorphism of $F(S)$ into G. For a in S, we have $\widetilde{\varphi}([a]) = \Phi(a) = \varphi(a)$. In other words, $\widetilde{\varphi}(\iota(a)) = \varphi(a)$. This completes the proof of existence. □

Proposition 7.2. Let S be a set, F be a group, and $\iota' : S \to F$ be a function. Suppose that the pair (F, ι') has the following universal mapping property: whenever G is a group and $\varphi : S \to G$ is a function, then there exists a unique group homomorphism $\widetilde{\varphi} : F \to G$ such that $\varphi = \widetilde{\varphi} \circ \iota'$. Then there exists a unique group homomorphism $\Phi : F(S) \to F$ such that $\iota' = \Phi \circ \iota$, and it is a group isomorphism.

REMARKS. Chapter VI is not a prerequisite for the present chapter. However, readers who have been through Chapter VI will recognize that Proposition 7.2 is a special case of Problem 19 at the end of that chapter.

PROOF. We apply the universal mapping property of $(F(S), \iota)$, as stated in Theorem 7.1, to the group $G = F$ and the function $\varphi = \iota'$, obtaining a group homomorphism $\Phi : F(S) \to F$ such that $\iota' = \Phi \circ \iota$. Then we apply the given universal mapping property of (F, ι') to the group $G = F(S)$ and the function $\varphi = \iota$, obtaining a group homomorphism $\Psi : F \to F(S)$ such that $\iota = \Psi \circ \iota'$.

The group homomorphism $\Psi \circ \Phi : F(S) \to F(S)$ has the property that $(\Psi \circ \Phi) \circ \iota = \Psi \circ (\Phi \circ \iota) = \Psi \circ \iota' = \iota$, and the identity $1_{F(S)}$ has this same property. By the uniqueness of the group homomorphism in Theorem 7.1, $\Psi \circ \Phi = 1_{F(S)}$.

Similarly the group homomorphism $\Phi \circ \Psi : F \to F$ has the property that $(\Phi \circ \Psi) \circ \iota' = \iota'$, and the identity 1_F has this same property. By the uniqueness of the group homomorphism in the assumed universal mapping property of F, $\Phi \circ \Psi = 1_F$.

Therefore Φ is a group isomorphism. We know that $\iota(S)$ generates $F(S)$. If $\Phi' : F(S) \to F$ is another group isomorphism with $\iota' = \Phi' \circ \iota$, then Φ' and Φ agree on $\iota(S)$ and therefore have to agree everywhere. Hence Φ is unique. □

Proposition 7.2 raises the question of recognizing candidates for the set $T = \iota'(S)$ in a given group F so as to be in a position to exhibit F as isomorphic to the free group $F(S)$. Certainly T has to generate F. But there is also an independence condition. The idea is that if we form words from the members of T, then two words are to lead to equal members of F only if they can be transformed into one another by the same rules that are allowed with free groups.

What this problem amounts to in the case that $F = F(S)$ is that we want a decision procedure for telling whether two given words are equivalent. This is the so-called **word problem** for the free group. If we think about the matter for a moment, not much is instantly obvious. If a_1 and a_2 are two members of S and if they are considered as words of length 1, are they equivalent? Equivalence allows for inserting pairs bb^{-1} with b in S', as well as deleting them. Might it be possible to do some complicated iterated insertion and deletion of pairs to transform a_1 into a_2? Although the negative answer can be readily justified in this situation by a parity argument, it can be justified even more easily by the universal mapping property: there exist groups G with more than one element; we can map a_1 to one element of G and a_2 to another element of G, extend to a homomorphism $\widetilde{\varphi} : F(S) \to G$, see that $\widetilde{\varphi}(\iota(a_1)) \neq \widetilde{\varphi}(\iota(a_2))$, and conclude that $\iota(a_1) \neq \iota(a_2)$. But what about the corresponding problem for two more-complicated words in a free group? Fortunately there is a decision procedure for the word problem in a free group. It involves the notion of "reduced" words. A word in $W(S')$ is said to be **reduced** if it contains no consecutive pair bb^{-1} with b in S'.

Proposition 7.3 (solution of the word problem for free groups). Let S be a set, let $S' = S \cup S^{-1}$, and let $W(S')$ be the corresponding set of words. Then each word in $W(S')$ is equivalent to one and only one reduced word.

REMARK. To test whether two words are equivalent, the proposition says to delete pairs bb^{-1} with $b \in S'$ as much as possible from each given word, and to check whether the resulting reduced words are identical.

PROOF. Removal of a pair bb^{-1} with $b \in S'$ decreases the length of a word by 2, and the length has to remain ≥ 0. Thus the process of successively removing such pairs has to stop after finitely many steps, and the result is a reduced word. This proves that each equivalence class contains a reduced word.

For uniqueness we shall associate to each word a finite sequence of reduced words such that the last member of the sequence is unchanged when we insert or delete within the given word any expression bb^{-1} with $b \in S'$. Specifically if $w = b_1 \cdots b_n$, with each b_i in S', is a given word, we associate to w the sequence of words $x_0, x_1, \ldots, x_n$ defined inductively by

$$\begin{aligned} &x_0 = 1, \\ &x_1 = b_1, \\ &x_i = \begin{cases} x_{i-1}b_i & \text{if } i \geq 2 \text{ and } x_{i-1} \text{ does not end in } b_i^{-1}, \\ y_{i-2} & \text{if } i \geq 2 \text{ and } x_{i-1} = y_{i-2}b_i^{-1}, \end{cases} \end{aligned} \tag{$*$}$$

and we define $r(w) = x_n$. Let us see, by induction on $i \geq 0$, that x_i is reduced. The base cases $i = 0$ and $i = 1$ are clear from the definition. Suppose that $i \geq 2$ and that $x_0, \ldots, x_{i-1}$ are reduced. If $x_{i-1} = y_{i-2}b_i^{-1}$ for some y_{i-2}, then x_{i-1} reduced forces y_{i-2} to be reduced, and hence $x_i = y_{i-2}$ is reduced. If x_{i-1} does not end in b_i^{-1}, then the last two symbols of $x_i = x_{i-1}b_i$ do not cancel, and no earlier pair can cancel since x_{i-1} is assumed reduced; hence x_i is reduced. This completes the induction and shows that x_i is reduced for $0 \leq i \leq n$.

If the word $w = b_1 \cdots b_n$ is reduced, then each x_i for $i \geq 2$ is determined by the first of the two choices in $(*)$, and hence $x_i = b_1 \cdots b_i$ for all i. Consequently $r(w) = w$ if w is reduced. If we can prove for a general word $b_1 \cdots b_n$ that

$$r(b_1 \cdots b_n) = r(b_1 \cdots b_k bb^{-1} b_{k+1} \cdots b_n), \tag{$**$}$$

then it follows that every word w' equivalent to a word w has $r(w') = r(w)$. Since $r(w) = w$ for w reduced, there can be only one reduced word in an equivalence class.

To prove $(**)$, let $x_0, \ldots, x_n$ be the finite sequence associated with $b_1 \cdots b_n$, and let $x'_0, \ldots, x'_{n+2}$ be the sequence associated with $b_1 \cdots b_k bb^{-1} b_{k+1} \cdots b_n$. Certainly $x_i = x'_i$ for $i \leq k$. Let us compute x'_{k+1} and x'_{k+2}. From $(*)$ we see that

$$x'_{k+1} = \begin{cases} x_k b & \text{if } x_k \text{ does not end in } b^{-1}, \\ y & \text{if } x_k = yb^{-1}. \end{cases}$$

In the first of these cases, x'_{k+1} ends in b, and $(*)$ says therefore that $x'_{k+2} = x_k$. In the second of the cases, the fact that x_k is reduced implies that y does not end in b; hence $(*)$ says that $x'_{k+2} = yb^{-1} = x_k$. In other words, $x'_{k+2} = x_k$ in both cases. Since the inductive definition of any x_i depends only on x_{i-1}, and similarly for x'_i, we see that $x'_{k+2+i} = x_{k+i}$ for $0 \le i \le n-k$. Therefore $x'_{n+2} = x_n$, and $(**)$ follows. This proves the proposition. □

Let us return to the problem of recognizing candidates for the set $T = \iota'(S)$ in a given group F so that the subgroup generated by T is a free group. Using the universal mapping property for the free group $F(T)$, we form the group homomorphism of $F(T)$ into F that extends the identity mapping on T. We want this homomorphism to be one-one, i.e., to have the property that the only way a word in F built from the members of T can equal the identity is if it comes from the identity. Because of Proposition 7.3 the only reduced word in $F(T)$ that yields the identity is the empty word. Thus the condition that the homomorphism be one-one is that the only image in F of a reduced word in $F(T)$ that can equal the identity is the image of the empty word. Making this condition into a definition, we say that a subset $S = \{g_t \mid t \in T\}$ of F not containing 1 is **free** if no nonempty product $h_1 h_2 \cdots h_m$ in which each h_i or h_i^{-1} is in S and each h_{i+1} is different from h_i^{-1} can be the identity. A free set in F that generates F is called a **free basis** for F.

EXAMPLE. Within the free group $F(\{x, y\})$ on two generators x and y, consider the subgroup generated by $u = x^2$, $v = y^2$, and $w = xy$. The claim is that the subset $\{u, v, w\}$ is free, so that the subgroup generated by u, v, and w is isomorphic to a free group $F(\{u, v, w\})$ on three generators. We are to check that no nonempty reduced word in $u, v, w, u^{-1}, v^{-1}, w^{-1}$ can reduce to the empty word after substitution in terms of x and y. We induct on the length of the u, v, w word, the base case being length 0. Suppose that $v = y^2$ occurs somewhere in our reduced u, v, w word that collapses to the empty word after substitution. Consider what is needed for the left-hand factor of y in the y^2 to cancel. The cancellation must result from the presence of some y^{-1}. Suppose that this y^{-1} occurs to the left of y^2. Since passing to a reduced word need involve only deletions and not insertions of pairs, everything between y^{-1} and y^2 must cancel. If the y^{-1} has resulted from $w^{-1} = y^{-1}x^{-1}$, then the number of x, y symbols between y^{-1} and y^2 is odd, and an odd number of factors can never cancel. So the y^{-1} must arise from the right-hand y^{-1} in a factor $v^{-1} = y^{-2}$. The symbols between y^{-2} and y^2 come from some reduced u, v, w word, and induction shows that this word must be trivial. Then y^{-2} and y^2 are adjacent, contradiction. Thus the left factor of y^2 must cancel because of some y^{-1} on the right of y^2. If the y^{-1} is part of $w^{-1} = y^{-1}x^{-1}$ or is the left y^{-1} in $v^{-1} = y^{-2}$, then the number of x, y

symbols between the left y and the y^{-1} is odd, and we cannot get cancellation. So the y^{-1} must be the right-hand y^{-1} in a factor y^{-2}. Then we have an expression $y(y\cdots y^{-1})y^{-1}$ in which the symbols in parentheses cancel. The symbols $\cdots$ must cancel also; since these represent some reduced u, v, w word, induction shows that $\cdots$ is empty. We conclude that y^2 and y^{-2} are adjacent, contradiction. Thus our reduced u, v, w word contains no factor v. Similarly examination of the right-hand factor x in an occurrence of x^2 shows that our reduced u, v, w word contains no factor u. It must therefore be a product of factors w or a product of factors w^{-1}. Substitution of $w = xy$ leads directly without any cancellation to an x, y reduced word, and we conclude that the u, v, w word is empty. Thus the subset $\{u, v, w\}$ is free.

If G is any group, the **commutator subgroup** G' of G is the subgroup generated by all elements $xyx^{-1}y^{-1}$ with $x \in G$ and $y \in G$.

Proposition 7.4. If G is a group, then the commutator subgroup is normal, and G/G' is abelian. If $\varphi : G \to H$ is any homomorphism of G into an abelian group H, then $\ker\varphi \supseteq G'$.

PROOF. The computation

$$axyx^{-1}y^{-1}a^{-1} = (axa^{-1})(aya^{-1})(axa^{-1})^{-1}(aya^{-1})^{-1}$$

shows that G' is normal. If $\psi : G \to G/G'$ is the quotient homomorphism, then $\psi(x)\psi(y) = xyG' = xy(y^{-1}x^{-1}yx)G' = yxG' = \psi(y)\psi(x)$, and therefore G/G' is abelian. Finally if $\varphi : G \to H$ is a homomorphism of G into an abelian group H, then the computation $\varphi(xyx^{-1}y^{-1}) = \varphi(x)\varphi(y)\varphi(x)^{-1}\varphi(y)^{-1} = \varphi(x)\varphi(x)^{-1}\varphi(y)\varphi(y)^{-1} = 1$ shows that $G' \subseteq \ker\varphi$. □

Corollary 7.5. If F is the free group on a set S and if F' is the commutator subgroup of F, then F/F' is isomorphic to the free abelian group $\bigoplus_{s\in S}\mathbb{Z}_s$.

PROOF. Let $H = \bigoplus_{s\in S}\mathbb{Z}_s$, and let $\varphi : S \to H$ be the function with $\varphi(s) = 1_s$, i.e., $\varphi(s)$ is to be the member of H that is 1 in the s^{th} coordinate and is 0 elsewhere. Application of the universal mapping property of F as given in Theorem 7.1 yields a group homomorphism $\widetilde{\varphi} : F \to H$ such that $\widetilde{\varphi} \circ \iota = \varphi$. Since the elements $\varphi(s)$, with s in S, generate H, $\widetilde{\varphi}$ carries F onto H. Since H is abelian, Proposition 7.4 shows that $\ker\widetilde{\varphi} \supseteq F'$. Proposition 4.11 shows that $\widetilde{\varphi}$ descends to a homomorphism $\widetilde{\varphi}_0 : F/F' \to H$, and $\widetilde{\varphi}_0$ has to be onto H.

To complete the proof, we show that $\widetilde{\varphi}_0$ is one-one. Let x be a member of F. Since the products of the elements $\iota(s)$ and their inverses generate F and since F/F' is abelian, we can write $xF' = s_{i_1}^{j_1}\cdots s_{i_n}^{j_n}F'$, where s_{i_1} occurs a total of j_1 times in $x, \ldots,$ and s_{i_n} occurs a total of j_n times in x; it is understood that

an occurrence of $s_{i_1}^{-1}$ is to contribute -1 toward j_1. Then we have $\widetilde{\varphi}_0(xF') = j_1\varphi(s_{i_1}) + \cdots + j_n\varphi(s_{i_n})$. If $\widetilde{\varphi}_0(xF') = 0$, we obtain $j_1\varphi(s_{i_1}) + \cdots + j_n\varphi(s_{i_n}) = 0$, and then $j_1 = \cdots = j_n = 0$ since the elements $\varphi(s_{i_1}), \ldots, \varphi(s_{i_n})$ are members of a $\mathbb{Z}$ basis of H. Hence $xF' = F'$, x is in F', and $\widetilde{\varphi}_0$ is one-one. □

Corollary 7.6. If F_1 and F_2 are isomorphic free groups on sets S_1 and S_2, respectively, then S_1 and S_2 have the same cardinality.

PROOF. Corollary 7.5 shows that an isomorphism of F_1 with F_2 induces an isomorphism of the free abelian groups $\bigoplus_{s\in S_1} \mathbb{Z}_{s_1}$ and $\bigoplus_{s\in S_2} \mathbb{Z}_{s_2}$. The rank of a free abelian group is a well-defined cardinal, and the result follows—almost.

We did not completely prove this fact about the rank of a free abelian group in Section IV.9. Theorem 4.53 did prove, however, that rank is well defined for finitely generated free abelian groups. Thus the corollary follows if S_1 and S_2 are finite. If S_1 or S_2 is uncountable, then the cardinality of the corresponding free abelian group matches the cardinality of its $\mathbb{Z}$ basis; hence the corollary follows if S_1 or S_2 is uncountable. The only remaining case to eliminate is that one of S_1 and S_2, say the first of them, has a countably infinite $\mathbb{Z}$ basis and the other has finite rank n. The first of the groups then has a linearly independent set of $n + 1$ elements, and Lemma 4.54 shows that the span of these elements cannot be isomorphic to a subgroup of a free abelian group of rank n. This completes the proof in all cases. □

Because of Corollary 7.6, it is meaningful to speak of the **rank** of a free group; it is the cardinality of any free basis. We shall see in the next section that any subgroup of a free group is free. In contrast to the abelian case, however, the rank may actually increase in passing from a free group to one of its subgroups: the example earlier in this section exhibited a free group of rank 3 as a subgroup of a free group of rank 2.

We turn to a way of describing general groups, particularly groups that are at most countable. The method uses "generators," which we already understand, and "relations," which are defined in terms of free groups. Let S be a set, let R be a subset of $F(S)$, and let $N(R)$ be the smallest *normal* subgroup of $F(S)$ containing R. The group $G = F(S)/N(R)$ is sometimes written as $G = \langle S;\ R\rangle$ or as

$$G = \langle \text{elements of } S;\ \text{elements of } R\rangle,$$

with the elements of S and R listed rather than grouped as a set. Either of these expressions is called a **presentation** of G. The set S is a set of **generators**, and the set R is the corresponding set of **relations**. The following result implicit in the universal mapping property of Theorem 7.1 shows the scope of this definition.

Proposition 7.7. Each group G is the homomorphic image of a free group.

PROOF. Let S be a set of generators for G; for example, S can be taken to be G itself. Let $\varphi : S \to G$ be the inclusion of the set of generators into G, and let $\widetilde{\varphi} : F(S) \to G$ be the group homomorphism of Theorem 7.1 such that $\widetilde{\varphi}(\iota(s)) = \varphi(s)$ for all s in S. The image of $\widetilde{\varphi}$ is a subgroup of G that contains the generating set S and is therefore equal to all of G. Thus $\widetilde{\varphi}$ is the required homomorphism. □

If G is any group and $\widetilde{\varphi} : F(S) \to G$ is the homomorphism given in Proposition 7.7, then the subgroup $R = \ker \widetilde{\varphi}$ has the property that $G \cong \langle S;\ R\rangle$. Consequently *every* group can be given by generators and relations.

For example the proof of the proposition shows that one possibility is to take $S = G$ and R equal to the set of all members of the multiplication table, but with the multiplication table entry $ss' = s''$ rewritten as the left side $ss'(s'')^{-1}$ of an equation $ss'(s'')^{-1} = 1$ specifying a combination of generators that maps to 1. This is of course not a very practical example. Generators and relations are most useful when S and R are fairly small. One says that G is **finitely generated** if S can be chosen to be finite, **finitely presented** if both S and R can be chosen to be finite.

A frequently used device in working with generators and relations is the following simple proposition.

Proposition 7.8. Let $G = \langle S;\ R\rangle$ be a group given by generators and relations, let G' be a second group, let φ be a one-one function φ from S onto a set of generators for G', and let $\Phi : F(S) \to G'$ be the extension of φ to a group homomorphism. If $\Phi(r) = 1$ for every member r of R, then Φ descends to a homomorphism of G onto G'. In particular, if $G = \langle S;\ R\rangle$ and $G' = \langle S;\ R'\rangle$ are groups given by generators and relations with $R \subseteq R'$, then the natural homomorphism of $F(S)$ onto G' descends to a homomorphism of G onto G'.

PROOF. The proposition follows immediately from the universal mapping property in Theorem 7.1 in combination with Proposition 4.11. □

Now let us consider some examples of groups given by generators and relations. The case of one generator is something we already understand: the group has to be cyclic. A presentation of $\mathbb{Z}$ is as $\langle a;\ \rangle$, and a presentation of C_n is as $\langle a;\ a^n\rangle$. But other presentations are possible with one generator, such as $\langle a;\ a^6, a^9\rangle$ for C_3. Here is an example with two generators.

EXAMPLE. Let us prove that $D_n \cong \langle x, y;\ x^n, y^2, (xy)^2 \rangle$, where D_n is the dihedral group of order $2n$. Concretely let us work with D_n as the group of 2-by-2 real matrices generated by $\begin{pmatrix} \cos 2\pi/n & -\sin 2\pi/n \\ \sin 2\pi/n & \cos 2\pi/n \end{pmatrix}$ and $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. The generated group indeed has order $2n$. If we identify

$$x \text{ with } \begin{pmatrix} \cos 2\pi/n & -\sin 2\pi/n \\ \sin 2\pi/n & \cos 2\pi/n \end{pmatrix} \qquad \text{and} \qquad y \text{ with } \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

then $y^2 = 1$, and the formula

$$\begin{pmatrix} \cos 2\pi/n & -\sin 2\pi/n \\ \sin 2\pi/n & \cos 2\pi/n \end{pmatrix}^k = \begin{pmatrix} \cos 2\pi k/n & -\sin 2\pi k/n \\ \sin 2\pi k/n & \cos 2\pi k/n \end{pmatrix}$$

shows that $x^n = 1$. In addition, $xy = \begin{pmatrix} \cos 2\pi/n & \sin 2\pi/n \\ \sin 2\pi/n & -\cos 2\pi/n \end{pmatrix}$, and the square of this is the identity. By Proposition 7.8, D_n is a homomorphic image of $\widetilde{D}_n = \langle x, y;\ x^n, y^2, (xy)^2 \rangle$. To complete the identification, it is enough to show that the order of $\widetilde{D}_n$ is $\leq 2n$ because the homomorphism of $\widetilde{D}_n$ onto D_n must then be one-one. In $\langle x, y;\ x^n, y^2, (xy)^2 \rangle$, we compute that $y^{-1} = y$ and that $x(yx)y = 1$ implies $yx = x^{-1}y^{-1} = x^{-1}y$. Induction then yields $yx^k = x^{-k}y$ for $k > 0$. Multiplying left and right by y gives $yx^{-k} = x^k y$ for $k > 0$. So $yx^l = x^{-l}y$ for every integer l. This means that every element is of the form x^m or $x^m y$, and we may take $0 \leq m \leq n-1$. Hence there are at most $2n$ elements.

Without trying to be too precise, let us mention that the **word problem** for finitely presented groups is to give an algorithm for deciding whether two words represent the same element of the group. It is known that there is no such algorithm applicable to all finitely presented groups. Of course, there can be such an algorithm for certain special classes of presentations. For example, if there are no relations in the presentation, then the group is a free group, and Proposition 7.3 gives a solution in this case. There tends to be a solution for a class of groups if the groups all correspond rather concretely to some geometric situation, such as a tiling of Euclidean space or some other space. The example above with D_n is of this kind.

By way of a concrete class of examples, one can identify any doubly generated group of the form $\langle x, y;\ x^a, y^b, (xy)^c \rangle$ if a, b, c are integers > 1, and one can describe what words represent what elements in these groups. These groups all correspond to tilings in 2 dimensions. In fact, let $\gamma = a^{-1} + b^{-1} + c^{-1}$. If $\gamma > 1$, the tiling is of the Riemann sphere, and the group is finite. If $\gamma = 1$, the tiling is of the Euclidean plane $\mathbb{R}^2$, and the group is infinite. If $\gamma < 1$, the tiling is of the hyperbolic plane, and the group is infinite. In all cases one starts from a triangle in the appropriate geometry with angles π/a, π/b, and π/c, and a basic tile consists of the double of this triangle obtained by reflecting the triangle about any of its

sides. The group elements x, y, and xy are rotations, suitably oriented, about the vertices of the triangle through respective angles $2\pi/a$, $2\pi/b$, and $2\pi/c$. Further information about the cases $\gamma > 1$ and $\gamma = 1$ is obtained in Problems 37–46 at the end of the chapter.

We conclude with one further example of a presentation whose group we can readily identify concretely.

Proposition 7.9. Let S be a set, and let $R = \{sts^{-1}t^{-1} \mid s \in S,\ t \in S\}$. Then the smallest normal subgroup of the free group $F(S)$ containing R is the commutator subgroup $F(S)'$, and therefore $\langle S;\ R\rangle$ is isomorphic to the free abelian group $\bigoplus_{s\in S} \mathbb{Z}_s$.

PROOF. The members of R are in $F(S)'$, the product of two members of $F(S)'$ is in $F(S)'$, and any conjugate of a member of $F(S)'$ is in $F(S)'$. Therefore the smallest normal subgroup $N(R)$ containing R has $N(R) \subseteq F(S)'$. Let $\varphi : F(S) \to F(S)/N(R)$ be the quotient homomorphism. Elements of the quotient $F(S)/N(R)$ may be expressed as words in the elements $\varphi(s)$ and $\varphi(s)^{-1}$ for s in S, and the factors commute because of the definition of R. Therefore $F(S)/N(R)$ is abelian. By Proposition 7.4, $N(R) \supseteq F(S)'$. Therefore $N(R) = F(S)'$. This proves the first conclusion, and the second conclusion follows from Corollary 7.5. □

2. Subgroups of Free Groups

The main result of this section is that any subgroup of a free group is a free group. An example in the previous section shows that the rank can actually increase in the process of passing to the subgroup.

The proof of the main result is ostensibly subtle but is relatively easy to understand in topological terms. Although we shall give the topological interpretation, we shall not pursue it further, and the proof that we give may be regarded as a translation of the topological proof into the language of algebra, combined with some steps of beautification.

For purposes of the topological argument, let us think of the given free group for the moment as finitely generated, and let us suppose that the subgroup has finite index. A free group on n symbols is the fundamental group of a bouquet of n circles, all joined at a single point, which we take as the base point. By the theory of covering spaces, any subgroup of index k is the fundamental group of some k-sheeted covering space of the bouquet of circles. This covering space is a 1-dimensional simplicial complex, and one can prove with standard tools that the fundamental group of any 1-dimensional simplicial complex is a free group. The theorem follows.

If the special hypotheses are dropped that the given free group is finitely generated and the subgroup has finite index, then the same proof is applicable as long as one allows a suitable generalization of the notion of simplicial complex. Thus the topological argument is completely general.

The theorem then is as follows.

Theorem 7.10 (Nielsen–Schreier Theorem). Every subgroup of a free group is a free group.

REMARKS. The algebraic proof will occupy the remainder of the section but will occasionally be interrupted by comments about the example in the previous section.

Let the given free group be F, let the subgroup be H, and form the right cosets Hg in F. Let C be a set of representatives for these cosets, with 1 chosen as the representative of the identity coset; we shall impose further conditions on C shortly.

EXAMPLE, CONTINUED. For the example in the previous section, we were given a free group F with two generators x, y, and the subgroup H is taken to have generators x^2, xy, y^2. In fact, one readily checks that H is the subgroup formed from all words of even length, and we shall think of it that way. The set C of coset representatives may be taken to be $\{1, x\}$ in this case. The argument we gave that H is free has points of contact with the proof we give of Theorem 7.10 but is not an exact special case of it. One point of contact is that within each generator of H that we identify, there is some particular factor that does not cancel when that generator appears in a word representing a member of the subgroup.

We define a function $\rho : F \to C$ by taking $\rho(x)$ to be the coset representative of the member x of F. This function has the property that $\rho(hx) = \rho(x)$ for all h in H and x in F. Also, $x \mapsto x\rho(x)^{-1}$ is a function from F to H, and it is the identity function on H. The first lemma shows that a relatively small subset of the elements $x\rho(x)^{-1}$ is a set of generators of H.

Lemma 7.11. Let S be the set of generators of F, and let $S' = S \cup S^{-1}$. Every element of H is a product of elements of the form $gb\rho(gb)^{-1}$ with g in C and b in S'. Furthermore the element $g' = \rho(gb)$ of C has the properties that $g = \rho(g'b^{-1})$ and that $gb^{-1}\rho(gb^{-1})^{-1}$ is of the form $\left(g'b\rho(g'b)^{-1}\right)^{-1}$. Consequently the elements $ga\rho(ga)^{-1}$ with g in C and a in S form a set of generators of H.

EXAMPLE, CONTINUED. In the example, we are taking $C = \{1, x\}$ and $S = \{x, y\}$. The elements $gb\rho(gb)^{-1}$ obtained with g=1 and b equal to x, y, x^{-1}, y^{-1} are 1, yx^{-1}, $x^{-1}x^{-1}$, and $y^{-1}x^{-1}$. The elements $gb\rho(gb)^{-1}$ obtained with $g = x$ and b equal to x, y, x^{-1}, y^{-1} are xx, xy, 1, and xy^{-1}. The lemma says that 1, yx^{-1}, xx, and xy form a set of generators of H and that the elements $x^{-1}x^{-1}$, $y^{-1}x^{-1}$, 1, and xy^{-1} are inverses of these generators in some order.

REMARK. The lemma needs no hypothesis that F is free. A nontrivial application of the lemma with F not free appears in Problem 43 at the end of the chapter.

PROOF. Any h in F can be written as a product $h = b_1 \cdots b_n$ with each b_j in S'. Define $r_0 = 1$ and $r_k = \rho(b_1 \cdots b_k)$ for $1 \le k \le n$. Then

$$hr_n^{-1} = (r_0 b_1 r_1^{-1})(r_1 b_2 r_2^{-1}) \cdots (r_{n-1} b_n r_n^{-1}). \qquad (*)$$

Since

$$r_k = \rho(b_1 \cdots b_k) = \rho(b_1 \cdots b_{k-1} b_k) = \rho(\rho(b_1 \cdots b_{k-1}) b_k) = \rho(r_{k-1} b_k),$$

we have $r_{k-1} b_k r_k^{-1} = gb\rho(gb)^{-1}$ with $g = r_{k-1}$ and $b = b_k$. Thus $(*)$ exhibits hr_n^{-1} as a product of elements as in the first conclusion of the lemma. Since $r_n = \rho(b_1 \cdots b_n) = \rho(h)$, $r_n = 1$ if h is in H. Therefore in this case, h itself is a product of elements as in the statement of that conclusion, and that conclusion is now proved.

For the other conclusion, let $gb^{-1}\rho(gb^{-1})^{-1}$ be given, and put $g' = \rho(gb^{-1})$, so that $gb^{-1}g'^{-1} = h$ is in H. This equation implies that $g'b = h^{-1}g$. Hence $\rho(g'b) = \rho(h^{-1}g) = \rho(g) = g$, and it follows that $gb^{-1}\rho(gb^{-1})^{-1} = gb^{-1}g'^{-1}$ $= (g'bg^{-1})^{-1} = \left(g'b\rho(g'b)^{-1}\right)^{-1}$. This proves the lemma. □

Lemma 7.12. With F free it is possible to choose the set C of coset representatives in such a way that all of its members have expansions in terms of S' as $g = b_1 \cdots b_n$ in which

(a) $g = b_1 b_2 \cdots b_n$ is a reduced word as written,
(b) $b_1 b_2 \cdots b_{n-1}$ is also a member of C.

REMARKS. It is understood from the case of $n = 1$ in (b) that 1 is the representative of the identity coset. When C is chosen as in this lemma, C is said to be a **Schreier set**. In the example, $C = \{1, x\}$ is a Schreier set. So is $C = \{1, y\}$, and hence the selection of a Schreier set may involve a choice.

PROOF. If S' is finite or countably infinite, we enumerate it. In the uncountable case (which is of less practical interest), we introduce a well ordering in S' by means of Zermelo's Well-Ordering Theorem as in Section A5 of the appendix.

The ordering of S' will be used to define a lexicographic ordering of the set of all reduced words in the members of S'. If

$$x = b_1 \cdots b_m \qquad \text{and} \qquad y = b'_1 \cdots b'_n \tag{$*$}$$

are reduced words with $m \leq n$, we say that $x < y$ if any of the following hold:

(i) $m < n$,
(ii) $m = n$ and $b_1 < b'_1$,
(iii) $m = n$, and for some $k < m$, $b_1 = b'_1, \ldots, b_k = b'_k$, and $b_{k+1} < b'_{k+1}$.

With this definition the set of reduced words is well ordered, and hence any nonempty subset of reduced words has a least element.

Let us observe that if x, y, z are reduced words with $x < y$ and if yz is reduced as written, then $xz < yz$ after xz has been reduced. In fact, let us assume that x and y are as in $(*)$ and that the length of z is r. The assumption is that yz has length $n + r$, and the length of xz is at most $m + r$. If $m < n$, then certainly $xz < yz$. If $m = n$ and xz fails to be reduced, then the length of xz is less than the length of yz, and $xz < yz$. If $m = n$ and xz is reduced, then the first inequality $b_k < b'_k$ with x and y shows that $xz < yz$.

To define the set C of coset representatives, let the representative of Hg be the least member of the set Hg, each element being written as a reduced word. Since the length of the empty word is 0, the representative of the identity coset H is 1 under this definition. Thus all we have to check is that an initial segment of a member of C is again in C.

Suppose that $b_1 \cdots b_n$ is in C, so that $b_1 \cdots b_n$ is the least element of $Hb_1 \cdots b_n$. Denote the least element of $Hb_1 \cdots b_{n-1}$ by g. If $g = b_1 \cdots b_{n-1}$, we are done. Otherwise $g < b_1 \cdots b_{n-1}$, and then the fact that $b_1 \cdots b_n$ is reduced implies that $gb_n < b_1 \cdots b_n$. But gb_n is in $Hb_1 \cdots b_n$, and this inequality contradicts the minimality of $b_1 \cdots b_n$ in that coset. Thus we conclude that $g = b_1 \cdots b_{n-1}$. This proves the lemma. □

For the remainder of the proof of Theorem 7.10, we assume, as we may by Lemma 7.12, that the set C of coset representatives is a Schreier set. Typical elements of S will be denoted by a, and typical elements of $S' = S \cup S^{-1}$ will be denoted by b. Let us write u for a typical element $ga\rho(ga)^{-1}$ with g in C, and let us write v for a typical element $gb\rho(gb)^{-1}$ with g in C. The elements u generate H by Lemma 7.11, and each element v is either an element u or the inverse of an element u, according to the lemma. We shall prove that the elements u not equal to 1 are distinct and form a free basis of H.

First we prove that each of the elements $v = gb\rho(gb)^{-1}$ either is reduced as written or is equal to 1. Put $g' = \rho(gb)$, so that $v = gbg'^{-1}$. Since g and g' are in the Schreier set C, they are reduced as written, and hence so are g and g'^{-1}. Thus

the only possible cancellation in v occurs because the last factor of g is b^{-1} or the last factor of g' is b. If the last factor of g is b^{-1}, then gb is an initial segment of g and hence is in the Schreier set C; thus $\rho(gb) = gb$ and $v = gb\rho(gb)^{-1} = 1$. Similarly if the last factor of g' is b, then $g'b^{-1}$ is an initial segment of g' and hence is in the Schreier set C; thus $\rho(g'b^{-1}) = g'b^{-1}$, and Lemma 7.11 gives $v^{-1} = \left(gb\rho(gb)^{-1}\right)^{-1} = g'b^{-1}\rho(g'b^{-1})^{-1} = 1$. Thus $v = gb\rho(gb)^{-1}$ either is reduced as written or is equal to 1.

Next let us see that the elements v other than 1 are distinct. Suppose that $v = gb\rho(gb)^{-1} = g'b'\rho(g'b')^{-1}$ is different from 1. Remembering that each of these expressions is reduced as written, we see that if g is shorter than g', then gb is an initial segment of g'. Since C is a Schreier set, gb is in C and $\rho(gb) = gb$; thus $v = gb\rho(gb)^{-1}$ equals 1, contradiction. Similarly g' cannot be shorter than g. So g and g' must have the same length l. In this case the first $l + 1$ factors must match in the two equal reduced words, and we conclude that $g = g'$ and $b = b'$. This proves the uniqueness.

We know that each v is either some u or some u'^{-1}, and this uniqueness shows that it cannot be both unless $v = 1$. Therefore the nontrivial u's are distinct, and the nontrivial v's consist of the u's and their inverses, each appearing once.

Since an element v not equal to 1 therefore determines its g and b, let us refer to the factor b of $v = gb\rho(gb)^{-1}$ as the **significant factor** of v. This is the part that will not cancel out when we pass from a product of v's to its reduced form.

Specifically suppose that we have $v = gb\rho(gb)^{-1}$ and $\bar{v} = \bar{g}\bar{b}\rho(\bar{g}\bar{b})^{-1}$, that neither of these is 1, and that $\bar{v} \neq v^{-1}$. Put $g' = \rho(gb)$ and $\bar{g}' = \rho(\bar{g}\bar{b})$. The claim is that the cancellation in forming $v\bar{v} = gbg'^{-1}\bar{g}\bar{b}\bar{g}'^{-1}$ does not extend to either of the significant factors b and $\bar{b}$. If it does, then one of three things happens:

(i) the b in bg'^{-1} gets canceled because the last factor of g' is b, in which case $g'b^{-1}$ is an initial segment of g', $g'b^{-1} = \rho(g'b^{-1}) = g$, and $v = gbg'^{-1} = 1$, or

(ii) the $\bar{b}$ in $\bar{g}\bar{b}$ gets canceled because the last factor of $\bar{g}$ is $\bar{b}^{-1}$, in which case $\bar{g}\bar{b}$ is an initial segment of $\bar{g}$, $\bar{g}\bar{b} = \rho(\bar{g}\bar{b}) = \bar{g}'$, and $\bar{v} = \bar{g}\bar{b}\bar{g}'^{-1} = 1$, or

(iii) $g'^{-1}\bar{g} = 1$ and $b\bar{b} = 1$, in which case $\bar{g} = g'$, $\bar{b} = b^{-1}$, and the middle conclusion of Lemma 7.11 allows us to conclude that $\bar{v} = v^{-1}$.

All three of these possibilities have been ruled out by our assumptions, and therefore neither of the significant factors in $v\bar{v}$ cancels.

As a consequence of this noncancellation, we can see that in any product $v_1 \cdots v_m$ of v's in which no v_k is 1 and no v_{k+1} equals v_k^{-1}, none of the significant factors cancel. In fact, the previous paragraph shows that the significant factors of v_1 and v_2 survive in forming v_1v_2, the significant factors of v_2 and v_3 survive in right multiplying by v_3, and so on. Since the nontrivial u's are distinct and

the nontrivial v's consist of the u's and their inverses, each appearing once, we conclude that the set of nontrivial u's is a free subset of F. Lemma 7.11 says that the u's generate H, and therefore the set of nontrivial u's is a free basis of H.

3. Free Products

The free abelian group on an index set S, as constructed in Section IV.9, has a universal mapping property that allows arbitrary functions from S into any target abelian group to be extended to homomorphisms of the free abelian group into the target group. The construction of free groups in Section 1 was arranged to adapt the construction so that the target group in the universal mapping property could be any group, abelian or nonabelian.

In this section we make a similar adaptation of the construction of a direct sum of abelian groups so that the result is applicable in a context of arbitrary groups. Proposition 4.17 gave the universal mapping property of the external direct sum $\bigoplus_{s\in S} G_s$ of a set of abelian groups with associated embedding homomorphisms $i_{s_0} : G_{s_0} \to \bigoplus_{s\in S} G_s$. The statement is that if H is any abelian group and $\{\varphi_s \mid s \in S\}$ is a system of group homomorphisms $\varphi_s : G_s \to H$, then there exists a unique group homomorphism $\varphi : \bigoplus_{s\in S} G_s \to H$ such that $\varphi \circ i_{s_0} = \varphi_{s_0}$ for all $s_0 \in S$. Example 2 of coproducts in Section IV.11 shows that direct sum is therefore the coproduct functor in the category of all abelian groups.

This universal mapping property of $\bigoplus_{s\in S} G_s$ fails when H is a nonabelian group such as the symmetric group $\mathfrak{S}_3$. In fact, $\mathfrak{S}_3$ has an element of order 2 and an element of order 3 and hence admits nontrivial homomorphisms $\varphi_2 : C_2 \to \mathfrak{S}_3$ and $\varphi_3 : C_3 \to \mathfrak{S}_3$. But there is no homomorphism $\varphi : C_2 \oplus C_3 \to \mathfrak{S}_3$ such that $\varphi \circ i_2 = \varphi_2$ and $\varphi \circ i_3 = \varphi_3$ because the image of φ has to be abelian but the images of φ_2 and φ_3 do not commute. Consequently direct sum cannot extend to a coproduct functor in the category of all groups.

Instead, the appropriate group constructed from C_2 and C_3 for this kind of universal mapping property is the "free product" of C_2 and C_3, denoted by $C_2 * C_3$. In this section we construct the free product of any set of groups, finite or infinite. Also, we establish its universal mapping property and identify it in terms of generators and relations. The prototype of a free product is the free group $F(S)$, which equals a free product of copies of $\mathbb{Z}$ indexed by S. A free product is always an infinite group if at least two of the factors are not 1-element groups.

An important application of free products occurs in the theory of the fundamental group in topology: if X is a topological space for which the theory of covering spaces is applicable, and if A and B are open subsets of X with $X = A \cup B$ such that $A \cap B$ is nonempty, connected, and simply connected, then the fundamental

group of X is the free product of the fundamental group of A and the fundamental group of B. This result, together with a generalization that no longer requires $A \cap B$ to be simply connected, is known as the **Van Kampen Theorem**.

Let S be a nonempty set of groups G_s for s in S. The set S is allowed to be infinite, but in practice it often has just two elements. We shall describe the group defined to be the free product $G = \ast_{s\in S} G_s$. We start from the set $W(\{G_s\})$ of all words built from the groups G_s. This consists of all finite sequences $g_1 \cdots g_n$ with each g_i in some G_s depending on i. The length of a word is the number of factors in it. The empty word is denoted by 1. We multiply two words by writing them end to end, and the resulting operation of multiplication is associative. A word is said to be equivalent to a second word if the first can be obtained from the second by a finite sequence of steps of the following kinds and their inverses:

(i) drop a factor for which g_i is the identity element of the group in which it lies,

(ii) collapse two factors $g_i g_{i+1}$ to a single one g_i^* if g_i and g_{i+1} lie in the same G_s and their product in that group is g_i^*.

The result is an equivalence relation, and the set of equivalence classes is the underlying set of $\ast_{s\in S} G_s$.

Theorem 7.13. If S is a nonempty set of groups G_s and $W(\{G_s\})$ is the set of all words from the groups G_s, then the product operation defined on $W(\{G_s\})$ descends in a well-defined fashion to the set $\ast_{s\in S} G_s$ of equivalence classes of members of $W(\{G_s\})$, and $\ast_{s\in S} G_s$ thereby becomes a group. For each s_0 in S, define $i_{s_0} : G_{s_0} \to \ast_{s\in S} G_s$ to be the group homomorphism obtained as the composition of the inclusion of G_{s_0} into words of length 1 followed by passage to equivalence classes. Then the pair $\left(\ast_{s\in S} G_s, \{i_s\}\right)$ has the following universal mapping property: whenever H is a group and $\{\varphi_s \mid s \in S\}$ is a system of group homomorphisms $\varphi_s : G_s \to H$, then there exists a unique group homomorphism $\varphi : \ast_{s\in S} G_s \to H$ such that $\varphi \circ i_{s_0} = \varphi_{s_0}$ for all $s_0 \in S$.

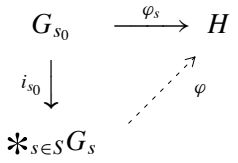


FIGURE 7.2. Universal mapping property of a free product.

REMARKS. The group $\ast_{s\in S} G_s$ is called the **free product** of the groups G_s. Figure 7.2 illustrates its universal mapping property. This universal mapping property actually characterizes $\ast_{s\in S} G_s$, as will be seen in Proposition 7.14. One

often writes $G_1 * \cdots * G_n$ when the set S is finite; the order of listing the groups is immaterial. The proof of Theorem 7.13 is rather similar to the proof of Theorem 7.1, and we shall skip some details.

PROOF. Let us write $\sim$ for the equivalence relation on words, and let us denote equivalence classes by brackets. We want to define multiplication in $\ast_{s\in S} G_s$ by $[w_1][w_2] = [w_1 w_2]$. To see that this formula makes sense in $\ast_{s\in S} G_s$, let x, x', and y be words in $W(\{G_s\})$, and suppose that x and x' differ by only one operation of type (i) or type (ii) as above. Then $x \sim x'$, and it is evident that $x'y \sim xy$ and $yx' \sim yx$. Iteration of this kind of relationship shows that $w_1' \sim w_1$ and $w_2' \sim w_2$ implies $w_1' w_2' \sim w_1 w_2$, and hence multiplication is well defined.

The associativity of multiplication in $W(\{G_s\})$ implies that multiplication in $\ast_{s\in S} G_s$ is associative, and [1] is a two-sided identity. We readily check that if $g = g_1 \cdots g_n$ is a word, then the word $g^{-1} = g_n^{-1} \cdots g_1^{-1}$ has the property that $[g^{-1}]$ is a two-sided inverse to $[g]$. Therefore $\ast_{s\in S} G_s$ is a group.

The uniqueness of the homomorphism φ in the universal mapping property is no problem since all words are products of words of length 1 and since the subgroups $i_{s_0}(G_{s_0})$ together generate $\ast_{s\in S} G_s$.

For existence of φ, we begin by defining a function $\Phi : W(\{G_s\}) \to H$ such that

$$\begin{aligned} \Phi(g_s) &= \varphi_s(g_s) && \text{for } g_s \text{ in } G_s \text{ when viewed as a word of length 1,} \\ \Phi(w_1 w_2) &= \Phi(w_1)\Phi(w_2) && \text{for } w_1 \text{ and } w_2 \text{ in } W(\{G_s\}). \end{aligned}$$

We take the formulas $\Phi(g_s) = \varphi(g_s)$ for g_s in G_s as a definition of Φ on words of length 1. Any member of $W(\{G_s\})$ can be written uniquely as $g_1 \cdots g_n$ with each g_i in G_{s_i}, and we set $\Phi(g_1 \cdots g_n) = \Phi(g_1) \cdots \Phi(g_n)$. (If $n = 0$, the understanding is that $\Phi(1) = 1$.) Then Φ has the required properties.

Let us show that $w' \sim w$ implies $\Phi(w') = \Phi(w)$. The questions are whether

(i) if $g_1, \ldots, g_n$ are in various G_s's with g_i equal to the identity 1_{s_i} of G_{s_i}, then

$$\Phi(g_1 \cdots g_{i-1} 1_{s_i} g_{i+1} \cdots g_n) \stackrel{?}{=} \Phi(g_1 \cdots g_{i-1} g_{i+1} \cdots g_n),$$

(ii) if $g_1, \ldots, g_n$ are in various G_s's with $G_{s_i} = G_{s_{i+1}}$ and if $g_i g_{i+1} = g_i^*$ in G_{s_i}, then

$$\Phi(g_1 \cdots g_{i-1} g_i g_{i+1} g_{i+2} \cdots g_n) \stackrel{?}{=} \Phi(g_1 \cdots g_{i-1} g_i^* g_{i+2} \cdots g_n).$$

In the case of (i), the question comes down to whether a certain $h\Phi(1_{s_i})h'$ in H equals hh', and this is true because $\Phi(1_{s_i}) = \varphi_{s_i}(1_{s_i})$ is the identity of H. In the case of (ii), the question comes down to whether $h\Phi(g_i)\Phi(g_{i+1})h'$ equals $h\Phi(g_i^*)h'$ if $G_{s_i} = G_{s_{i+1}}$ and $g_i g_{i+1} = g_i^*$ in G_{s_i}, and this is true because $\Phi(g_i)\Phi(g_{i+1}) = \varphi_{s_i}(g_i)\varphi_{s_i}(g_{i+1}) = \varphi_{s_i}(g_i g_{i+1}) = \varphi_{s_i}(g_i^*) = \Phi(g_i^*)$. We conclude that $w' \sim w$ implies $\Phi(w') = \Phi(w)$.

We may therefore define $\varphi([w]) = \Phi(w)$ for $[w]$ in $F(\{G_s\})$, and φ is a homomorphism of $F(\{G_s\})$ into H as a consequence of the property $\Phi(w_1 w_2) = \Phi(w_1)\Phi(w_2)$ of Φ on $W(\{G_s\})$. For g_s in G_s, we have $\varphi([g_s]) = \Phi(g_s) = \varphi_s(g_s)$, i.e., $\varphi(i(g_s)) = \varphi_s(g_s)$. This completes the proof of existence. □

Proposition 7.14. Let S be a nonempty set of groups G_s. Suppose that G' is a group and that $i'_s : G_s \to G'$ for $s \in S$ is a system of group homomorphisms with the following universal mapping property: whenever H is a group and $\{\varphi_s \mid s \in S\}$ is a system of group homomorphisms $\varphi_s : G_s \to H$, then there exists a unique group homomorphism $\varphi : G' \to H$ such that $\varphi \circ i'_s = \varphi_s$ for all $s \in S$. Then there exists a unique group homomorphism $\Phi : \ast_{s\in S} G_s \to G'$ such that $i'_s = \Phi \circ i_s$ for all $s \in S$. Moreover, Φ is a group isomorphism, and the homomorphisms $i'_s : G_s \to G'$ are one-one.

REMARKS. As was true with Proposition 7.2, readers who have been through Chapter VI will recognize that Proposition 7.14 is a special case of Problem 19 at the end of that chapter.

PROOF. Put $G = \ast_{s\in S} G_s$. In the universal mapping property of Theorem 7.13, let $H = G'$ and $\varphi_s = i'_s$, and let $\Phi : G \to G'$ be the homomorphism φ produced by that theorem. Then Φ satisfies $\Phi \circ i_s = i'_s$ for all s. Reversing the roles of G and G', we obtain a homomorphism $\Phi' : G' \to G$ with $\Phi' \circ i'_s = i_s$ for all s. Therefore $(\Phi' \circ \Phi) \circ i_s = \Phi' \circ i'_s = i_s$.

Comparing $\Phi' \circ \Phi$ with the identity 1_G and applying the uniqueness in the universal mapping property for G, we see that $\Phi' \circ \Phi = 1_G$. Similarly the uniqueness in the universal mapping property of G' gives $\Phi \circ \Phi' = 1_{G'}$. Thus Φ is a group isomorphism. It is uniquely determined by the given properties since the various subgroups $i_s(G_s)$ generate G. Since $i'_s = \Phi \circ i_s$ and since Φ and i_s are one-one, i'_s is one-one. □

As was the case for free groups, we want a decision procedure for telling whether two given words in $W(\{G_s\})$ are equivalent. This is the so-called **word problem** for the free product. Solving it allows us to use free products concretely, just as Proposition 7.3 allowed us to use free groups concretely. A word in $W(\{G_s\}))$ is said to be **reduced** if it

(i) contains no factor for which g_i is the identity element of the group G_s in which it lies,

(ii) contains no two consecutive factors g_i and g_{i+1} taken from the same group G_s.

Proposition 7.15. (solution of the word problem for free products). If S is a nonempty set of groups G_s and $W(\{G_s\})$ is the set of all words from the groups G_s, then each word in $W(\{G_s\})$ is equivalent to one and only one reduced word.

EXAMPLE. Consider the free product $C_2 * C_2$ of two cyclic groups, one with x as generator and the other with y as generator. Words consist of a finite sequence of factors of x, y, the identity of the first factor, and the identity of the second factor. A word is reduced if no factor is an identity and if no two x's are adjacent and no two y's are adjacent. Thus the reduced words consist of finite sequences whose terms are alternately x and y. Those of length ≤ 3 are $1, x, y, xy, yx, xyx, yxy$, and in general there are two of each length > 0. The proposition tells us that all these reduced words give distinct members of $C_2 * C_2$. In particular, the group is infinite.

REMARK. More generally, to test whether two words are equivalent, the proposition says to eliminate factors of the identity and multiply consecutive factors in each word when they come from the same group, and repeat these steps until it is no longer possible to do either of these operations on either word. Then each of the given words has been replaced by a reduced word, and the two given words are equivalent if and only if the two reduced words are identical. Problems 37–46 at the end of the chapter concern $C_2 * C_3$, and some of these problems make use of the result of this proposition—that distinct reduced words are inequivalent.

PROOF OF PROPOSITION 7.15. Both operations—eliminating factors of the identity and multiplying consecutive factors in each word when they come from the same group—reduce the length of a word. Since the length has to remain ≥ 0, the process of successively carrying out these two operations as much as possible has to stop after finitely many steps, and the result is a reduced word. This proves that each equivalence class of words contains a reduced word.

For uniqueness of the reduced word in an equivalence class, we proceed somewhat as with Proposition 7.3, associating to each word a finite sequence of reduced words such that the last member of the sequence is unchanged when we apply an operation to the word that preserves equivalence. However, there are considerably more details to check this time.

If $w = g_1 \cdots g_n$ is a given word with each g_i in G_{s_i}, then we associate to w the sequence of reduced words $x_0, x_1, \ldots, x_n$ defined inductively by

$$x_0 = 1,$$

$$x_1 = \begin{cases} g_1 & \text{if } g_1 \text{ is not the identity of } G_{s_1}, \\ 1 & \text{if } g_1 \text{ is the identity of } G_{s_1}, \end{cases}$$

and the following formula for $i \geq 2$ if x_{i-1} is of the reduced form $h_1 \cdots h_k$ with h_j in G_{t_j}:

$$x_i = \begin{cases} h_1 \cdots h_k g_i & \text{if } G_{s_i} \neq G_{t_k} \text{ and } g_i \text{ is not the identity } 1_{G_{s_i}} \text{ of } G_{s_i}, \\ h_1 \cdots h_k & \text{if } g_i \text{ is the identity } 1_{G_{s_i}} \text{ of } G_{s_i}, \\ h_1 \cdots h_{k-1} & \text{if } G_{t_k} = G_{s_i} \text{ with } h_k g_i = 1_{G_{s_i}}, \\ h_1 \cdots h_{k-1} g_i^* & \text{if } G_{t_k} = G_{s_i} \text{ with } h_k g_i = g_i^* \neq 1_{G_{s_i}}. \end{cases}$$

Put $r(w) = x_n$. We check inductively for $i \geq 0$ that each x_i is reduced. In fact, x_i for $i \geq 2$ begins in every case with $h_1 \cdots h_{k-1}$, which is assumed reduced. The only possible reduction for x_i thus comes from factors that are adjoined or from interference with h_{k-1}, and all possibilities are addressed in the above choices. Thus $r(w) = x_n$ is necessarily reduced for each word w.

If $g_1 \cdots g_n$ is reduced as given, then x_i is determined by the first possible choice $h_1 \cdots h_k g_i$ every time, and hence $x_i = g_1 \cdots g_i$ for all i. Therefore we obtain $r(w) = w$ if w is reduced.

Now consider the equivalent words

$$w = g_1 \cdots g_j g_{j+1} \cdots g_n \qquad \text{and} \qquad w' = g_1 \cdots g_j 1_{G_s} g_{j+1} \cdots g_n.$$

Form $x_0, \ldots, x_n$ for w and $x'_0, \ldots, x'_{n+1}$ for w'. Then we have $x'_j = x_j$; let $h_1 \cdots h_k$ be a reduced form of x'_j. The formula for x'_{j+1} is governed by the second choice in the display, and $x'_{j+1} = h_1 \cdots h_k = x_j$. Then $x'_{j+i+1} = x_{j+i}$ for $1 \leq i \leq n - j$ as well. Hence $x'_{n+1} = x_n$, and $r(w') = r(w)$.

Next suppose that $g_j^* = g_j g_{j+1}$ in G_{s_j}, and consider the equivalent words

$$w = g_1 \cdots g_{j-1} g_j^* g_{j+2} \cdots g_n \qquad \text{and} \qquad w' = g_1 \cdots g_{j-1} g_j g_{j+1} g_{j+2} \cdots g_n.$$

As above, form $x_0, \ldots, x_n$ for w and $x'_0, \ldots, x'_{n+1}$ for w'. Then we have $x_{j-1} = x'_{j-1}$, and we let $h_1 \cdots h_k$ be a reduced form of x_{j-1}. There are cases, subcases, and subsubcases.

First assume $G_{t_k} \neq G_{s_j}$. Then x_j equals $h_1 \cdots h_k g_j^*$ or $h_1 \cdots h_k$ in the two subcases $g_j^* \neq 1_{G_{s_j}}$ and $g_j^* = 1_{G_{s_j}}$. In the first subcase, we have $g_j^* \neq 1_{G_{s_j}}$ and $x_j = h_1 \cdots h_k g_j^*$. Then x'_j equals $h_1 \cdots h_k g_j$ or $h_1 \cdots h_k$ in the two subsubcases $g_j \neq 1_{G_{s_j}}$ and $g_j = 1_{G_{s_j}}$. In the first subsubcase, $x'_{j+1} = h_1 \cdots h_k g_j^* = x_j$ whether or not $g_{j+1} = 1_{G_{s_j}}$. In the second subsubcase, $g_j^* = g_j g_{j+1}$ cannot be $1_{G_{s_j}}$, and therefore $x'_{j+1} = h_1 \cdots h_k g_j^* = x_j$.

In the second subcase of the case $G_{t_k} \neq G_{s_j}$, we have $g_j^* = 1_{G_{s_j}}$ and $x_j = x_{j-1} = h_1 \cdots h_k$. Then x'_j equals $h_1 \cdots h_k g_j$ or $h_1 \cdots h_k$ in the two subsubcases $g_j \neq 1_{G_{s_j}}$ and $g_j = 1_{G_{s_j}}$. In both subsubcases, $x'_{j+1} = h_1 \cdots h_k$, so that $x'_{j+1} = x_j$.

Now assume $G_{t_k} = G_{s_j}$. Then x_j equals $h_1 \cdots h_{k-1} h_k^*$ or $h_1 \cdots h_{k-1}$ in the two subcases $h_k g_j^* = h_k^* \neq 1_{G_{s_j}}$ and $h_k g_j^* = 1_{G_{s_j}}$. In the first subcase, we have $h_k g_j^* = h_k^* \neq 1_{G_{s_j}}$ and $x_j = h_1 \cdots h_{k-1} h_k^*$. Then x_j' equals $h_1 \cdots h_{k-1} h_k'$ or $h_1 \cdots h_{k-1}$ in the two subsubcases $h_k g_j = h_k' \neq 1_{G_{s_j}}$ and $h_k g_j = 1_{G_{s_j}}$. In the first subsubcase, $h_k' g_{j+1} = h_k g_j g_{j+1} = h_k g_j^* = h_k^*$ implies $x_{j+1}' = h_1 \cdots h_{k-1} h_k^* = x_j$. In the second subsubcase, we know that h_k^* cannot be $1_{G_{s_i}}$ and hence that $g_{j+1} = h_k g_j g_{j+1} = h_k g_j^* = h_k^*$ cannot be $1_{G_{s_j}}$; thus $x_{j+1}' = h_1 \cdots h_{k-1} h_k^* = x_j$.

In the second subcase of the case $G_{t_k} = G_{s_j}$, we have $h_k g_j^* = 1_{G_{s_j}}$ and $x_j = h_1 \cdots h_{k-1}$. Then x_j' equals $h_1 \cdots h_{k-1} h_k^{*\prime}$ or $h_1 \cdots h_{k-1}$ in the two subsubcases $h_k g_j = h_k^{*\prime} \neq 1_{G_{s_j}}$ and $h_k g_j = 1_{G_{s_j}}$. In the first subsubcase, g_{j+1} cannot be $1_{G_{s_j}}$ but $h_k^{*\prime} g_{j+1} = h_k g_j g_{j+1} = h_k g_j^* = 1_{G_{s_j}}$; hence $x_{j+1}' = h_1 \cdots h_{k-1} = x_j$. In the second subsubcase, $x_j' = h_1 \cdots h_{k-1}$ and $g_{j+1} = 1_{G_{s_j}}$, so that $x_{j+1}' = h_1 \cdots h_{k-1} = x_j$.

We conclude that $x_{j+1}' = x_j$ in all cases. Hence $x_{j+i+1}' = x_{j+i}$ for $0 \leq i \leq n - j$, $x_{n+1}' = x_n$, and $r(w') = r(w)$. Consequently the only reduced word that is equivalent to w is $r(w)$. □

Proposition 7.16. Let $\mathcal{S}$ be a nonempty set of groups G_s, and suppose that $\langle S_s;\ R_s \rangle$ is a presentation of G_s, the sets S_s being understood to be disjoint for $s \in \mathcal{S}$. Then $\left\langle \bigcup_{s\in\mathcal{S}} S_s;\ \bigcup_{s\in\mathcal{S}} R_s \right\rangle$ is a presentation of the free product $\ast_{s\in\mathcal{S}} G_s$.

REMARK. One effect of this proposition is to make Proposition 7.8 available as a tool for use with free products. Using Proposition 7.8 may be easier than appealing to the universal mapping property in Theorem 7.13.

PROOF. Put $S = \bigcup_{s\in\mathcal{S}} S_s$ and $R = \bigcup_{s\in\mathcal{S}} R_s$, and define G to be a group given by generators and relations as $G = \langle S;\ R \rangle$. Consider the function from S_s into the quotient group $G = F(S)/N(R)$ given by carrying x in S_s into the word x in S and then passing to $F(S)$ and its quotient G. Because of the universal mapping property of free groups, this function extends to a group homomorphism $\widetilde{i}_s : F(S_s) \to G$. If r is a reduced word relative to S_s representing a member of R_s, then r is carried by $\widetilde{i}_s$ into a member of the larger set R and then into the identity of G. Since $\ker \widetilde{i}_s$ is normal in $F(S_s)$, $\ker \widetilde{i}_s$ contains the smallest normal subgroup $N(R_s)$ in $F(S_s)$ that contains R_s. Proposition 4.11 shows that $\widetilde{i}_s$ descends to a group homomorphism $i_s : G_s \to G$.

We shall prove that G and the system $\{i_s\}$ have the universal mapping property of Proposition 7.14 that characterizes a free product. Then it will follow from that proposition that $G \cong \ast_{s\in S} G_s$, and the proof will be complete.

Thus let H be a group, and let $\{\varphi_s \mid s \in \mathcal{S}\}$ be a system of group homomorphisms $\varphi_s : G_s \to H$. We are to produce a homomorphism $\Phi : G \to H$ such that $\Phi \circ i_s = \varphi_s$ for all s, and we are to prove that such a homomorphism

is unique. Let $q_s : F(S_s) \to G_s$ be the quotient homomorphism, and define $\widetilde{\varphi}_s : F(S_s) \to H$ by $\widetilde{\varphi}_s = \varphi_s \circ q_s$. Now define $\widetilde{\Phi} : S \to H$ as follows: if x is in S, then x is in a set S_s for a unique s and thereby defines a member of $F(S_s)$ for that unique s; $\widetilde{\Phi}(x)$ is taken to be $\widetilde{\varphi}_s(x)$. The universal mapping property of the free group $F(S)$ allows us to extend $\widetilde{\Phi}$ to a group homomorphism, which we continue to call $\widetilde{\Phi}$, of $F(S)$ into H. Let r be a nontrivial relation in $R \subseteq F(S)$. Then r, by hypothesis of disjointness for the sets S_s, lies in a unique R_s. Hence $\widetilde{\Phi}(r) = \widetilde{\varphi}_s(r) = \varphi_s(q_s(r)) = \varphi_s(1_s) = 1_H$. Consequently the kernel of $\widetilde{\Phi}$ contains the smallest normal subgroup $N(R)$ of $F(S)$ containing R, and $\widetilde{\Phi}$ descends to a homomorphism $\Phi : G \to H$. This Φ satisfies

$$\Phi \circ i_s \circ q_s = \Phi \circ \widetilde{i}_s = \widetilde{\Phi}\big|_{F(S_s)} = \widetilde{\varphi}_s = \varphi_s \circ q_s.$$

Since the quotient homomorphism q_s is onto G_s, we obtain $\Phi \circ i_s = \varphi_s$, and existence of the homomorphism Φ is established.

For uniqueness, we observe that the identities $\Phi \circ i_s = \varphi_s$ imply that Φ is uniquely determined on the subgroup of G generated by the images of all i_s. Since q_s is onto G_s, this subgroup is the same as the subgroup generated by the images of all $\widetilde{i}_s$. This subgroup contains the image in G of every generator of $F(S)$ and hence is all of G. Thus Φ is uniquely determined. □

4. Group Representations

Group representations were defined in Section IV.6 as group actions on vector spaces by invertible linear functions. The underlying field of the vector space will be taken to be $\mathbb{C}$ in this section and the next, and the theory will then be especially tidy. The subject of group representations is one that uses a mix of linear algebra and group theory to reveal hidden structure within group actions. It has broad applications to algebra and analysis, but we shall be most interested in an application to finite groups known as Burnside's Theorem that will be proved in the next section.

Let us begin with the abelian case, taking G for the moment to be a finite abelian group. A **multiplicative character** of G is a homomorphism $\chi : G \to S^1 \subseteq \mathbb{C}^\times$ of G into the multiplicative group of complex numbers of absolute value 1. The multiplicative characters form an abelian group $\widehat{G}$ under pointwise multiplication of their complex values: $(\chi\chi')(g) = \chi(g)\chi'(g)$. The identity of $\widehat{G}$ is the multiplicative character that is identically 1 on G, and the inverse of χ is the complex conjugate of χ.

The notion of multiplicative character adapts to the case of a finite group the familiar exponential functions $x \mapsto e^{inx}$ on the line, which can be regarded as multiplicative characters of the additive group $\mathbb{R}/2\pi\mathbb{Z}$ of real numbers modulo 2π. These functions have long been used to resolve a periodic function of

time into its component frequencies: The device is the Fourier series of the function f. If f is periodic of period 2π, then the **Fourier coefficients** of f are $c_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}\,dx$, and the **Fourier series** of f is the infinite series $\sum_{n=-\infty}^{\infty} c_n e^{inx}$. A portion of the subject of Fourier series looks for senses in which $f(x)$ is actually equal to the sum of its Fourier series. This is the problem of **Fourier inversion**.

A similar problem can be formulated when $\mathbb{R}/2\pi\mathbb{Z}$ is replaced by the finite abelian group G. The exponential functions are replaced by the multiplicative characters. One can form an analog of Fourier coefficients for the vector space $C(G, \mathbb{C})$ of complex-valued functions[2] defined on G, and then one can form the analog of the Fourier series of the function. The problem of Fourier inversion becomes one of linear algebra, once we take into account the known structure of all finite abelian groups (Theorem 4.56). The result is as follows.

Theorem 7.17 (Fourier inversion formula for finite abelian groups). Let G be a finite abelian group, and introduce an inner product on the complex vector space $C(G, \mathbb{C})$ of all functions from G to $\mathbb{C}$ by the formula

$$\langle F, F'\rangle = \sum_{g\in G} F(g)\overline{F'(g)},$$

the corresponding norm being $\|F\| = \langle F, F\rangle^{1/2}$. Then the members of $\widehat{G}$ form an orthogonal basis of $C(G, \mathbb{C})$, each χ in $\widehat{G}$ satisfying $\|\chi\|^2 = |G|$. Consequently $|\widehat{G}| = |G|$, and any function $F : G \to \mathbb{C}$ is given by the "sum of its Fourier series":

$$F(g) = \frac{1}{|G|}\sum_{\chi\in\widehat{G}} \Big(\sum_{h\in G} F(h)\overline{\chi(h)}\Big)\chi(g).$$

REMARKS. This theorem is one of the ingredients in the proof in Chapter I of *Advanced Algebra* of Dirichlet's theorem that if a and b are positive relatively prime integers, then there are infinitely many primes of the form $an + b$. In applications to engineering, the ordinary Fourier transform on the line is often approximated, for computational purposes, by a Fourier series on a large cyclic group, and then Theorem 7.17 is applicable. Such a Fourier series can be computed with unexpected efficiency using a special grouping of terms; this device

[2]The notation $C(G, \mathbb{C})$ is to be suggestive of what happens for $G = S^1$ and for $G = \mathbb{R}^1$, where one works in part with the space of *continuous* complex-valued functions vanishing off a bounded set. In any event, pointwise multiplication makes $C(G, \mathbb{C})$ into a commutative ring. Later in the section we introduce a second multiplication, called "convolution," that makes $C(G, \mathbb{C})$ into a ring in a different way. In Chapter VIII we shall introduce the "complex group algebra" $\mathbb{C}G$ of G. The vector space $C(G, \mathbb{C})$ is the dual vector space of $\mathbb{C}G$. However, $C(G, \mathbb{C})$ and $\mathbb{C}G$ are canonically isomorphic because they have distinguished bases, and the isomorphism respects the multiplication structures—convolution in $C(G, \mathbb{C})$ and the group-algebra multiplication in $\mathbb{C}G$.

is called the **fast Fourier transform** and is described in Problems 29–31 at the end of the chapter.

PROOF. For orthogonality let χ and χ' be distinct members of $\widehat{G}$, and put $\chi'' = \chi\overline{\chi'} = \chi\chi'^{-1}$. Choose g_0 in G with $\chi''(g_0) \neq 1$. Then

$$\chi''(g_0)\big(\textstyle\sum_{g\in G}\chi''(g)\big) = \sum_{g\in G}\chi''(g_0 g) = \sum_{g\in G}\chi''(g),$$

so that $$[1-\chi''(g_0)]\textstyle\sum_{g\in G}\chi''(g) = 0$$

and therefore $$\textstyle\sum_{g\in G}\chi''(g) = 0.$$

Consequently $\langle\chi,\chi'\rangle = \sum_{g\in G}\chi(g)\overline{\chi'(g)} = \sum_{g\in G}\chi''(g) = 0.$

The orthogonality implies that the members of $\widehat{G}$ are linearly independent, and we obtain $|\widehat{G}| \le \dim C(G,\mathbb{C}) = |G|$. Certainly $\|\chi\|^2 = \sum_{g\in G}|\chi(g)|^2 = \sum_{g\in G} 1 = |G|$.

To see that the members of $\widehat{G}$ are a basis of $C(G,\mathbb{C})$, we write G as a direct sum of cyclic groups, by Theorem 4.56. A summand $\mathbb{Z}/m\mathbb{Z}$ has at least m distinct multiplicative characters, given by $j \bmod m \mapsto e^{2\pi ijr/m}$ for $0 \le r \le m-1$, and these characters extend to G as 1 on the other direct summands of G. Taking products of such multiplicative characters from the different summands of G, we see that $|\widehat{G}| \ge |G|$. Therefore $|\widehat{G}| = |G|$, and $\widehat{G}$ is an orthogonal basis by Corollary 2.4. The formula for $F(g)$ in the statement of the theorem follows by applying Theorem 3.11c. □

Now suppose that the finite group G is not necessarily abelian. Since S^1 is abelian, Proposition 7.4 shows that χ takes the value 1 on every member of the commutator subgroup G' of G. Consequently there is no way that the multiplicative characters can form a basis for the vector space $C(G,\mathbb{C})$ of complex-valued functions on G. The above analysis thus breaks down, and some adjustment is needed in order to extend the theory.

The remedy is to use representations, as defined in Section IV.6, on complex vector spaces of dimension > 1. We shall assume in the text that the vector space is finite-dimensional. The sense in which representations extend the theory of multiplicative characters is that any multiplicative character χ gives a representation R on the 1-dimensional vector space $\mathbb{C}$ by $R(g)(z) = \chi(g)z$ for g in G and z in $\mathbb{C}$. Conversely any 1-dimensional representation gives a multiplicative character: if R is the representation on the 1-dimensional vector space V and if $v_0 \neq 0$ is in V, then $\chi(g)$ is the scalar such that $R(g)v_0 = \chi(g)v_0$. It is enough to observe that the only elements of finite order in the multiplicative group $\mathbb{C}^\times$ are certain members of the circle S^1, and then it follows that χ takes values in S^1.

In the higher-dimensional case, the analog of the multiplicative character χ in passing to a 1-dimensional representation R is a "matrix representation." A **matrix representation** of G is a function $g \mapsto [\rho(g)_{ij}]$ from G into invertible square matrices of some given size such that $\rho(g_1g_2)_{ij} = \sum_{k=1}^{n} \rho(g_1)_{ik}\rho(g_2)_{kj}$. If a representation R acts on the finite-dimensional complex vector space V, then the choice of an ordered basis Γ for V leads to a matrix representation by the formula

$$[\rho(g)_{ij}] = \begin{pmatrix} R(g) \\ \Gamma\Gamma \end{pmatrix}.$$

Conversely if a matrix representation $g \mapsto [\rho(g)_{ij}]$ and an ordered basis Γ of V are given, then the same formula may be used to obtain a representation R of G on V.

In contrast to the 1-dimensional case, the matrices that occur with a matrix representation of dimension > 1 need not be unitary. The correspondence between unitary linear maps and unitary matrices was discussed in Chapter III. When the finite-dimensional vector space V has an inner product, a linear map was defined to be unitary if it satisfies the equivalent conditions of Proposition 3.18. A complex square matrix A was defined to be unitary if $A^*A = I$. The matrix of a unitary linear map relative to an ordered orthonormal basis is unitary, and conversely when a unitary matrix and an ordered orthonormal basis are given, the associated linear map is unitary. We can thus speak of **unitary representations** and **unitary matrix representations**.

Some examples of representations appear in Section IV.6. One further pair of examples will be of interest to us. With the finite group G fixed but not necessarily abelian, we continue to let $C(G, \mathbb{C})$ be the complex vector space of all functions $f : G \to \mathbb{C}$. We define two representations of G on $C(G, \mathbb{C})$: the **left regular representation** ℓ given by $(\ell(g)f)(x) = f(g^{-1}x)$ and the **right regular representation** r given by $(r(g)f)(x) = f(xg)$. The reason for the presence of an inverse in one case and not the other was discussed in Section IV.6. Relative to the inner product

$$(f_1, f_2) = \sum_{x \in G} f_1(x)\overline{f_2(x)},$$

both ℓ and r are unitary. The argument for ℓ is that

$$(\ell(g)f_1, \ell(g)f_2) = \sum_{x \in G} (\ell(g)f_1)(x)\overline{(\ell(g)f_2)(x)} = \sum_{x \in G} f_1(g^{-1}x)\overline{f_2(g^{-1}x)}$$

$$\overset{\text{under } y = g^{-1}x}{=} \sum_{y \in G} f_1(y)\overline{f_2(y)} = (f_1, f_2),$$

and the argument for r is completely analogous.

It will be convenient to abbreviate "representation R on V" as "representation (R, V)." Let (R, V) be a representation of the finite group G on a finite-dimensional complex vector space. An **invariant subspace** U of V is a vector subspace such that $R(g)U \subseteq U$ for all g in G. The representation is **irreducible** if $V \neq 0$ and if V has no invariant subspaces other than 0 and V.

Two representations (R_1, V_1) and (R_2, V_2) on finite-dimensional complex vector spaces are **equivalent** if there exists a linear invertible function $A : V_1 \to V_2$ such that $AR_1(g) = R_2(g)A$ for all g in G. In the terminology of Section IV.11, "equivalent" is the notion of "is isomorphic to" in the category of all finite-dimensional representations of G.

In more detail a morphism from (R_1, V_1) to (R_2, V_2) in this category is an **intertwining operator**, namely a linear map $A : V_1 \to V_2$ such that $AR_1(g) = R_2(g)A$ for all g in G. The condition for this equality to hold is that the diagram in Figure 7.3 commute.

$$\begin{array}{ccc} V_1 & \xrightarrow{A} & V_2 \\ {\scriptstyle R_1(g)}\downarrow & & \downarrow{\scriptstyle R_2(g)} \\ V_1 & \xrightarrow{A} & V_2 \end{array}$$

FIGURE 7.3. An intertwining operator for two representations, i.e., a morphism in the category of finite-dimensional representations of G.

An example of a pair of representations that are equivalent is the left and right regular representations of G on $C(G, \mathbb{C})$: in fact, if we define $(Af)(x) = f(x^{-1})$, then

$$(\ell(g)Af)(x) = (Af)(g^{-1}x) = f(x^{-1}g) = (r(g)f)(x^{-1}) = (Ar(g)f)(x).$$

Proposition 7.18 (Schur's Lemma). If (R_1, V_1) and (R_2, V_2) are irreducible representations of the finite group G on finite-dimensional complex vector spaces and if $A : V_1 \to V_2$ is an intertwining operator, then A is invertible (and hence exhibits R_1 and R_2 as equivalent) or else $A = 0$. If $(R_1, V_1) = (R_2, V_2)$ and $A : V_1 \to V_2$ is an intertwining operator, then A is scalar.

REMARK. The conclusion that A is scalar makes essential use of the fact that the underlying field is $\mathbb{C}$.

PROOF. The equality $R_2(g)Av_1 = AR_1(g)v_1$ shows that $\ker A$ and image A are invariant subspaces. By the assumed irreducibility, $\ker A$ equals 0 or V_1, and image A equals 0 or V_2. The first statement follows. When $(R_1, V_1) = (R_2, V_2)$, the identity $I : V_1 \to V_2$ is an intertwining operator. If λ is an eigenvalue of A, then $A - \lambda I$ is another intertwining operator. Since $A - \lambda I$ is not invertible when λ is an eigenvalue of A, $A - \lambda I$ must be 0. □

Corollary 7.19. Every irreducible finite-dimensional representation of a finite abelian group G is 1-dimensional.

PROOF. If (R, V) is given, then the linear map $A = R(g)$ satisfies $AR(x) = R(gx) = R(xg) = R(x)A$ for all x in G. By Schur's Lemma (Proposition 7.18), $A = R(g)$ is scalar. Since g is arbitrary, every vector subspace of V is invariant. Irreducibility therefore implies that V is 1-dimensional. □

Let R be a representation of the finite group G on the finite-dimensional complex vector space V, let $(\cdot\,,\,\cdot)_0$ be any inner product on V, and define

$$(v_1, v_2) = \sum_{x \in G} (R(x)v_1, R(x)v_2)_0.$$

Then we have

$$\begin{aligned}(R(g)v_1, R(g)v_2) &= \sum_{x \in G} (R(x)R(g)v_1, R(x)R(g)v_2)_0 \\ &= \sum_{x \in G} (R(xg)v_1, R(xg)v_2)_0 \\ &= \sum_{y \in G} (R(y)v_1, R(y)v_2)_0 \qquad \text{by the change } y = xg \\ &= (v_1, v_2).\end{aligned}$$

With respect to the inner product $(\cdot\,,\,\cdot)$, the representation (R, V) is therefore unitary. In other words, we are always free to introduce an inner product to make a given finite-dimensional representation unitary. The significance of this construction is noted in the following proposition.

Proposition 7.20. If (R, V) is a finite-dimensional representation of the finite group G and if an inner product is introduced in V that makes the representation unitary, then the orthogonal complement of an invariant subspace is invariant.

PROOF. Let U be an invariant subspace. If u is in U and $u^\perp$ is in $U^\perp$, then $(R(g)u^\perp, u) = (R(g)^{-1}R(g)u^\perp, R(g)^{-1}u) = (u^\perp, R(g)^{-1}u) = 0$. Thus $u^\perp$ in $U^\perp$ implies $R(g)u^\perp$ is in $U^\perp$. □

Corollary 7.21. Any finite-dimensional representation of the finite group G is a direct sum of irreducible representations.

REMARK. That is, we can find a system of invariant subspaces such that the action of G is irreducible on each of these subspaces and such that the whole vector space is the direct sum of these subspaces.

PROOF. This is immediate by induction on the dimension. For dimension 0, the representation is the empty direct sum of irreducible representations. If the decomposition is known for dimension $< n$ and if U is an invariant subspace under R of smallest possible dimension ≥ 1, then U is irreducible under R, and Proposition 7.20 says that the subspace $U^\perp$, which satisfies $V = U \oplus U^\perp$, is invariant. It is therefore enough to decompose $U^\perp$, and induction achieves such a decomposition. □

Proposition 7.22 (Schur orthogonality). For finite-dimensional representations of a finite group G in which inner products have been introduced to make the representations unitary,

(a) if (R_1, V_1) and (R_2, V_2) are inequivalent and irreducible, then

$$\sum_{x\in G}(R_1(x)v_1, v_1')\overline{(R_2(x)v_2, v_2')} = 0 \quad \text{for all } v_1, v_1' \in V_1 \text{ and } v_2, v_2' \in V_2.$$

(b) if (R, V) is irreducible, then

$$\sum_{x\in G}(R(x)v_1, v_1')\overline{(R(x)v_2, v_2')} = \frac{|G|(v_1, v_2)\overline{(v_1', v_2')}}{\dim V} \quad \text{for } v_1, v_2, v_1', v_2' \in V.$$

REMARKS. If G is abelian, then V_1 and V_2 in (a) are 1-dimensional, and the conclusion of (a) reduces to the statement that the multiplicative characters are orthogonal. Conclusion (b) in this case reduces to a trivial statement.

PROOF. For (a), let $l : V_2 \to V_1$ be any linear map, and form the linear map

$$L = \sum_{x\in G} R_1(x) l R_2(x^{-1}).$$

Multiplying on the left by $R_1(g)$ and on the right by $R_2(g^{-1})$ and changing variables in the sum, we obtain $R_1(g)LR_2(g^{-1}) = L$, so that $R_1(g)L = LR_2(g)$ for all $g \in G$. By Schur's Lemma (Proposition 7.18) and the assumed irreducibility and inequivalence, $L = 0$. Thus $(Lv_2', v_1') = 0$. For the particular choice of l as $l(w_2) = (w_2, v_2)v_1$, we have

$$\begin{aligned} 0 = (Lv_2', v_1') &= \sum_{x\in G} (R_1(x) l R_2(x^{-1})v_2', v_1') \\ &= \sum_{x\in G} \big(R_1(x)(R_2(x^{-1})v_2', v_2)v_1, v_1'\big) = \sum_{x\in G} (R_1(x)v_1, v_1')(R_2(x^{-1})v_2', v_2), \end{aligned}$$

and (a) results since $(R_2(x^{-1})v_2', v_2) = \overline{(R_2(x)v_2, v_2')}$.

For (b), we proceed in the same way, starting from $l : V \to V$, and we obtain $L = \lambda I$ from Schur's Lemma. Taking the trace of both sides, we find that

$$\lambda \dim V = \operatorname{Tr} L = |G| \operatorname{Tr} l.$$

Therefore $\lambda = |G|(\operatorname{Tr} l)\big/ \dim V$. Since $L = \lambda I$,

$$(Lv_2', v_1') = \frac{|G| \operatorname{Tr} l}{\dim V} \overline{(v_1', v_2')}.$$

Again we make the particular choice of l as $l(w_2) = (w_2, v_2)v_1$. Since $\operatorname{Tr} l = (v_1, v_2)$, we obtain

$$\begin{aligned}
\frac{(v_1, v_2)\overline{(v_1', v_2')}}{\dim V} &= \frac{\operatorname{Tr} l}{\dim V} \overline{(v_1', v_2')} = |G|^{-1}(Lv_2', v_1') \\
&= |G|^{-1} \sum_{x \in G} (R(x) l R(x^{-1}) v_2', v_1') \\
&= |G|^{-1} \sum_{x \in G} \big(R(x)(R(x^{-1})v_2', v_2)v_1, v_1'\big) \\
&= |G|^{-1} \sum_{x \in G} (R(x)v_1, v_1')(R(x^{-1})v_2', v_2),
\end{aligned}$$

and (b) results since $(R(x^{-1})v_2', v_2) = \overline{(R(x)v_2, v_2')}$. □

Let us interpret Proposition 7.22 as a statement about the left and right regular representations ℓ and r of G on the inner-product space $C(G, \mathbb{C})$, the inner product being $\langle f, f' \rangle = \sum_{g \in G} f(g)\overline{f'(g)}$. Let R be an irreducible representation of G on the finite-dimensional vector space V, and introduce an inner product to make it unitary. A member of $C(G, \mathbb{C})$ of the form $g \mapsto (R(g)v, v')$ is called a **matrix coefficient** of R. Let $v_1, \dots, v_n$ be an orthonormal basis of V. The matrix representation of G that corresponds to R and this choice of orthonormal basis has $\rho(g)_{ij} = (R(g)v_j, v_i)$, and hence the entries of $[\rho(g)_{ij}]$, as functions on G, provide examples of matrix coefficients. These particular matrix coefficients are orthogonal, according to Proposition 7.22b, with

$$\sum_{g \in G} |\rho(g)_{ij}|^2 = \sum_{g \in G} (R(g)v_j, v_i)\overline{(R(g)v_j, v_i)} = \frac{|G|(v_j, v_j)\overline{(v_i, v_i)}}{\dim V} = \frac{|G|}{\dim V}.$$

Thus the functions $\sqrt{|G|^{-1} \dim V}\ \rho(x)_{ij}$ form an orthonormal basis of an n^2-dimensional subspace V_R of $C(G, \mathbb{C})$, where $n = \dim V$. The vector subspace V_R has the following properties:

(i) All matrix coefficients of R are in V_R, as is seen by expanding $v = \sum_j c_j v_j$ and $v' = \sum_i d_i v_i$ and obtaining $(R(g)v, v') = \sum_{i,j} c_j \bar{d}_i (R(g)v_j, v_i) = \sum_{i,j} c_j \bar{d}_i \rho(g)_{ij}$.

(ii) V_R is invariant under ℓ and r because

$$\ell(g)(R(\cdot)v, v')(x) = (R(g^{-1}x)v, v') = (R(x)v, R(g)v'),$$
$$r(g)(R(\cdot)v, v')(x) = (R(xg)v, v') = (R(x)R(g)v, v').$$

(iii) Any representation R' equivalent to R has $V_{R'} = V_R$.

Let us see how V_R decomposes into irreducible subspaces under r. The computation with r in (ii) above shows, for each i, that the vector space of all functions $x \mapsto (R(x)v, v_i)$ for $v \in V$ is invariant under r. This is the linear span of the matrix coefficients obtained from the i^{th} row of $[\rho(x)_{ij}]$. Define a linear map A from V into this vector space by $Av = (R(\cdot)v, v_i)$. It is evident that A is one-one onto, and moreover $AR(g)v = (R(\cdot)R(g)v, v_i) = r(g)(R(\cdot)v, v_i) = r(g)Av$. Thus A exhibits this space, with r as representation, as equivalent to (R, V). The space V_R is the direct sum of these spaces on i, and the summands are orthogonal, according to Proposition 7.22b. Thus V_R decomposes under r as the direct sum of $\dim V$ irreducible subspaces, each one equivalent to (R, V).

One can make a similar analysis with ℓ, using columns in place of rows. However, this analysis is a little more subtle since V_R, acted upon by ℓ, is the direct sum of $\dim V$ copies of the "contragredient" of (R, V), rather than (R, V) itself. The details are left to Problems 32–36 at the end of the chapter.

As R varies over inequivalent representations, these vector spaces V_R are orthogonal, according to Proposition 7.22a. The claim is that their direct sum is the space $C(G, \mathbb{C})$ of all functions on G. We argue by contradiction. The sum is invariant under r, and if it is not all of $C(G, \mathbb{C})$, then we can find a nonzero vector subspace $U = \{f(\cdot)\}$ of $C(G, \mathbb{C})$ orthogonal to all the spaces V_R such that U is invariant and irreducible under r. Let $u_1, \ldots, u_m$ be an orthonormal basis of U. Then each function $x \mapsto (r(x)u_j, u_i)$ is orthogonal to U by construction, i.e.,

$$0 = \sum_{x \in G} (r(x)u_j, u_i)\overline{f(x)} \qquad \text{for all } f \text{ in } U.$$

Applying the Riesz Representation Theorem (Theorem 3.12), choose a member e of U such that $f(1) = (f, e)$ for all f in U. By definition of $r(x)$ and e, we find that

$$u(x) = (r(x)u)(1) = (r(x)u, e)$$

for all u in U. Substitution and use once more of Proposition 7.22b gives

$$0 = \sum_{x \in G} (r(x)u_j, u_i)\overline{(r(x)u, e)} = \frac{|G|(u_j, u)\overline{(u_i, e)}}{\dim U}$$

for all i and j. Since we can take $u = u_j = u_1$ and since i is arbitrary, this equation forces $e = 0$ and gives a contradiction. We conclude that the sum of all the spaces V_R is all of $C(G, \mathbb{C})$. Let us state the result as a theorem.

Theorem 7.23. For the finite group G, let $\{(R_\alpha, U_\alpha)\}$ be a complete set of inequivalent irreducible finite-dimensional representations of G, and let V_{R_α} be the linear span of the matrix coefficients of R_α. Then

(a) the spaces V_{R_α} are mutually orthogonal and are invariant under the left and right regular representations ℓ and r,
(b) the representation (r, V_{R_α}) is equivalent to the direct sum of $\dim U_\alpha$ copies of (R_α, U_α),
(c) the direct sum of the spaces V_{R_α} is the space $C(G, \mathbb{C})$ of all complex-valued functions on G.

Moreover,

(d) the number of R_α's is finite,
(e) $\dim V_{R_\alpha} = (\dim U_\alpha)^2$,
(f) any irreducible subspace of $(r, C(G, \mathbb{C}))$ that is equivalent to (R_α, U_α) is contained in V_{R_α}.

Corollary 7.24. Let $\{(R_\alpha, U_\alpha)\}$ be a complete set of inequivalent irreducible finite-dimensional representations of the finite group G, and let $d_\alpha = \dim U_\alpha$. In each U_α, introduce an inner product making (R_α, U_α) unitary. For each α, let $\{u_1^{(\alpha)}, \dots, u_{d_\alpha}^{(\alpha)}\}$ be an orthonormal basis of U_α. Then the functions in $C(G, \mathbb{C})$ given by $\sqrt{|G|^{-1} d_\alpha}\ \big(R_\alpha(x) v_j^{(\alpha)}, v_i^{(\alpha)}\big)$ form an orthonormal basis of $C(G, \mathbb{C})$. Consequently every f in $C(G, \mathbb{C})$ satisfies

$$f(x) = \frac{1}{|G|} \sum_\alpha d_\alpha \sum_{i,j} \Big(\sum_{y \in G} f(y) \overline{\big(R_\alpha(y) v_j^{(\alpha)}, v_i^{(\alpha)}\big)} \Big) \big(R_\alpha(x) v_j^{(\alpha)}, v_i^{(\alpha)}\big)$$

and

$$\sum_{x \in G} |f(x)|^2 = \frac{1}{|G|} \sum_\alpha d_\alpha \sum_{i,j} \Big| \sum_{y \in G} f(y) \overline{\big(R_\alpha(y) v_j^{(\alpha)}, v_i^{(\alpha)}\big)} \Big|^2.$$

REMARKS. The first displayed formula is the **Fourier inversion formula** for an arbitrary finite group G and generalizes Theorem 7.17, which gives the result in the abelian case; in the abelian case all the dimensions d_α equal 1, and the functions $\big(R_\alpha(x) v_j^{(\alpha)}, v_i^{(\alpha)}\big)$ are just the multiplicative characters of G. The second displayed formula is known as the **Plancherel formula**, a result incorporating the conclusion about norms in Parseval's equality (Theorem 3.11d).

PROOF. This follows form (a), (c), and (e) in Theorem 7.23, together with Theorem 3.11 and the remarks made before the statement of Theorem 7.23. □

Corollary 7.25. Let $\{(R_\alpha, U_\alpha)\}$ be a complete set of inequivalent irreducible finite-dimensional representations of the finite group G, and let $d_\alpha = \dim U_\alpha$. Then $\sum_\alpha d_\alpha^2 = |G|$.

PROOF. This follows by counting the number of members listed in the orthonormal basis of $C(G, \mathbb{C})$ given in Corollary 7.24. □

We shall make use of a second multiplication on the vector space $C(G, \mathbb{C})$ besides the pointwise multiplication that itself makes $C(G, \mathbb{C})$ into a ring. The new multiplication is called **convolution** and is defined by

$$(f_1 * f_2)(x) = \sum_{y \in G} f_1(y) f_2(y^{-1}x) = \sum_{y \in G} f_1(xy^{-1}) f_2(y),$$

the two expressions on the right being equal by a change of variables. The first of the expressions on the right equals the value of the function $\sum_{y \in G} f_1(y)\ell(y) f_2$ at x and shows that the convolution is an average of the left translates of f_2 weighted by f_1. Convolution is associative because

$$\begin{aligned}(f_1 * (f_2 * f_3))(x) &= \sum_y f_1(y)(f_2 * f_3)(y^{-1}x) = \sum_{y,z} f_1(y) f_2(y^{-1}xz^{-1}) f_3(z) \\ &= \sum_z (f_1 * f_2)(xz^{-1}) f_3(z) = ((f_1 * f_2) * f_3)(x),\end{aligned}$$

and one readily checks that $C(G, \mathbb{C})$ becomes a ring when convolution is used as the multiplication.

For any finite-dimensional representation (R, V) and any v in V, let us define $R(f)v = \sum_{x \in G} f(x)R(x)v$. Convolution has the property that

$$R(f_1 * f_2) = R(f_1)R(f_2)$$

because

$$\begin{aligned}R(f_1 * f_2)v &= \textstyle\sum_x (f_1 * f_2)(x)R(x)v = \sum_{x,y} f_1(xy^{-1}) f_2(y)R(x)v \\ &= \textstyle\sum_{x,y} f_1(x) f_2(y)R(xy)v = \sum_x f_1(x)R(x)\big(\sum_y f_2(y)R(y)v\big) \\ &= \textstyle\sum_x f_1(x)R(x)R(f_2)v = R(f_1)R(f_2)v.\end{aligned}$$

We shall combine the notion of convolution with the notion of a "character." If (R, V) is a finite-dimensional representation of G, then the **character** of (R, V) is the function χ_R given by

$$\chi_R(x) = \operatorname{Tr} R(x),$$

with Tr denoting the trace. Equivalent representations have the same character since $\operatorname{Tr}(AR(x)A^{-1}) = \operatorname{Tr} R(x)$ if A is invertible. Characters have the additional properties that

(i) $\chi_R(gxg^{-1}) = \chi_R(x)$ because $\operatorname{Tr} R(gxg^{-1}) = \operatorname{Tr}(R(g)R(x)R(g)^{-1}) = \operatorname{Tr} R(x)$,

(ii) $\chi_{R_1 \oplus \cdots \oplus R_n} = \chi_{R_1} + \cdots + \chi_{R_n}$ since the trace of a block-diagonal matrix is the sum of the traces of the blocks.

The character of a 1-dimensional representation is the associated multiplicative character. Here is an example of a character for a representation on a space of dimension more than 1; its values are not all in S^1.

EXAMPLE. The dihedral group D_n with $2n$ elements, defined in Section IV.1, is isomorphic to the matrix group generated by

$$x = \begin{pmatrix} \cos 2\pi/n & -\sin 2\pi/n \\ \sin 2\pi/n & \cos 2\pi/n \end{pmatrix} \qquad \text{and} \qquad y = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

The map carrying each matrix of the group to itself is a representation of D_n on $\mathbb{C}^2$. The value of the character of this representation is $2\cos 2\pi k/n$ on x^k for $0 \le k \le n-1$, and the value of the character is 0 on y and on the remaining $n-1$ elements of the group.

Computations with characters are sometimes aided by the use of inner products. If an inner product is imposed on a finite-dimensional complex vector space V and if $\{v_i\}$ is an orthonormal basis, then the trace of a linear $A : V \to V$ is given by $\operatorname{Tr} A = \sum_i (Av_i, v_i)$. If R is a representation on V, we consequently have $\chi_R(x) = \sum_i (R(x)v_i, v_i)$.

Proposition 7.26. Let R, R_1, and R_2 be irreducible finite-dimensional representations of a finite group G. Then their characters satisfy

(a) $\sum_{x\in G} |\chi_R(x)|^2 = |G|$,
(b) $\sum_{x\in G} \chi_{R_1}(x)\overline{\chi_{R_2}(x)} = 0$ if R_1 and R_2 are inequivalent.

PROOF. These follow from Schur orthogonality (Proposition 7.22): For (a), let R act on the vector space V, let $d = \dim V$, introduce an inner product with respect to which R is unitary, and let $\{v_i\}$ be an orthonormal basis of V. Then Proposition 7.22b gives

$$\begin{aligned} \sum_x |\chi_R(x)|^2 &= \sum_x \Big(\sum_i (R(x)v_i, v_i)\Big)\overline{\Big(\sum_j (R(x)v_j, v_j)\Big)} \\ &= \sum_{i,j} \sum_x (R(x)v_i, v_i)\overline{(R(x)v_j, v_j)} \\ &= \sum_{i,j} |G|d^{-1}\delta_{ij}\delta_{ij} = \sum_i |G|d^{-1} = |G|. \end{aligned}$$

Part (b) is proved in the same fashion, using Proposition 7.22a. □

Let us now bring together the notions of convolution and character. A **class function** on G is a function f in $C(G, \mathbb{C})$ with $f(gxg^{-1}) = f(x)$ for all g and x in G. That is, class functions are the ones that are constant on each conjugacy class of the group. Every character is an example of a class function. The class

functions form a vector subspace of $C(G, \mathbb{C})$, and the dimension of this vector subspace equals the number of conjugacy classes in G. Class functions are closed under convolution because if f_1 and f_2 are class functions, then

$$\begin{aligned}(f_1 * f_2)(gxg^{-1}) &= \textstyle\sum_y f_1(gxg^{-1}y^{-1}) f_2(y) = \sum_y f_1(xg^{-1}y^{-1}g) f_2(g^{-1}yg) \\ &= \textstyle\sum_z f_1(xz^{-1}) f_2(z) = (f_1 * f_2)(x).\end{aligned}$$

On an abelian group every member of $C(G, \mathbb{C})$ is a class function.

Theorem 7.27 (Fourier inversion formula for class functions). For the finite group G, let $\{(R_\alpha, U_\alpha)\}$ be a complete set of inequivalent irreducible finite-dimensional representations of G. If f is a class function on G, then

$$f(x) = \frac{1}{|G|} \sum_\alpha \Big(\sum_{y \in G} f(y) \overline{\chi_{R_\alpha}(y)} \Big) \chi_{R_\alpha}(x).$$

REMARK. This result may be regarded as a second way (besides the one in Corollary 7.24) of generalizing Theorem 7.17 to the nonabelian case.

PROOF. Using the result and notation of Corollary 7.24, we have

$$f(x) = |G|^{-1} \sum_\alpha d_\alpha \sum_{i,j} \Big(\sum_{y \in G} f(y) \overline{(R_\alpha(y) v_i^{(\alpha)}, v_j^{(\alpha)})} \Big) (R_\alpha(x) v_i^{(\alpha)}, v_j^{(\alpha)}).$$

Replace $f(y)$ by $f(gyg^{-1})$ since f is a class function, and then change variables and sum over g in G to see that $|G| f(x)$ is equal to

$$|G|^{-1} \sum_\alpha d_\alpha \sum_{i,j} \Big(\sum_{g,y} f(y) \overline{(R_\alpha(y) R_\alpha(g) v_i^{(\alpha)}, R_\alpha(g) v_j^{(\alpha)})} \Big) (R_\alpha(x) v_i^{(\alpha)}, v_j^{(\alpha)}).$$

Within this expression we have

$$\begin{aligned}&\sum_g \overline{(R_\alpha(y) R_\alpha(g) v_i^{(\alpha)}, R_\alpha(g) v_j^{(\alpha)})} \\ &\quad = \sum_{g,k} \overline{\big(R_\alpha(y)(R_\alpha(g) v_i^{(\alpha)}, v_k^{(\alpha)}) v_k^{(\alpha)}, R_\alpha(g) v_j^{(\alpha)} \big)} \\ &\quad = \sum_{g,k} \overline{(R_\alpha(g) v_i^{(\alpha)}, v_k^{(\alpha)})} (R_\alpha(g) v_j^{(\alpha)}, R_\alpha(y) v_k^{(\alpha)}) \\ &\quad = \tfrac{|G|}{d_\alpha} \sum_k (v_j^{(\alpha)}, v_i^{(\alpha)}) \overline{(R_\alpha(y) v_k^{(\alpha)}, v_k^{(\alpha)})} \qquad \text{by Schur orthogonality} \\ &\quad = \tfrac{|G|}{d_\alpha} (v_j^{(\alpha)}, v_i^{(\alpha)}) \overline{\chi_{R_\alpha}(y)} \\ &\quad = \tfrac{|G|}{d_\alpha} \delta_{ij} \overline{\chi_{R_\alpha}(y)}.\end{aligned}$$

Substituting, we obtain the formula of the theorem. □

Corollary 7.28. If G is a finite group, then the number of irreducible finite-dimensional representations of G, up to equivalence, equals the number of conjugacy classes of G.

PROOF. Theorem 7.27 shows that the irreducible characters span the vector space of class functions. Proposition 7.26b shows that the irreducible characters are orthogonal and hence are linearly independent. Thus the number of irreducible characters equals the dimension of the space of class functions, which equals the number of conjugacy classes. □

EXAMPLE. The above information already gives us considerable control over finding a complete set of inequivalent irreducible finite-dimensional representations of elementary groups. We know that the number of such representations equals the number of conjugacy classes and that the sum of the squares of their dimensions equals $|G|$. For the symmetric group $\mathfrak{S}_3$ of order 6, for example, the conjugacy classes are given by the cycle structures of the possible permutations, namely the cycle structures of (1), (1 2), and (1 2 3). Hence there are three inequivalent irreducible representations. The sum of the squares of the three dimensions is to be 6; thus we have two of dimension 1 and one of dimension 2. The multiplicative characters 1 and sgn are the two of dimension 1, and the one of dimension 2 can be taken to be the 2-dimensional representation of D_3 whose character was computed in the example preceding Proposition 7.26.

One final constraint on the dimensions of the irreducible representations of a finite group G is as follows.

Proposition 7.29. If G is a finite group and (R, V) is an irreducible finite-dimensional representation of G, then $\dim V$ divides $|G|$.

For example, if $|G| = p^2$ with p prime, then it follows from Propositions 7.29 and 7.25 that every irreducible finite-dimensional representation of G has dimension 1, and one can easily conclude from this fact that G is abelian. (See Problem 14 at the end of the chapter.) Thus we recover as an immediate consequence the conclusion of Corollary 4.39 that groups of order p^2 are abelian.

The proof of Proposition 7.29 is surprisingly subtle. We shall obtain the theorem as a consequence of Theorem 7.31 below, a theorem that will be used also in the proof of Burnside's Theorem in the next section. Theorem 7.31 gives a little taste of the usefulness of algebraic number theory, and we shall see more of this usefulness in Chapter IX. The application to Burnside's Theorem will use the Fundamental Theorem of Galois Theory, whose proof is deferred to Chapter IX.

An **algebraic integer** is any complex number that is a root of a monic polynomial with coefficients in $\mathbb{Z}$. For example, $\sqrt{2}$ and $\frac{1}{2}(1 + i\sqrt{3})$ are algebraic

integers because they are roots of $X^2 - 2$ and $X^2 - X + 1$, respectively. Any root of unity is an algebraic integer, being a root of some polynomial $X^n - 1$. The set of algebraic integers will be denoted in this chapter by $\mathcal{O}$. Before stating Theorem 7.31, let us establish two elementary facts about $\mathcal{O}$.

Lemma 7.30. The set $\mathcal{O}$ of algebraic integers is a ring, and $\mathcal{O} \cap \mathbb{Q} = \mathbb{Z}$.

PROOF. Suppose that x and y are complex numbers satisfying the polynomial equations $x^m + a_{m-1}x^{m-1} + \cdots + a_1 x + a_0 = 0$ and $y^n + b_{n-1}y^{n-1} + \cdots + b_1 y + b_0 = 0$, each with integer coefficients. Form the subset of $\mathbb{C}$ given by

$$M = \sum_{k=0}^{m-1} \sum_{l=0}^{n-1} \mathbb{Z} x^k y^l.$$

This is a finitely generated subgroup of the abelian group $\mathbb{C}$ under addition. It satisfies

$$\begin{aligned} xM &= \sum_{k=1}^{m} \sum_{l=0}^{n-1} \mathbb{Z} x^k y^l \subseteq M + \sum_{l=0}^{n-1} \mathbb{Z} y^l x^m \\ &= M + \sum_{l=0}^{n-1} \mathbb{Z} y^l (-a_{m-1}x^{m-1} - \cdots - a_1 x - a_0) \subseteq M, \end{aligned}$$

and similarly $yM \subseteq M$. Hence $(x \pm y)M \subseteq M$ and $xy \subseteq M$.

To prove that $\mathcal{O}$ is a ring, it is enough to show that if N is a nonzero finitely generated subgroup of the abelian group $\mathbb{C}$ under addition and if z is a complex number with $zN \subseteq N$, then z is an algebraic integer. By Theorem 4.56, N is a direct sum of cyclic groups. Since every nonzero member of $\mathbb{C}$ has infinite order additively, these cyclic groups must be copies of $\mathbb{Z}$. So N is free abelian. Let $z_1, \ldots, z_n$ be a $\mathbb{Z}$ basis of N. Here $n > 0$. Since $zN \subseteq N$, we can find unique integers c_{ij} such that

$$z z_i = \sum_{j=1}^{n} c_{ij} z_j \qquad \text{for } 1 \le i \le n.$$

This equation says that the matrix $C = [c_{ij}]$ has $\begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}$ as an eigenvector with eigenvalue z. Therefore the matrix $zI - C$ is singular, and $\det(zI - C) = 0$. Since $\det(zI - C)$ is a monic polynomial expression in z with integer coefficients, z is an algebraic integer.

To see that $\mathcal{O} \cap \mathbb{Q} = \mathbb{Z}$, let p and q be relatively prime integers with $q > 0$, and suppose that p/q is a root of $X^n + a_{n-1}X^{n-1} + \cdots + a_1 X + a_0$ with $a_{n-1}, \ldots, a_0$ in $\mathbb{Z}$. Substituting p/q for X, setting the expression equal to 0, and clearing fractions, we obtain $p^n + a_{n-1}p^{n-1}q + \cdots + a_1 pq^{n-1} + a_0 q^n = 0$. Since q divides every term here after the first, we conclude that q divides p^n. Since $\mathrm{GCD}(p, q) = 1$, we conclude that $q = 1$. Thus p/q is in $\mathbb{Z}$. □

Lemma 7.30 allows us to see that if G is a finite group and χ is the irreducible character corresponding to an irreducible finite-dimensional representation R, then $\chi(x)$ is an algebraic integer for each x in G. In fact, the subgroup H of G generated by x is cyclic and is in particular abelian. Corollary 7.21 says that $R\big|_H$ is the direct sum of irreducible representations of H, and Corollary 7.19 says that each such irreducible representation is 1-dimensional. Thus in a suitable basis, $R\big|_H$ is diagonal. The diagonal entries must be roots of unity (in fact, N^{th} roots of unity if x has order N), and $\chi(x)$ is thus a sum of roots of unity. By Lemma 7.30, $\chi(x)$ is an algebraic integer.

Theorem 7.31. Let G be a finite group, (R, V) be an irreducible finite-dimensional representation of G, χ be the character of R, and C be a conjugacy class in G. Denote by $\chi(C)$ the constant value of χ on the conjugacy class C. Then $|C|\chi(C)\big/\dim V$ is an algebraic integer.

PROOF. If f is any class function on G, then $R(f)$ commutes with each $R(x)$ for x in G because $R(f) = \sum_y f(y)R(y)$ yields

$$\begin{aligned} R(x)R(f)R(x)^{-1} &= \sum_y f(y)R(x)R(y)R(x)^{-1} = \sum_y f(y)R(xyx^{-1}) \\ &= \sum_z f(x^{-1}zx)R(z) = \sum_z f(z)R(z) = R(f). \end{aligned}$$

By Schur's Lemma (Proposition 7.18), $R(f)$ is scalar. If C is a conjugacy class, then the function I_C that is 1 on C and is 0 elsewhere is a class function, and hence $R(I_C)$ is a scalar λ_C. As C varies, the functions I_C form a vector-space basis of the space of class functions. The formula $(I_C * I_{C'})(x) = \sum_y I_C(y)I_{C'}(y^{-1}x)$ shows that $I_C * I_{C'}$ is integer-valued, and we have seen that the convolution of two class functions is a class function. Therefore $I_C * I_{C'} = \sum_{C''} n_{CC'C''}I_{C''}$ for suitable integers $n_{CC'C''}$. Application of R gives $\lambda_C\lambda_{C'} = \sum_{C''} n_{CC'C''}\lambda_{C''}$. If we fix C and let A be the square matrix with entries $A_{C'C''} = n_{CC'C''}$, we obtain

$$\lambda_C\lambda_{C'} = \sum_{C''} A_{C'C''}\lambda_{C''}.$$

This equation says that the matrix A has the column vector with entries $\lambda_{C''}$ as an eigenvector with eigenvalue λ_C. Therefore the matrix $\lambda_C I - A$ is singular, and $\det(\lambda_C I - A) = 0$. Since $\det(\lambda_C I - A)$ is a monic polynomial expression in λ_C with integer coefficients, λ_C is an algebraic integer. Taking the trace of the equation $R(I_C) = \lambda_C I$, we obtain $\sum_{x\in C}\chi(x) = \lambda_C \dim V$. Since $\chi(x) = \chi(C)$ for x in C, the result is that $|C|\chi(C)/\dim V = \lambda_C$. Since λ_C is an algebraic integer, $|C|\chi(C)/\dim V$ is an algebraic integer. □

PROOF THAT THEOREM 7.31 IMPLIES PROPOSITION 7.29. Proposition 7.26a gives

$$\frac{|G|}{\dim V} = \frac{\sum_{x\in G} |\chi(x)|^2}{\dim V} = \frac{\sum_C \sum_{x\in C} |\chi(x)|^2}{\dim V} = \sum_C \Big(\frac{|C|\chi(C)}{\dim V}\Big)\overline{\chi(C)}.$$

Each term in parentheses on the right side is an algebraic integer, according to Theorem 7.31, and therefore Lemma 7.30 shows that $|G|/\dim V$ is an algebraic integer. Since $|G|/\dim V$ is in $\mathbb{Q}$, Lemma 7.30 shows that $|G|/\dim V$ is in $\mathbb{Z}$. □

5. Burnside's Theorem

The theorem of this section is as follows.

Theorem 7.32 (Burnside's Theorem). If G is a finite group of order $p^a q^b$ with p and q prime and with $a + b > 1$, then G has a nontrivial normal subgroup.

The argument will use the result Theorem 7.31 from algebraic number theory, and also it will make use of a special case of the Fundamental Theorem of Galois Theory, whose proof is deferred to Chapter IX. That special case is the following statement, whose context was anticipated in Section IV.1, where groups of automorphisms of certain fields were discussed briefly. Since the set $\{1, e^{2\pi i/n}, e^{2\cdot 2\pi i/n}, e^{3\cdot 2\pi i/n}, \dots\}$ is linearly dependent over $\mathbb{Q}$, Proposition 4.1 in that section implies that the subring $\mathbb{Q}[e^{2\pi i/n}]$ of $\mathbb{C}$ generated by $\mathbb{Q}$ and $e^{2\pi i/n}$ is a subfield and is a finite-dimensional vector space over $\mathbb{Q}$. According to Example 9 of that section, the group $\Gamma = \mathrm{Gal}(\mathbb{Q}[e^{2\pi i/n}]/\mathbb{Q})$ of automorphisms of $\mathbb{Q}[e^{2\pi i/n}]$ fixing every element of $\mathbb{Q}$ is a finite group.

Proposition 7.33 (special case of the Fundamental Theorem of Galois Theory). Let $n > 0$ be an integer, and put $K = \mathbb{Q}[e^{2\pi i/n}]$. Let Γ be the finite group of field automorphisms of K fixing every element of $\mathbb{Q}$. Then the only members β of K such that $\sigma(\beta) = \beta$ for every σ in Γ are the members of $\mathbb{Q}$.

Lemma 7.34. Let G be a finite group, (R, V) be an irreducible finite-dimensional representation of G, χ be the character of R, and C be a conjugacy class in G. If $\mathrm{GCD}(|C|, \dim V) = 1$ and if x is in C, then either $R(x)$ is scalar or $\chi(x) = 0$.

PROOF. Define $\chi(C)$ to be the constant value of χ on C, and put $\alpha = \chi(x)/\dim V = \chi(C)/\dim V$. Since $\mathrm{GCD}(|C|, \dim V) = 1$, we can choose integers m and n with $m|C| + n\dim V = 1$. Multiplication by α yields

$$\frac{m|C|\chi(C)}{\dim V} + n\chi(C) = \alpha.$$

Theorem 7.31 shows that the coefficients $\frac{|C|\chi(C)}{\dim V}$ and $\chi(C)$ of m and n on the left side are algebraic integers, and therefore α is an algebraic integer. As we observed toward the end of the previous section, $\chi(x) = \chi(C)$ is the sum of $\dim V$ roots of unity. Since $\alpha = \chi(C)/\dim V$, we see that $|\alpha| \le 1$ with equality only if all the roots of unity are equal, in which case $R(x)$ is scalar. In view of the hypothesis, we may assume that $|\alpha| < 1$. We shall show that $\alpha = 0$.

Let $K = \mathbb{Q}[e^{2\pi i/|G|}]$ be the smallest subfield of $\mathbb{C}$ containing $\mathbb{Q}$ and the complex number $e^{2\pi i/|G|}$, and let Γ be the group of field automorphisms of K that fix every element of $\mathbb{Q}$. We know that K is finite-dimensional over $\mathbb{Q}$ and that Γ is a finite group, and Proposition 7.33 shows that the only members of K fixed by every element of Γ are the members of $\mathbb{Q}$.

Our element x of G has $x^{|G|} = 1$. Thus every root of unity contributing to $\chi(x)$ is a $|G|^{\text{th}}$ root of unity and is in K. Therefore the algebraic integer α is in K. If σ is in Γ, each of the $|G|^{\text{th}}$ roots of unity is mapped by σ to some complex number x satisfying $x^{|G|} = 1$, and hence the member $\sigma(\alpha)$ of K satisfies $|\sigma(\alpha)| \le 1$. Also, $\sigma(\alpha)$ is an algebraic integer, as we see by applying σ to the monic equation with integer coefficients satisfied by α, and we are assuming that $|\alpha| < 1$. Consequently $\beta = \prod_{\sigma\in\Gamma} \sigma(\alpha)$ is an algebraic integer and has absolute value < 1. A change of variables in the product shows that β is fixed by every member of Γ, and we see from the previous paragraph that β is in $\mathbb{Q}$. By Lemma 7.30, β is in $\mathbb{Z}$. Being of absolute value less than 1, it is 0. Thus $\alpha = 0$, and $\chi(x) = 0$. □

Lemma 7.35. Let G be a finite group, and let C be a conjugacy class in G such that $|C| = p^k$ for some prime p and some integer $k > 0$. Then there exists an irreducible finite-dimensional representation $R \ne 1$ of G with $R(x)$ scalar for every x in C. Consequently G is not simple.

PROOF. The conjugacy class C cannot be $\{1\}$ because $|\{1\}| \ne p^k$ with $k > 0$. Let χ_{reg} be the character of the right regular representation r of G on $C(G, \mathbb{C})$. If I_g denotes the function that is 1 at g and is 0 elsewhere, then the functions I_g form an orthonormal basis of $C(G, \mathbb{C})$, and therefore $\chi_{\text{reg}}(x) = \sum_{g\in G} (r(x)I_g, I_g) = \sum_{g\in G} (I_{gx^{-1}}, I_g)$. Every term on the right side is 0 if $x \ne 1$, and thus Theorem 7.23 gives

$$0 = \chi_{\text{reg}}(x) = 1 + \sum_{\chi\ne 1} d_\chi \chi(x) \qquad \text{for } x \in C, \tag{$*$}$$

the sum being taken over all irreducible characters other than 1, with d_χ being the dimension of an irreducible representation corresponding to χ. Let R_χ be an irreducible representation with character χ. Any χ such that p does not divide d_χ has $\text{GCD}(|C|, d_\chi) = 1$ since $|C|$ is assumed to be a power of p. Arguing by

contradiction, we may assume that no such χ has $R_\chi(x)$ scalar, and then Lemma 7.34 says that $\chi(x) = 0$ for all such χ. Hence $(*)$ simplifies to

$$0 = 1 + \sum_{\chi \neq 1,\ p \text{ divides } d_\chi} d_\chi \chi(x) \qquad \text{for } x \in C. \tag{**}$$

Since $\chi(x)$ is an algebraic integer, Lemma 7.30 shows that this equation is of the form $1 + p\beta = 0$, where β is an algebraic integer. Then $\beta = -1/p$ shows that $-1/p$ is an algebraic integer. Since $-1/p$ is in $\mathbb{Q}$, Lemma 7.30 shows that it must be in $\mathbb{Z}$, and we have arrived at a contradiction. Thus there must have been some χ with $R_\chi(x)$ scalar for x in C.

The set of g in G for which this R_χ has $R_\chi(g)$ scalar is a normal subgroup of G that contains x and cannot therefore be $\{1\}$. Assume by way of contradiction that G is simple. Then $R_\chi(g)$ is scalar for all g in G. Since R_χ is irreducible, R_χ is 1-dimensional. Then the commutator subgroup G' of G is contained in the kernel of R_χ. Since $R_\chi \neq 1$, G' is not all of G. Since G' is normal, $G' = \{1\}$, and we conclude that G is abelian. But the given G has a conjugacy class with more than one element, and we have arrived at a contradiction. □

PROOF OF THEOREM 7.32. Corollary 4.38 shows that a group of prime-power order has a center different from $\{1\}$, and we may therefore assume that $p \neq q$, $a > 0$, and $b > 0$. Let H be a Sylow q-subgroup. Applying Corollary 4.38, let x be a member of the center Z_H of H other than 1. The centralizer $Z_G(\{x\})$ is a subgroup containing H, and it therefore has order $p^{a'}q^b$. If $a' = a$, then x is in the center of G, and the powers of x form the desired proper normal subgroup of G. Thus $a' < a$. By Proposition 4.37 the conjugacy class C of x has $|G|/p^{a'}q^b = p^{a-a'}$ elements with $a - a' > 0$. By Lemma 7.35, G is not simple. □

6. Extensions of Groups

In Section IV.8 we examined composition series for finite groups. For a given finite group, a composition series consists of a decreasing sequence of subgroups starting with the whole group and ending with $\{1\}$, each normal in the next larger one, such that the successive quotient groups are simple. The Jordan–Hölder Theorem (Corollary 4.50) assured us that the set of successive quotients, up to isomorphism, is independent of the choice of composition series. This theorem raises the question of reconstructing the whole group from data of this kind. Consider a single step of the process. If we know the normal subgroup and the simple quotient that it yields at a certain stage, what are the possibilities for the next-larger subgroup? We study this question and some of its ramifications in this section, dropping any hypotheses that are not helpful in the analysis. Here is an example that we shall carry along.

EXAMPLE. Suppose that the normal subgroup is the cyclic group C_4 and that the quotient is the cyclic group C_2. The whole group has to be of order 8, and the classification of groups of order 8 done in Problems 39–44 at the end of Chapter IV tells us that there are four different possibilities for the whole group: the abelian groups $C_4 \times C_2$ and C_8, the dihedral group D_4, and the quaternion group H_8.

Let us establish a framework for the general problem. We start with a group E, a normal subgroup N, and the quotient $G = E/N$. We seek data that determine the group law in E in terms of N and G. For each member u of G, fix a coset representative $\bar{u}$ in E such that $\bar{u}N = u$. Since N is normal, the element $\bar{u}$ of E yields an automorphism $(\,\cdot\,)^u$ of N defined by $x^u = \bar{u}x\bar{u}^{-1}$. In addition, the fact that G is a group says that any two of our representatives $\bar{u}$ and $\bar{v}$ have

$$\bar{u}\bar{v} = a(u, v)\overline{uv} \qquad \text{for some unique } a(u, v) \text{ in } N.$$

The set of all elements $a(u, v)$ for this choice of coset representatives is called a **factor set**, and E is called a **group extension** of N by the group[3] G.

The automorphisms and the factor set constructed above have to satisfy two compatibility conditions, as follows:

(i) $(x^v)^u = a(u, v)x^{uv}a(u, v)^{-1}$ because $(x^u)^v = \bar{u}(x^v)\bar{u}^{-1} = \bar{u}\bar{v}x\bar{v}^{-1}\bar{u}^{-1}$ $= (a(u, v)\overline{uv})x(a(u, v)\overline{uv})^{-1} = a(u, v)x^{uv}a(u, v)^{-1}$,

(ii) $a(v, w)^u a(u, vw) = a(u, v)a(uv, w)$ because $(\bar{u}\bar{v})\bar{w} = a(u, v)\overline{uv}\bar{w}$ $= a(u, v)a(uv, w)\overline{uvw}$ and $\bar{u}(\bar{v}\bar{w}) = \bar{u}a(v, w)\overline{vw} = a(v, w)^u\bar{u}\overline{vw} =$ $a(v, w)^u a(u, vw)\overline{uvw}$.

Then the multiplication law in E is given in terms of the automorphisms and the factor set by the formula

(iii) $(x\bar{u})(y\bar{v}) = xy^u a(u, v)\overline{uv}$ by the computation $(x\bar{u})(y\bar{v}) = xy^u\bar{u}\bar{v} =$ $xy^u a(u, v)\overline{uv}$.

Conversely, according to the proposition below, such data determine a group E with a normal subgroup isomorphic to N and a quotient E/N isomorphic to G.

Proposition 7.36 (Schreier). Let two groups N and G be given, along with a family of automorphisms $x \mapsto x^u$ of N parametrized by u in G, as well as a function $a : G \times G \to N$ such that

(a) $(x^v)^u = a(u, v)x^{uv}a(u, v)^{-1}$ for all u and v in G,

(b) $a(v, w)^u a(u, vw) = a(u, v)a(uv, w)$ for all u, v, w in G.

Then the set $N \times G$ becomes a group E under the multiplication

(c) $(x, u)(y, v) = (xy^u a(u, v), uv)$,

[3] *Warning:* Some authors say "group extension of G by N."

and this group has a normal subgroup isomorphic to N with quotient group isomorphic to G. More particularly, the identity of E is $(a(1,1)^{-1},1)$, the map $x \mapsto (xa(1,1)^{-1},1)$ of N into E is a one-one homomorphism that exhibits N as a normal subgroup of E, and the map $(x,u) \mapsto u$ of E onto G is a homomorphism that exhibits G as isomorphic to E/N.

PROOF. Reverting to the earlier notation, let us write $x\bar{u}$ in place of (x,u) for elements of E. Associativity of multiplication follows from the computation

$$\begin{aligned}
(x\bar{u}y\bar{v})(z\bar{w}) &= \big(xy^{u}a(u,v)\overline{uv}\big)z\bar{w} && \text{by (c)}\\
&= xy^{u}a(u,v)z^{uv}a(uv,w)\overline{uvw} && \text{by (c)}\\
&= xy^{u}a(u,v)z^{uv}a(u,v)^{-1}a(u,v)a(uv,w)\overline{uvw}\\
&= xy^{u}a(u,v)z^{uv}a(u,v)^{-1}a(v,w)^{u}a(u,vw)\overline{uvw} && \text{by (b)}\\
&= x\big(yz^{v}a(v,w)\big)^{u}a(u,vw)\overline{uvw} && \text{by (a)}\\
&= (x\bar{u})\big(yz^{v}a(v,w)\overline{vw}\big) && \text{by (c)}\\
&= (x\bar{u})(y\bar{v}z\bar{w}) && \text{by (c).}
\end{aligned}$$

The identity is to be $\bar{1}a(1,1)^{-1}$. Before checking this assertion, we prove three preliminary identities. Setting $u = v = 1$ in (a) and replacing x^1 by x gives[4]

$$x^1 = a(1,1)xa(1,1)^{-1} \qquad \text{for all } x \in N. \tag{$*$}$$

Setting $v = w = 1$ in (b) gives $a(1,1)^{u}a(u,1) = a(u,1)a(u,1)$ and hence

$$a(1,1)^{u} = a(u,1) \qquad \text{for all } u \in G. \tag{$\dagger$}$$

Meanwhile, setting $u = v = 1$ in (b) gives $a(1,w)^{1}a(1,w) = a(1,1)a(1,w)$ and hence $a(1,w)^{1} = a(1,1)$ for all $w \in G$. The left side $a(1,w)^{1}$ of this last equality is equal to $a(1,1)a(1,w)a(1,1)^{-1}$ by $(*)$; canceling $a(1,1)$ yields

$$a(1,w) = a(1,1) \qquad \text{for all } w \in G. \tag{$\dagger\dagger$}$$

Using these identities, we check that $a(1,1)^{-1}\bar{1}$ is a two-sided identity by making the computations

$$\begin{aligned}
(x\bar{u})(a(1,1)^{-1}\bar{1}) &= x(a(1,1)^{-1})^{u}a(u,1)\bar{u} && \text{by (c)}\\
&= x(a(1,1)^{-1})^{u}a(1,1)^{u}\bar{u} && \text{by } (\dagger)\\
&= x\bar{u}
\end{aligned}$$

[4]The effect of the automorphism $x \mapsto x^1$ is not necessarily trivial since the coset representative $\bar{1}$ of 1 is not assumed to be the identity. Thus we must distinguish between x^1 and x.

and

$$\begin{aligned}(a(1,1)^{-1}\bar{1})(y\bar{v}) &= a(1,1)^{-1}y^1a(1,v)\bar{v} && \text{by (c)}\\ &= ya(1,1)^{-1}a(1,v)\bar{v} && \text{by } (*)\\ &= y\bar{v} && \text{by } (\dagger\dagger).\end{aligned}$$

Let us check that a left inverse for $x\bar{u}$ is $a(1,1)^{-1}a(u^{-1},u)^{-1}(x^{u^{-1}})^{-1}\overline{u^{-1}}$. In fact,

$$\begin{aligned}&\big(a(1,1)^{-1}a(u^{-1},u)^{-1}(x^{u^{-1}})^{-1}\overline{u^{-1}}\big)(x\bar{u})\\ &\qquad = a(1,1)^{-1}a(u^{-1},u)^{-1}(x^{u^{-1}})^{-1}x^{u^{-1}}a(u^{-1},u)\bar{1} && \text{by (c)}\\ &\qquad = a(1,1)^{-1}\bar{1},\end{aligned}$$

as required. Thus multiplication is associative, there is a two-sided identity, and every element has a left inverse. It follows that E is a group.

The map $x\bar{u} \mapsto u$ of E into G is a homomorphism by (c), and it is certainly onto G. Its kernel is evidently the subgroup of all elements $xa(1,1)^{-1}\bar{1}$ in E. Since

$$\begin{aligned}\big(xa(1,1)^{-1}\bar{1}\big)\big(ya(1,1)^{-1}\bar{1}\big) &= xa(1,1)^{-1}(ya(1,1)^{-1})^1a(1,1)\bar{1} && \text{by (c)}\\ &= xa(1,1)^{-1}a(1,1)(ya(1,1)^{-1})\bar{1} && \text{by } (*)\\ &= xya(1,1)^{-1}\bar{1},\end{aligned}$$

the one-one map $x \mapsto xa(1,1)^{-1}\bar{1}$ of N onto the kernel respects the group structures and is therefore an isomorphism. In other words, the embedded version of N is the kernel. Being a kernel, it is a normal subgroup. □

EXAMPLE, CONTINUED. Let $N = C_4 = \{1, r, r^2, r^3\}$ and $G = C_2 = \{1, u_0\}$ with $u_0^2 = 1$. The group N has two automorphisms, the nontrivial one fixing 1 and r^2 while interchanging r and r^3. The automorphism of N from $1 \in G$ has to be trivial, while the automorphism of N from $u_0 \in G$ can be trivial or nontrivial. In fact,

$$\text{the automorphism is } \begin{cases}\text{trivial} & \text{for } E = C_4 \times C_2 \text{ and } E = C_8,\\ \text{nontrivial} & \text{for } E = D_4 \text{ and } E = H_8.\end{cases}$$

In each case the automorphism does not depend on the choice of coset representatives. The factor sets do depend on the choice of representatives, however. Let us fix $\bar{1}$ as the identity of E and make a particular choice of $\overline{u_0}$ for each E. Then

the definition of factor set shows that $a(1, 1) = a(u_0, 1) = a(1, u_0) = 1$, and the only part of the factor set yet to be determined is $a(u_0, u_0)$. Let us consider matters group by group. For $C_4 \times C_2$, we can take $\overline{u_0}$ to be the generator of the C_2 factor; this has square 1, and hence $a(u_0, u_0) = 1$. For $C_8 = \{1\,\theta, \theta^2, \ldots, \theta^7\}$, let us think of N as embedded in E with $r = \theta^2$. The element $\overline{u_0}$ can be any odd power of θ; if we take $\overline{u_0} = \theta$, then $(\overline{u_0})^2 = \theta^2 = r$, and hence $a(u_0, u_0) = r$. For $E = D_4$, the example following Proposition 7.8 shows that we may view the elements as the rotations $1, r, r^2, r^3$ and the reflections s, rs, r^2s, r^3s for particular choices of r and s. We can take $\overline{u_0}$ to be any of the reflections, and then $(\overline{u_0})^2 = 1$ and $a(u_0, u_0) = 1$. Finally for $E = H_8 = \{\pm 1, \pm\mathbf{i}, \pm\mathbf{j}, \pm\mathbf{k}\}$, let us say that N is embedded as $\{\pm 1, \pm\mathbf{i}\}$. Then $\overline{u_0}$ can be any of the four elements $\pm\mathbf{j}$ and $\pm\mathbf{k}$. Each of these has square -1, and hence $a(u_0, u_0) = -1$. For the choices we have made, we therefore have

$$a(u_0, u_0) = \begin{cases} 1 & \text{for } E = C_4 \times C_2 \text{ and } E = D_4, \\ r & \text{for } E = C_8, \\ -1 & \text{for } E = H_8. \end{cases}$$

The formula of Proposition 7.36a reduces to $(x^v)^u = x^{uv}$ since N is abelian, and it is certainly satisfied. The formula for Proposition 7.36b is $a(v, w)^u a(u, vw) = a(u, v)a(uv, w)$. This is satisfied for $E = C_4 \times C_2$ and $E = D_4$ since $a(\,\cdot\,,\,\cdot\,)$ is identically 1. For the other two cases the values of $a(\,\cdot\,,\,\cdot\,)$ lie in the 2-element subgroup of N that is fixed by the nontrivial automorphism, and hence $a(v, w)^u = a(v, w)$ in every case. The formula to be checked reduces to $a(v, w)a(1, 1) = a(1, 1)a(v, w)$ by (††) if $u = 1$, to $a(1, 1)a(u, w) = a(1, 1)a(u, w)$ by (†) and (††) if $v = 1$, and to $a(1, 1)a(u, v) = a(u, v)a(1, 1)$ by (†) if $w = 1$. Thus all that needs checking is the case that $u = v = w = u_0$, and then the formula in question reduces to $a(u_0, u_0)a(1, 1) = a(u_0, u_0)a(1, 1)$ by (†) and (††).

Let us examine for a particular extension the dependence of the automorphisms and factor set on the choice of coset representatives. Returning to our original construction, suppose that we change the coset representatives of the members of G, associating a member $\widetilde{u}$ to $u \in G$ in place of $\bar{u}$. We then obtain a new automorphism of N corresponding to u, and we write it as $x \mapsto x^{u^*} = \widetilde{u}x\widetilde{u}^{-1}$ instead of $x \mapsto x^u = \bar{u}x\bar{u}^{-1}$. To quantify matters, we observe that $\widetilde{u}$ lies in the same coset of N as does $\bar{u}$. Thus $\widetilde{u} = \alpha(u)\bar{u}$ for some function $\alpha : G \to N$, and the function α can be absolutely arbitrary. In terms of this function α, the two automorphisms are related by

$$x^{u^*} = \widetilde{u}x\widetilde{u}^{-1} = \alpha(u)\bar{u}x\bar{u}^{-1}\alpha(u)^{-1} = \alpha(u)x^u\alpha(u)^{-1}.$$

If the factor set for the system $\{\widetilde{u}\}$ of coset representatives is denoted by $\{b(u, v)\}$, then we have $b(u, v)\alpha(uv)\overline{uv} = b(u, v)\widetilde{uv} = \widetilde{u}\,\widetilde{v} = \alpha(u)\bar{u}\alpha(v)\bar{v} =$

$\alpha(u)\alpha(v)^u a(u, v)\overline{uv}$. Equating coefficients of $\overline{uv}$, we obtain

$$b(u, v) = \alpha(u)\alpha(v)^u a(u, v)\alpha(uv)^{-1}.$$

Accordingly we say that a group extension of N by G determined by automorphisms $x \mapsto x^u$ and a factor set $a(u, v)$ is **equivalent**, or **isomorphic**, to a group extension of N by G determined by automorphisms $x \mapsto x^{u^*}$ and a factor set $b(u, v)$ if there is a function $\alpha : G \to N$ such that

$$x^{u^*} = \alpha(u)x^u\alpha(u)^{-1} \qquad \text{and} \qquad b(u, v) = \alpha(u)\alpha(v)^u a(u, v)\alpha(uv)^{-1}$$

for all u and v in G. It is immediate that equivalence of group extensions is an equivalence relation.

Proposition 7.37. Suppose that E_1 and E_2 are group extensions of N by G with respective inclusions $i_1 : N \to E_1$ and $i_2 : N \to E_2$ and with respective quotient homomorphisms $\varphi_1 : E_1 \to G$ and $\varphi_2 : E_2 \to G$. If there exists a group isomorphism $\Phi : E_1 \to E_2$ such that the two squares in Figure 7.4 commute, then the two group extensions are equivalent. Conversely if the two group extensions are equivalent, then there exists a group isomorphism $\Phi : E_1 \to E_2$ such that the two squares in Figure 7.4 commute.

$$\begin{array}{ccccc} N & \xrightarrow{i_1} & E_1 & \xrightarrow{\varphi_1} & G \\ \| & & \Phi\downarrow & & \| \\ N & \xrightarrow{i_2} & E_2 & \xrightarrow{\varphi_2} & G \end{array}$$

FIGURE 7.4. Equivalent group extensions.

REMARKS. The commutativity of the squares is important. Just because two group extensions of N by G are isomorphic as groups does not imply that they are equivalent group extensions. An example is given in Problem 19 at the end of the chapter.

PROOF. For the direct part, suppose that Φ exists. For each u in G, select $\bar{u}$ in E_1 with $\varphi_1(\bar{u}) = u$. Then we can form the extension data $\{x \mapsto x^u\}$ and $\{a(u, v)\}$ for E_1 relative to the normal subgroup $i_1(N)$ and the system $\{\bar{u} \mid u \in G\}$ of coset representatives. When reinterpreted in terms of N, E_1, and G, these data become $\{i_1^{-1}(x) \mapsto i_1^{-1}(x^u)\}$ and $\{i_1^{-1}(a(u, v))\}$.

Application of Φ to the coset $i_1(N)\bar{u}$ yields $i_2(N)\Phi(\bar{u})$ since $\Phi\, i_1 = i_2$, and $\Phi(\bar{u})$ is a member of E_2 with $\varphi_2(\Phi(\bar{u})) = \varphi_1(\bar{u}) = u$. Setting $\widetilde{u} = \Phi(\bar{u})$, we see that $\Phi(i_1(N)\bar{u})$ is the coset $i_2(N)\widetilde{u}$ of $i_2(N)$ in E_2. Thus we can determine

extension data for E_2 relative to $i_2(N)$ and the system $\{\widetilde{u} \mid u \in G\}$, and we can transform them by i_2^{-1} to obtain data relative to N, E_2, and G.

The claim is that the data relative to N, E_2, and G match those for N, E_1, and G. The automorphisms of N from E_2 are the maps $i_2^{-1}(x') \mapsto i_2^{-1}(x'^{u^*})$, where $x'^{u^*} = \widetilde{u}x'\widetilde{u}^{-1}$. From $i_2 = \Phi i_1$ and the fact that each of these maps is one-one, we obtain $i_2^{-1} = i_1^{-1}\Phi^{-1}$ on $i_2(N)$. Substitution shows that the automorphisms of N from E_2 are

$$\begin{aligned} i_1^{-1}(\Phi^{-1}(x')) \mapsto i_1^{-1}(\Phi^{-1}(x'^{u^*})) &= i_1^{-1}(\Phi^{-1}(\widetilde{u}x'\widetilde{u}^{-1})) \\ &= i_1^{-1}(\bar{u}\Phi^{-1}(x')\bar{u}^{-1}) = i_1^{-1}((\Phi^{-1}(x'))^u). \end{aligned}$$

If we set $x' = \Phi(x)$ with x in $i_1(N)$, then the automorphisms of N from E_2 take the form $i_1^{-1}(x) \mapsto i_1^{-1}(x^u)$. Thus they match the automorphisms of N from E_1.

In the case of the factor sets, we have $\bar{u}\bar{v} = a(u, v)\overline{uv}$. Application of Φ gives $\widetilde{u}\,\widetilde{v} = \Phi(a(u, v))\widetilde{uv}$. Thus the factor set for E_2 relative to N is $\{i_2^{-1}\Phi(a(u, v))\}$. Since $i_2^{-1}\Phi = i_1^{-1}$, this matches the factor set for E_1 relative to N.

We turn to the converse part. Suppose that the multiplication law in E_1 is $(i_1(x)\bar{u})(i_1(y)\bar{v}) = i_1(x)i_1(y)^u i_1(a(u, v))\overline{uv}$ for x and y in N, and that the multiplication law in E_2 is $(i_2(x)\widetilde{u})(i_2(y)\widetilde{v}) = i_2(x)i_2(y)^{u^*} i_2(b(u, v))\widetilde{uv}$. Here $\bar{u}$ and $\bar{v}$ are preimages of u and v under φ_1, and $\widetilde{u}$ and $\widetilde{v}$ are preimages of u and v under φ_2. Define automorphisms of N by $x^u = i_1^{-1}(i_1(x)^u)$ and $x^{u^*} = i_2^{-1}(i_2(x)^{u^*})$. We can then rewrite the multiplication laws as

$$(i_1(x)\bar{u})(i_1(y)\bar{v}) = i_1(xy^u a(u, v))\overline{uv}$$

and

$$(i_2(x)\widetilde{u})(i_2(y)\widetilde{v}) = i_2(xy^{u^*} b(u, v))\widetilde{uv}.$$

The assumption that E_1 is equivalent to E_2 as an extension of N by G means that there exists a function $\alpha : G \to N$ such that

$$x^{u^*} = \alpha(u)x^u\alpha(u)^{-1} \qquad \text{and} \qquad b(u, v) = \alpha(u)\alpha(v)^u a(u, v)\alpha(uv)^{-1}$$

for all u and v in G. Define $\Phi : E_1 \to E_2$ by

$$\Phi(i_1(x)\bar{u}) = i_2(x\alpha(u)^{-1})\widetilde{u}.$$

Certainly Φ is one-one onto. It remains to check that Φ is a group homomorphism and that the squares commute in Figure 7.4.

To check that $\Phi : E_1 \to E_2$ is a group homomorphism, we compare

$$\Phi(i_1(x)\bar{u}i_1(y)\bar{v}) = \Phi(i_1(xy^u a(u, v))\overline{uv} = i_2(xy^u a(u, v)\alpha(uv)^{-1})\widetilde{uv}$$

with the product

$$\begin{aligned}\Phi(i_1(x)\bar{u})\Phi(i_1(y)\bar{v}) &= i_2(x\alpha(u)^{-1})\widetilde{u}i_2(y\alpha(v)^{-1})\widetilde{v}\\ &= i_2(x\alpha(u)^{-1}(y\alpha(v)^{-1})^{u^*}b(u,v))\widetilde{uv}.\end{aligned}$$

Since

$$\begin{aligned}\alpha(u)^{-1}(y\alpha(v)^{-1})^{u^*}b(u,v) &= \alpha(u)^{-1}(y\alpha(v)^{-1})^{u^*}\alpha(u)\alpha(v)^u a(u,v)\alpha(uv)^{-1}\\ &= (y\alpha(v)^{-1})^u\alpha(v)^u a(u,v)\alpha(uv)^{-1}\\ &= y^u a(u,v)\alpha(uv)^{-1},\end{aligned}$$

these expressions are equal, and Φ is a group homomorphism. Thus Φ is a group isomorphism.

Now we check the commutativity of the squares. The computation

$$\varphi_2\Phi(i_1(x)\bar{u}) = \varphi_2(i_2(x\alpha(u)^{-1})\widetilde{u}) = u = \varphi_1(i_1(x)\bar{u})$$

shows that the right-hand square commutes.

For the left-hand square we use the fact recorded in the statement of Proposition 7.36 that $i_1(a(1,1)^{-1})\bar{1}$ is the identity of E_1 and $i_2(b(1,1)^{-1})\widetilde{1}$ is the identity of E_2. Therefore $\Phi\, i_1(x) = \Phi(i_1(xa(1,1)^{-1})\bar{1}) = i_2(xa(1,1)^{-1}\alpha(1)^{-1})\widetilde{1}$. Since $i_2(x) = xb(1,1)^{-1}\widetilde{1}$, the left-hand square commutes if $b(1,1) = \alpha(1)a(1,1)$. This formula follows from $(*)$ in the proof of Proposition 7.36 by the computation

$$b(1,1) = \alpha(1)\alpha(1)^1 a(1,1)\alpha(1)^{-1} = \alpha(1)a(1,1)\alpha(1)\alpha(1)^{-1} = \alpha(1)a(1,1),$$

and thus the left-hand square indeed commutes. □

For the remainder of this section, *let us assume that N is abelian.* In this case Proposition 7.36a reduces to the identity $(x^v)^u = x^{uv}$ for all u and v in G independently of the choice of representatives, just as it does in the example we studied with $N = C_4$ and $G = C_2$. In the terminology of Section IV.7, G acts on N by automorphisms.[5] Suppose we fix such an action $\tau : G \to \operatorname{Aut} N$ by automorphisms and consider all extensions of N by G built from τ. In our example we are thus to consider E equal to $C_4 \times C_2$ or C_8, which are built with the trivial τ, or else E equal to D_4 or H_8, which are built with the nontrivial τ (in which the nontrivial element of G acts by the nontrivial automorphism of N).

Since N is abelian, let us switch to additive notation for N and to ordinary function notation for $\tau(w)$, rewriting the formula of Proposition 7.36b as

$$\tau(u)a(v,w) + a(u,vw) = a(u,v) + a(uv,w).$$

[5]The formula $(x^v)^u = x^{uv}$ correctly corresponds to a group action with the group on the left as in Section IV.7.

This condition is preserved under addition of factor sets as long as τ does not change, it is satisfied by the 0 factor set, and the negative of a factor set is again a factor set. Therefore the factor sets for this τ form an abelian group.

Two factor sets for this τ are equivalent (in the sense of yielding equivalent group extensions) if and only if their difference is equivalent to 0, and $a(u, v)$ is equivalent to 0 if and only if

$$a(u, v) = \alpha(uv) - \alpha(u) - \tau(u)\alpha(v)$$

for some function $\alpha : G \to N$. The set of factor sets for this τ that are equivalent to 0 is thus a subgroup,[6] and we arrive at the following result.

Proposition 7.38. Let G and N be groups with N abelian, and suppose that $\tau : G \to \operatorname{Aut} N$ is a homomorphism. Then the set of equivalence classes of group extensions of N by G corresponding to the action $\tau : G \to \operatorname{Aut} N$ is parametrized by the quotient of the abelian group of factor sets by the subgroup of factor sets equivalent to 0.

The extension E corresponding to the 0 factor set is of special interest. In this case the multiplication law for the coset representatives is $\bar{u}\bar{v} = \overline{uv}$ since the member $a(u, v) = 0$ of N is to be interpreted multiplicatively in this product formula. Consequently the map $u \mapsto \bar{u}$ of G into E is a group homomorphism, necessarily one-one, and we can regard G as a subgroup of E. Proposition 4.44 allows us to conclude that E is the semidirect product $G \times_\tau N$. The multiplication law for general elements of E, with multiplicative notation used for N, is

$$(x\bar{u})(y\bar{v}) = x(\tau(u)y)\overline{uv}.$$

It is possible also to describe explicitly the extension one obtains from the sum of two factor sets corresponding to the same τ, but we leave this matter to Problems 20–23 at the end of the chapter. The operation on extensions that corresponds to addition of factor sets in this way is called **Baer multiplication**. What we saw in the previous paragraph says that the group identity under Baer multiplication is the semidirect product.

The two conditions, the compatibility condition on a factor set given in Proposition 7.36b and the condition with α in it for equivalence to 0, are of a combinatorial type that occurs in many contexts in mathematics and is captured by the ideas of "homology" and "cohomology." For the current situation the notion is that of **cohomology of groups**, and we shall define it now. The subject of homological

[6]One can legitimately ask whether an arbitrary $\alpha : G \to N$ leads to a factor set under the definition $a(u, v) = \alpha(uv) - \tau(v)\alpha(u) - \alpha(v)$, and one easily checks that the answer is yes. Alternatively, one can refer to the case $n = 2$ in the upcoming Proposition 7.39.

algebra, which is developed in Chapter IV of *Advanced Algebra*, puts cohomology of groups in a wider context and explains some of its mystery.

We fix an abelian group N, a group G, and a group action τ of G on N by automorphisms. It is customary to suppress τ in the notation for the group action, and we shall follow that convention. For integers $n \geq 0$, one begins with the abelian group $C^n(G, N)$ of n-**cochains** of G with coefficients in N. This is defined by

$$C^n(G, N) = \begin{cases} N & \text{if } n = 0, \\ \{f : \prod_{k=1}^n G \to N\} & \text{if } n > 0. \end{cases}$$

In words, $C^n(G, N)$ is the set of all functions into N from the n-fold direct product of G with itself. The **coboundary map** $\delta_n : C^n(G, N) \to C^{n+1}(G, N)$ is the homomorphism of abelian groups defined by

$$(\delta_0 f)(g_1) = g_1 f - f$$

and by

$$\begin{aligned}(\delta_n f)(g_1, \ldots, g_{n+1}) = {} & g_1(f(g_2, \ldots, g_{n+1})) \\ & + \sum_{i=1}^{n} (-1)^i f(g_1, \ldots, g_{i-1}, g_i g_{i+1}, g_{i+2}, \ldots, g_{n+1}) \\ & + (-1)^{n+1} f(g_1, \ldots, g_n)\end{aligned}$$

for $n > 0$. We postpone to the end of this section the proof of the following result.

Proposition 7.39. $\delta_n \delta_{n-1} = 0$ for all $n \geq 1$.

It follows from Proposition 7.39 that image $\delta_{n-1} \subseteq \ker \delta_n$ for all $n \geq 1$. Thus if we define abelian groups by

$$Z^n(G, N) = \ker \delta_n,$$

$$B^n(G, N) = \begin{cases} 0 & \text{for } n = 0, \\ \text{image } \delta_{n-1} & \text{for } n > 0, \end{cases}$$

then $B^n(G, N) \subseteq Z^n(G, N)$ for all n, and it makes sense to define the abelian groups

$$H^n(G, N) = Z^n(G, N)/B^n(G, N) \qquad \text{for } n \geq 0.$$

The elements of $Z^n(G, N)$ are called n-**cocycles**, the elements of $B^n(G, N)$ are called n-**coboundaries**, and $H^n(G, N)$ is called the n^{th} **cohomology group** of G with coefficients in N.

EXAMPLES IN LOW DEGREE.

DEGREE 0. Here $(\delta_0 f)(u) = uf - f$ with f in N and u in G. The cocycle condition is that this is 0 for all u. Thus f is to be fixed by G. We say that an f

fixed by G is an **invariant** of the group action. The space of invariants is denoted by N^G. By convention above, we are taking $B^0(G, N) = 0$. Thus

$$H^0(G, N) = N^G.$$

DEGREE 1. Here $(\delta_1 f)(u, v) = u(f(v)) - f(uv) + f(u)$ with f a function from G to N. The cocycle condition is that

$$f(uv) = f(u) + u(f(v)) \qquad \text{for all } u, v \in G.$$

A function f satisfying this condition is called a **crossed homomorphism** of G into N. A coboundary is a function $f : G \to N$ of the form $f(u) = (\delta_0 x)(u) = ux - x$ for some $x \in N$. Then $H^1(G, N)$ is the quotient of the group of crossed homomorphisms by this subgroup. In the special case that the action of G on N is trivial, the crossed homomorphisms reduce to ordinary homomorphisms of G into N, and every coboundary is 0. Thus $H^1(G, N)$ is the group of homomorphisms of G into N if G acts trivially on N.

DEGREE 2. Here f is a function from $G \times G$ into N, and

$$(\delta_2 f)(u, v, w) = u(f(v, w)) - f(uv, w) + f(u, vw) - f(u, v).$$

The cocycle condition is that

$$u(f(v, w)) + f(u, vw) = f(uv, w) + f(u, v) \qquad \text{for all } u, v, w \in G.$$

This is the same as the condition that $\{f(u, v)\}$ be a factor set for extensions of N by G relative to the given action of G on N by automorphisms. A coboundary is a function $f : G \times G \to N$ of the form

$$f(u, v) = (\delta_0 \alpha)(u, v) = u(\alpha(v)) - \alpha(uv) + \alpha(u) \qquad \text{for some } \alpha : G \to N.$$

This is the same as the condition that $\{-f(u, v)\}$ be a factor set equivalent to 0. Thus we can restate Proposition 7.38 as follows.

Proposition 7.40. Let G and N be groups with N abelian, and suppose that $\tau : G \to \operatorname{Aut} N$ is a homomorphism. Then the set of equivalence classes of group extensions of N by G corresponding to the action $\tau : G \to \operatorname{Aut} N$ is parametrized by $H^2(G, N)$.

Since group extensions have such a nice interpretation in terms of cohomology groups H^2, it is reasonable to look for a nice interpretation for H^1 as well. Indeed, H^1 has an interpretation in terms of uniqueness up to inner isomorphisms for semidirect-product decompositions. We continue with the abelian group N, a group G, and a group action τ of G on N by automorphisms. A semidirect product $E = G \times_\tau N$ is an allowable extension. Since G embeds as a subgroup of E, we are given a one-one group homomorphism $u \mapsto \bar{u}$ of G into E. The construction at the beginning of this section works with the set $\bar{u}$ of coset representatives, and they have $\bar{u}\bar{v} = \overline{uv}$.

Suppose that the semidirect product can be formed by a second one-one group homomorphism $u \mapsto \widetilde{u}$ of G into E. If we write $\widetilde{u} = \alpha(u)\bar{u}$ for a function $\alpha : G \to N$, then we know from earlier in the section that the extensions formed from $\{\bar{u}\}$ and from $\{\widetilde{u}\}$ are equivalent. Because G maps homomorphically into E for both systems, the factor sets are 0 in both cases. Consequently the function α must satisfy

$$\alpha(uv) - \alpha(u) - \tau(u)\alpha(v) = 0.$$

This is exactly the condition that $\alpha : G \to N$ be a 1-cocyle. Thus the group $Z^1(G, N)$ parametrizes all ways that we can embed G as a complementary subgroup to N in the semidirect product $E = G \times_\tau N$.

A relatively trivial way to construct a one-one group homomorphism $u \mapsto \widetilde{u}$ from $u \mapsto \bar{u}$ is to form, in the usual multiplicative notation, $\widetilde{u} = x_0^{-1}\bar{u}x_0$ for some $x_0 \in N$. Then $\widetilde{u} = x_0^{-1}\bar{u}x_0\bar{1} = x_0^{-1}(\tau(u)(x_0))\bar{u}$, and the additive notation for $\alpha(u)$ has $\alpha(u) = \tau(u)(x_0) - x_0$. Referring to our earlier computations in degree 1, we see that α is in the group $B^1(G, N)$ of coboundaries.

The conclusion is that $H^1(G, N)$ parametrizes all ways, modulo relatively trivial ways, that we can embed G as a complementary subgroup to N in the semidirect product $E = G \times_\tau N$.

As promised, we now return to the proof of Proposition 7.39.

PROOF OF PROPOSITION 7.39. For $n = 1$, we have

$$\begin{aligned}(\delta_1\delta_0 f)(u, v) &= u((\delta_0 f)(v)) - (\delta_0 f)(uv) + (\delta_0 f)(u)\\ &= u(vf - f) - (uvf - f) + (uf - f) = 0.\end{aligned}$$

For $n > 1$, we begin with

$$\begin{aligned}(\delta_n\delta_{n-1}f)(g_1, \ldots, g_{n+1}) &= g_1((\delta_{n-1}f)(g_2, \ldots, g_{n+1}))\\ &\quad + \sum_{i=1}^{n}(-1)^i(\delta_{n-1}f)(g_1, \ldots, g_ig_{i+1}, \ldots, g_{n+1})\\ &\quad + (-1)^{n+1}(\delta_{n-1}f)(g_1, \ldots, g_n)\\ &= \mathrm{I} + \mathrm{II} + \mathrm{III}.\end{aligned}$$

Here

$$\begin{aligned}\mathrm{I} &= g_1g_2(f(g_3, \ldots, g_{n+1})) + \sum_{i=2}^{n}(-1)^{i-1}g_1(f(g_2, \ldots, g_ig_{i+1}, \ldots, g_{n+1}))\\ &\quad + (-1)^n g_1(f(g_2, \ldots, g_n)) = \mathrm{IA} + \mathrm{IB} + \mathrm{IC},\end{aligned}$$

$$\begin{aligned}\mathrm{II} &= -(\delta_{n-1}f)(g_1g_2, g_3, \ldots, g_n) + \sum_{i=2}^{n}(-1)^i(\delta_{n-1}f)(g_1, \ldots, g_ig_{i+1}, \ldots, g_{n+1})\\ &= \mathrm{IIA} + \mathrm{IIB},\end{aligned}$$

$$\begin{aligned}\text{III} &= (-1)^{n+1} g_1(f(g_2,\ldots,g_n)) + (-1)^{n+1}(-1) f(g_1g_2, g_3,\ldots,g_n)\\ &\quad + (-1)^{n+1}\sum_{i=2}^{n-1}(-1)^i f(g_1,\ldots,g_ig_{i+1},\ldots,g_n)\\ &\quad + (-1)^{n+1}(-1)^n f(g_1,\ldots,g_{n-1})\\ &= \text{IIIA} + \text{IIIB} + \text{IIIC} + \text{IIID}.\end{aligned}$$

Terms IIA and IIB decompose further as

$$\begin{aligned}\text{IIA} &= -g_1g_2(f(g_3,\ldots,g_{n+1})) + f(g_1g_2g_3, g_4,\ldots,g_{n+1})\\ &\quad - \sum_{i=3}^{n}(-1)^{i+1} f(g_1g_2,\ldots,g_ig_{i+1},\ldots,g_{n+1}) - (-1)^n f(g_1g_2, g_3,\ldots,g_n)\\ &= \text{IIAa} + \text{IIAb} + \text{IIAc} + \text{IIAd},\end{aligned}$$

$$\begin{aligned}\text{IIB} &= \sum_{i=2}^{n}(-1)^i g_1(f(g_2,\ldots,g_ig_{i+1},\ldots,g_{n+1}))\\ &\quad + (-1)^2(-1) f(g_1g_2g_3, g_4,\ldots,g_{n+1})\\ &\quad + \sum_{i=3}^{n}(-1)^i(-1) f(g_1g_2,\ldots,g_ig_{i+1},\ldots,g_{n+1})\\ &\quad + \sum_{i=2}^{n}(-1)^i\sum_{j=2}^{i-2}(-1)^j f(g_1,\ldots,g_jg_{j+1},\ldots,g_ig_{i+1},\ldots,g_{n+1})\\ &\quad + \sum_{i=3}^{n}(-1)^i(-1)^{i-1} f(g_1,\ldots,g_{i-1}g_ig_{i+1},\ldots,g_{n+1})\\ &\quad + \sum_{i=2}^{n-1}(-1)^i(-1)^i f(g_1,\ldots,g_ig_{i+1}g_{i+2},\ldots,g_{n+1})\\ &\quad + \sum_{i=2}^{n-2}(-1)^i\sum_{j=i+2}^{n}(-1)^{j-1} f(g_1,\ldots,g_ig_{i+1},\ldots,g_jg_{j+1},\ldots,g_{n+1})\\ &\quad + \sum_{i=2}^{n-1}(-1)^i(-1)^n f(g_1,\ldots,g_ig_{i+1},\ldots,g_n)\\ &\quad + (-1)^n(-1)^n f(g_1,\ldots,g_{n-1})\\ &= \text{IIBa} + \text{IIBb} + \text{IIBc} + \text{IIBd} + \text{IIBe} + \text{IIBf} + \text{IIBg} + \text{IIBh} + \text{IIBi}.\end{aligned}$$

Inspection shows that we have cancellation between term IA and term IIAa, term IB and term IIBa, term IC and term IIIA, term IIAb and term IIBb, term IIAc and term IIBc, term IIAd and term IIIB, term IIBd and term IIBg, term IIBe and term IIBf, term IIBh and term IIIC, and term IIBi and term IIID. All the terms cancel, and we conclude that $\delta_n\delta_{n-1}f = 0$. □

7. Problems

1. Using Burnside's Theorem and Problem 34 at the end of Chapter IV, show that 60 is the smallest possible order of a nonabelian simple group.
2. A **commutator** in a group is any element of the form $xyx^{-1}y^{-1}$.
   (a) Prove that the inverse of a commutator is a commutator.
   (b) Prove that any conjugate of a commutator is a commutator.
3. Let a and b be elements of a group G. Prove that the subgroup generated by a and b is the same as the subgroup generated by bab^2 and bab^3.
4. A subgroup H of a group G is said to be **characteristic** if it is carried into itself by every automorphism of G.
   (a) Prove that characteristic implies normal.
   (b) Prove that the center Z_G of G is a characteristic subgroup.
   (c) Prove that the commutator subgroup G' of G is a characteristic subgroup.
5. In the terminology of the previous problem, which subgroups of the quaternion subgroup H_8 are characteristic?
6. Is every finite group finitely presented? Why or why not?
7. Let $G = \mathrm{SL}(2, \mathbb{R})$, and let G' be the commutator subgroup.
   (a) Prove that every element $\begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$ is in G'.
   (b) Prove that $G' = G$.
   (c) Prove that $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ is not a commutator even though it is in G'.
8. Problem 53 at the end of Chapter IV produced a group G of order 27 generated by two elements a and b satisfying $a^9 = b^3 = b^{-1}aba^{-4} = 1$. Prove that G is given by generators and relations as
$$G = \langle a, b;\ a^9, b^3, b^{-1}aba^{-4} \rangle.$$
9. Let G_n be given by generators and a single relation as
$$G_n = \langle x_1, y_1, \ldots, x_n, y_n;\ x_1 y_1 x_1^{-1} y_1^{-1} \cdots x_n y_n x_n^{-1} y_n^{-1} \rangle.$$
   Prove that G_n / G_n' is free abelian of rank $2n$, and conclude that the groups G_n are mutually nonisomorphic as n varies. (Educational note related to topology: The group G_n may be shown to be the fundamental group of a compact orientable 2-dimensional manifold without boundary and with n handles.)
10. Prove that a free group of finite rank n cannot be generated by fewer than n elements.

11. Let F be the free group on generators a, b, c, and let H be the subgroup generated by all words of length 2.
    (a) Find coset representatives g such that G is the disjoint union of the cosets Hg.
    (b) Find a free basis of H.
12. For the free group on generators x and y, prove that the elements y, xyx^{-1}, x^2yx^{-2}, x^3yx^{-3}, ... , constitute a free basis of the subgroup that they generate. Conclude that a free group of rank 2 has a free subgroup of infinite rank.
13. Let $G = C_2 * C_2$. Prove that the only quotient groups of G, up to isomorphism, are G itself, $\{1\}$, C_2, $C_2 \times C_2$, and the dihedral groups D_n for $n \geq 3$.
14. Prove that if every irreducible finite-dimensional representation of a finite group G is 1-dimensional, then G is abelian.
15. Let G be a finitely generated group, and let H be a subgroup of finite index. Prove that H is finitely generated.
16. Let N be an abelian group, let G be a group, let τ be an action of G on N by automorphisms, and let $n > 0$ be an integer.
    (a) Prove that if every element of N has finite order dividing an integer m, then every member of $H^n(G, N)$ has finite order dividing m.
    (b) Suppose that G is finite and that f is an n-cocycle. Define an $(n-1)$-cochain F by
    $$F(g_1, \ldots, g_{n-1}) = \sum_{g \in G} f(g_1, \ldots, g_{n-1}, g).$$
    By summing the cocycle condition for f over the last variable, express $|G| f(g_1, \ldots, g_n)$ in terms of F, and deduce that $|G| f$ is a coboundary. Conclude that every member of $H^n(G, N)$ has order dividing $|G|$.
17. Let G be a finite group. Suppose that G has a normal abelian subgroup N, and suppose that $\mathrm{GCD}(|N|, |G/N|) = 1$. Prove that there exists a subgroup H of G such that G is the semidirect product of H and N.
18. Let N be the cyclic group C_2, and let G be an arbitrary group of order 4. Identify up to equivalence all group extensions of N by G.
19. Let $N = C_2$, and let $E = \bigoplus_{n=1}^{\infty} (C_2 \oplus C_4)$. Regard E as an extension of N in two ways—first by embedding N as one of the summands C_2 of E and then by embedding N as a subgroup of one of the summands C_4 of E. Show that the quotient groups E/N in the two cases are isomorphic, that E/N acts trivially on N in both cases, and that the two group extensions are not equivalent.

Problems 20–23 concern **Baer multiplication** of extensions. Let N be an abelian group, let G be a group, let τ be an action of G on N by automorphisms, and let E_1 and E_2 be two extensions of N by G relative to τ. Write $\varphi_1 : E_1 \to G$ and $\varphi_2 : E_2 \to G$ for the quotient mappings. Let (E, E') denote the subgroup of all

members (e_1, e_2) of $E_1 \times E_2$ for which $\varphi_1(e_1) = \varphi_2(e_2)$. Writing the operation in N multiplicatively, let $Q = \{(x, x^{-1}) \in E_1 \times E_2 \mid x \in N\}$. The Baer product of E_1 and E_2 is defined to be the quotient $(E_1, E_2)/Q$. A typical coset of the Baer product will be denoted by $(e_1, e_2)Q$.

20. Prove that the homomorphism $x \mapsto (x, 1)Q$ is one-one from N into $(E_1, E_2)/Q$, that the homomorphism $\varphi : (E_1, E_2) \to G$ defined by $\varphi(e_1, e_2) = \varphi_1(e_1)$ has image G and descends to the quotient $(E_1, E_2)/Q$, and that the kernel of the descended φ is the embedded copy of N. (Therefore $(E_1, E_2)/Q$ is an extension of N by G, evidently relative to τ.)

21. For each $u \in G$, select $\bar{u} \in E_1$ and $\widetilde{u} \in E_2$ with $\varphi_1(\bar{u}) = u = \varphi_2(\widetilde{u})$, and define $a(u, v)$ and $b(u, v)$ for u and v in G by $(x\bar{u})(y\bar{v}) = a(u, v)\overline{uv}$ and $(x\widetilde{u})(y\widetilde{v}) = b(u, v)\widetilde{b}(u, v)$. Show that $(\bar{u}, \widetilde{u})Q$ has $\varphi((\bar{u}, \widetilde{u})Q) = u$ and that the associated 2-cocyle for $(E_1, E_2)/Q$ is $a(u, v)b(u, v)$ if the group operation in N is written multiplicatively.

22. Prove that Baer multiplication descends to a well-defined multiplication of equivalence classes of extensions of N by G relative to τ, in the following sense: Suppose that E_1 and E_1' are equivalent extensions and that E_2 and E_2' are equivalent extensions. Let $(E_1, E_2)/Q$ and $(E_1', E_2')/Q'$ be the Baer products. Then $(E_1, E_2)/Q$ is equivalent to $(E_1', E_2')/Q'$. Conclude that if Baer multiplication is imposed on equivalence classes of extensions of N by G relative to τ, then the correspondence stated in Proposition 7.40 of equivalence classes to members of $H^2(G, N)$ is a group isomorphism.

Problems 23–24 derive the Poisson summation formula for finite abelian groups. If G is a finite abelian group and $\widehat{G}$ is its group of multiplicative characters, then the **Fourier coefficient** at $\chi \in \widehat{G}$ of a function f in $C(G, \mathbb{C})$ is $\widehat{f}(\chi) = \sum_{g \in G} f(g)\overline{\chi(g)}$. The Fourier inversion formula in Theorem 7.17 says that $f(g) = |G|^{-1} \sum_{\chi \in \widehat{G}} \widehat{f}(\chi)\chi(g)$.

23. Let G be a finite abelian group, let H be a subgroup, and let G/H be the quotient group. If t is in G, write $\dot{t}$ for the coset of t in G/H. Let f be in $C(G, \mathbb{C})$ and define $F(\dot{t}) = \sum_{h \in H} f(t + h)$ as a function on G/H. Suppose that χ is a member of $\widehat{G}$ that is identically 1 on H, so that χ descends to a member $\dot{\chi}$ of $\widehat{G/H}$. Prove that $\widehat{f}(\chi) = \widehat{F}(\dot{\chi})$.

24. **(Poisson summation formula)** With f and F as in the previous problem, apply the Fourier inversion formula for G/H to the function F, and derive the formula

$$\sum_{h \in H} f(t + h) = \frac{1}{|G/H|} \sum_{\omega \in \widehat{G},\ \omega|_H = 1} \widehat{f}(\omega)\omega(t).$$

(Educational note: This formula is often applied with $t = 0$, in which case it reduces to $\sum_{h \in H} f(h) = \frac{1}{|G/H|} \sum_{\omega \in \widehat{G},\ \omega|_H = 1} \widehat{f}(\omega)$.)

Problems 25–28 continue the introduction to error-correcting codes begun in Problems 63–73 at the end of Chapter IV, combining those results with the Poisson summation formula in the problems above and with notions from Section VI.1. Let $\mathbb{F}$ be the field $\mathbb{Z}/2\mathbb{Z}$, and form the Hamming space $\mathbb{F}^n$. Define a nondegenerate bilinear form on $\mathbb{F}^n$ by $(a, c) = \sum_{i=1}^n a_i c_i$ for a and c in $\mathbb{F}^n$. Recall from Chapter IV that a linear code C is a vector subspace of $\mathbb{F}^n$. For such a C, let $C^\perp$ as in Section VI.1 be the set of all $a \in \mathbb{F}^n$ such that $(a, c) = 0$ for all $c \in C$; the linear code $C^\perp$ is called the **dual code**. A linear code is **self dual** if $C^\perp = C$.

25. (a) Show that the codes 0 and $\mathbb{F}^n$ are dual to each other.
    (b) Show that the repetition code and the parity-check code are dual to each other.
    (c) Show that the Hamming code of order 8 is self dual.
    (d) Show that any self-dual linear code C has $\dim C = n/2$, and conclude that the Hamming code of order 2^r with $r > 3$ is not self dual.
    (e) Show that any member c of a self-dual linear code C has even weight.
    (f) Show that if a linear code C has $C \subseteq C^\perp$ and if every member c of C has even weight, then $c \mapsto \frac{1}{2}\mathrm{wt}(c) \bmod 2$ is a group homomorphism of C into $\mathbb{Z}/2\mathbb{Z}$. Here $\mathrm{wt}(c)$ denotes the weight of c.

26. Regard $\mathbb{F}^n$ as an additive group G to which the Fourier inversion formula of Section 4 can be applied.
    (a) Show that one can map $\widehat{G}$ to $\mathbb{F}^n$ by $\chi \mapsto a_\chi$ with $\chi(c) = (-1)^{(a_\chi, c)}$ and that the result is a group isomorphism. (Therefore if f is in $C(\mathbb{F}^n, \mathbb{C})$, we can henceforth regard $\widehat{f}$ as a function on $\mathbb{F}^n$.)
    (b) Show under the identification in (a) that if f is in $C(\mathbb{F}^n, \mathbb{C})$, then $\widehat{f}(a) = \sum_{c \in \mathbb{F}^n} f(c)(-1)^{(a,c)}$ for a in $\mathbb{F}^n$.
    (c) Suppose that the function $f \in C(\mathbb{F}^n, \mathbb{C})$ is of the special form $f(c) = \prod_{i=1}^n f_i(c_i)$ whenever $c = (c_1, \dots, c_n)$. Here each f_i is a function on the 2-element group $\mathbb{F}$. Prove that $\widehat{f}(a) = \prod_{i=1}^n \widehat{f_i}(a_i)$ whenever $a = (a_1, \dots, a_n)$. Here $\widehat{f_i}$ is given by the formula of (b) for the case $n = 1$: $\widehat{f_i}(a_i) = \sum_{c_i \in \mathbb{F}} f_i(c_i)(-1)^{a_i c_i}$.

27. Fix two complex numbers x and y. Define $f_0 : \mathbb{F} \to \mathbb{C}$ to be the function with $f_0(0) = x$ and $f_0(1) = y$. Define $f : \mathbb{F} \to \mathbb{C}$ to be the function with $f(c) = \prod_{i=1}^n f_0(c_i) = x^{n - \mathrm{wt}(c)} y^{\mathrm{wt}(c)}$ where $\mathrm{wt}(c)$ is the weight of c.
    (a) Show that $\widehat{f_0}(0) = x + y$ and $\widehat{f_0}(1) = x - y$.
    (b) Show that $\widehat{f}(a) = (x + y)^{n - \mathrm{wt}(a)}(x - y)^{\mathrm{wt}(a)}$.

28. Let C be a linear code in $\mathbb{F}^n$. Take G to be the additive group of $\mathbb{F}^n$ and H to be the additive group of C. Regard $C^\perp$ as an additive group also.
    (a) Map $\widehat{G/H}$ to $C^\perp$ by $\chi \mapsto a_\chi$ with $\chi(c) = (-1)^{(a_\chi, c)}$. Show that this mapping is a group isomorphism.

(b) Applying the Poisson summation formula of Problem 24, prove that

$$\sum_{h\in C} f(h) = \frac{1}{|C^\perp|} \sum_{a\in C^\perp} \widehat{f}(a)$$

for all f in $C(\mathbb{F}^n, \mathbb{C})$.

(c) **(MacWilliams identity)** Let $W_C(X, Y) = \sum_{k=0}^{n} N_k(C) X^{n-k} Y^k$, where $N_k(C)$ is the number of members of C with weight k, be the weight-enumerator polynomial of C, and let $W_{C^\perp}(X, Y)$ be defined similarly. By applying (b) to the function f in the previous problem, prove that $W_C(x, y) = |C^\perp|^{-1} W_{C^\perp}(x + y, x - y)$ for each x and y. Conclude from Corollary 4.32 that weight-enumerator polynomials satisfy $W_C(X, Y) = |C^\perp|^{-1} W_{C^\perp}(X + Y, X - Y)$.

(d) The polynomials $W_C(X, Y)$ were seen in Chapter IV to be X^n for the 0 code, $(X + Y)^n$ for the code $\mathbb{F}^n$, $X^n + Y^n$ for the repetition code, $\frac{1}{2}((X+Y)^n+(X-Y)^n)$ for the parity-check code, and $X^8+14X^4Y^4+Y^8$ for the Hamming code of order 8. Using relationships established in Problem 25, verify the result of (c) for each of these codes.

(e) Suppose that C is a self-dual linear code. Applying (c) in this case, exhibit $W_C(X, Y)$ as being invariant under a copy of the dihedral group D_8 of order 16. (Educational note: If the polynomial $W_C(X, Y)$ is invariant also under $X \mapsto iX$, as is true for the Hamming code of order 8, then $W_C(X, Y)$ is invariant under the group generated by D_8 and this transformation, which can be shown to have order 192.)

Problems 29–31 concern an unexpectedly fast method of computation of Fourier coefficients in the context of finite abelian groups, particularly in the context of cyclic groups. They show for a cyclic group of order $m = pq$ that the use of the idea behind the Poisson summation formula of Problem 24 makes it possible to compute the Fourier coefficients of a function in about $pq(p+q)$ steps rather than the expected $m^2 = p^2q^2$ steps. This savings may be iterated in the case of a cyclic group of order 2^n so that the Fourier coefficients are computed in about $n2^n$ steps rather than the expected 2^{2n} steps. An organized algorithm to implement this method of computation is known as the **fast Fourier transform**. Write the cyclic group C_m as the set $\{0, 1, 2, \ldots, m-1\}$ of integers modulo m under addition, and let $\zeta_m = e^{2\pi i/m}$. For k in C_m define a multiplicative character χ_n of C_m by $\chi_n(k) = (\zeta_m^n)^k$. The resulting m multiplicative characters satisfy $\chi_n\chi_{n'} = \chi_{n+n'}$, and they exhaust $\widehat{C_m}$ since distinct multiplicative characters are orthogonal. It will be convenient to identify χ_n with $\chi_n(1) = \zeta_m^n$.

29. In the setting of Problem 23, suppose that $G = C_m$ with $m = pq$; here p and q need not be relatively prime. Let $H = \{0, q, 2q, \ldots, (p-1)q\}$ be the subgroup of G isomorphic to C_p, so that $G/H = \{0, 1, 2, \ldots, q-1\}$ is isomorphic to C_q. Prove that the characters χ of G identified with $\zeta_m^0, \zeta_m^p, \zeta_m^{2p}, \ldots, \zeta_m^{(q-1)p}$

are the ones that are identically 1 on H and therefore descend to characters of G/H. Verify that the descended characters $\dot{\chi}$ are the ones identified with $\zeta_q^0, \zeta_q^1, \zeta_q^2, \ldots, \zeta_q^{q-1}$. Consequently the formula $\widehat{f}(\chi) = \widehat{F}(\dot{\chi})$ of Problem 23 provides a way of computing $\widehat{f}$ at $\zeta_m^0, \zeta_m^p, \zeta_m^{2p}, \ldots, \zeta_m^{(q-1)p}$ from the values of $\widehat{F}$. Show that if $\widehat{F}$ is computed from the definition of Fourier coefficients, then the number of steps involved in its computation is about q^2, apart from a constant factor. Show therefore that the total number of steps in computing $\widehat{f}$ at these special values of χ is therefore on the order of $q^2 + pq$.

30. In the previous problem show for each k with $0 \le k \le p-1$ that the value of $\widehat{f}$ at $\zeta_m^k, \zeta_m^{p+k}, \zeta_m^{2p+k}, \ldots, \zeta_m^{(q-1)p+k}$ can be handled in the same way with a different F by replacing f by a suitable variant of f. Doing so for each k requires p times the number of steps detected in the previous problem, and therefore all of $\widehat{f}$ can be computed in about $p(q^2 + pq) = pq(p + q)$ steps.

31. Show how iteration of this process to compute the Fourier coefficients of each F, together with further iteration of this process, allows one to compute the Fourier coefficients for a function on $C_{m_1 m_2 \cdots m_r}$ in about $m_1 m_2 \cdots m_r (m_1 + m_2 + \cdots + m_r)$ steps.

Problems 32–36 concern contragredient representations and the decomposition of the left regular representation of a finite group G. They make use of Problems 24–28 in Chapter III, which introduce the complex conjugate $\overline{V}$ of a complex vector space V. In the case that V is an inner-product space, those problems define $(u, v)_{\overline{V}} = (v, u)_V$, and they show that if $\ell_v \in V'$ is given by $\ell_v(u) = (u, v)_V = (v, u)_{\overline{V}}$, then the mapping $\ell_v \leftrightarrow v$ is an isomorphism of V' with $\overline{V}$.

32. Show that the definition $(\ell_{v_1}, \ell_{v_2})_{V'} = (v_1, v_2)_{\overline{V}}$ makes the isomorphism of V' with $\overline{V}$ preserve inner products.

33. If R is a unitary representation of G on the finite-dimensional complex vector space V, define the **contragredient** representation R^c of G on V' by $R^c(x) = R(x^{-1})^t$. Prove that $R^c(x)\ell_v = \ell_{R(x)v}$ and that R^c is unitary on V'.

34. Show that the matrix coefficients of R^c are the complex conjugates of those of R and that the characters satisfy $\chi_{R^c} = \overline{\chi_R}$.

35. Give an example of an irreducible representation of a finite group G that is not equivalent to its contragredient.

36. Let ℓ be the left regular representation of G on $C(G, \mathbb{C})$, and let V_R be the linear span in $C(G, \mathbb{C})$ of the matrix coefficients of an irreducible representation R of dimension d. Prove that the representation (ℓ, V_R) of G is equivalent to the direct sum of d copies of the contragredient R^c.

Problems 37–46 concern the free product $C_2 * C_3$ and its quotients. The problems make use of the group of matrices $\mathrm{SL}(2, \mathbb{Z}/m\mathbb{Z})$ of determinant 1 over the commutative ring $\mathbb{Z}/m\mathbb{Z}$, as discussed in Section V.2. One of the quotients of $C_2 * C_3$

will be $\mathrm{PSL}(2,\mathbb{Z}) = \mathrm{SL}(2,\mathbb{Z})/\{\text{scalar matrices}\}$, and these problems show that the quotient mapping can be arranged to be an isomorphism. Other quotients will be the groups $G_m = \langle X, Y;\ X^2, Y^3, (XY)^m\rangle$ with $m \geq 2$. These arise in connection with tilings in 2-dimensional geometry. The isomorphism $C_2 * C_3 \cong \mathrm{PSL}(2,\mathbb{Z})$ leads to a homomorphism that will be called σ_m carrying G_m onto $\mathrm{PSL}(2,\mathbb{Z}/m\mathbb{Z}) = \mathrm{SL}(2,\mathbb{Z}/m\mathbb{Z})/\{\text{scalar matrices}\}$, the image group being finite. The problems show that the homomorphism $\sigma_m : G_m \to \mathrm{PSL}(2,\mathbb{Z}/m\mathbb{Z})$ is an isomorphism for the cases in which G_m arises from spherical geometry, namely for $2 \leq m \leq 5$, and that the homomorphism is not an isomorphism for $m = 6$, the case in which G_m arises from Euclidean geometry.

37. Show that the elements $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 0 & 1 \\ -1 & -1 \end{smallmatrix}\right)$ generate $\mathrm{SL}(2,\mathbb{Z})$ by arguing as follows: if the subgroup Γ of $\mathrm{SL}(2,\mathbb{Z})$ generated by these two elements is not $\mathrm{SL}(2,\mathbb{Z})$, choose an element $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ outside Γ having $\max(|a|,|b|)$ as small as possible, and derive a contradiction by showing that a suitable right multiple of it by elements of Γ is in Γ.

38. By mapping $X \mapsto x = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right) \bmod \pm I$ and $Y \mapsto y = \left(\begin{smallmatrix} 0 & 1 \\ -1 & -1 \end{smallmatrix}\right) \bmod \pm I$, produce a group homomorphism Φ of $C_2 * C_3 = \langle X, Y;\ X^2, Y^3\rangle$ onto $\mathrm{PSL}(2,\mathbb{Z})$.

39. Let x, y, and $\Phi : C_2 * C_3 \to \mathrm{PSL}(2,\mathbb{Z})$ be as in the previous problem.
    (a) For any member $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \bmod \pm I$ of $\mathrm{PSL}(2,\mathbb{Z})$, define $\mu\left(\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \bmod \pm I\right) = \max(|a|,|b|)$ and $\nu\left(\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \bmod \pm I\right) = \max(|c|,|d|)$. Prove that if $z = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \bmod \pm I$ in $\mathrm{PSL}(2,\mathbb{Z})$ has $ab \leq 0$, then $\mu(zyx) \geq \mu(z)$ and $\mu(zy^{-1}x) \geq \mu(z)$, while if $cd \leq 0$, then $\nu(zyx) \geq \nu(z)$ and $\nu(zy^{-1}x) \geq \nu(z)$.
    (b) Prove that $\mu(zx) = \mu(z)$ and $\nu(zx) = \nu(z)$ for all z in $\mathrm{PSL}(2,\mathbb{Z})$.
    (c) Show that there are only 10 members z of $\mathrm{PSL}(2,\mathbb{Z})$ for which the two conditions $\mu(z) = 1$ and $\nu(z) = 1$ both hold.
    (d) A reduced word in $C_2 * C_3$ is a finite sequence of factors X, Y, and Y^{-1}, with no two consecutive factors equal and with no two consecutive factors YY^{-1} or $Y^{-1}Y$. Prove for any reduced word $a_1 \cdots a_n$ in $C_2 * C_3$, where each a_j is one of X, Y, and Y^{-1}, that $\mu(\Phi(a_1 \cdots a_n)) \geq \mu(\Phi(a_1 \cdots a_{n-1}))$ and that $\nu(\Phi(a_1 \cdots a_n)) \geq \nu(\Phi(a_1 \cdots a_{n-1}))$.
    (e) Deduce that the homomorphism Φ is an isomorphism.

40. Let $\Gamma(m)$ be the group of all matrices M in $\mathrm{SL}(2,\mathbb{Z})$ such that every entry of $M - I$ is divisible by m.
    (a) Prove that passage from a matrix in $\mathrm{SL}(2,\mathbb{Z})$ to the same matrix with its entries considered modulo m gives a homomorphism $\widetilde{\sigma}_m : \mathrm{SL}(2,\mathbb{Z}) \to \mathrm{SL}(2,\mathbb{Z}/m\mathbb{Z})$ with $\ker \widetilde{\sigma}_m = \Gamma(m)$.

(b) Prove that if α, β, and m are positive integers with $\mathrm{GCD}(\alpha, \beta, m) = 1$, then there exists an integer r such that $\mathrm{GCD}(\alpha + mr, \beta) = 1$. (One way of proceeding is to use Dirichlet's theorem on primes in arithmetic progressions.)

(c) Prove that image $\widetilde{\sigma}_m = \mathrm{SL}(2, \mathbb{Z}/m\mathbb{Z})$, i.e., $\widetilde{\sigma}_m$ is onto.

41. Let $\Phi_m : C_2 * C_3 \to G_m$ be the homomorphism defined by the conditions $X \mapsto X$ and $Y \mapsto Y$. Let H_m be the smallest normal subgroup of $\mathrm{PSL}(2, \mathbb{Z})$ containing $(xy)^m \bmod \pm I$. Let $\widetilde{\sigma}_m : \mathrm{SL}(2, \mathbb{Z}) \to \mathrm{SL}(2, \mathbb{Z}/m\mathbb{Z})$ be the homomorphism of the previous problem.

(a) Why is Φ_m well defined?

(b) Why is $H_m = \Phi(\ker \Phi_m)$?

(c) Define $\mathrm{PSL}(\mathbb{Z}/m\mathbb{Z}) = \mathrm{SL}(2, \mathbb{Z}/m\mathbb{Z})/\{\text{scalar matrices}\}$. Why does the composition of $\widetilde{\sigma}_m$ followed by passage to the quotient descend to a homomorphism σ_m of $\mathrm{PSL}(2, \mathbb{Z})$ onto $\mathrm{PSL}(2, \mathbb{Z}/m\mathbb{Z})$?

(d) If $K \subseteq \mathrm{PSL}(2, \mathbb{Z})$ is the kernel of σ_m, why is $H_m \subseteq K_m$?

(e) Show that if t is any integer, then the following members of K_m lie in the subgroup H_m: $\begin{pmatrix} 1 & tm \\ 0 & 1 \end{pmatrix} \bmod \pm I$, $\begin{pmatrix} 1 & 0 \\ tm & 1 \end{pmatrix} \bmod \pm I$, $\begin{pmatrix} 1+tm & tm \\ -tm & 1-tm \end{pmatrix} \bmod \pm I$, and $\begin{pmatrix} 1+tm & -tm \\ tm & 1-tm \end{pmatrix} \bmod \pm I$.

42. With G_m defined as above, exhibit homomorphisms of various groups G_m onto the following finite groups:

(a) $\mathfrak{S}_3$ when $m = 2$ by sending $X \mapsto (1\ 2)$ and $Y \mapsto (1\ 2\ 3)$.

(b) $\mathfrak{A}_4$ when $m = 3$ by sending $X \mapsto (1\ 2)(3\ 4)$ and $Y \mapsto (1\ 2\ 3)$.

(c) $\mathfrak{S}_4$ when $m = 4$ by sending $X \mapsto (1\ 2)$ and $Y \mapsto (2\ 3\ 4)$.

(d) $\mathfrak{A}_5$ when $m = 5$ by sending $X \mapsto (1\ 2)(3\ 4)$ and $Y \mapsto (1\ 3\ 5)$.

43. This problem shows how to prove that $H_m = K_m$ for $2 \le m \le 5$, and it asks that the steps be carried out for $m = 2$ and $m = 3$. Recall from the remark with Lemma 7.11 that Lemma 7.11 is valid for *all* groups in determining a set of generators of a subgroup from generators of the whole group and a system of coset representatives. The lemma is to be applied to the group $\mathrm{PSL}(2, \mathbb{Z})$ and the subgroup K_m. Generators of $\mathrm{PSL}(2, \mathbb{Z})$ are taken as $b_1 = x \bmod \pm I$ and $b_2 = y \bmod \pm I$.

(a) For the case $m = 2$, find members $g_1, \dots, g_6$ of $\mathrm{PSL}(2, \mathbb{Z})$ such that the six cosets of $\mathrm{PSL}(2, \mathbb{Z})/K_2$ are exactly $K_2 g_1, \dots, K_2 g_6$.

(b) Still for the case $m = 2$, find $g_j b_i \rho(g_j b_i)^{-1}$ for $1 \le i \le 2$ and $1 \le j \le 6$. Lemma 7.11 says that these 12 elements generate K_2.

(c) Using Problem 41e and any necessary variations of it, show that each of the 12 generators of K_2 in (b) lies in the subgroup H_2, and conclude that $H_2 = K_2$.

(d) Repeat steps (a), (b), and (c) for $m = 3$. There are 12 cosets $K_3 g_j$ of $\mathrm{PSL}(2, \mathbb{Z})/K_3$. (Educational note: There are 24 cosets for $\mathrm{PSL}(2, \mathbb{Z})/K_4$ and 60 cosets for $\mathrm{PSL}(2, \mathbb{Z})/K_5$.)

44. Take for granted that $H_m = K_m$ for $2 \le m \le 5$. Deduce the isomorphisms
    (a) $G_2 \cong \mathrm{PSL}(2, \mathbb{Z}/2\mathbb{Z}) \cong \mathfrak{S}_3$.
    (b) $G_3 \cong \mathrm{PSL}(2, \mathbb{Z}/3\mathbb{Z}) \cong \mathfrak{A}_4$. (This group is called the **tetrahedral group**.)
    (c) $G_4 \cong \mathrm{PSL}(2, \mathbb{Z}/4\mathbb{Z}) \cong \mathfrak{S}_4$. (This group is called the **octahedral group**.)
    (d) $G_5 \cong \mathrm{PSL}(2, \mathbb{Z}/5\mathbb{Z}) \cong \mathfrak{A}_5$. (This group is called the **icosahedral group**.)

45. A translation in the Euclidean plane $\mathbb{R}^2$ is any function $T_{(a,b)}(x, y) = (a + x, b + y)$, the rotation about the origin clockwise through the angle θ is the linear map R_θ given by the matrix $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$, and the rotation about (x_0, y_0) clockwise through the angle θ is the linear map given by $(x, y) \mapsto R_\theta(x - x_0, y - y_0) + (x_0, y_0)$.
    (a) Prove that $R_\theta T_{(a,b)} R_\theta^{-1} = T_{R_\theta(a,b)}$.
    (b) Prove that the union of the set of translations and all the sets of rotations about points of $\mathbb{R}^2$ is a group by showing that it is the semidirect product of the subgroup of rotations about the origin and the normal subgroup of translations.

46. Fix a triangle T in the Euclidean plane with vertices arranged counterclockwise at a, b, c and with angles $\pi/2$ at a, $\pi/3$ at b, and $\pi/6$ at c. Let r_a be rotation clockwise through π at a, r_b be rotation clockwise through $2\pi/3$ at b, and r_c be rotation *counter*clockwise through $\pi/3$ at c.
    (a) Show that $r_a^2 = 1, r_b^3 = 1, r_c^6 = 1$, and $r_c = r_a r_b$.
    (b) Show that the member $r_b r_a r_b r_a r_b$ of the group generated by r_a and r_b is a nontrivial translation and therefore that the generated group is infinite.
    (c) Conclude that $G_6 \ncong \mathrm{PSL}(2, \mathbb{Z}/6\mathbb{Z})$. (Educational note: If $\widetilde{T}$ denotes the union of T and the reflection of T in one of the sides of T, it can be shown that the group generated by r_a and r_b is isomorphic to G_6 and tiles the plane with copies of $\widetilde{T}$.)

Problems 47–52 establish a harmonic analysis for arbitrary representations of finite groups on complex vector spaces, whether finite-dimensional or infinite-dimensional. Let G be a finite group, and let V be a complex vector space. For any representation R of G on V, one defines $R(f)v = \sum_{x \in G} f(x) R(x) v$ for f in $C(G, \mathbb{C})$ and v in V, just as in the case that V is finite-dimensional. The same computation as in Section VII.4 shows that the formula $R(f_1 * f_2) = R(f_1)R(f_2)$ remains valid when V is infinite-dimensional.

47. Let (R_1, V_1) and (R_2, V_2) be irreducible finite-dimensional representations of G on complex vector spaces, and let χ_{R_1} and χ_{R_2} be their characters. Using Schur orthogonality, prove that
    (a) $\chi_{R_1} * \chi_{R_2} = 0$ if R_1 and R_2 are inequivalent,

(b) $\chi_{R_1} * \chi_{R_1} = |G| d_{R_1}^{-1} \chi_{R_1}$, where $d_{R_1} = \dim V_R$.

48. With (R, V) given, let (R_α, V_α) be any irreducible finite-dimensional representation of G, and define $E_\alpha : V \to V$ by $E_\alpha = |G|^{-1} d_\alpha R(\overline{\chi_\alpha})$, where χ_α is the character of R_α and where $d_\alpha = \dim V_\alpha$.
    (a) Prove that $E_\alpha^2 = E_\alpha$.
    (b) Prove that $E_\alpha E_\beta = E_\beta E_\alpha = 0$ if (R_β, V_β) is an irreducible finite-dimensional representation of G such that R_α and R_β are inequivalent.

49. Observe for each v in V that $\{R(x)v \mid x \in G\}$ spans a finite-dimensional invariant subspace of V. By Corollary 7.21, each v in V lies in a finite direct sum of finite-dimensional invariant subspaces of V on each of which R acts irreducibly. Using Zorn's Lemma, prove that V is the direct sum of finite-dimensional subspaces on each of which R acts irreducibly. (If V is infinite-dimensional, there will of course be infinitely many such subspaces.)

50. Suppose that V_0 is a finite-dimensional invariant subspace of V such that $R\big|_{V_0}$ is equivalent to some R_α, where R_α is as in Problem 48. Prove that E_α is the identity on V_0.

51. Deduce that if $\{(R_\beta, V_\beta)\}$ is a maximal collection of inequivalent finite-dimensional irreducible representations of G, then $\sum_\beta E_\beta = I$ on V and the image of E_α is the set of all sums of vectors in V lying in some finite-dimensional invariant subspace V_0 of V such that $R\big|_{V_0}$ is equivalent to R_α. (Educational note: Consequently V is exhibited as the finite direct sum of the spaces image E_α, each space image E_α is the direct sum of finite-dimensional irreducible invariant subspaces, and the restriction of R to any finite-dimensional irreducible invariant subspace of image E_α is equivalent with R_α.

52. Suppose that (R_α, V_α) is a 1-dimensional representation of G given by a multiplicative character ω. Prove that the image of E_α consists of all vectors v in V such that $R(x)v = \omega(x)v$ for all x in G.