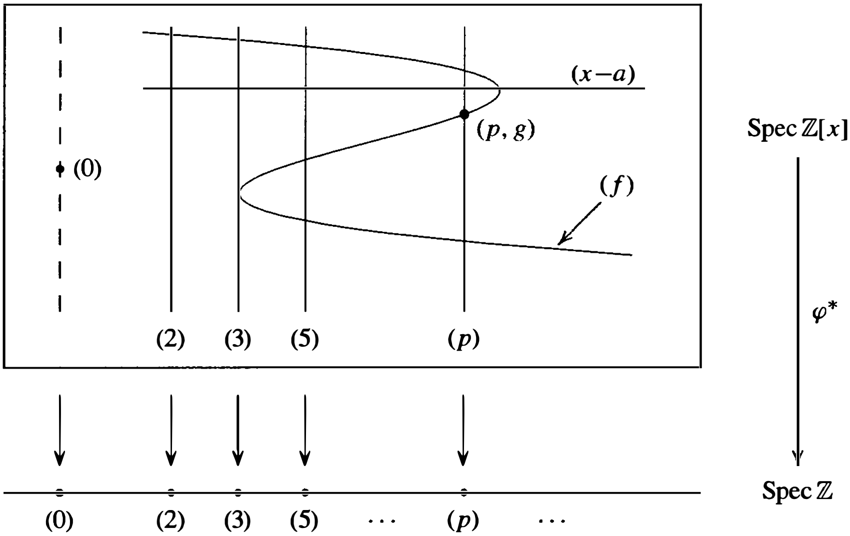


over $\mathbb{F}_p$ factors into irreducibles as follows:

$$f \equiv x^3 + x + 1 \bmod 2$$
$$f \equiv x(x+1)^2 \bmod 3$$
$$f \equiv (x+1)(x+2)(x+3) \bmod 5.$$

There is one point in the fiber over (2) intersecting (f), namely the closed point $(2, x^3 + x + 1)$. There are two closed points in the fiber over (3) given by $(3, x)$ and $(3, x+1)$ (with some "multiplicity" at the latter point). Over (5) there are three closed points: $(5, x+1)$, $(5, x+2)$, and $(5, x+3)$. For the diagram above, the prime p might be $p = 53$, since this is the first prime p greater than 5 for which this polynomial has three irreducible factors mod p. Note that while the prime (f) is drawn as a smooth curve in this diagram to emphasize the geometric similarity with the structure of Spec $k[x, y]$ in the previous example, the fibers above the primes in Spec $\mathbb{Z}$ are discrete, so some care should be exercised. For example, since f factors as $(x+2)(x^2+x+6)$ mod 7, the intersection of (f) with the fiber above (7) contains only the two points $(7, x+2)$ and $(7, x^2+x+6)$, each with multiplicity one.

The possible number of closed points in (f) lying in a fiber over $(p) \in$ Spec $\mathbb{Z}$ is controlled by the Galois group of the polynomial f over $\mathbb{Q}$ (cf. Section 14.8). For example, $f = x^4 + 1$ has one closed point in the fiber above (2) and either two or four closed points in a fiber above (p) for p odd (cf. Exercise 8).

The space Spec R together with its Zariski topology gives a geometric generalization for arbitrary commutative rings of the points in a variety V. We now consider the question of generalizing the ring of rational functions on V.

When V is a variety over the algebraically closed field k the elements in the quotient field $k(V)$ of the coordinate ring $k[V]$ define the rational functions on V. Each element α in $k(V)$ can in general be written as a quotient a/f of elements $a, f \in k[V]$ in many different ways. The set of points U at which α is regular is an open subset of V; by definition, it consists of all the points $v \in V$ where α can be represented by some quotient a/f with $f(v) \neq 0$, and then the representative a/f defines an element in the local ring $\mathcal{O}_{v,V}$. Note also that the same representative a/f defines α not only at v, but also at all the other points where f is nonzero, namely on the open subset $V_f = \{w \in V \mid f(w) \neq 0\}$ of V. These open sets V_f (called principal open sets, cf. Exercise 21 in Section 2) for the various possible representatives a/f for α give an open cover of U. The example of the function $\alpha = \bar{x}/\bar{y}$ for $V = \mathcal{Z}(xz - yw) \subset \mathbb{A}^4$ preceding Proposition 51 shows that in general a single representative for α does not suffice to determine all of U — for this example, $U = V_{\bar{y}} \cup V_{\bar{z}}$, and U is not covered by any single V_f (cf. Exercise 25 of Section 4).

This interpretation of rational functions as functions that are regular on open subsets of V can be generalized to Spec R. We first define the analogues X_f in $X = \text{Spec } R$ of the sets V_f and establish their basic properties.

Definition. For any $f \in R$ let X_f denote the collection of prime ideals in $X = \text{Spec } R$ that do not contain f. Equivalently, X_f is the set of points of Spec R at which the value of $f \in R$ is nonzero. The set X_f is called a *principal* (or *basic*) *open set* in Spec R.

Since X_f is the complement of the Zariski closed set $\mathcal{Z}(f)$ it is indeed an open set in Spec R as the name implies. Some basic properties of the principal open sets are indicated in the next proposition. Recall that a map between topological spaces is a *homeomorphism* if it is continuous and bijective with continuous inverse.

Proposition 56. Let $f \in R$ and let X_f be the corresponding principal open set in $X = \text{Spec } R$. Then

(1) $X_f = X$ if and only if f is a unit, and $X_f = \emptyset$ if and only if f is nilpotent,

(2) $X_f \cap X_g = X_{fg}$,

(3) $X_f \subseteq X_{g_1} \cup \cdots \cup X_{g_n}$ if and only if $f \in \text{rad}(g_1, \ldots, g_n)$; in particular $X_f = X_g$ if and only if $\text{rad}(f) = \text{rad}(g)$,

(4) the principal open sets form a basis for the Zariski topology on Spec R, i.e., every Zariski open set in X is the union of some collection of principal open sets X_f,

(5) the natural map from R to R_f induces a homeomorphism from Spec R_f to X_f, where R_f is the localization of R at f,

(6) the spectrum of any ring is quasicompact (i.e., every open cover has a finite subcover); in particular, X_f is quasicompact, and

(7) if $\varphi : R \to S$ is any homomorphism of rings (with $\varphi(1_R) = 1_S$) then under the induced map $\varphi^* : Y = \text{Spec } S \to \text{Spec } R$ the full preimage of the principal open set X_f in X is the principal open set $Y_{\varphi(f)}$ in Y.

Proof: Parts (1), (2) and (7) are left as easy exercises. For (3), observe that, by definition, $X_{g_1} \cup \cdots \cup X_{g_n}$ consists of the primes P not containing at least one of $g_1, \ldots, g_n$. Hence $X_{g_1} \cup \cdots \cup X_{g_n}$ is the complement of the closed set $\mathcal{Z}((g_1, \ldots, g_n))$ consisting of the primes P that contain the ideal generated by $g_1, \ldots, g_n$. If $(g_1, \ldots, g_n) = R$ then $X_{g_1} \cup \cdots \cup X_{g_n} = X$ and there is nothing to prove. Otherwise, $X_f \subseteq X_{g_1} \cup \cdots \cup X_{g_n}$ if and only if every prime P with $f \notin P$ also satisfies $P \notin \mathcal{Z}((g_1, \ldots, g_n))$. This latter condition is equivalent to the statement that if the prime P contains the ideal

$(g_1, \ldots, g_n)$ then P also contains f, i.e., f is contained in the intersection of all the prime ideals P containing $(g_1, \ldots, g_n)$. Since this intersection is $\operatorname{rad}(g_1, \ldots, g_n)$ by Proposition 12, this proves (3).

If $U = X - \mathcal{Z}(I)$ is a Zariski open subset of X, then U is the union of the sets X_f with $f \in I$, which proves (4).

The natural ring homomorphism from R to the localization R_f establishes a bijection between the prime ideals in R_f and the prime ideals in R not containing (f) (Proposition 38). The corresponding Zariski continuous map from Spec R_f to Spec R is therefore continuous and bijective. Since every ideal of R_f is the extension of some ideal of R (cf. Proposition 38(1)), it follows that the inverse map is also continuous, which proves (5).

In (6), every open set is the union of principal open sets by (4), so it suffices to prove that if X is covered by principal open sets X_{g_i} (for i in some index set $\mathcal{J}$) then X is a finite union of some of the X_{g_i}. If the ideal I generated by the g_i were a proper ideal in R, then I would be contained in some maximal ideal P. But in this case the element P in $X =$ Spec R would not be contained in any principal open set X_{g_i}, contradicting the assumption that X is covered by the X_{g_i}. Hence $I = R$ and so $1 \in R$ can be written as a finite sum $1 = a_1 g_{i_1} + \cdots + a_n g_{i_n}$ with $i_1, \ldots, i_n \in \mathcal{J}$. Consider the finite union $X_{g_1} \cup \cdots \cup X_{g_n}$. Any point P in X not contained in this union would be a prime in R that contains $g_{i_1}, \ldots, g_{i_n}$, hence would contain 1, a contradiction. It follows that $X = X_{g_1} \cup \cdots \cup X_{g_n}$ as needed. The second part of (6) follows from (5).

We now define an analogue for $X =$ Spec R of the rational functions on a variety V. As we observed, for the variety V a rational function $\alpha \in k(V)$ is a regular function on some open set U. At each point $v \in U$ there is a representative a/f for α with $f(v) \neq 0$, and this representative is an element in the localization $\mathcal{O}_{v,V} = k[V]_{\mathcal{I}(v)}$. In this way the regular function α on U can be considered as a function from U to the disjoint union of these localizations: the point $v \in U$ is mapped to the representative $a/f \in k[V]_{\mathcal{I}(v)}$. Furthermore the same representative can be used simultaneously not only at v but on the whole Zariski neighborhood V_f of v (so, "locally near v," α is given by a single quotient of elements from $k[V]$). Note that a/f is an element in the localization $k[V]_f$, which is contained in each of the localizations $k[V]_{\mathcal{I}(w)}$ for $w \in V_f$.

We now generalize this to Spec R by considering the collection of functions s from the Zariski open subset U of Spec R to the disjoint union of the localizations R_P for $P \in U$ such that $s(P) \in R_P$ and such that s is given locally by quotients of elements of R. More precisely:

Definition. Suppose U is a Zariski open subset of Spec R. If $U = \emptyset$, define $\mathcal{O}(U) = 0$. Otherwise, define $\mathcal{O}(U)$ to be the set of functions $s : U \to \bigsqcup_{Q \in U} R_Q$ from U to the disjoint union of the localizations R_Q for $Q \in U$ with the following two properties:

(1) $s(Q) \in R_Q$ for every $Q \in U$, and

(2) for every $P \in U$ there is an open neighborhood $X_f \subseteq U$ of P in U and an element a/f^n in the localization R_f defining s on X_f, i.e., $s(Q) = a/f^n \in R_Q$ for every $Q \in X_f$.

If s, t are elements in $\mathcal{O}(U)$ then $s + t$ and st are also elements in $\mathcal{O}(U)$ (cf. Exercise 18), so each $\mathcal{O}(U)$ is a ring. Also, every $a \in R$ gives an element in $\mathcal{O}(U)$

defined by $s(Q) = a \in R_Q$, and in particular $1 \in R$ gives an identity for the ring $\mathcal{O}(U)$. If U' is an open subset of U, then there is a natural restriction map from $\mathcal{O}(U)$ to $\mathcal{O}(U')$ which is a homomorphism of rings (cf. Exercise 19).

Definition. Let R be a commutative ring with 1, and let $X = \operatorname{Spec} R$.

(1) The collection of rings $\mathcal{O}(U)$ for the Zariski open sets of X together with the restriction maps $\mathcal{O}(U) \to \mathcal{O}(U')$ for $U' \subseteq U$ is called the *structure sheaf* on X, and is denoted simply by $\mathcal{O}$ (or $\mathcal{O}_X$).

(2) The elements s of $\mathcal{O}(U)$ are called the *sections of* $\mathcal{O}$ over U. The elements of $\mathcal{O}(X)$ are called the *global sections* of $\mathcal{O}$.

The next proposition generalizes the result of Proposition 51 that the only rational functions on a variety V that are regular everywhere are the elements of the coordinate ring $k[V]$.

Proposition 57. Let $X = \operatorname{Spec} R$ and let $\mathcal{O} = \mathcal{O}_X$ be its structure sheaf. The global sections of $\mathcal{O}$ are the elements of R, i.e., $\mathcal{O}(X) \cong R$. More generally, if X_f is a principal open set in X for some $f \in R$, then $\mathcal{O}(X_f)$ is isomorphic to the localization R_f.

Proof: Suppose that a/f^n is an element of the localization R_f. Then the map defined by $s(Q) = a/f^n \in R_Q$ for $Q \in X_f$ gives an element in $\mathcal{O}(X_f)$, and it is immediate that the resulting map ψ from R_f to $\mathcal{O}(X_f)$ is a ring homomorphism. Suppose that $a/f^n = b/f^m$ in R_Q for every $Q \in X_f$, i.e., $g(af^m - bf^n) = 0$ in R for some $g \notin Q$. If I is the ideal in R of elements $r \in R$ with $r(af^m - bf^n) = 0$, it follows from $g \in I$ that I is not contained in Q for any $Q \in X_f$. Put another way, every prime ideal of R containing I also contains f. Hence f is contained in the intersection of all the prime ideals of R containing I, which is to say that $f \in \operatorname{rad} I$. Then $f^N \in I$ for some integer $N \geq 0$, and so $f^N(af^m - bf^n) = 0$ in R. But this shows that $a/f^n = b/f^m$ in R_f and so the map ψ is injective. Suppose now that $s \in \mathcal{O}(X_f)$. Then by definition X_f can be covered by principal open sets X_{g_i} on which $s(Q) = a_i/g_i^{n_i} \in R_Q$ for every $Q \in X_{g_i}$. By (6) of Proposition 56, we may take a finite number of the g_i and then by taking different a_i we may assume all the n_i are equal (since $a_i/g_i^{n_i} = (a_i g_i^{n-n_i})/g_i^n$ if n is the maximum of the n_i). Since $s(Q) = a_i/g_i^n = a_j/g_j^n$ in R_Q for all $Q \in X_{g_i g_j} = X_{g_i} \cap X_{g_j}$, the injectivity of ψ (applied to $R_{g_i g_j}$) shows that $a_i/g_i^n = a_j/g_j^n$ in $R_{g_i g_j}$. This means that $g_i g_j{}^N(a_i g_j^n - a_j g_i^n) = 0$, i.e.,

$$a_i g_i^N g_j{}^{n+N} = a_j g_i^{n+N} g_j{}^N$$

in R for some $N \geq 0$, and we may assume N sufficiently large that this holds for every i and j. Since X_f is the union of the $X_{g_i} = X_{g_i^{n+N}}$, f is contained in the radical of the ideal generated by the g_i^n by (3) of Proposition 56, say

$$f^M = \sum_i b_i g_i^{n+N}$$

for some $M \geq 1$ and $b_i \in R$. Define $a = \sum b_i a_i g_i^N \in R$. Then

$$g_j^N a_j f^M = \sum_i b_i (a_j g_i^{n+N} g_j{}^N) = \sum_i b_i (a_i g_i^N g_j{}^{n+N}) = g_j{}^{n+N} a.$$

It follows that $a/f^M = a_j/g_j^n$ in R_{g_j}, and so the element in $\mathcal{O}(X_f)$ defined by a/f^M in R_f agrees with s on every X_{g_j}, and so on all of X_f since these open sets cover X_f. Hence the map ψ gives an isomorphism $R_f \cong \mathcal{O}(X_f)$. Taking $f = 1$ gives $R \cong \mathcal{O}(X)$, completing the proof.

In the case of affine varieties V the local ring $\mathcal{O}_{v,V}$ at the point $v \in V$ is the collection of all the rational functions in $k(V)$ that are defined at v. Put another way, $\mathcal{O}_{v,V}$ is the union of the rings of regular functions on U for the open sets U containing P, where this union takes place in the function field $k(V)$ of V. In the more general case of $X = \operatorname{Spec} R$, the rings $\mathcal{O}(U)$ for the open sets containing $P \in \operatorname{Spec} R$ are not contained in such an obvious common ring. In this case we proceed by considering the collection of pairs (s, U) with U an open set of X containing P and $s \in \mathcal{O}(U)$. We identify two pairs (s, U) and (s', U') if there is an open set $U'' \subseteq U \cap U'$ containing P on which s and s' restrict to the same element of $\mathcal{O}(U'')$. In the situation of affine varieties, this says that two functions defined in Zariski neighborhoods of the point v define the same regular function at v if they agree in some common neighborhood of v. The collection of equivalence classes of pairs (s, U) defines the *direct limit* of the rings $\mathcal{O}(U)$, and is denoted $\varinjlim \mathcal{O}(U)$ (cf. Exercise 8 in Section 7.6).

Definition. If $P \in X = \operatorname{Spec} R$, then the direct limit, $\varinjlim \mathcal{O}(U)$, of the rings $\mathcal{O}(U)$ for the open sets U of X containing P is called the *stalk* of the structure sheaf at P, and is denoted $\mathcal{O}_P$.

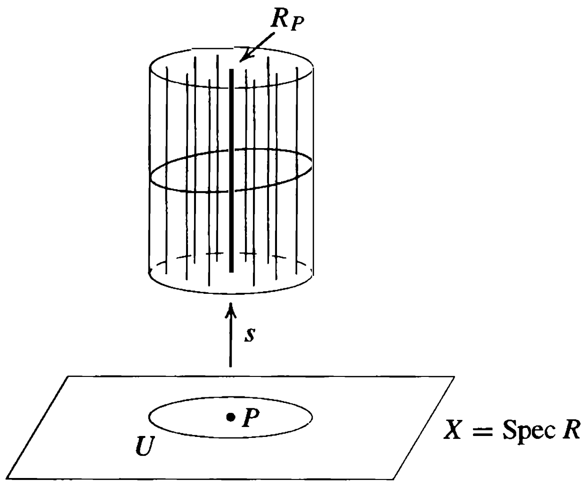
Proposition 58. Let $X = \operatorname{Spec} R$ and let $\mathcal{O} = \mathcal{O}_X$ be its structure sheaf. The stalk of $\mathcal{O}$ at the point $P \in X$ is isomorphic to the localization R_P of R at P: $\mathcal{O}_P \cong R_P$. In particular, the stalk $\mathcal{O}_P$ is a local ring.

Proof: If (s, U) represents an element in the stalk $\mathcal{O}_P$, then $s(P)$ is an element of the localization R_P. By the definition of the direct limit, this element does not depend on the choice of representative (s, U), and so gives a well defined ring homomorphism φ from $\mathcal{O}_P$ to R_P. If $a, f \in R$ with $f \notin P$, then the map $s(Q) = a/f \in R_Q$ defines an element in $\mathcal{O}(X_f)$. Then the class of (s, X_f) in the stalk $\mathcal{O}_P$ is mapped to a/f in R_P by φ, so φ is a surjective map. To see that φ is also injective, suppose that the classes of (s, U) and (s', U') in $\mathcal{O}_P$ satisfy $s(P) = s'(P)$ in R_P. By definition of $\mathcal{O}(U)$, $s = a/g^n$ on X_g for some $g \notin P$. Similarly, $s' = b/(g')^m$ on $X_{g'}$ for some $g' \notin P$. Since $a/g^n = b/(g')^m$ in R_P, there is some $h \notin P$ with $h(a(g')^m - bg^n) = 0$ in R. If $Q \in X_{gg'h} = X_g \cap X_{g'} \cap X_h$ this last equality shows that $a/g^n = b/(g')^m$ in R_Q, so that s and s' agree when restricted to $X_{gg'h}$. By definition of the direct limit, (s, U) and (s', U') define the same element in the stalk $\mathcal{O}_P$, which proves that φ is injective and establishes the proposition.

Proposition 58 shows that the algebraically defined localization R_P for $P \in \operatorname{Spec} R$ plays the role of the local ring $\mathcal{O}_{v,V}$ of regular functions at v for the affine variety V. If $\mathfrak{m}_P$ denotes the maximal ideal PR_P in R_P and $k(P) = R_P/\mathfrak{m}_P$ denotes the corresponding quotient field (which by Proposition 46(1) is also the fraction field of R/P), then the *tangent space* at P is defined to be the $k(P)$-vector space dual of $\mathfrak{m}_P/\mathfrak{m}_P^2$.

This is an algebraic definition that generalizes the definition of the tangent space $\mathbb{T}_{v,V}$ to a variety V at a point v (by Proposition 52). This can now be used to define what it means for a point in Spec R to be nonsingular: the point $P \in \operatorname{Spec} R$ is *nonsingular* or *smooth* if the local ring R_P is what is called a "regular local ring" (cf. Section 16.2).

Proposition 58 also suggests a nice geometric view of the structure sheaf on Spec R. If we view each point $P \in \operatorname{Spec} R$ as having the local ring R_P above it, then above the open set U in $X = \operatorname{Spec} R$ is a "sheaf" (in the sense of a "bundle") of these "stalks" (in the sense of a "stalk of wheat"), which helps explain some of the terminology. A section s in the structure sheaf $\mathcal{O}(U)$ is a map from U to this bundle of stalks. The image of U under such a section s is indicated by the shaded region in the following figure.



Definition. Let R be a commutative ring with 1. The pair (Spec R, $\mathcal{O}_{\operatorname{Spec} R}$), consisting of the space Spec R with the Zariski topology together with the structure sheaf $\mathcal{O}_{\operatorname{Spec} R}$, is called an *affine scheme*.

The notion of an affine scheme gives a completely algebraic generalization of the geometry of affine algebraic sets valid for arbitrary commutative rings, and is the starting point for modern algebraic geometry.

Examples

(1) If F is any field then $X = \operatorname{Spec} F = \{(0)\}$. In this case there are only two open sets X and $\emptyset$, both of which are principal open sets: $X = X_1$ and $\emptyset = X_0$. The global sections are $\mathcal{O}(X) = F$. There is only one stalk: $\mathcal{O}_{(0)} = F_0 = F$.

(2) If $R = \mathbb{Z}$ then because R is a P.I.D. every open set in $X = \operatorname{Spec} \mathbb{Z}$ is principal open:

$$X_n = \{(p) \mid p \nmid n\} \qquad \text{and}$$

$$\mathcal{O}(X_n) = \mathbb{Z}_n = \mathbb{Z}[1/n] = \{a/b \in \mathbb{Q} \mid \text{if the prime } p \mid b \text{ then } p \mid n\}.$$

For nonzero p the stalk at (p) is the local ring $\mathbb{Z}_{(p)}$, and the stalk at (0) is $\mathbb{Q}$. All the restriction maps as well as the maps from sections to stalks are the natural inclusions.

(3) For a general integral domain R with quotient field F the stalks and sections are

$$\mathcal{O}(U) = \{a/b \in F \mid b \notin P \text{ for all } P \in U\}$$
$$\mathcal{O}_P = R_P = \{a/b \in F \mid b \notin P\}$$

where the stalk at (0) is F, i.e., $\mathcal{O}_{(0)} = F$. Again, the restriction maps and the maps to the stalks are all inclusions.

(4) For the local ring $R = \mathbb{Z}_{(2)} = \{a/b \in \mathbb{Q} \mid b \text{ odd}\}$ we have $\operatorname{Spec} R = \{(0), (2)\}$ with (2) the only closed point and $\{(0)\} = X_2$ a principal open set. The sections $\mathcal{O}(\{(0)\})$ are $R_2 = \mathbb{Q}$, and the stalks are $\mathcal{O}_{(0)} = R_{(0)} = \mathbb{Q}$ and $\mathcal{O}_{(2)} = R_{(2)} = R$.

We next consider the relationship of the affine schemes corresponding to rings R and S with respect to a ring homomorphism from R to S.

Suppose that $\varphi : R \to S$ is a ring homomorphism. We have already seen in Proposition 56(7) that there is an induced continuous map φ^* from $Y = \operatorname{Spec} S$ to $X = \operatorname{Spec} R$ and that under this map the full preimage of the principal open set X_g for $g \in R$ is the principal open set $Y_{\varphi(g)}$. It follows that φ also induces a map on corresponding sections, as follows. Let $Q' \in Y$ be any element in Spec S and let $Q = \varphi^*(Q') = \varphi^{-1}(Q') \in X$ be the corresponding element in Spec R. If U is a Zariski open set in X containing Q, then $U' = (\varphi^*)^{-1}(U)$ is a Zariski open set in Y containing Q'. Note that φ induces a natural ring homomorphism, φ_Q say, from the localization R_Q to the localization $S_{Q'}$ defined by $\varphi_Q(a/f) = \varphi(a)/\varphi(f) \in S_{Q'}$ for $f \notin Q$. Let $s \in \mathcal{O}_X(U)$ be a section of the structure sheaf of X given locally in the neighborhood X_g of $P \in X$ by a/g^n. It is easy to check that the composite

$$s' : U' \xrightarrow{\varphi^*} U \xrightarrow{s} \bigsqcup_{Q \in U} R_Q \xrightarrow{\varphi} \bigsqcup_{Q' \in U} S_{Q'}$$

defines a map given locally in the neighborhood $Y_{\varphi(g)}$ by the element $\varphi(a)/\varphi(g)^n$, so that $s' \in \mathcal{O}_Y(U')$ is a section of the structure sheaf of Y. It is then straightforward to check that the resulting map $\varphi^\# : \mathcal{O}_X(U) \to \mathcal{O}_Y(U')$ is a ring homomorphism (mapping $1 \in \mathcal{O}_X(U)$ to $1 \in \mathcal{O}_Y(U')$) that is compatible with the restriction maps on $\mathcal{O}_X$ and $\mathcal{O}_Y$ (cf. Exercise 20). It also follows that there is an induced ring homomorphism on the stalks: $\varphi^\# : \mathcal{O}_{X,P} \to \mathcal{O}_{Y,P'}$ for any point $P' \in \operatorname{Spec} S$ and corresponding point $P = \varphi^*(P') \in$ Spec R. Under the isomorphism in Proposition 58, the homomorphism $\varphi^\#$ from $R_P \cong \mathcal{O}_{X,P}$ to $S_{P'} \cong \mathcal{O}_{Y,P'}$ is just the natural ring homomorphism φ_P on the localizations induced by the homomorphism φ. In particular, the inverse image under $\varphi^\#$ of the maximal ideal in the local ring $\mathcal{O}_{Y,P'}$ is the maximal ideal in the local ring $\mathcal{O}_{X,P}$.

Definition. Suppose (Spec R, $\mathcal{O}_{\operatorname{Spec} R}$) and (Spec S, $\mathcal{O}_{\operatorname{Spec} S}$) are two affine schemes. A *morphism of affine schemes* from (Spec S, $\mathcal{O}_{\operatorname{Spec} S}$) to (Spec R, $\mathcal{O}_{\operatorname{Spec} R}$) is a pair $(\varphi^*, \varphi^\#)$ such that

(1) $\varphi^* : \operatorname{Spec} S \to \operatorname{Spec} R$ is Zariski continuous,

(2) there are ring homomorphisms $\varphi^\# : \mathcal{O}(U) \to \mathcal{O}(\varphi^{*-1}(U))$ for every Zariski open subset U in Spec R that commute with the restriction maps, and

(3) if $P' \in \operatorname{Spec} S$ with corresponding point $P = \varphi^*(P) \in \operatorname{Spec} R$, then under the induced homomorphism on stalks $\varphi^{\#} : \mathcal{O}_{\operatorname{Spec} R, P} \to \mathcal{O}_{\operatorname{Spec} S, P'}$ the preimage of the maximal ideal of $\mathcal{O}_{\operatorname{Spec} S, P'}$ is the maximal ideal of $\mathcal{O}_{\operatorname{Spec} R, P}$.

A homomorphism $\psi : A \to B$ from the local ring A to the local ring B with the property that the preimage of the maximal ideal of B is the maximal ideal of A is called a *local homomorphism* of local rings. The third condition in the definition is then the statement that the induced homomorphism on stalks is required to be a local homomorphism.

With this terminology, the discussion preceding the definition shows that a ring homomorphism $\varphi : R \to S$ induces a morphism of affine schemes from $(\operatorname{Spec} S, \mathcal{O}_{\operatorname{Spec} S})$ to $(\operatorname{Spec} R, \mathcal{O}_{\operatorname{Spec} R})$.

Conversely, suppose $(\varphi^*, \varphi^{\#})$ is a morphism of affine schemes from $(\operatorname{Spec} S, \mathcal{O}_{\operatorname{Spec} S})$ to $(\operatorname{Spec} R, \mathcal{O}_{\operatorname{Spec} R})$. Then in particular, for $U = \operatorname{Spec} R$, $(\varphi^*)^{-1}(U) = \operatorname{Spec} S$, so by assumption there is a ring homomorphism $\varphi^{\#} : \mathcal{O}_{\operatorname{Spec} R}(\operatorname{Spec} R) \to \mathcal{O}_{\operatorname{Spec} S}(\operatorname{Spec} S)$ defined on the global sections. By Proposition 57, we have $\mathcal{O}_{\operatorname{Spec} R}(\operatorname{Spec} R) \cong R$ and $\mathcal{O}_{\operatorname{Spec} S}(\operatorname{Spec} S) \cong S$ as rings. Composing with these isomorphisms shows that $\varphi^{\#}$ gives a ring homomorphism $\varphi : R \to S$. By Proposition 58 we have a local homomorphism $\varphi^{\#} : R_P \to S_{P'}$, and by the compatibility with the restriction homomorphisms it follows that the diagram

$$\begin{tikzcd} R \arrow[r, "\varphi"] \arrow[d] & S \arrow[d] \\ R_P \arrow[r, "\varphi^{\#}"] & S_{P'} \end{tikzcd}$$

commutes, where the two vertical maps are the natural localization homomorphisms. Since $\varphi^{\#}$ is assumed to be a local homomorphism, $(\varphi^{\#})^{-1}(P'S_{P'}) = PR_P$, from which it follows that $\varphi^{-1}(P') = P$. Hence the continuous map from $\operatorname{Spec} S$ to $\operatorname{Spec} R$ induced by φ is the same as φ^*, and it follows easily that φ also induces the homomorphism $\varphi^{\#}$. This shows that there is a ring homomorphism $\varphi : R \to S$ inducing both φ^* and $\varphi^{\#}$ as before.

We summarize this in the following proposition:

Theorem 59. Every ring homomorphism $\varphi : R \to S$ induces a morphism

$$(\varphi^*, \varphi^{\#}) : (\operatorname{Spec} S, \mathcal{O}_{\operatorname{Spec} S}) \to (\operatorname{Spec} R, \mathcal{O}_{\operatorname{Spec} R})$$

of affine schemes. Conversely, every morphism of affine schemes arises from such a ring homomorphism φ.

Theorem 59 is the analogue for $\operatorname{Spec} R$ of Theorem 6, which converted geometric questions relating to affine algebraic sets to algebraic questions for their coordinate rings.

The condition that the homomorphism on stalks be a local homomorphism in the definition of a morphism of affine schemes is necessary: a continuous map on the spectra together with a set of compatible ring homomorphisms on sections (hence also on stalks) is not sufficient to force these maps to come from a ring homomorphism.

Example

Let $R = \mathbb{Z}_{(2)}$ and $S = \mathbb{Q}$ as in the preceding set of examples. Define $\varphi^* : \operatorname{Spec}\mathbb{Q} \to \operatorname{Spec}\mathbb{Z}_{(2)}$ by $\varphi^*((0)) = (2)$ (which is Zariski continuous). Define $\varphi^\# : \mathcal{O}(\operatorname{Spec} R) \to \mathcal{O}(\operatorname{Spec} S)$ to be the inclusion map $\mathbb{Z}_{(2)} \hookrightarrow \mathbb{Q}$ and define $\varphi^\#$ for all other $U \subseteq \operatorname{Spec} R$ simply to be the zero map. It is straightforward to check that these homomorphisms commute with the restriction maps. This family of maps does *not* arise from a ring homomorphism, however, because on the stalks for $(0) \in \operatorname{Spec} S$ and $\varphi^*((0)) = (2) \in \operatorname{Spec} R$ the induced homomorphism

$$\varphi^\# : \mathcal{O}_{\operatorname{Spec} R,(2)} \hookrightarrow \mathcal{O}_{\operatorname{Spec} S,(0)}$$

is the injection $\mathbb{Z}_{(2)} \hookrightarrow \mathbb{Q}$, which is not a *local* homomorphism (the inverse image of (0) is (0) and not the maximal ideal $2\mathbb{Z}_{(2)}$).

The proof of Theorem 59 shows that a morphism $(\varphi^*, \varphi^\#)$ of affine schemes necessarily comes from the ring homomorphism defined by $\varphi^\#$ on global sections. In this example, the homomorphism on global sections is the inclusion map of R into S. The inclusion map from R to S defines a map from $\operatorname{Spec} S$ to $\operatorname{Spec} R$ that maps $(0) \in \operatorname{Spec} S$ to $(0) \in \operatorname{Spec} R$ and not to $(2) \in \operatorname{Spec} R$, so this map does not agree with the original map φ^*.

The previous example shows that the converse in Theorem 59 would not be true without the third (local homomorphism) condition in the definition of a morphism of affine schemes. As a result, Theorem 59 shows that the appropriate place to view affine schemes is in the category of *locally ringed spaces*. Roughly speaking, a locally ringed space is a topological space X together with a collection of rings $\mathcal{O}(U)$ for each open subset of X (with a compatible set of homomorphisms from $\mathcal{O}(U)$ to $\mathcal{O}(U')$ if $U' \subseteq U$ and with some local conditions on the sections) such that the stalks $\mathcal{O}_P = \varinjlim \mathcal{O}(U)$ for $P \in U$ are local rings. The morphisms in this category are continuous maps between the topological spaces together with ring homomorphisms between corresponding $\mathcal{O}(U)$ with precisely the same conditions as imposed in the definition of a morphism of affine schemes.

A *scheme* is a locally ringed space in which each point lies in a neighborhood isomorphic to an affine scheme (with some compatibility conditions between such neighborhoods), and is a fundamental object of study in modern algebraic geometry. The affine schemes considered here form the building blocks that are "glued together" to define general schemes in the same way that ordinary Euclidean spaces form the building blocks that are "glued together" to define manifolds in analysis.

EXERCISES

All rings are assumed commutative with identity, and all ring homomorphisms are assumed to map identities to identities.

1. If N is the nilradical of R, prove that $\operatorname{Spec} R$ and $\operatorname{Spec} R/N$ are homeomorphic. [Show that the natural homomorphism from R to R/N induces a Zariski continuous isomorphism from $\operatorname{Spec} R/N$ to $\operatorname{Spec} R$.]
2. Let I be an ideal in the ring R. Prove that the continuous map from $\operatorname{Spec} R/I$ to $\operatorname{Spec} R$ induced by the canonical projection homomorphism $R \to R/I$ maps $\operatorname{Spec} R/I$ homeomorphically onto the closed set $\mathcal{Z}(I)$ in $\operatorname{Spec} R$.

3. Prove that two elements $f, g \in R$ have the same values at all elements P in Spec R if and only if $f - g$ is contained in the nilradical of R. In particular, prove that an element in an affine k-algebra is uniquely determined by its values.

4. Let k be an arbitrary field, not necessarily algebraically closed. Prove that the prime ideals in $k[x, y]$ (i.e., the elements of Spec $k[x, y]$) are
   (i) (0),
   (ii) (f) where f is an irreducible polynomial in $k[x, y]$, and
   (iii) $(p(x), g(x, y))$ where $p(x)$ is an irreducible polynomial in $k[x]$ and $g(x, y)$ is an irreducible polynomial in $k[x, y]$ that is irreducible modulo $p(x)$, i.e., $g(x, y)$ remains irreducible in the quotient $k[x, y]/(p(x))$.

   Prove that mSpec $k[x, y]$ consists of the primes in (iii). [Use Exercise 20 in Section 1.]

5. Let $\mathfrak{m} = (p(x), g(x, y))$ be a maximal ideal in $k[x, y]$ as in the previous exercise. Show that $K = k[x, y]/\mathfrak{m}$ is an algebraic field extension of k, so that $k[x, y]$ can also be viewed as a subring of $K[x, y]$. If x, y are mapped to $\alpha, \beta \in K$, respectively, under the canonical homomorphism $k[x, y] \to k[x, y]/\mathfrak{m}$, prove that $\mathfrak{m} = k[x, y] \cap (x - \alpha, y - \beta) \subseteq K[x, y]$.

6. Describe the elements in Spec $\mathbb{R}[x]$ and Spec $\mathbb{C}[x]$. Describe the elements in Spec $\mathbb{Z}_{(2)}[x]$ where $\mathbb{Z}_{(2)} = \{a/b \in \mathbb{Q} \mid b \text{ is odd}\}$ is the localization of $\mathbb{Z}$ at the prime (2).

7. Let $(f) = (x^5 + x + 1)$ in Spec $\mathbb{Z}[x]$ viewed as fibered over Spec $\mathbb{Z}$ as in Example 3 following Proposition 55. Show that there are two closed points in the fiber over (2), three closed points in the fiber over (5), four closed points in the fiber over (19), and five closed points in the fiber over (211).

8. Let $(f) = (x^4 + 1)$ in Spec $\mathbb{Z}[x]$ viewed as fibered over Spec $\mathbb{Z}$ as in Example 3 following Proposition 55. Prove that there is one closed point in the fiber over (2), four closed points in the fiber over p for p odd, $p \equiv 1 \bmod 8$, and two closed points in the fiber over p for all other odd primes p (cf. Corollary 16 in Section 3 of Chapter 14).

9. Prove that the elements in the fiber over (p) of the Zariski continuous map from Spec $\mathbb{Z}[x]$ to Spec $\mathbb{Z}$ are homeomorphic with the elements in Spec$(\mathbb{Z}[x] \otimes_{\mathbb{Z}} \mathbb{F}_p)$.

10. Let $X = \operatorname{Spec} R$ and let X_f be the principal open set corresponding to $f \in R$. Prove that $X_f \cap X_g = X_{fg}$. Prove that $X_f = X$ if and only if f is a unit in R, and that $X_f = \emptyset$ if and only if f is nilpotent.

11. If X_f and X_g are principal open sets in $X = \operatorname{Spec} R$, prove that the open set $X_f \cup X_g$ is the complement of the closed set $\mathcal{Z}(I)$ where $I = (f, g)$ is the ideal in R generated by f and g.

12. Prove that a Zariski open subset U of $X = \operatorname{Spec} R$ is quasicompact if and only if U is a finite union of principal open subsets. Give an example of a ring R, a Zariski open subset U of Spec R, and a Zariski open covering of U that cannot be reduced to a finite subcovering.

13. Let $\varphi : R \to S$ be a homomorphism of rings. Prove that under the induced map φ^* from $Y = \operatorname{Spec} S$ to $X = \operatorname{Spec} R$ the full preimage of the principal open set X_f in X is the principal open set $Y_{\varphi(f)}$ in Y.

14. Suppose that $R = R_1 \times R_2$ is the direct product of the rings R_1 and R_2. Prove that $X = \operatorname{Spec} R$ is the disjoint union of open subspaces X_1, X_2 (which are therefore also closed), where X_1 is homeomorphic to Spec R_1 and X_2 is homeomorphic to Spec R_2.

15. Prove that $X = \operatorname{Spec} R$ is not connected if and only if R is the direct product of two nonzero rings if and only if R contains an idempotent e with $e \neq 0, 1$ (cf. the previous exercise).

16. Prove that $X = \operatorname{Spec} R$ is irreducible (i.e., any two nonempty open subsets have a nontrivial intersection) if and only if $X_f \cap X_g \neq \emptyset$ for any two nonempty principal open sets X_f and X_g. Deduce that $X = \operatorname{Spec} R$ is irreducible if and only if the nilradical of R is a prime ideal. [Use Exercise 10.]

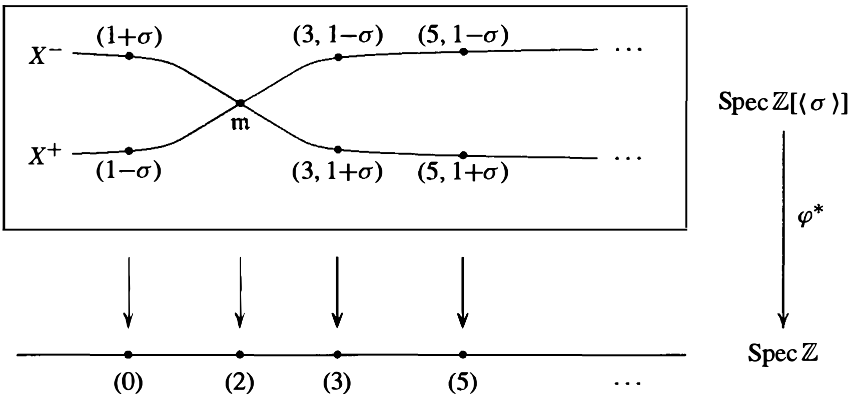
17. Let $G = \langle \sigma \rangle$ be a group of order 2, let $R = \mathbb{Z}[G] = \{a + b\sigma \mid a, b \in \mathbb{Z}\}$ be the corresponding group ring, and let $X = \operatorname{Spec} R$.

(a) Prove that the nilradical of R is (0) but is not a prime ideal. Prove that $X = X^+ \cup X^-$ where $X^+ = \mathcal{Z}(1 - \sigma)$ and $X^- = \mathcal{Z}(1 + \sigma)$. [Use $(1 + \sigma)(1 - \sigma) = 0$.]

(b) Prove that the homomorphism $\mathbb{Z}[G] \to \mathbb{Z}$ defined by mapping σ to 1 induces a homeomorphism of X^+ with $\operatorname{Spec} \mathbb{Z}$, and the homomorphism mapping σ to -1 induces a homeomorphism of X^- with $\operatorname{Spec} \mathbb{Z}$.

(c) Prove that $X^+ \cap X^-$ consists of the single element $\mathfrak{m} = (1 + \sigma, 1 - \sigma) = (2, 1 - \sigma)$ and that this is a closed point in X.

(d) Show that $(1 - \sigma)$ and $(1 + \sigma)$ are the unique non-closed points in X, with closures X^+ and X^-, respectively. Describe the closed points, mSpec R, in X and prove that $\operatorname{Spec} \mathbb{Z}[\langle \sigma \rangle]$ can be pictured as follows:



18. Let $\mathcal{O}$ be the structure sheaf on $X = \operatorname{Spec} R$, let U be an open set in X, and suppose $s, t \in \mathcal{O}(U)$. If $s = a/f_1^n$ on X_{f_1} and $t = b/f_2^m$ on X_{f_2}, show that

$$st = (abf_1^m f_2^n)/(f_1 f_2)^{n+m} \quad \text{and} \quad s + t = (af_1^m f_2^{m+n} + bf_1^{m+n} f_2^n)/(f_1 f_2)^{n+m}$$

on $X_{f_1 f_2}$. Deduce that $\mathcal{O}(U)$ is a commutative ring with identity.

19. Let $\mathcal{O}$ be the structure sheaf on $X = \operatorname{Spec} R$, let $V \subseteq U$ be open sets in X, and let $s \in \mathcal{O}(U)$. Suppose $P \in V$ and that $s = a/f^n$ on $X_f \subseteq U$.

(a) Show that there is a principal open set $X_{f'} \subseteq V \cap X_f$ containing P.

(b) Show that $(f')^m = bf$ for some $b \in R$.

(c) Show that $s = (ab^n)/(f')^{mn}$ on $X_{f'}$ and conclude that restricting s to V gives a well defined ring homomorphism from $\mathcal{O}(U)$ to $\mathcal{O}(V)$.

20. Let $\varphi : R \to S$ be a homomorphism of rings, let $X = \operatorname{Spec} R$, $Y = \operatorname{Spec} S$, and let $V \subseteq U$ be Zariski open subsets of X. Set $V' = (\varphi^*)^{-1}(V)$ and $U' = (\varphi^*)^{-1}(U)$, the corresponding Zariski open subsets of Y with respect to the continuous map $\varphi^* : Y \to X$ induced by φ. Prove that the induced map $\varphi^\# : \mathcal{O}_X(U) \to \mathcal{O}_Y(U')$ on sections is a ring homomorphism. Prove that $V' \subseteq U'$ and that $\varphi^\#$ is compatible with restriction i.e., that the diagram

$$\begin{array}{ccc} \mathcal{O}_X(U) & \xrightarrow{\varphi^\#} & \mathcal{O}_Y(U') \\ \downarrow & & \downarrow \\ \mathcal{O}_X(V) & \xrightarrow{\varphi^\#} & \mathcal{O}_Y(V') \end{array}$$

is commutative, where the vertical maps are the restriction homomorphisms.

21. Suppose D is a multiplicatively closed subset of R. Show that the localization homomorphism $R \to D^{-1}R$ induces a homeomorphism from $\operatorname{Spec}(D^{-1}R)$ to the collection of prime ideals P of R with $P \cap D = \emptyset$.

22. Show that $\operatorname{Spec} k[x, y]/(xy)$ is connected but is the union of two proper closed subsets each homeomorphic to $\operatorname{Spec} k[x]$, hence is not irreducible (cf. Exercise 16).

23. For each of the following rings R exhibit the elements of $\operatorname{Spec} R$, the open sets U in $\operatorname{Spec} R$, the sections $\mathcal{O}(U)$ of the structure sheaf for $\operatorname{Spec} R$ for each open U, and the stalks $\mathcal{O}_P$ at each point $P \in \operatorname{Spec} R$:
(a) $\mathbb{Z}/4\mathbb{Z}$ **(b)** $\mathbb{Z}/6\mathbb{Z}$ **(c)** $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z}$ **(d)** $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$.

24. **(a)** If every ideal of R is principal, show every open set in $\operatorname{Spec} R$ is a principal open set.
(b) Show that if $R = \mathbb{Z}[x]/(4, x^2)$ then R contains a nonprincipal ideal, but every open set in $\operatorname{Spec} R$ is a principal open set.

25. **(a)** If M is an R-module prove that $\operatorname{Supp}(M)$ is a Zariski closed subset of $\operatorname{Spec} R$. [Use Exercise 33 of Section 4.]
(b) If M is a finitely generated R-module prove that $\operatorname{Supp}(M) = \mathcal{Z}(\operatorname{Ann}(M)) \subseteq \operatorname{Spec} R$. [Use Exercise 34 of Section 4.]

26. Suppose M is a finitely generated module over the Noetherian ring R.
(a) Prove that there are finitely many minimal primes $*P_1, \dots, P_n$ containing $\operatorname{Ann}(M)$. [Use Corollary 22.]
(b) Prove that $\{P_1, \dots, P_n\}$ is also the set of minimal primes in $\operatorname{Ass}_R(M)$ and that $\operatorname{Supp}(M)$ is the union of the Zariski closed sets $\mathcal{Z}(P_1), \dots, \mathcal{Z}(P_n)$ in $\operatorname{Spec} R$. [Use the previous exercise and Exercise 40 in Section 4.]

The previous exercise gives a geometric view of a finitely generated module M over a Noetherian ring R: over each point P in $\operatorname{Spec} R$ is the localization M_P (the *stalk* over P). The stalk is nonzero precisely over the points in the Zariski closed subsets $\mathcal{Z}(P_1), \dots, \mathcal{Z}(P_n)$ where the P_i are the minimal primes in $\operatorname{Ass}_R(M)$. These ideas lead to the notion of the *(coherent) module sheaf on* Spec R *associated to* M (with a picture similar to that of the structure sheaf following Proposition 58), which is a powerful tool in modern algebraic geometry.

27. Let $R = k[x, y]$ and let M be the ideal (x, y) in R. Prove that $\operatorname{Supp}(M) = \operatorname{Spec} R$ and $\operatorname{Ass}_R(M) = \{0\}$.

The next two exercises show that the associated primes for an ideal I in a Noetherian ring R in the sense of primary decomposition are the associated primes for I in the sense of $\operatorname{Ass}_R(R/I)$.

28. This exercise proves that the ideal Q in a Noetherian ring R is P-primary if and only if $\operatorname{Ass}_R(R/Q) = \{P\}$.
(a) Suppose Q is a P-primary ideal and let M be the R-module R/Q. If $0 \neq m \in M$, show that $Q \subseteq \operatorname{Ann}(m) \subseteq P$ and that $\operatorname{rad}\operatorname{Ann}(m) = P$. Deduce that if $\operatorname{Ann}(m)$ is a prime ideal then it is equal to P and hence that $\operatorname{Ass}_R(R/Q) = \{P\}$. [Use Exercise 33 in Section 1.]

(b) For any ideal Q of R, let $0 \neq M \subseteq R/Q$. Prove that the radical of $\text{Ann}(M)$ is the intersection of the prime ideals in $\text{Supp}(M)$. [Use Proposition 12 and Exercise 25.]

(c) For M as in (b), prove that the radical of $\text{Ann}M$ is also the intersection of the prime ideals in $\text{Ass}_R(M)$. [Use Exercise 26(b).]

(d) If Q is an ideal of R with $\text{Ass}_R(R/Q) = \{P\}$ prove that $\text{rad}\, Q = P$. [Use the fact that $Q = \text{Ann}(R/Q)$ and (c).]

(e) If Q is an ideal of R with $\text{Ass}_R(R/Q) = \{P\}$ prove that Q is P-primary. [If $ab \in Q$ with $a \notin Q$ consider $0 \neq M = (Ra + Q)/Q \subseteq R/Q$ and show that b is contained in $\text{Ann}M \subseteq \text{rad}\,\text{Ann}(M)$. Use Exercises 33–34 in Section 1, to show that $\text{Ass}_R(M) = \{P\}$, then use (c) to show that $\text{rad}\,\text{Ann}(M) = P$, and conclude finally that $b \in P$.]

29. Suppose $I = Q_1 \cap \cdots \cap Q_n$ is a minimal primary decomposition of the ideal I in the Noetherian ring R with $P_i = \text{rad}\, Q_i$, $i = 1, \ldots, n$. This exercise proves that $\text{Ass}_R(R/I) = \{P_1, \ldots, P_n\}$.

(a) Prove that the natural projection homomorphisms induce an injection of R/I into $R/Q_1 \oplus \cdots \oplus R/Q_n$ and deduce that $\text{Ass}_R(R/I) \subseteq \{P_1, \ldots, P_n\}$. [Use Exercise 34 in Section 1 and the previous exercise.]

(b) Let $Q_i' = \cap_{j \neq i} Q_j$. Show that the minimality of the decomposition implies that $0 \neq Q_i'/I = (Q_i' + Q_i)/Q_i \subseteq R/Q_i$. Deduce that $\text{Ass}_R(Q_i'/I) = \{P_i\}$. [Use Exercises 33–34 in Section 1 and the previous exercise.] Deduce that $\{P_i\} \in \text{Ass}_R(R/I)$, so that $\text{Ass}_R(R/I) = \{P_1, \ldots, P_n\}$. [Use $Q_i'/I \subseteq R/I$ and Exercise 34 in Section 1.]

30. Let I be the ideal (x^2, xy, xz, yz) in $R = k[x, y, z]$. Prove that $\text{Ass}_R(R/I)$ consists of the primes $\{(x, y), (x, z), (x, y, z)\}$.

31. (Spec *for Quadratic Integer Rings)* Let R be the ring of integers in the quadratic field $K = \mathbb{Q}(\sqrt{D})$ where D is a squarefree integer and let P be a nonzero prime ideal in R. This exercise shows how the prime ideals in R are determined explicitly from the primes (p) in $\mathbb{Z}$, giving in particular a description of Spec R fibered over Spec $\mathbb{Z}$.

As in the discussion and example following Theorem 29, we have $R = \mathbb{Z}[\omega]$ where $\omega = \sqrt{D}$ if $D \equiv 2, 3 \bmod 4$ (respectively, $\omega = (1+\sqrt{D})/2$ if $D \equiv 1 \bmod 4$), with minimal polynomial $m_\omega(x) = x^2 - D$ (respectively, $m_\omega(x) = x^2 - x + (1-D)/4$), and $P \cap \mathbb{Z} = p\mathbb{Z}$ is a nonzero prime ideal of $\mathbb{Z}$.

(a) For any prime p in $\mathbb{Z}$ show that $R/pR \cong \mathbb{Z}[x]/(p, m_\omega(x)) \cong \mathbb{F}_p[x]/(\overline{m}_\omega(x))$ as rings, where $\overline{m}_\omega(x)$ is the reduction of $m_\omega(x)$ modulo p. Deduce that there is a prime ideal P in R with $P \cap \mathbb{Z} = (p)$ (this gives an alternate proof of Theorem 26(2) in this case).

(b) Use the isomorphism in (a) to prove that P is determined explicitly by the factorization of $m_\omega(x)$ modulo p:

(i) If $\overline{m}_\omega(x) \equiv (x-a)^2 \bmod p$ where $a \in \mathbb{Z}$ then $P = (p, \omega - a)$ and $pR = P^2$. Show that this case occurs only for the finitely many primes p dividing the discriminant of $m_\omega(x)$.

(ii) If $\overline{m}_\omega(x) \equiv (x-a)(x-b) \bmod p$ with integers $a, b \in \mathbb{Z}$ that are distinct modulo p then P is either $P_1 = (p, \omega - a)$ or $P_2 = (p, \omega - b)$ and P_1, P_2 are distinct prime ideals in R with $pR = P_1P_2$.

(iii) If $\overline{m}_\omega(x)$ is irreducible modulo p then $P = pR$.

(c) Show that the picture for Spec R over Spec $\mathbb{Z}$ for any D is similar to that for the case $R = \mathbb{Z}[i]$ when $D = -1$: there is precisely one nonclosed point $(0) \in$ Spec R over $(0) \in$ Spec $\mathbb{Z}$, precisely one closed point $P \in$ Spec R over each of the primes (p) in Spec $\mathbb{Z}$ in (i) (called *ramified* primes) and over the primes in (iii) (called *inert* primes), and precisely two closed points over the primes in (ii) (called *split* primes).

CHAPTER 16

Artinian Rings, Discrete Valuation Rings, and Dedekind Domains

Throughout this chapter R will denote a commutative ring with $1 \neq 0$.

16.1 ARTINIAN RINGS

In this section we shall study the basic theory of commutative rings that satisfy the descending chain condition (D.C.C.) on ideals, the Artinian rings (named after E. Artin). While one might at first expect that these rings have properties analogous to those for the commutative rings satisfying the ascending chain condition (the Noetherian rings), in fact this is not the case. The structure of Artinian rings is very restricted; for example an Artinian ring is necessarily also Noetherian (Theorem 3). Noncommutative Artinian rings play a central role in Representation Theory (cf. Chapters 18 and 19).

Definition. For any commutative ring R the *Krull dimension* (or simply the *dimension*) of R is the maximum possible length of a chain $P_0 \subset P_1 \subset P_2 \subset \cdots \subset P_n$ of distinct prime ideals in R. The dimension of R is said to be infinite if R has arbitrarily long chains of distinct prime ideals.

A ring with finite dimension must satisfy both the ascending and descending chain conditions on prime ideals (although not necessarily on all ideals). A field has dimension 0 and a Principal Ideal Domain that is not a field has dimension 1.

We shall see shortly that rings with D.C.C. on ideals always have dimension 0 (i.e., primes are maximal). If R is an integral domain that is also a finitely generated k–algebra over a field k, then the dimension of R is equal to the transcendence degree over k of the field of fractions of R (cf. Exercise 11). In particular, the Krull dimension agrees with the definition introduced earlier for the dimension of an affine variety. The advantage of the definition above is that it does not refer to any k–algebra structure and applies to arbitrary commutative rings R.

Definition. The *Jacobson radical* of R is the intersection of all maximal ideals of R and is denoted by Jac R.

The Jacobson radical is analogous to the Frattini subgroup of a group, and it enjoys some corresponding properties (cf. Exercise 24 in Section 6.1):

Proposition 1. Let $\mathcal{J}$ be the Jacobson radical of the commutative ring R.

(1) If I is a proper ideal of R, then so is $(I, \mathcal{J})$, the ideal generated by I and $\mathcal{J}$.

(2) The Jacobson radical contains the nilradical of R: $\operatorname{rad} 0 \subseteq \operatorname{Jac} R$.

(3) An element x belongs to $\mathcal{J}$ if and only if $1 - rx$ is a unit for all $r \in R$.

(4) *(Nakayama's Lemma)* If M is any finitely generated R-module and $\mathcal{J}M = M$, then $M = 0$.

Proof: If I is a proper ideal in R, then $I \subseteq M$ for some maximal ideal M. Since $\mathcal{J} \subseteq M$, also $(I, \mathcal{J}) \subseteq M$, which proves (1).

Part (2) follows from the definitions of the two radicals and Proposition 12 in Section 15.2 since maximal ideals are prime.

Suppose $1 - rx$ is not a unit and let M be a maximal ideal containing $1 - rx$. Since $1 \notin M, rx \notin M$, so x cannot belong to $\mathcal{J}$ because $\mathcal{J} \subseteq M$. Conversely, suppose $x \notin \mathcal{J}$, i.e., there is a maximal ideal M with $x \notin M$. Then $R = (x, M)$, hence $1 = rx + y$ for some $y \in M$. Thus $1 - rx = y \in M$ and so $1 - rx$ is not a unit, which proves (3).

To prove (4), assume $M \neq 0$ and let n be the smallest integer such that M is generated by n elements, say $m_1, \dots, m_n$. Since $M = \mathcal{J}M$ we have

$$m_n = r_1 m_1 + r_2 m_2 + \cdots + r_n m_n \qquad \text{for some } r_1, r_2, \dots, r_n \in \mathcal{J}.$$

Thus $(1 - r_n)m_n = r_1 m_1 + \cdots + r_{n-1} m_{n-1}$. By (3), $1 - r_n$ is a unit, so m_n lies in the module generated by $m_1, \dots, m_{n-1}$, contradicting the minimality of n. Hence $M = 0$, completing the proof.

Definition. A commutative ring R is said to be *Artinian* or to satisfy the *descending chain condition on ideals* (or *D.C.C. on ideals*) if there is no infinite decreasing chain of ideals in R, i.e., whenever $I_1 \supseteq I_2 \supseteq I_3 \supseteq \cdots$ is a decreasing chain of ideals of R, then there is a positive integer m such that $I_k = I_m$ for all $k \geq m$. Similarly, an R-module M is said to be Artinian if it satisfies D.C.C. on submodules.

It is immediate from the Lattice Isomorphism Theorem that every quotient R/I of an Artinian ring R by an ideal I is again an Artinian ring.

The following result for Artinian rings is parallel to results in Theorem 15.2. The proof is completely analogous, and so is left as an exercise.

Proposition 2. The following are equivalent:

(1) R is an Artinian ring.

(2) Every nonempty set of ideals of R contains a minimal element under inclusion.

The next result gives the main structure theorem for Artinian rings.

Theorem 3. Let R be an Artinian ring.

(1) There are only finitely many maximal ideals in R.

(2) The quotient $R/(\text{Jac } R)$ is a direct product of a finite number of fields. More precisely, if $M_1, \ldots, M_n$ are the finitely many maximal ideals in R then

$$R/(\text{Jac } R) \cong k_1 \times \cdots \times k_n,$$

where k_i is the field R/M_i for $1 \leq i \leq n$.

(3) Every prime ideal of R is maximal, i.e., R has Krull dimension 0. The Jacobson radical of R equals the nilradical of R and is a nilpotent ideal: $(\text{Jac } R)^m = 0$ for some $m \geq 1$.

(4) The ring R is isomorphic to the direct product of a finite number of Artinian local rings.

(5) Every Artinian ring is Noetherian.

Proof: To prove (1), let $\mathcal{S}$ be the set of all ideals of R that are the intersection of a finite number of maximal ideals. By Proposition 2, $\mathcal{S}$ has a minimal element, say $M_1 \cap M_2 \cap \cdots \cap M_n$. Then for any maximal ideal M we have

$$M \cap M_1 \cap M_2 \cap \cdots \cap M_n = M_1 \cap M_2 \cap \cdots \cap M_n,$$

so $M \supseteq M_1 \cap M_2 \cap \cdots \cap M_n$. By Exercise 11 in Section 7.4, $M \supseteq M_i$ for some i. Thus $M = M_i$ and so $M_1, \ldots, M_n$ are all the maximal ideals of R.

The proof of (2) is immediate from the Chinese Remainder Theorem (Section 7.6) applied to $M_1, \ldots, M_n$, since these maximal ideals are clearly pairwise comaximal and their intersection is Jac R.

For (3), we first prove $\mathcal{J} = \text{Jac } R$ is nilpotent. By D.C.C. there is some $m > 0$ such that $\mathcal{J}^m = \mathcal{J}^{m+i}$ for all positive i. By way of contradiction assume $\mathcal{J}^m \neq 0$. Let $\mathcal{S}$ be the set of proper ideals I such that $I\mathcal{J}^m \neq 0$, so $\mathcal{J} \in \mathcal{S}$. Let I_0 be a minimal element of $\mathcal{S}$. There is some $x \in I_0$ such that $x\mathcal{J}^m \neq 0$, so by minimality we must have $I_0 = (x)$. But now $((x)\mathcal{J})\mathcal{J}^m = x\mathcal{J}^{m+1} = x\mathcal{J}^m$, so it follows by minimality of (x) that $(x) = (x)\mathcal{J}$. By Nakayama's Lemma above, $(x) = 0$, a contradiction. This proves Jac R is nilpotent.

Since Jac R is nilpotent, in particular Jac $R \subseteq \text{rad } 0$, so these two ideals are equal by the second statement in Proposition 1.

Every prime ideal P in R contains the nilradical of R, hence contains Jac R by what has already been proved,. The image of P is a prime ideal in the quotient ring $R/(\text{Jac } R) = k_1 \times \cdots \times k_n$. But in a direct product of rings $R_1 \times R_2$ (where each R_i has a 1) every ideal is of the form $I_1 \times I_2$, where I_j is an ideal of R_j for $j = 1, 2$ (cf. Exercise 3 in Section 7.6). It follows that a prime ideal in $k_1 \times \cdots \times k_n$ consists of the elements that are 0 in one of the components. In particular, such a prime ideal is also a maximal ideal in $k_1 \times \cdots \times k_n$ and it follows that P was a maximal ideal in R, which finishes the proof of (3).

Let $M_1, \ldots, M_n$ be all the distinct maximal ideals of R and let $(\text{Jac } R)^m = 0$ as in (3). Then

$$\prod_{i=1}^{n} M_i^m \subseteq \left(\prod_{i=1}^{n} M_i\right)^m \subseteq (\text{Jac } R)^m = 0.$$

By the Chinese Remainder Theorem it follows that

$$R \cong (R/M_1^m) \times (R/M_2^m) \times \cdots \times (R/M_n^m),$$

and each R/M_i^m is an Artinian ring with unique maximal ideal M_i/M_i^m, proving (4).

To prove (5), it suffices by (4) to prove that an Artinian local ring is Noetherian, so assume R is Artinian with unique maximal ideal M. In this case we have $M = \operatorname{Jac} R$, so $M^m = (\operatorname{Jac} R)^m = 0$ for some positive m. Then $R \cong R/M^m$, and in this case it is an exercise to see that R/M^m is Noetherian if and only if it is Artinian (cf. Exercise 8).

Corollary 4. The ring R is Artinian if and only if R is Noetherian and has Krull dimension 0.

Proof: The forward implication was proved in Theorem 3. Suppose now that R is Noetherian and that R has Krull dimension 0, i.e., that prime ideals of R are maximal. Since R is Noetherian, by Corollary 22(3) in Section 15.2, the ideal $(0) = P_1 \cdots P_n$ is the product of (not necessarily distinct) prime ideals, and these prime ideals are then maximal since R has dimension 0. By the Chinese Remainder Theorem, R is isomorphic to the direct product of a finite number of Noetherian rings of the form R/M^m where M is a maximal ideal in R. As in the proof of (5) of the theorem, R/M^m is Artinian, and it follows that R is Artinian.

Examples

(1) Let $n > 1$ be an integer. Since the ring $R = \mathbb{Z}/n\mathbb{Z}$ is finite, it is Artinian. If $n = p_1^{a_1} p_2^{a_2} \cdots p_s^{a_s}$ is the unique factorization of n into distinct prime powers, then

$$\mathbb{Z}/n\mathbb{Z} \cong (\mathbb{Z}/p_1^{a_1}\mathbb{Z}) \times (\mathbb{Z}/p_2^{a_2}\mathbb{Z}) \times \cdots \times (\mathbb{Z}/p_s^{a_s}\mathbb{Z}).$$

Each $\mathbb{Z}/p_i^{a_i}\mathbb{Z}$ is an Artinian local ring with unique maximal ideal $(p_i)/(p_i^{a_i})$, so this is the decomposition of $\mathbb{Z}/n\mathbb{Z}$ given by Theorem 3(4). The Jacobson radical of R is the ideal generated by $p_1 p_2 \cdots p_s$, the squarefree part of n and $R/(\operatorname{Jac} R) \cong (\mathbb{Z}/p_1\mathbb{Z}) \times \cdots \times (\mathbb{Z}/p_s\mathbb{Z})$ is a direct product of fields. The ideals generated by p_i for $i = 1, \ldots, s$ are the maximal ideals of R.

(2) For any field k, a k-algebra R that is finite dimensional as a vector space over k is Artinian because ideals in R are in particular k-subspaces of R, hence the length of any chain of ideals in R is bounded by $\dim_k R$.

(3) Suppose f is a nonzero polynomial in $k[x]$ where k is a field. Then the quotient ring $R = k[x]/(f(x))$ is Artinian by the previous example. The decomposition of R as a direct product of Artinian local rings is given by

$$k[x]/(f(x)) \cong k[x]/(f_1(x)^{a_1}) \times \cdots \times k[x]/(f_s(x)^{a_s})$$

where $f(x) = f_1(x)^{a_1} \cdots f_s(x)^{a_s}$ is the factorization of $f(x)$ into powers of distinct irreducibles in $k[x]$ (cf. Proposition 16 in Section 9.5). The Jacobson radical of R is the ideal generated by the squarefree part of $f(x)$ and the maximal ideals of R are the ideals generated by the irreducible factors $f_i(x)$ for $i = 1, \ldots, s$ similar to Example 1.

EXERCISES

Let R be a commutative ring with 1 and let $\mathcal{J}$ be its Jacobson radical.

1. Suppose R is an Artinian ring and I is an ideal in R. Prove that R/I is also Artinian.
2. Show that every finite commutative ring with 1 is Artinian.
3. Prove that an integral domain of Krull dimension 0 is a field.
4. Prove that an Artinian integral domain is a field.
5. Suppose I is a nilpotent ideal in R and $M = IM$ for some R-module M. Prove that $M = 0$.
6. Suppose that $0 \longrightarrow M' \longrightarrow M \longrightarrow M'' \longrightarrow 0$ is an exact sequence of R-modules. Prove that M is an Artinian R-module if and only if M' and M'' are Artinian R-modules.
7. Suppose $R = F$ is a field. Prove that an R-module M is Artinian if and only if it is Noetherian if and only if M is a finite dimensional vector space over F.
8. Let M be a maximal ideal of the ring R and suppose that $M^n = 0$ for some $n \geq 1$. Prove that R is Noetherian if and only if R is Artinian. [Observe the each successive quotient $M^i/M^{i+1}, i = 0, \ldots, n-1$ in the filtration $R \supseteq M \supseteq \cdots \supseteq M^{n-1} \supseteq M^n = 0$ is a module over the field $F = R/M$. Then use the previous two exercises and Exercise 6 of Section 15.1.]
9. Let M be a finitely generated R-module. Prove that if $x_1, \ldots, x_n$ are elements of M whose images in $M/\mathcal{J}M$ generate $M/\mathcal{J}M$, then they generate M. Deduce that if R is Noetherian and the images of $a_1, \ldots, a_n$ in $\mathcal{J}/\mathcal{J}^2$ generate $\mathcal{J}/\mathcal{J}^2$, then $\mathcal{J} = (a_1, \ldots, a_n)$. [Let N be the submodule generated by $x_1, \ldots, x_n$ and apply Nakayama's Lemma to the module $A = M/N$.]
10. Let $R = \mathbb{Z}_{(2)}$ be the localization of $\mathbb{Z}$ at the prime ideal (2). Prove that $\operatorname{Jac} R = (2)$ is the ideal generated by 2. If $M = \mathbb{Q}$, prove that $M/2M$ is a finitely generated R-module but that M is not finitely generated over R. Why doesn't this contradict the previous exercise? [Note the hypotheses in Nakayama's Lemma.]
11. Let V be an affine variety over a field k and let $R = k[V]$ be its coordinate ring. Let $d_t(R)$ denote the transcendence degree of the field of fractions $k(V)$ over k, and let $d_p(R)$ be the Krull dimension of R defined in terms of chains of prime ideals. This exercise shows $d_t(R) = d_p(R)$. By Noether's Normalization Lemma there is a polynomial subring $R_1 = k[y_1, \ldots, y_m]$ of R such that R is integral over R_1.
    - (a) Show that $d_t(R_1) = d_t(R) = m$ and that $d_p(R_1) = d_p(R)$. Deduce that we may assume $R = R_1$. [Use the Going-up and Going-down Theorems (cf. Theorem 26, Section 15.3) to prove the second equality.]
    - (b) When $R = R_1$ show that $d_p(R) \geq d_t(R)$ by exhibiting an explicit chain of prime ideals of length m.
    - (c) When $R = R_1$ show that any nonzero prime ideal of R contains an element f such that $R(f)$ is transcendental over R of transcendence degree 1. Use induction to show that $d_p(R) \leq d_t(R)$, and deduce that $d_p(R) = d_t(R)$.
12. Let R be a Noetherian local ring with maximal ideal M.
    - (a) The quotient M/M^2 is a module (i.e., vector space) over the field R/M. Prove that $d = \dim_{R/M}(M/M^2)$ is finite.
    - (b) Prove that M can be generated as an ideal in R by d elements and by no fewer. [Use Exercise 9.]
    - (c) Let $R = k[x_1, \ldots, x_n]_{(x_1,\ldots,x_n)}$ be the localization of the polynomial ring $k[x_1, \ldots, x_n]$ over the field k at the maximal ideal $(x_1, \ldots, x_n)$, and let M be the maximal ideal in

R. Prove that $\dim_{R/M}(M/M^2) = n = \dim R$. [Cf. the previous exercise.]

It can be shown that $\dim_{R/M}(M/M^2) \geq \dim R$ for any Noetherian local ring R with maximal ideal M. A Noetherian local ring R is called a *regular local ring* if $\dim_{R/M}(M/M^2) = \dim R$. It is a fact that a regular local ring is necessarily an integral domain and is also integrally closed.

13. If R is a Noetherian ring, prove that the Zariski topology on Spec R is discrete (i.e., every subset is Zariski open and also Zariski closed) if and only if R is Artinian.

14. Suppose I is the ideal $(x_1, x_2^2, x_3^3, \dots)$ in the polynomial ring $k[x_1, x_2, x_3, \dots]$ where k is a field and let R be the quotient ring $k[x_1, x_2, x_3, \dots]/I$. Prove that the image of the ideal $(x_1, x_2, x_3, \dots)$ in R is the unique prime ideal in R but is not finitely generated. Deduce that R is a local ring of Krull dimension 0 but is not Artinian.

16.2 DISCRETE VALUATION RINGS

In the previous section we showed that the Artinian rings are the Noetherian rings having Krull dimension 0. We now consider the easiest Noetherian rings of dimension 1, the Discrete Valuation Rings first introduced in Section 8.1:

Definition.

(1) A *discrete valuation* on a field K is a function $\nu : K^\times \to \mathbb{Z}$ satisfying

- **(i)** ν is surjective,
- **(ii)** $\nu(xy) = \nu(x) + \nu(y)$ for all $x, y \in K^\times$,
- **(iii)** $\nu(x + y) \geq \min\{\nu(x), \nu(y)\}$ for all $x, y \in K^\times$ with $x + y \neq 0$.

The subring $\{x \in K \mid \nu(x) \geq 0\} \cup \{0\}$ is called the *valuation ring* of ν.

(2) An integral domain R is called a *Discrete Valuation Ring* (D.V.R.) if R is the valuation ring of a discrete valuation ν on the field of fractions of R.

The valuation ν is often extended to all of K by defining $\nu(0) = +\infty$, in which case (ii) and (iii) hold for all $a, b \in K$.

Examples

(1) The localization $\mathbb{Z}_{(p)}$ of $\mathbb{Z}$ at any nonzero prime ideal (p) is a D.V.R. with respect to the discrete valuation ν_p on $\mathbb{Q}$ defined as follows (cf. Exercise 27, Section 7.1). Every element $a/b \in \mathbb{Q}^\times$ can be written uniquely in the form $p^n(a_1/b_1)$ where $n \in \mathbb{Z}$, $a_1/b_1 \in \mathbb{Q}^\times$ and both a_1 and b_1 are relatively prime to p. Define

$$\nu_p\left(\frac{a}{b}\right) = \nu_p\left(p^n\frac{a_1}{b_1}\right) = n.$$

One easily checks that the axioms for a D.V.R. are satisfied. We call ν_p the *p-adic valuation* on $\mathbb{Q}$. The corresponding valuation ring is the set of rational numbers with $n \geq 0$ together with 0, i.e., the rational numbers a/b where b is not divisible by p, which is $\mathbb{Z}_{(p)}$.

(2) For any field F, let f be an irreducible polynomial in $F[x]$. Every nonzero element in the field $F(x)$ can be written uniquely in the form $f^n(a/b)$ where $n \in \mathbb{Z}$, $a/b \in F[x]^\times$ and both a and b are relatively prime to f. Then

$$\nu_f\left(f^n\frac{a}{b}\right) = n$$

defines a valuation on $F(x)$ and the corresponding valuation ring is the localization $F[x]_f$ of $F[x]$ at f consisting of the rational functions in $F(x)$ whose denominator is not divisible by f. When $f = x - \alpha$ is a polynomial of degree 1 in $F[x]$, the valuation ν_f gives the *order of the zero* (if $n \geq 0$) or *pole* (if $n < 0$) of the element in $F(x)$ at $x = \alpha$.

(3) The ring of formal Laurent series $F((x))$ with coefficients in the field F has a discrete valuation ν defined by

$$\nu\left(\sum_{i \geq n}^{\infty} a_i x^i\right) = n$$

(cf. Exercise 5, Section 7.2). The corresponding D.V.R. is the ring $F[[x]]$ of power series in x with coefficients in F.

Note that $\nu(1) = \nu(1) + \nu(1)$ implies that $\nu(1) = 0$, so every Discrete Valuation Ring R is a ring with identity $1 \neq 0$. Since R is a subring of a field by definition, R is in particular an integral domain. It is easy to see that a D.V.R. is a Euclidean Domain (cf. Example 4 in Section 8.1), so in particular is also a P.I.D. and a U.F.D. In fact the factorization and ideal structure of a D.V.R. is very simple, as the next proposition shows.

Proposition 5. Suppose R is a Discrete Valuation Ring with respect to the valuation ν, and let t be any element of R with $\nu(t) = 1$. Then

(1) A nonzero element $u \in R$ is a unit if and only if $\nu(u) = 0$.

(2) Every nonzero element $r \in R$ can be written in the form $r = ut^n$ for some unit $u \in R$ and some $n \geq 0$. Every nonzero element x in the field of fractions of R can be written in the form $x = ut^n$ for some unit $u \in R$ and some $n \in \mathbb{Z}$.

(3) Every nonzero ideal of R is a principal ideal of the form (t^n) for some $n \geq 0$. In particular, R is a Noetherian ring.

Proof: If u is a unit, then $uv = 1$ for some $v \in R$ and then $\nu(u)+\nu(v) = \nu(uv) = 1$ with $\nu(u) \geq 0$ and $\nu(v) \geq 0$ shows that $\nu(u) = 0$. Conversely, if u is nonzero and $\nu(u) = 0$ then $u^{-1} \in K$ satisfies $\nu(u^{-1}) + \nu(u) = \nu(1) = 0$. Hence $\nu(u^{-1}) = 0$ and $u^{-1} \in R$, so u is a unit. This proves (1).

For (2), note that if $\nu(x) = n$ then $\nu(xt^{-n}) = 0$, so $xt^{-n} = u$ is a unit in R by (1). Hence $x = ut^n$, where $x \in R$ if and only if $n = \nu(x) \geq 0$.

If I is a nonzero ideal in R, let $r \in I$ be an element with $\nu(r)$ minimal. If $\nu(r) = n$, then r differs from t^n by a unit by (2), so $t^n \in I$ and $(t^n) \subseteq I$. If now a is any nonzero element of I, then $\nu(a) \geq n$ by choice of n. Then $\nu(at^{-n}) \geq 0$ and so $at^{-n} \in R$, which shows that $a \in (t^n)$. Hence $I = (t^n)$, proving the first statement in (3). It is then clear that ascending chains of ideals in R are finite, proving that R is Noetherian and completing the proof.

Definition. If R is a D.V.R. with valuation ν, then an element t of R with $\nu(t) = 1$ is called a *uniformizing* (or *local*) *parameter* for R.

Corollary 6. Let R be a Discrete Valuation Ring.

(1) The ring R is an integrally closed local ring with unique maximal ideal given by the elements with strictly positive valuation: $M = \{r \in R \mid \nu(r) > 0\}$. Every nonzero ideal in R is of the form M^n for some integer $n \geq 0$.

(2) The only prime ideals of R are M and 0, i.e., Spec $R = \{0, M\}$. In particular, a D.V.R. has Krull dimension 1.

Proof: Any U.F.D. is integrally closed in its fraction field (Example 3 in Section 15.3), so R is integrally closed. The remainder of the statements follow immediately from the description of the ideals of R in Proposition 5.

The definition of a Discrete Valuation Ring is extremely explicit in terms of a valuation on the fraction field, and as a result it appears that it might be difficult to recognize whether a given ring R is a D.V.R. from purely "internal" algebraic properties of R. In fact, the ring-theoretic properties in Proposition 5 and Corollary 6 characterize Discrete Valuation Rings. The following theorem gives several alternate algebraic descriptions of Discrete Valuation Rings in which there is no explicit mention of the valuation.

Theorem 7. The following properties of a ring R are equivalent:

(1) R is a Discrete Valuation Ring,

(2) R is a P.I.D. with a unique maximal ideal $P \neq 0$,

(3) R is a U.F.D. with a unique (up to associates) irreducible element t,

(4) R is a Noetherian integral domain that is also a local ring whose unique maximal ideal is nonzero and principal,

(5) R is a Noetherian, integrally closed, integral domain that is also a local ring of Krull dimension 1 i.e., R has a unique nonzero prime ideal: Spec $R = \{0, M\}$.

Proof: That (1) implies each of the other properties was proved above.

If (2) holds then (3) is immediate since irreducible elements generate prime ideals in a U.F.D. (Proposition 12, Section 8.3).

If (3) holds, then every nonzero element in R can be written uniquely in the form ut^n for some unit u and some $n \geq 0$. Then every nonzero element in the fraction field of R can be written uniquely in the form ut^n for some unit u and some $n \in \mathbb{Z}$. It is now straightforward to check that the map $\nu(ut^n) = n$ is a discrete valuation on the field of fractions of R, and R is the valuation ring of ν, and (1) holds.

Suppose (4) holds, let $M = (t)$ be the unique maximal ideal of R, and let $M_0 = \cap_{i=1}^{\infty} M^i$. Then $M_0 = MM_0$, and since R is Noetherian M_0 is finitely generated. By hypothesis $M = \text{Jac } R$, so by Nakayama's Lemma $M_0 = 0$. If I is any proper, nonzero ideal of R then there is some $n \geq 0$ such that $I \subseteq M^n$ but $I \not\subseteq M^{n+1}$. Let $a \in I - M^{n+1}$ and write $a = t^n u$ for some $u \in R$. Then $u \notin M$, and so u is a unit in the local ring R. Thus $(a) = (t^n) = M^n$ for every $a \in I - M^{n+1}$. This shows that $I = (t^n)$, and so every ideal of R is principal, which shows that (2) holds.

We have shown that (1), (2), (3) and (4) are equivalent, and that each of these implies (5). To complete the proof we show that (5) implies (4), which amounts to showing that the ideal M in (5) is a principal ideal. Since $0 \neq M = \text{Jac } R$ and M is finitely generated because R is Noetherian, by Nakayama's Lemma (Proposition 1(4)), $M \neq M^2$. Let $t \in M - M^2$. We argue that $M = (t)$. By Proposition 12 in Section 15.2, the assumption that M is the unique nonzero prime ideal in R implies that $M = \operatorname{rad}(t)$, and then Proposition 14 in Section 15.2 implies that some power of M is contained in (t). Proceeding by way of contradiction, assume $(t) \neq M$, so that $M^n \subseteq (t)$ but $M^{n-1} \not\subseteq (t)$ for some $n \geq 2$. Then there is an element $x \in M^{n-1} - (t)$ such that $xM \subseteq (t)$. Note that $t \neq 0$ so $y = x/t$ belongs to the field of fractions of R. Also, $y \notin R$ because $x = ty \notin (t)$. However, by choice of x we have $yM \subseteq R$, and then one checks that yM is an ideal in R. If $yM = R$ then $1 = ym$ for some $m \in M$. This leads to a contradiction because we would then have $t = xm \in M^2$, contrary to the choice of t. Thus yM is a proper ideal, hence is contained in the unique maximal ideal of R, namely $yM \subseteq M$. Now M is a finitely generated R-module on which y acts by left multiplication as an R-module homomorphism. By the same (determinant) method as in the proof of Proposition 23 in Section 15.3 there is a monic polynomial p with coefficients in R such that $p(y)m = 0$ for all $m \in M$. Since $p(y)$ is an element of a field containing R and M, we must have $p(y) = 0$. Hence y is integral over R. Since R is integrally closed by assumption, it follows that $y \in R$, a contradiction. Hence $M = (t)$ is principal, so (5) implies (4), completing the proof of the theorem.

Corollary 8. If R is any Noetherian, integrally closed, integral domain and P is a minimal nonzero prime ideal of R, then the localization R_P of R at P is a Discrete Valuation Ring.

Proof: By results in Section 15.4, the localization R_P is a Noetherian (Proposition 38(4)), integrally closed (Proposition 49), integral domain (Proposition 46(2)), that is a local ring with unique nonzero prime ideal (Proposition 46(4)), so R_P satisfies (5) in the theorem.

Examples

(1) If R is any Principal Ideal Domain then every localization R_P of R at a nonzero prime ideal $P = (p)$ is a Discrete Valuation Ring. This follows immediately from Corollary 8 since R is integrally closed (being a U.F.D., cf. Example 3 in Section 15.3) and nonzero prime ideals in a P.I.D. are maximal (Proposition 8.7). Note that the quotient field K of R_P is the same as the quotient field of R, so each nonzero prime p in R produces a valuation ν_p on K, given by the formula

$$\nu\left(p^n \frac{a}{b}\right) = n$$

where a and b are elements of R not divisible by p. This generalizes both Examples 1 and 2 above.

(2) The ring $\mathbb{Z}_p$ of p-adic integers is a Discrete Valuation Ring since it is a P.I.D. with unique maximal ideal $p\mathbb{Z}_p$ (cf. Exercise 11, Section 7.6). The fraction field of $\mathbb{Z}_p$ is called the *field of p-adic numbers* and is denoted $\mathbb{Q}_p$. The element p is a uniformizing parameter for $\mathbb{Z}_p$, so every nonzero element in $\mathbb{Q}_p$ can be written uniquely in the form $p^n u$ for some $n \in \mathbb{Z}$ and unit $u \in \mathbb{Z}_p^\times$, (where $u = a_0 + a_1 p + a_2 p^2 + \dots$ with $0 < a_0 < p$ as in Exercise 11(c), Section 7.6). The corresponding *p-adic valuation* ν_p on $\mathbb{Q}_p$ is then given by $\nu_p(p^n u) = n$.

A discrete valuation ν on a field K defines an associated *metric* (or "distance function"), d_ν, on K as follows: fix any real number $\beta > 1$ (the actual value of β does not matter for verifying the axioms of a metric), and for all $a, b \in K$ define

$$d_\nu(a, b) = ||a - b||_\nu \qquad \text{where} \qquad ||a||_\nu = \beta^{-\nu(a)}$$

and where we set $d_\nu(a, a) = 0$. It is easy to check that d_ν satisfies the three axioms for a metric:

(i) $d_\nu(a, b) \geq 0$, with equality holding if and only if $a = b$,
(ii) $d_\nu(a, b) = d_\nu(b, a)$, i.e., d_ν is symmetric,
(iii) $d_\nu(a, b) \leq d_\nu(a, c) + d_\nu(c, b)$, for all $a, b, c \in K$, i.e., d_ν satisfies the "triangle inequality."

The triangle inequality is a consequence of axiom (iii) of the discrete valuation. Indeed, a stronger version of the triangle inequality holds:

(iii)′ $d_\nu(a, b) \leq \max\{d_\nu(a, c), d_\nu(c, b)\}$, for all $a, b, c \in K$.

For this reason d_ν is sometimes called an *ultrametric*. One may now use Cauchy sequences to form the *completion* of K with respect to d_ν, denoted by K_ν, in the same way that the real numbers $\mathbb{R}$ are constructed from the rational numbers $\mathbb{Q}$. It is not difficult to show that K_ν is also a field with a discrete valuation that agrees with ν on the dense subset K of K_ν.

Examples

(1) Consider the p-adic valuation ν_p on $\mathbb{Q}$ and take $\beta = p$. Write $||a||_p$ for $||a||_{\nu_p}$, so that for a, b relatively prime to p,

$$||p^n \frac{a}{b}||_p = p^{-n}.$$

Note that integers (or rational numbers) have small p-adic absolute value if they are divisible by a large power of p. For example, the sequence $1, p, p^2, p^3, \ldots$ converges to zero in the p-adic metric.

It is not too difficult to see that the completion of $\mathbb{Q}$ with respect to the p-adic metric is the field $\mathbb{Q}_p$ of p-adic numbers, and the completion of $\mathbb{Z}$ is the ring $\mathbb{Z}_p$ of p-adic integers. One way to see this is to check that each element a of the completion may be represented as a *p-adic Laurent series*:

$$a = \sum_{n=n_0}^{\infty} a_i p^i \qquad \text{where } n_0 \in \mathbb{Z} \text{ and } a_i \in \{0, 1, \ldots, p-1\} \text{ for all } i,$$

and then use Example 2 previously. In terms of this expansion, the p-adic valuation is given by $\nu_p(a) = n_0$ (when $a_{n_0} \neq 0$).

(2) In a similar way, the completion of $F(x)$ with respect to the valuation ν_x in Example 2 at the beginning of this section gives the field $F((x))$ with corresponding valuation ring $F[[x]]$ in Example 3 in the same set of examples.

The completion of a field K with respect to a discrete valuation ν is a field K_ν in which the elements can be easily described in terms of a uniformizing parameter. In addition, K_ν is a topological space where the topology is defined by the metric d_ν. Furthermore, Cauchy sequences of elements in K_ν converge to elements of K_ν (i.e., K_ν

is *complete* in the ν-adic topology). This is similar to the situation of the completion $\mathbb{R}$ of $\mathbb{Q}$ with respect to the usual Euclidean metric. This allows the application of ideas from analysis to the study of such rings, and is an important tool in the study of algebraic number fields and in algebraic geometry.

Fractional Ideals

We complete our discussion of Discrete Valuation Rings by giving another characterization of D.V.R.s in terms of "fractional ideals," which can be defined for any integral domain:

Definition. For any integral domain R with fraction field K, a *fractional ideal* of R is an R-submodule A of K such that $dA \subseteq R$ for some nonzero $d \in R$ (equivalently, a submodule of the form $d^{-1}I$ for some nonzero $d \in R$ and ideal I of R).

The equivalence of these two definitions follows from the observation that dA is an R-submodule (i.e., an ideal) of R.

The notion of a fractional ideal in K depends on the ring R. Loosely speaking, a fractional ideal is an ideal of R up to a fixed "denominator" d. The ideals of R are also fractional ideals of R (with denominator $d = 1$) and are the fractional ideals that are contained in R. For clarity these are occasionally called the *integral ideals* of R. When R is a Noetherian integral domain, a fractional ideal of R is the same as a finitely generated R-submodule of K (cf. Exercise 6).

For any $x \in K$ the (cyclic) R-module $Rx = \{rx \mid r \in R\}$ is called the *principal fractional ideal* generated by x.

If A and B are fractional ideals, their product, AB, is defined to be the set of all finite sums of elements of the form ab where $a \in A$ and $b \in B$. If $A = d^{-1}I$ and $B = (d')^{-1}J$ for ideals I, J in R and nonzero $d, d' \in R$, then $AB = (dd')^{-1}IJ$ where IJ is the usual product ideal. In particular, this shows that the product of two fractional ideals is a fractional ideal.

Definition. The fractional ideal A is said to be *invertible* if there exists a fractional ideal B with $AB = R$, in which case B is called the *inverse* of A and denoted A^{-1}.

If A is an invertible fractional ideal, the fractional ideal B with $AB = R$ is unique: $AB = AC = R$ implies $B = B(AC) = (BA)C = C$.

Proposition 9. Let R be an integral domain and let A be a fractional ideal of R.

(1) If A is a nonzero principal fractional ideal then A is invertible.

(2) If A is nonzero then the set $A' = \{x \in K \mid xA \subseteq R\}$ is a fractional ideal of R. In general we have $AA' \subseteq R$ and $AA' = R$ if and only if A is invertible, in which case $A^{-1} = A'$.

(3) If A is an invertible fractional ideal of R then A is finitely generated.

(4) The set of invertible fractional ideals is an abelian group under multiplication with identity R. The set of nonzero principal fractional ideals is a subgroup of the invertible fractional ideals.

Proof: If $A = xR$ is a nonzero principal fractional ideal, then taking $B = x^{-1}R$ shows that A is invertible, proving (1).

One easily sees that A' is an R-submodule of K. If A is a nonzero fractional ideal there is some nonzero element $d \in R$ such that $dA \subseteq R$, so A contains nonzero elements of R. Let a be any nonzero element of A contained in R. Then by definition of A' we have $aA' \subseteq R$, so A' is a fractional ideal. Also by definition, $AA' \subseteq R$. If $AA' = R$ then A is invertible with inverse $A^{-1} = A'$. Conversely, if $AB = R$, then $B \subseteq A'$ by definition of A'. Then $R = AB \subseteq AA' \subseteq R$, showing that $AA' = R$, proving (2).

If A is invertible, then $AA' = R$ by (2) and so $1 = a_1a_1' + \cdots + a_na_n'$ for some $a_1, \ldots, a_n \in A$ and $a_1', \ldots, a_n' \in A'$. If $a \in A$, then $a = (aa_1')a_1 + \cdots + (aa_n')a_n$, where each $aa_i' \in R$ by definition of A'. It follows that A is generated over R by $a_1, \ldots, a_n$ and so A is finitely generated, proving (3).

Finally, it is clear that the product of two invertible fractional ideals is again invertible. This product is commutative, associative, and $RA = A$ for any fractional ideal. The inverse of an invertible fractional ideal is an invertible fractional ideal by definition, proving the first statement in (4). The second statement in (4) is immediate since the product of xR and yR is $(xy)R$ and the inverse of xR is $x^{-1}R$.

Definition. If R is an integral domain, then the quotient of the group of invertible fractional ideals of R by the subgroup of nonzero principal fractional ideals of R is called the *class group* of R. The order of the class group of R is called the *class number* of R.

The class group of R is the trivial group and the class number of R is 1 if and only if R is a P.I.D. The class group of R measures how close the ideals of R are to being principal.

Whether a fractional ideal A of R is invertible is also related to whether A is *projective* as an R-module. Recall that an R-module M is projective over R if and only if M is a direct summand of a free module (Proposition 30, Section 10.5). Equivalently, M is projective if and only if there is a free R-module F and R-module homomorphisms $f : F \to M$ and $g : M \to F$ with $f \circ g = 1$ (Proposition 25, Section 10.5).

Proposition 10. Let R be an integral domain with fraction field K and let A be a nonzero fractional ideal of R. Then A is invertible if and only if A is a projective R-module.

Proof: Assume first that A is invertible, so $\sum_{i=1}^{n} a_ia_i' = 1$ for some $a_i \in A$ and $a_i' \in A'$ as in (2) of Proposition 9. Let F be the free R-module on $y_1, \ldots, y_n$. Define $f : F \to A$ by $f(\sum_{i=1}^{n} r_iy_i) = \sum_{i=1}^{n} r_ia_i$ and $g : A \to F$ by $f(c) = \sum_{i=1}^{n}(ca_i')y_i$. It is immediate that both f and g are R-module homomorphisms (note that $ca_i' \in R$ by definition of A'). Since

$$(f \circ g)(c) = f\left(\sum_{i=1}^{n}(ca_i')y_i\right) = \sum_{i=1}^{n}(ca_i')a_i = c\left(\sum_{i=1}^{n} a_ia_i'\right) = c,$$

so $f \circ g = 1$ and A is a direct summand of F, hence is projective.

Conversely, suppose that A is nonzero and projective, so there is a free R-module F and R-homomorphisms $f : F \to A$ and $g : A \to F$ with $f \circ g = 1$. Fix any $0 \neq a \in A$ and suppose $g(a) = \sum_{i=1}^{n} \tilde{a}_i y_i$ where $\tilde{a}_i \in R$ and $y_1, \ldots, y_n$ is part of a set of free generators for F. Define $a_i = f(y_i)$ and $a_i' = \tilde{a}_i / a \in K$ for $i = 1, \ldots, n$. For any $b \in A$ we have $bg(a) = ag(b) = g(ab)$ since g is an R-module homomorphism. Write $g(b) = \sum_{i=1}^{n} \tilde{b}_i y_i + \sum_{j \in \mathcal{J}} \tilde{b}_j y_j$ where $\{y_j\}$ for $j \in \mathcal{J}$ are the remaining elements in the set of free generators for F. Then

$$\sum_{i=1}^{n} (b\tilde{a}_i) y_i = \sum_{i=1}^{n} (a\tilde{b}_i) y_i + \sum_{j \in \mathcal{J}} (a\tilde{b}_j) y_j .$$

We may equate coefficients of the elements in the free R-module basis for F in this equation and it follows that $g(b) = \sum_{i=1}^{n} \tilde{b}_i y_i$ where $\tilde{b}_i \in R$ and that $b\tilde{a}_i = a\tilde{b}_i$ for $i = 1, \ldots, n$. In particular, it follows from the definition of a_i' that $ba_i' = b(\tilde{a}_i / a) = \tilde{b}_i$ is an element of R for every element b of A. This shows that $a_i' \in A'$ for $i = 1, \ldots, n$. Since $f \circ g = 1$, we have

$$a = f \circ g(a) = f\left(\sum_{i=1}^{n} \tilde{a}_i y_i\right) = \sum_{i=1}^{n} \tilde{a}_i a_i = \sum_{i=1}^{n} (aa_i') a_i = a\left(\sum_{i=1}^{n} a_i a_i'\right) ,$$

and so $\sum_{i=1}^{n} a_i a_i' = 1$. It follows that $AA' = R$ and so A is invertible by Proposition 9, completing the proof.

The next result shows that if the integral domain R is also a local ring, then whether fractional ideals are invertible determines whether R is a D.V.R.

Proposition 11. Suppose the integral domain R is a local ring that is not a field. Then R is a Discrete Valuation Ring if and only if every nonzero fractional ideal of R is invertible.

Proof: If R is a D.V.R. with uniformizing parameter t, then by Proposition 5 every nonzero ideal of R is of the form (t^n) for some $n \geq 0$ and every element d in R can be written in the form ut^m for some unit $u \in R$ and some $m \geq 0$. It follows that every nonzero fractional ideal of R is of the form $t^N R$ for some $N \in \mathbb{Z}$, so is a principal fractional ideal and hence invertible by the previous proposition.

Conversely, suppose that every nonzero fractional ideal of R is invertible. Then every nonzero ideal of R is finitely generated by (3) of Proposition 9, so R is Noetherian. Let M be the unique maximal ideal of R. If $M = M^2$ then $M = 0$ by Nakayama's Lemma, and then R would be a field, contrary to hypothesis. Hence there is an element t with $t \in M - M^2$. By assumption M is invertible, and since $t \in M$, the fractional ideal tM^{-1} is a nonzero ideal in R. If $tM^{-1} \subseteq M$, then $t \in M^2$, contrary to the choice of t. Hence $tM^{-1} = R$, so $(t) = M$, and M is a nonzero principal ideal. It follows by the equivalent condition 4 of Theorem 7 that R is a D.V.R., completing the proof.

We end this section with an application to algebraic geometry.

Nonsingularity and Local Rings of Affine Plane Curves

Let k be an algebraically closed field and let C be an irreducible affine *curve* over k. In other words, C is an affine algebraic set whose coordinate ring $k[C]$ is an integral domain and whose field of rational functions $k(C)$ has transcendence degree 1 over k (cf. Section 15.4).

Recall that, by definition, the point v on C is nonsingular if $\mathfrak{m}_{v,C}/\mathfrak{m}_{v,C}^2$ is a 1-dimensional vector space over k, where $\mathfrak{m}_{v,C}$ is the unique maximal ideal in the local ring $\mathcal{O}_{v,C}$ of rational functions on C defined at v.

Proposition 12. Let v be a point on the irreducible affine curve C over k. Then C is nonsingular at v if and only if the local ring $\mathcal{O}_{v,C}$ is a Discrete Valuation Ring.

Proof: Suppose first that v is nonsingular. Then $\dim_k(\mathfrak{m}_{v,C}/\mathfrak{m}_{v,C}^2) = 1$, and since $\mathcal{O}_{v,C}$ is Noetherian, it follows from Exercise 12 in Section 1 that $\mathfrak{m}_{v,C}$ is principal. Hence $\mathcal{O}_{v,C}$ is a D.V.R. by Theorem 7(4). Conversely, suppose $\mathcal{O}_{v,C}$ is a D.V.R. and t is a uniformizing element for $\mathcal{O}_{v,C}$. Then every element in $\mathfrak{m}_{v,C}$ can be written uniquely in the form at for some a in $\mathcal{O}_{v,C}$. The map from $\mathfrak{m}_{v,C}$ to $\mathcal{O}_{v,C}/\mathfrak{m}_{v,C}$ defined by mapping at to $a \bmod \mathfrak{m}_{v,C}$ is easily checked to be a surjective $\mathcal{O}_{v,C}$-module homomorphism with kernel $\mathfrak{m}_{v,C}^2$. Hence $\mathfrak{m}_{v,C}/\mathfrak{m}_{v,C}^2$ is isomorphic as an $\mathcal{O}_{v,C}/\mathfrak{m}_{v,C}$-module to $\mathcal{O}_{v,C}/\mathfrak{m}_{v,C}$. Since $\mathcal{O}_{v,C}/\mathfrak{m}_{v,C} \cong k$ (Proposition 46(5) in Section 15.4), it follows that $\dim_k(\mathfrak{m}_{v,C}/\mathfrak{m}_{v,C}^2) = 1$, and so v is a nonsingular point on C.

Definition. If v is a nonsingular point on C with corresponding discrete valuation ν_v defined on $k(C)$, then $\nu_v(f) = n$ for $f \in k(V)$ is the *order of zero of f at v* (if $n \geq 0$) or the *order of the pole of f at v* (if $n < 0$).

Using the criterion for nonsingularity for points on curves in Proposition 12 we can prove a result first mentioned in Section 15.4:

Corollary 13. An irreducible affine curve C over an algebraically closed field k is smooth if and only if its coordinate ring $k[C]$ is integrally closed.

Proof: The curve C is smooth if and only if every localization $\mathcal{O}_{v,C}$ is a D.V.R. Since $k[C]$ has Krull dimension 1 (Exercise 11 in Section 1), the same is true for each $\mathcal{O}_{v,C}$. It then follows by Theorem 7(5) that every localization $\mathcal{O}_{v,C}$ is a D.V.R. if and only if $\mathcal{O}_{v,C}$ is integrally closed. By Proposition 49 in Section 15.4, this in turn is equivalent to the statement that $k[C]$ is integrally closed, which proves the corollary.

EXERCISES

1. Suppose R is a Discrete Valuation Ring with respect to the valuation ν on the fraction field K of R. If $x, y \in K$ with $\nu(x) < \nu(y)$ prove that $\nu(x + y) = \min(\nu(x), \nu(y))$. [Note that $x + y = x(1 + y/x)$.]
2. Suppose R is a Discrete Valuation Ring with unique maximal ideal M and quotient $F = R/M$. For any $n \geq 0$ show that M^n/M^{n+1} is a vector space over F and that $\dim_F(M^n/M^{n+1}) = 1$.

3. Suppose R is an integral domain that is also a local ring whose unique maximal ideal $M = (t)$ is nonzero and principal, and suppose that $\bigcap_{n \geq 1}(t^n) = 0$. Prove that R is a Discrete Valuation Ring. [Show that every nonzero element in R can be written in the form ut^n for some unit $u \in R$ and some $n \geq 0$.]

4. Suppose R is a Noetherian local ring whose unique maximal ideal $M = (t)$ is principal. Prove that either R is a Discrete Valuation Ring or $t^n = 0$ for some $n \geq 0$. In the latter case show that R is Artinian.

5. Suppose that R is a Noetherian integral domain that is also a local ring of Krull dimension 1. Let M be the unique maximal ideal of R and let $F = R/M$, so that M/M^2 is a vector space over F.
   (a) Prove that if $\dim_F(M/M^2) = 1$ then R is a Discrete Valuation Ring.
   (b) If every nonzero ideal of R is a power of M prove that R is a Discrete Valuation Ring.

6. Let R be an integral domain with fraction field K. Prove that every finitely generated R-submodule of K is a fractional ideal of R. If R is Noetherian, prove that A is a fractional ideal of R if and only if R is a finitely generated R-submodule of K.

7. If R is an integral domain and A is a fractional ideal of R, prove that if A is projective then A is finitely generated. Conclude that every integral domain that is not Noetherian contains an ideal that is not projective.

8. Suppose R is a Noetherian integral domain that is also a local ring with nonzero maximal ideal M. Prove that R is a D.V.R. if and only if the only M-primary ideals in R are the powers of M.

9. Let $C = \mathcal{Z}(xz - y^2, yz - x^3, z^2 - x^2y) \subset \mathbb{A}^3$ over the algebraically closed field k. If $v = (0, 0, 0) \in C$, prove that $\dim_k(\mathfrak{m}_{v,C}/\mathfrak{m}_{v,C}^2) = 3$ so that v is singular on C. Conclude that $k[C]$ is not integrally closed in $k(C)$ and determine its integral closure. [cf. Exercise 27, Section 15.4.]

16.3 DEDEKIND DOMAINS

In the previous section we showed that Discrete Valuation Rings are the local rings that are integrally closed Noetherian integral domains of Krull dimension 1. In this section we consider the effect of relaxing the condition that the ring be a local ring:

Definition. A *Dedekind Domain* is a Noetherian, integrally closed, integral domain of Krull dimension 1.

Equivalently, R is a Dedekind Domain if R is a Noetherian, integrally closed, integral domain that is not a field in which every nonzero prime ideal is maximal.

The first result shows that Dedekind Domains are a generalization of the class of Principal Ideal Domains. We shall see later (Theorem 22) that there is a structure theorem for finitely generated modules over a Dedekind Domain extending the corresponding result for P.I.D.s proved in Section 12.1.

Proposition 14.
(1) Every Principal Ideal Domain is a Dedekind Domain.
(2) The ring of integers in an algebraic number field is a Dedekind Domain.

Proof: A P.I.D. is clearly Noetherian, is integrally closed since it is a U.F.D. (Example 3, Section 15.3), and nonzero prime ideals are maximal (Proposition 7 in Section 8.2), which proves (1). Let $\mathcal{O}_K$ be the ring of integers in the number field K, i.e., the integral closure of $\mathbb{Z}$ in K. Then Corollary 25 in Section 15.3 shows that $\mathcal{O}_K$ is integrally closed, $\mathcal{O}_K$ is Noetherian by Theorem 29 in Section 15.3, and the fact that nonzero prime ideals in $\mathcal{O}_K$ are maximal was proved in the discussion following the same theorem. This proves (2).

The following theorem gives a number of important equivalent characterizations of Dedekind Domains. Recall that the basic properties of fractional ideals were developed in the previous section.

Theorem 15. Suppose R is an integral domain with fraction field $K \neq R$. The following are equivalent conditions for R to be a Dedekind Domain:

(1) The ring R is Noetherian, integrally closed, and every nonzero prime ideal is maximal.

(2) The ring R is Noetherian and for each nonzero prime P of R the localization R_P is a Discrete Valuation Ring.

(3) Every nonzero fractional ideal of R in K is invertible.

(4) Every nonzero fractional ideal of R in K is a projective R-module.

(5) Every nonzero proper ideal I of R can be written as a finite product of prime ideals: $I = P_1 P_2 \cdots P_n$ (not necessarily distinct).

When the condition in (5) holds, the set of primes $\{P_1, \dots, P_n\}$ is uniquely determined and so every nonzero proper ideal I of R can be written uniquely (up to order) as a product of powers of prime ideals.

Proof: If R satisfies (1), then R_P is a D.V.R. by Corollary 8, so (1) implies (2). Conversely, assume each R_P is a D.V.R. Then R is integrally closed by Proposition 49 in Section 15.4 and every nonzero prime ideal is maximal by Proposition 46(3) in Section 15.4, so (2) implies (1).

Suppose now that (1) is satisfied and that A is a nonzero fractional ideal of R. Let $A' = \{x \in K \mid xA \subseteq R\}$ as in Proposition 9. For any prime ideal P of R the behavior of R-modules under localization shows that $(AA')_P = A_P(A')_P = A_P(A_P)'$ (cf. Exercise 4). Since R_P is a D.V.R. by what has already been shown, $A_P(A_P)' = R_P$ by Proposition 11. Hence $(AA')_P = R_P$ for all nonzero primes P of R, so $AA' = R$ (Exercise 13 in Section 15.4), and A is invertible, showing (1) implies (3). Conversely, suppose every nonzero fractional ideal of R is invertible. Then every ideal in R is finitely generated by Proposition 9(3), so R is Noetherian. Every localization R_P of R at a nonzero prime P is a local ring in which the nonzero fractional ideals are invertible (cf. Exercise 4), hence is a D.V.R. by Proposition 11. Hence (3) implies (2) and so (1), (2) and (3) are equivalent. The equivalence of these with (4) is given by Proposition 10.

Suppose now that (1) is satisfied, and let I be any nonzero proper ideal in R. Since R is Noetherian, I has a minimal primary decomposition $I = Q_1 \cap \cdots \cap Q_n$ as in Theorem 21 of Section 15.2. The associated primes $P_i = \operatorname{rad} Q_i$ for $i = 1, \dots, n$ are all distinct, and since primes are maximal in R by hypothesis, the associated primes are all pairwise comaximal, and it follows easily that the same is true for the Q_i (Exercise

5). It follows that $Q_1 \cap \cdots \cap Q_n = Q_1 \cdots Q_n$ (Theorem 17 in Section 7.6) so that I is the product of primary ideals. The P-primary ideals of R correspond bijectively with the PR_P-primary ideals in the localization R_P (Proposition 42(3) in Section 15.4), and since R_P is a D.V.R. (because (1) implies (2)), it follows from Corollary 6 that if Q is a P-primary ideal in R then $Q = P^m$ for some integer $m \geq 1$. Applying this to Q_i, $i = 1, \dots, n$ shows that I is the product of powers of prime ideals, which gives the first implication in (5).

Conversely, suppose that all the nonzero proper ideals of R can be written as a product of prime ideals. We first show for any integral domain that a factorization of an ideal into *invertible* prime ideals is unique, i.e., if $P_1 \cdots P_n = \tilde{P}_1 \cdots \tilde{P}_m$ are two factorizations of I into invertible prime ideals then $n = m$ and the two sets of primes $\{P_1, \dots, P_n\}$ and $\{\tilde{P}_1, \dots, \tilde{P}_m\}$ are equal. Suppose $\tilde{P}_1$ is a minimal element in the set $\{\tilde{P}_1, \dots, \tilde{P}_m\}$. Since $P_1 \cdots P_n \subseteq \tilde{P}_1$, the prime ideal $\tilde{P}_1$ contains one of the primes $P_1, \dots, P_n$, say $P_1 \subseteq \tilde{P}_1$. Similarly P_1 contains $\tilde{P}_i$ for some $i = 1, \dots, m$. Then $\tilde{P}_i \subseteq P_1 \subseteq \tilde{P}_1$ and by the minimality of $\tilde{P}_1$ it follows that $\tilde{P}_i = P_1 = \tilde{P}_1$, so the factorization becomes $P_1 P_2 \cdots P_n = P_1 \tilde{P}_2 \cdots \tilde{P}_m$. Since P_1 is invertible, multiplying by the inverse ideal shows that $P_2 \cdots P_n = \tilde{P}_2 \cdots \tilde{P}_m$ and an easy induction finishes the proof. In particular, the uniqueness statement in (5) now follows from the first statement in (5) since in a Dedekind domain every fractional ideal, in particular every prime ideal of R, is invertible.

We next show that *invertible* primes in R are maximal. Suppose then that P is an invertible prime ideal in R and take $a \in R$, $a \notin P$. We want to show that $P + aR = R$. By assumption, the two ideals $P + aR$ and $P + a^2R$ can be written as a product of prime ideals, say $P + aR = P_1 \cdots P_n$ and $P + a^2R = \tilde{P}_1 \cdots \tilde{P}_m$. Note that $P \subseteq P_i$ for $i = 1, \dots, n$ and also $P \subseteq \tilde{P}_j$ for $j = 1, \dots, m$. In the quotient R/P, which is an integral domain, we have the factorization $(\bar{a}) = (P_1/P) \cdots (P_n/P)$, and each P_i/P is a prime ideal in R/P. Since the product is a principal ideal, each P_i/P is also an invertible R/P-ideal (cf. Exercise 2). Similarly, $(\bar{a}^2) = (\tilde{P}_1/P) \cdots (\tilde{P}_m/P)$ is a factorization into a product of invertible prime ideals. Then $(\bar{a})^2 = (P_1/P)^2 \cdots (P_n/P)^2 = (\tilde{P}_1/P) \cdots (\tilde{P}_m/P)$ give two factorizations into a product of invertible prime ideals in the integral domain R/P, so by the uniqueness result in the previous paragraph, $m = 2n$ and $\{P_1/P, P_1/P, \dots, P_n/P, P_n/P\} = \{\tilde{P}_1/P, \dots, \tilde{P}_m/P\}$. It follows that the set of primes $\tilde{P}_1, \dots, \tilde{P}_m$ in R consists of the primes $P_1, \dots, P_n$, each repeated twice. This shows that $P + a^2R = (P + aR)^2$. Since $P \subseteq P + a^2R$ and $(P + aR)^2 \subseteq P^2 + aR$, we have $P \subseteq P^2 + aR$, so every element x in P can be written in the form $x = y + az$ where $y \in P^2$ and $z \in R$. Then $az = x - y \in P$ and since $a \notin P$, we have $z \in P$, which shows that $P \subseteq P^2 + aP$. Clearly $P^2 + aP \subseteq P$ and so $P = P^2 + aP = P(P + aR)$. Since P is assumed invertible, it follows that $R = P + aR$ for any $a \in R - P$, which proves that P is a maximal ideal.

We now show that every nonzero prime ideal is invertible. If P is a nonzero prime ideal, let a be any nonzero element in P. By assumption, $Ra = P_1 \cdots P_n$ can be written as a product of prime ideals, and $P_1, \dots, P_n$ are invertible since their product is principal (by Exercise 2 again). Since $P_1 \cdots P_n = Ra \subseteq P$, the prime ideal P contains P_i for some $1 \leq i \leq n$. Since P_i is maximal by the previous paragraph, it follows that

$P = P_i$ is invertible.

Finally, since every nonzero proper ideal of R is a product of prime ideals, it follows that every nonzero ideal of R is invertible, and since every fractional ideal of R is of the form $(d^{-1})I$ for some ideal in R, also every fractional ideal of R is invertible. This proves that (5) implies (3), and complete the proof of the theorem.

The following corollary follows immediately from Proposition 14:

Corollary 16. If $\mathcal{O}_K$ is the ring of integers in an algebraic number field K then every nonzero ideal I in $\mathcal{O}_K$ can be written uniquely as the product of powers of distinct prime ideals:

$$I = P_1^{e_1} P_2^{e_2} \cdots P_n^{e_n},$$

where $P_1, \ldots, P_n$ are distinct prime ideals and $e_i \geq 1$ for $i = 1, \ldots, n$.

Remark: The development of Dedekind Domains given here reverses the historical development. As mentioned in Section 9.3, the unique factorization of nonzero *ideals* into a product of prime *ideals* replaces the failure of unique factorization of nonzero *elements* into products of prime *elements* in rings of integers of number fields. This property of rings of integers in Corollary 16 is what led originally to the definition of an ideal, and Dedekind originally defined what we now call Dedekind Domains by property 5 in Theorem 15. It was Noether who observed that they can also be characterized by property (1), which we have taken as the initial definition of a Dedekind Domain.

The unique factorization into prime ideals in Dedekind Domains can be used to explicitly define the valuations ν_P on R with respect to which the valuation rings are the localizations R_P in Theorem 15(2) (cf. Exercise 6). We now indicate how unique factorization for ideals can be used to define a divisibility theory for ideals similar to the divisibility of integers in $\mathbb{Z}$.

Definition. If A and B are ideals in the integral domain R then B is said to *divide* A (and A is *divisible by* B) if there is an ideal C in R with $A = BC$.

If B divides A then certainly $A \subseteq B$. If R is a Dedekind Domain, the converse is true: $A \subseteq B$ implies $C = AB^{-1} \subseteq BB^{-1} = R$ so C is an ideal in R with $BC = A$.

We can also define the notion of the *greatest common divisor* (A, B) of two ideals A and B: (A, B) divides both A and B and any ideal dividing both A and B divides (A, B). The second statement in the next proposition shows that this greatest common divisor always exists for integral ideals in a Dedekind Domain and gives a formula for it similar to the formula for the greatest common divisor of two integers.

Proposition 17. Suppose R is a Dedekind Domain and A, B are two nonzero ideals in R, with prime ideal factorizations $A = P_1^{e_1} \cdots P_n^{e_n}$ and $B = P_1^{f_1} \cdots P_n^{f_n}$ (where $e_i, f_i \geq 0$ for $i = 1, \ldots, n$). Then

(1) $A \subseteq B$ if and only if B divides A (i.e., "to contain is to divide") if and only if $f_i \leq e_i$ for $i = 1, \ldots, n$,

(2) $A + B = (A, B) = P_1^{\min(e_1, f_1)} \cdots P_n^{\min(e_n, f_n)}$, so in particular A and B are relatively prime, $A + B = R$, if and only if they have no prime ideal factors in common.

Proof: We proved the first statement in (1) above. If each $f_i \leq e_i$, then taking $C = P_1^{e_1 - f_1} \cdots P_n^{e_n - f_n} \subseteq R$ shows that B divides A. Conversely, if B divides A, then writing C as a product of prime ideals in $A = BC$ shows that $f_i \leq e_i$ for all i, which proves all of (1). Since $A + B$ is the smallest ideal containing both A and B, (2) now follows from (1).

Proposition 18. *(Chinese Remainder Theorem)* Suppose R is a Dedekind Domain, $P_1, P_2, \ldots, P_n$ are distinct prime ideals in R and $a_i \geq 0$ are integers, $i = 1, \ldots, n$. Then

$$R/P_1^{a_1} \cdots P_n^{a_n} \cong R/P_1^{a_1} \times R/P_2^{a_2} \times \cdots \times R/P_n^{a_n}.$$

Equivalently, for any elements $r_1, r_2, \ldots, r_n \in R$ there exists an element $r \in R$, unique up to an element in $P_1^{a_1} \cdots P_n^{a_n}$, with

$$r \equiv r_1 \bmod P_1^{a_1}, \quad r \equiv r_2 \bmod P_2^{a_2}, \quad \ldots, \quad r \equiv r_n \bmod P_n^{a_n}.$$

Proof: This is immediate from Theorem 17 in Section 7.6 since the previous proposition shows that the $P_i^{a_i}$ are pairwise comaximal ideals.

Corollary 19. Suppose I is an ideal in the Dedekind Domain R. Then

(1) there is an ideal J of R relatively prime to I such that the product $IJ = (a)$ is a principal ideal,

(2) if I is nonzero then every ideal in the quotient R/I is principal; equivalently, if I_1 is an ideal of R containing I then $I_1 = I + Rb$ for some $b \in R$, and

(3) every ideal in R can be generated by two elements; in fact if I is nonzero and $0 \neq a \in I$ then $I = Ra + Rb$ for some $b \in I$.

Proof: Suppose $I = P_1^{e_1} \cdots P_n^{e_n}$ is the prime ideal factorization of I in R. For each $i = 1, \ldots, n$, let r_i be an element of $P_i^{e_i} - P_i^{e_i+1}$. By the proposition, there is an element $a \in R$ with $a \equiv r_i \bmod P_i^{e_i+1}$ for all i. Hence $a \in P_i^{e_i} - P_i^{e_i+1}$ for all i, so the power of P_i in prime ideal factorization of (a) is precisely e_i by (1) of Proposition 17:

$$(a) = P_1^{e_1} \cdots P_n^{e_n} P_{n+1}^{e_{n+1}} \cdots P_m^{e_m}$$

for some prime ideals $P_{n+1}, \ldots, P_m$ distinct from $P_1, \ldots, P_n$. Letting $J = P_{n+1}^{e_{n+1}} \cdots P_m^{e_m}$ gives (1). For (2), by the Chinese Remainder Theorem it suffices to prove that every ideal in R/P^m is principal in the case of a power of a prime ideal P, and this is immediate since $R/P^m \cong R_P/P^m R_P$ and the localization R_P is a P.I.D. Finally, (3) follows from (2) by taking $I = Ra$.

The first statement in Corollary 19 shows that there is an integral ideal J relatively prime to I lying in the inverse class of I in the class group of R. One can even impose additional conditions on J, cf. Exercise 11.

Corollary 20. If R is a Dedekind Domain then R is a P.I.D. (i.e., R has class number 1) if and only if R is a U.F.D.

Proof: Every P.I.D. is a U.F.D., so suppose that R is a U.F.D. and let P be any prime ideal in R. Then $P = Ra + Rb$ for some $a \neq 0$ and b in R by Corollary 19. We have $(a') \subseteq P$ for one of the irreducible factors a' of a since their product is an element in the prime P, and then P divides (a') in R by Proposition 17(1). It follows that $P = (a')$ is principal since (a') is a prime ideal (Proposition 12 in Section 8.3). Since every ideal in R is a product of prime ideals, every ideal of R is principal, i.e., R is a P.I.D.

Corollary 20 shows that the class number of a Dedekind domain R gives a measure of the failure of unique factorization of elements. It is a fundamental result in algebraic number theory that the class number of the ring of integers of an algebraic number field is finite. For general Dedekind Domains, however, the class number need not be finite. In fact, for any abelian group A (finite or infinite) there is a Dedekind Domain whose class group is isomorphic to A.

Modules over Dedekind Domains and the Fundamental Theorem of Finitely Generated Modules

We turn next to the consideration of modules over Dedekind Domains R. Every fractional ideal of R is an R-module and the first statement in the following proposition shows that two fractional ideals of R are isomorphic as R-modules if and only if they represent the same element in the class group of R.

Proposition 21. Let R be a Dedekind Domain with fraction field K.

(1) Suppose I and J are two fractional ideals of R. Then $I \cong J$ as R-modules if and only if I and J differ by a nonzero principal ideal: $I = (a)J$ for some $0 \neq a \in K$.

(2) More generally, suppose $I_1, I_2, \dots, I_n$ and $J_1, J_2, \dots, J_m$ are nonzero fractional ideals in the fraction field K of the Dedekind Domain R. Then

$$I_1 \oplus I_2 \oplus \cdots \oplus I_n \cong J_1 \oplus J_2 \oplus \cdots \oplus J_m$$

as R-modules if and only if $n = m$ and the product ideals $I_1 I_2 \cdots I_n$ and $J_1 J_2 \cdots J_n$ differ by a principal ideal:

$$I_1 I_2 \cdots I_n = (a) J_1 J_2 \cdots J_n$$

for some $0 \neq a \in K$.

(3) In particular,

$$I_1 \oplus I_2 \oplus \cdots \oplus I_n \cong \underbrace{R \oplus \cdots \oplus R}_{n-1 \text{ factors}} \oplus (I_1 I_2 \cdots I_n)$$

and $R^n \oplus I \cong R^n \oplus J$ if and only if I and J differ by a principal ideal: $I = (a)J$, $a \in K$.

Proof: Multiplication by $0 \neq a \in K$ gives an R-module isomorphism from J to $(a)J$, so if $I = (a)J$ we have $I \cong J$ as R-modules. For the converse, observe that we may assume $J \neq 0$ and then $I \cong J$ implies $R \cong J^{-1}I$. But this says that $J^{-1}I = aR$ is principal (with generator a given by the image of $1 \in R$), i.e., $I = (a)J$, proving (1).

We next show that for any nonzero fractional ideals I and J that $I \oplus J \cong R \oplus IJ$. Replacing I and J by isomorphic R-modules aI and bJ, if necessary, we may assume that I and J are integral ideals that are relatively prime (cf. Exercise 12), so that $I + J = R$ and $I \cap J = IJ$. It is easy to see that the map from $I \oplus J$ to $I + J = R$ defined by mapping (x, y) to $x+y$ is a surjective R-module homomorphism with kernel $I \cap J = IJ$, so we have an exact sequence

$$0 \longrightarrow IJ \longrightarrow I \oplus J \longrightarrow R \longrightarrow 0$$

of R-modules. This sequence splits since R is free, so $I \oplus J \cong R \oplus IJ$, as claimed.

The first statement in (3) now follows by induction, and combining this statement with (1) shows that if $I_1 \cdots I_n = (a)J_1 \cdots J_n$ for some nonzero $a \in K$ then $I_1 \oplus \cdots \oplus I_n$ is isomorphic to $J_1 \oplus \cdots \oplus J_n$. This proves the "if" statement in (2). It remains to prove the "only if" statement in (2) since the corresponding statement in (3) is a special case. So suppose $I_1 \oplus I_2 \oplus \cdots \oplus I_n \cong J_1 \oplus J_2 \oplus \cdots \oplus J_m$ as R-modules.

Since $I \otimes_R K$ is the localization of the ideal I in K (cf. Proposition 41 in Section 15.4) it follows that $I \otimes_R K \cong K$ for any nonzero fractional ideal I of K. Since tensor products commute with direct sums, $(I_1 \oplus \cdots \oplus I_n) \otimes_R K \cong K^n$ is an n-dimensional vector space over K. Similarly, $J_1 \oplus \cdots \oplus J_m \otimes_R K \cong K^m$, from which it follows that $n = m$.

Note that replacing I_1 by the isomorphic fractional ideal $a_1{}^{-1}I_1$ for any nonzero element $a_1 \in I_1$ does not effect the validity of the statements in (2). Hence we may assume I_1 contains R, and similarly we may assume that each of the fractional ideals in (2) contains R. Let φ denote the R-module isomorphism from $I_1 \oplus \cdots \oplus I_n$ to $J_1 \oplus \cdots \oplus J_n$. For $i = 1, 2, \dots, n$ define

$$\varphi((0, \dots, 0, 1, 0, \dots, 0)) = (a_{1,i}, a_{2,i}, \dots, a_{n,i}) \in J_1 \oplus J_2 \oplus \cdots \oplus J_n$$

where $1 \in I_i$ on the left hand side occurs in position i. Since φ is an R-module homomorphism it follows that

$$J_j = a_{j,1}I_1 + a_{j,2}I_2 + \cdots + a_{j.i}I_i + \cdots + a_{j,n}I_n$$

for each $j = 1, 2, \dots, n$. Taking the product of these ideals for $j = 1, 2, \dots, n$ it follows that

$$(a_{j_1,1}a_{j_2,2} \cdots a_{j_n,n})I_1I_2 \cdots I_n \subseteq J_1J_2 \cdots J_n$$

for any permutation $\{j_1, j_2, \dots, j_n\}$ of $\{1, 2, \dots, n\}$. Hence

$$dI_1I_2 \cdots I_n \subseteq J_1J_2 \cdots J_n$$

where d is the determinant of the matrix $(a_{i,j})$, since the determinant is the sum of terms $\epsilon(\sigma)a_{1.\sigma(1)} \cdots a_{n,\sigma(n)}$ where $\epsilon(\sigma)$ is the sign of the permutation σ of $\{1, 2, \dots, n\}$. Similarly, for $j = 1, \dots, n$, define

$$\varphi^{-1}((0, \dots, 0, 1, 0, \dots, 0)) = (b_{1,j}, b_{2,j}, \dots, b_{n,j}) \in I_1 \oplus I_2 \oplus \cdots \oplus I_n$$

where $1 \in J_j$ on the left hand side occurs in position j. The product of the two matrices $(a_{i,j})$ and $(b_{i,j})$ is just the identity matrix, so $d \neq 0$ and the determinant of the matrix $(b_{i.j})$ is d^{-1}. As above we have

$$d^{-1}J_1J_2 \cdots J_n \subseteq I_1I_2 \cdots I_n,$$

which shows that $I_1 I_2 \cdots I_n = (a) J_1 J_2 \cdots J_n$, where $0 \neq a = d^{-1} \in K$, completing the proof of the proposition.

We now consider finitely generated modules over Dedekind Domains and prove a structure theorem for such modules extending the results in Chapter 12 for finitely generated modules over P.I.D.s.

Recall that the *rank* of M is the maximal number of R-linearly independent elements in M, or, equivalently, the dimension of $M \otimes_R K$ as a K-vector space, where K is the fraction field of R (cf. Exercises 1–4, 20 in Section 12.1).

Theorem 22. Suppose M is a finitely generated module over the Dedekind Domain R. Let $n \geq 0$ denote the rank of M and let $\mathrm{Tor}(M)$ be the torsion submodule of M. Then

$$M \cong \underbrace{R \oplus R \oplus \cdots \oplus R \oplus I}_{n \text{ factors}} \oplus \mathrm{Tor}(M)$$

for some ideal I of R, and

$$\mathrm{Tor}(M) \cong R/P_1^{e_1} \times R/P_2^{e_2} \times \cdots \times R/P_s^{e_s}$$

for some $s \geq 0$ and powers $P_i^{e_i}$, $e_1 \geq 1$, of (not necessarily distinct) prime ideals. The ideals $P_i^{e_i}$ for $i = 1, \ldots, s$ are unique and the ideal I is unique up to multiplication by a principal ideal.

Proof: Suppose first that M is a finitely generated torsion free module over R, i.e., $\mathrm{Tor}(M) = 0$. Then the natural R-module homomorphism from M to $M \otimes_R K$ is injective, so we may view M as an R-submodule of the vector space $M \otimes_R K$. If M has rank n over R, then $M \otimes_R K$ is a vector space over K of dimension n. Let $x_1, \ldots, x_n$ be a basis for $M \otimes_R K$ over K and let $m_1, \ldots, m_s$ be R-module generators for M. Each m_i, $i = 1, \ldots, s$ can be written as a K-linear combination of $x_1, \ldots, x_n$. Let $0 \neq d \in R$ be a common denominator for all the coefficients in K of these linear combinations, and set $y_i = x_i/d$, $i = 1, \ldots, n$. Then

$$M \subseteq R y_1 + \cdots + R y_n \subset K x_1 + \cdots + K x_n$$

which shows that M is contained in a *free* R-submodule of rank n and every element m in M can be written uniquely in the form

$$m = a_1 y_1 + \cdots + a_n y_n$$

with $a_1, \ldots, a_n \in R$. The map $\varphi : M \to R$ defined by $\varphi(a_1 y_1 + \cdots + a_n y_n) = a_n$ is an R-module homomorphism, so we have an exact sequence

$$0 \longrightarrow \ker \varphi \longrightarrow M \xrightarrow{\varphi} I_1 \longrightarrow 0$$

where I_1 is the image of φ in R, hence is an ideal in R. The submodule $\ker \varphi$ is also a torsion free R-module whose rank is at most $n - 1$ (since it is contained in $R y_1 + \cdots + R y_{n-1}$), and it follows by comparing ranks that I_1 is nonzero and that $\ker \varphi$ has rank precisely $n - 1$. By (4) of Theorem 15, I_1 is a projective R-module, so this sequence splits:

$$M \cong I_1 \oplus (\ker \varphi).$$

By induction on the rank, we see that a finitely generated torsion free R-module is isomorphic to the direct sum of n nonzero ideals of R:

$$M \cong I_1 \oplus I_2 \oplus \cdots \oplus I_n.$$

Since $I_1, \ldots, I_n$ are each projective R-modules, it follows that any finitely generated torsion free R-module is projective.

If now M is any finitely generated R-module, the quotient $M/\mathrm{Tor}(M)$ is finitely generated and torsion free, hence projective by what was just proved. The exact sequence

$$0 \longrightarrow \mathrm{Tor}(M) \longrightarrow M \longrightarrow M/\mathrm{Tor}(M) \longrightarrow 0$$

therefore splits, and so

$$M \cong \mathrm{Tor}(M) \oplus (M/\mathrm{Tor}(M)).$$

By the results in the previous paragraph $M/\mathrm{Tor}(M)$ is isomorphic to a direct sum of n nonzero ideals of R, and by Proposition 21 we obtain

$$M \cong \underbrace{R \oplus R \oplus \cdots \oplus R \oplus I}_{n \text{ factors}} \oplus \mathrm{Tor}(M)$$

for some ideal I of R. The uniqueness statement regarding the ideal I is also immediate from the uniqueness statement in Proposition 21(3).

It remains to prove the statements regarding the torsion submodule $\mathrm{Tor}(M)$. Suppose then that N is a finitely generated torsion R-module. Let $I = \mathrm{Ann}(N)$ be the annihilator of N in R and suppose $I = P_1^{e_1} \cdots P_t^{e_t}$ is the prime ideal factorization of I in R, where $P_1, \ldots, P_t$ are distinct prime ideals. Then N is a module over R/I, and

$$R/I \cong R/P_1^{e_1} \times R/P_2^{e_2} \times \cdots \times R/P_t^{e_t}.$$

It follows that

$$N \cong (N/P_1^{e_1}N) \times (N/P_2^{e_2}N) \times \cdots \times (N/P_t^{e_t}N)$$

as R-modules. Each N/P^eN is a finitely generated module over $R/P^e \cong R_P/P^eR_P$ where R_P is the localization of R at the prime P, i.e., is a finitely generated module over R_P that is annihilated by P^eR_P. Since R is a Dedekind Domain, each R_P is a P.I.D. (even a D.V.R.), so we may apply the Fundamental Theorem for Finitely Generated Modules over a P.I.D. to see that each N/P^eN is isomorphic as an R_P-module to a direct sum of finitely many modules of the form $R_P/P^f R_P$ where $f \leq e$. It follows that each N/P^eN is isomorphic as an R-module to a direct sum of finitely many modules of the form $R/P^f R$ where $f \leq e$. This proves that N is isomorphic to the direct sum of finitely many modules of the form $R/P_i^{f_i}$ for various prime ideals P_i. Hence $\mathrm{Tor}(M)$ can be decomposed into a direct sum as in the statement in the theorem.

Finally, it remains to prove that the ideals $P_i^{e_i}$ for $i = 1, \ldots, s$ in the decomposition of $\mathrm{Tor}(M)$ are unique. This is similar to the uniqueness argument in the proof of Theorem 10 in Section 12.1 (cf. also Exercises 11–12 in Section 12.1): for any prime ideal P of R, the quotient $P^{i-1}M/P^iM$ is a vector space over the field R/P and the difference $\dim_{R/P} P^{i-1}M/P^iM - \dim_{R/P} P^iM/P^{i+1}M$ is the number of direct summands of M isomorphic to R/P^i, hence is uniquely determined by M. This concludes the proof of the theorem.

If M is a finitely generated module over the Dedekind Domain R as in Theorem 22, then the isomorphism type of M as an R-module is determined by the *rank* n, the prime powers $P_i^{e_i}$ for $i = 1, \dots, s$ (called the *elementary divisors* of M, and the class of the ideal I in the class group of R (called the *Steinitz class* of M). Note that a P.I.D. is the same as a Dedekind Domain whose class number is 1, in which case every nonzero ideal I of R is isomorphic as an R-module simply to R. In this case, Theorem 22 reduces to the elementary divisor form of the structure theorem for finitely generated modules over P.I.D.s in Chapter 12. There is also an invariant factor version of the description of the torsion R-modules in Theorem 22 (cf. Exercise 14).

The next result extends the characterization of finitely generated projective modules over P.I.D.s (Exercise 21 in Section 12.1) to Dedekind Domains.

Corollary 23. A finitely generated module over a Dedekind Domain is projective if and only if it is torsion free.

Proof: We showed that a finitely generated torsion free R-module is projective in the proof of Theorem 22, so by the decomposition of M in Theorem 22, M is projective if and only if $\operatorname{Tor}(M)$ is projective (cf. Exercise 3 in Section 10.5). To complete the proof it suffices to show that no nonzero torsion R-module is projective, which is left as an exercise (cf. Exercise 15).

EXERCISES

1. If R is an integral domain, show that every fractional ideal of R is invertible if and only if every integral ideal of R is invertible.
2. Suppose R is an integral domain with fraction field K and $A_1, A_2, \dots, A_n$ are fractional ideals of R whose product is a nonzero principal fractional ideal: $A_1 A_2 \cdots A_n = Rx$ for some $0 \neq x \in K$. For each $i = 1, \dots, n$ prove that A_i is an invertible fractional ideal with inverse $(x^{-1}) A_1 \cdots A_{i-1} A_{i+1} \cdots A_n$.
3. Suppose R is an integral domain with fraction field K and P is a nonzero prime ideal in R. Show that the fractional ideals of R_P in K are the R_P-modules of the form AR_P where A is a fractional ideal of R.
4. Suppose R is an integral domain with fraction field K and A is a fractional ideal of R in K. Let $A' = \{x \in K \mid xA \subseteq R\}$ as in Proposition 9.
   (a) For any prime ideal P in R prove that the localization $(A')_P$ of A' at P is a fractional ideal of R_P in K.
   (b) If A is a finitely generated R-module, prove that $(A')_P = (A_P)'$ where $(A_P)'$ is the fractional R_P ideal $\{x \in K \mid xA_P \subseteq R_P\}$ corresponding to the localization A_P.
5. If Q_1 is a P_1-primary ideal and Q_2 is a P_2-primary ideal where P_1 and P_2 are comaximal ideals in a Noetherian ring R, prove that Q_1 and Q_2 are also comaximal. [Use Proposition 14 in Section 15.2.]
6. Suppose R is a Dedekind Domain with fraction field K.
   (a) Prove that every nonzero fractional ideal of R in K can be written uniquely as the product of distinct prime powers $P_1^{a_1} \cdots P_n^{a_n}$ where the a_i are nonzero integers, possibly negative.

(b) If $0 \neq x \in K$, let $P^{\nu_P(x)}$ be the power of the prime P in the factorization of the principal ideal (x) as in (a) (where $\nu_P(x) = 0$ if P is not one of the primes occurring). Prove ν_P is a valuation on K with valuation ring R_P, the localization of R at P.

7. Suppose R is a Noetherian integral domain that is not a field. Prove that R is a Dedekind Domain if and only if for every maximal ideal M of R there are no ideals I of R with $M^2 \subset I \subset M$. [Use Exercise 12 in Section 1 and Theorems 7 and 15.]

8. Suppose R is a Noetherian integral domain with Krull dimension 1. Prove that every nonzero ideal I in R can be written uniquely as a product of primary ideals whose radicals are all distinct. [Cf. the proof of Theorem 15. Use the uniqueness of the primary components belonging to the isolated primes in a minimal primary decomposition (Theorem 21 in Section 15.2).]

9. Suppose R is an integral domain. Prove that R_P is a D.V.R. for every nonzero prime ideal P if and only if R_M is a D.V.R. for every nonzero maximal ideal.

10. Suppose R is a Noetherian integral domain that is not a field. Prove that R is a Dedekind Domain if and only if nonzero primes M are maximal and every M-primary ideal is a power of M.

11. If I and J are nonzero ideals in the Dedekind Domain R show there exists an integral ideal I_1 in R that is relatively prime to both I and J such that $I_1 I$ is a principal ideal in R.

12. If I and J are nonzero fractional ideals for the Dedekind Domain R prove there are elements $\alpha, \beta \in K$ such that αI and βJ are nonzero integral ideals in R are relatively prime.

13. Suppose I and J are nonzero ideals in the Dedekind Domain R. Prove that there is an ideal $I_1 \cong I$ that is relatively prime to J. [Use Corollary 19 to find an ideal I_2 with $I_2 I = (a)$ and $(I_2, J) = R$. If $I_2 = P_1^{e_1} \cdots P_n^{e_n}$, choose $b \in R$ with $b \in P_i^{e_i} - P_i^{e_i+1}$ and $b \equiv 1 \bmod P$ for every prime P dividing J. Show that $(b) = I_2 I_1$ for some ideal I_1 and consider $(a)I_1$ to prove that $I_1 \cong I$.]

14. Prove that every finitely generated torsion module over a Dedekind Domain R is isomorphic to a direct sum $R/I_1 \oplus R/I_2 \oplus \cdots \oplus R/I_n$ with unique nonzero ideals $I_1, \ldots, I_n$ of R satisfying $I_1 \subseteq I_2 \subseteq \cdots \subseteq I_n$ (called the *invariant factors* of M). [cf. Section 12.1.]

15. If P is a nonzero prime ideal in the Dedekind Domain R prove that R/P^n is not a projective R-module for any $n \geq 1$. [Consider the exact sequence $0 \to P^n/P^{n+1} \to R/P^{n+1} \to R/P^n \to 0$.] Conclude that if $M \neq 0$ is a finitely generated torsion R-module then M is not projective. [cf. Exercise 3, Section 10.5.]

16. Prove that the class number of the Dedekind Domain R is 1 if and only if every finitely generated projective R-module is free.

17. Suppose R is a Dedekind Domain.

(a) Show that $I \sim J$ if and only if $I \cong J$ as R-modules defines an equivalence relation on the set of nonzero fractional ideals of R. Let $C(R)$ be the corresponding set of R-module isomorphism classes and let $[I] \in C(R)$ denote the equivalence class containing the fractional ideal I of R.

(b) Show that the multiplication $[I][J] = [I \oplus J]$ gives a well defined binary operation with respect to which $C(R)$ is an abelian group with identity $1 = [R]$.

(c) Prove that the abelian group $C(R)$ in (b) is isomorphic to the class group of R.

18. If R is a Dedekind Domain and I is any nonzero ideal, prove that R/I contains only finitely many ideals. In particular, show that R/I is an Artinian ring.

19. Suppose I is a nonzero fractional ideal in the Dedekind Domain R. Explicitly exhibit I as a direct summand of a free R-module to show that I is projective. [Consider $I \oplus I^{-1}$

and use Proposition 21.]

20. Suppose I and J are two nonzero fractional ideals in the Dedekind Domain R and that $I^n = J^n$ for some $n \neq 0$. Prove that $I = J$.

21. Suppose K is an algebraic number field and $\mathcal{O}_K$ is the ring of integers in K. If P is a nonzero prime ideal in $\mathcal{O}_K$ prove that $P = (p, \pi)$ for some prime $p \in \mathbb{Z}$ and algebraic integer $\pi \in \mathcal{O}_K$.

22. Suppose $K = \mathbb{Q}(\sqrt{D})$ is a quadratic extension of $\mathbb{Q}$ where D is a squarefree integer and $\mathcal{O}_K$ is the ring of integers in K.

(a) Prove that $|\mathcal{O}_K/(p)| = p^2$. [Observe that $\mathcal{O}_K \cong \mathbb{Z}^2$ as an abelian group.]

(b) Use Corollary 16 to show that there are 3 possibilities for the prime ideal factorization of (p) in $\mathcal{O}_K$:

(i) $(p) = P$ is a prime ideal with $|\mathcal{O}_K/P| = p^2$,

(ii) $(p) = P_1P_2$ with distinct prime ideals P_1, P_2 and $|\mathcal{O}_K/P_1| = |\mathcal{O}_K/P_2| = p$,

(iii) $(p) = P^2$ for some prime ideal P with $|\mathcal{O}_K/P| = p$.

(In cases (i), (ii), and (iii) the prime p is said to be *inert*, *split*, or *ramified* in $\mathcal{O}_K$, respectively. The set of ramified primes is finite: the primes p dividing D if $D \equiv 1, 2 \bmod 4$; $p = 2$ and the primes p dividing D if $D \equiv 3 \bmod 4$. Cf. Exercise 31 in Section 15.5.)

(c) Determine the prime ideal factorizations of the primes $p = 2, 3, 5, 7, 11$ in the ring of integers $\mathcal{O}_K = \mathbb{Z}[\sqrt{-5}]$ of $K = \mathbb{Q}(\sqrt{-5})$.

23. Let $\mathcal{O}$ be the ring of integers in the algebraic closure $\overline{\mathbb{Q}}$ of $\mathbb{Q}$.

(a) Show that the infinite sequence of ideals in $\mathcal{O}$ $(2) \subseteq (\sqrt{2}) \subseteq (\sqrt[4]{2}) \subseteq (\sqrt[8]{2}) \subseteq \cdots$ is strictly increasing, and so $\mathcal{O}$ is not Noetherian.

(b) Show that $\mathcal{O}$ has Krull dimension 1. [Use Theorem 26 in Section 15.3.]

(c) Let K be a number field and let I be any ideal in $\mathcal{O}_K$. Show that there is some finite extension L of K such that I becomes principal when extended to $\mathcal{O}_L$, i.e., the ideal $I\mathcal{O}_L$ is principal (where L depends on I)—you may use the theorem that the class group of K is a finite group. [cf. Exercise 20.]

(d) Prove that $\mathcal{O}$ is a Bezout Domain (cf. Section 8.1).

24. Suppose F and K are algebraic number fields with $\mathbb{Q} \subseteq F \subseteq K$, with rings of integers $\mathcal{O}_F$ and $\mathcal{O}_K$, respectively. Since $\mathcal{O}_F \subseteq \mathcal{O}_K$, the ring $\mathcal{O}_K$ is naturally a module over $\mathcal{O}_F$.

(a) Prove $\mathcal{O}_K$ is a torsion free $\mathcal{O}_F$-module of rank $n = [K : F]$. [Compute ranks over $\mathbb{Z}$.] If $\mathcal{O}_K$ is *free* over $\mathcal{O}_F$ then $\mathcal{O}_K$ is said to have a *relative integral basis* over $\mathcal{O}_F$.

(b) Prove that if F has class number 1 then $\mathcal{O}_K$ has a relative integral basis over $\mathcal{O}_F$.

If $K = \mathbb{Q}(\sqrt{-5}, \sqrt{2})$ then the ring of integers $\mathcal{O}_K$ is given by

$$\mathcal{O}_K = \mathbb{Z} + \mathbb{Z}\sqrt{-5} + \mathbb{Z}\sqrt{-10} + \mathbb{Z}\omega \qquad \text{where } \omega = (\sqrt{-10} + \sqrt{2})/2.$$

(c) If $F_1 = \mathbb{Q}(\sqrt{2})$ prove that $\mathcal{O}_K$ has a relative integral basis over $\mathcal{O}_{F_1}$ and find an explicit basis $\{\alpha, \beta\}$: $\mathcal{O}_K = \mathcal{O}_{F_1} \cdot \alpha + \mathcal{O}_{F_1} \cdot \beta$.

(d) If $F_2 = \mathbb{Q}(\sqrt{-5})$, show that $P_3 = (3, 1 + \sqrt{-5}) = (3, 5 - \sqrt{-5})$ is a prime ideal of $\mathcal{O}_{F_2}$ that is not principal and that $\mathcal{O}_K = \mathcal{O}_{F_2} \cdot 1 + (1/3)P_3 \cdot \omega$. [Check that $\sqrt{-10} = -(5 - \sqrt{-5})\omega/3$.] Conclude that the Steinitz class of $\mathcal{O}_K$ as a module over $\mathcal{O}_{F_2}$ is the nontrivial class of P_3 in the class group of $\mathcal{O}_{F_2}$ and so there is no relative integral basis of $\mathcal{O}_K$ over $\mathcal{O}_{F_2}$.

(e) Determine whether $\mathcal{O}_K$ has a relative integral basis over the ring of integers of the remaining quadratic subfield $F_3 = \mathbb{Q}(\sqrt{-10})$ of K.

25. Suppose C is a nonsingular irreducible affine curve over an algebraically closed field k. Prove that the coordinate ring $k[C]$ is a Dedekind Domain.

CHAPTER 17

Introduction to Homological Algebra and Group Cohomology

Let R be a ring with 1. In Section 10.5 we saw that a short exact sequence

$$0 \longrightarrow L \xrightarrow{\psi} M \xrightarrow{\varphi} N \longrightarrow 0 \tag{17.1}$$

of R-modules gives rise to an exact sequence of abelian groups

$$0 \longrightarrow \operatorname{Hom}_R(N, D) \xrightarrow{\varphi'} \operatorname{Hom}_R(M, D) \xrightarrow{\psi'} \operatorname{Hom}_R(L, D) \tag{17.2}$$

for any R-module D and that the homomorphism ψ' is in general not surjective so this sequence cannot always be extended to a short exact sequence. Equivalently, homomorphisms from L to D cannot in general be lifted to homomorphisms from M into D. In this chapter we introduce some of the techniques of "homological algebra," which provide a method of extending some exact sequences in a natural way. For the situation above one obtains an infinite exact sequence involving the "cohomology groups" $\operatorname{Ext}^n_R(\_\_, D)$ (cf. Theorem 8), and these groups provide a measure of the set of homomorphisms from L into D that cannot be extended to M. We then consider the analogous questions for the other two functors considered in Section 10.5, namely taking homomorphisms *from* D into the terms of the sequence (1) and tensoring the sequence (1) with D.

In the subsequent sections we concentrate on an important special case of this general type of homological construction—the "cohomology of finite groups." We make explicit the computations in this case and indicate some applications of these techniques to establish some new results in group theory. In this sense, Sections 2–4 may be considered as an explicit "example" illustrating some uses of the general theory in Section 1.

Cohomology and homology groups occur in many areas of mathematics. The formal notions of homology and cohomology groups and the general area of homological algebra arose from algebraic topology around the middle of the 20th century in the study of the relation between the higher homotopy groups and the fundamental group of a topological space, although the study of certain specific cohomology groups, such as Schur's work on group extensions (described in Section 4), predates this by half a century. As with much of algebra, the ideas common to a number of different areas were abstracted into general theories. Much of the language of homology and cohomology reflects its topological origins: homology groups, chains, cycles, boundaries, etc.

In this section we describe some general terminology and results in homological algebra leading to the so called Long Exact Sequence in Cohomology. We then define certain (cohomology) groups associated to the sequence (2) and apply the general homological results to obtain a long exact sequence extending this sequence at the right end. We then indicate the corresponding development for sequences obtained by taking homomorphisms from D to the terms in (1) or by tensoring the terms with D.

We begin with a generalization of the notion of an exact sequence, namely a sequence of abelian group homomorphisms where successive maps compose to zero (i.e., the image of one map is contained in the kernel of the next):

Definition. Let C be a sequence of abelian group homomorphisms:

$$0 \longrightarrow C^0 \xrightarrow{d_1} C^1 \longrightarrow \cdots \longrightarrow C^{n-1} \xrightarrow{d_n} C^n \xrightarrow{d_{n+1}} \cdots. \tag{17.3}$$

(1) The sequence C is called a *cochain complex* if the composition of any two successive maps is zero: $d_{n+1} \circ d_n = 0$ for all n.
(2) If C is a cochain complex, its n^{th} *cohomology group* is the quotient group $\ker d_{n+1}/\operatorname{image} d_n$, and is denoted by $H^n(C)$.

There is a completely analogous "dual" version in which the homomorphisms are between groups in *decreasing* order, in which case the sequence corresponding to (3) is written $\cdots \xrightarrow{d_{n+1}} C_n \xrightarrow{d_n} \cdots \xrightarrow{d_1} C_0 \to 0$. Then if the composition of any two successive homomorphisms is zero, the complex is called a *chain complex*, and its *homology groups* are defined as $H_n(C) = \ker d_n/\operatorname{image} d_{n+1}$. For chain complexes the notation is often chosen so that the indices appear as subscripts and are decreasing, whereas for cochain complexes the indices are superscripts and are increasing. We shall instead use a uniform notation for the maps on both, since it will be clear from the context whether we are dealing with a chain or a cochain complex.

Chain complexes were the first to arise in topological settings, with cochain complexes soon following. With our applications in Section 2 in mind, we shall concentrate on cochains and cohomology, although all of the general results in this section have similar statements for chains and homology. We shall also be interested in the situation where each C^n is an R-module and the homomorphisms d_n are R-module homomorphisms (referred to simply as a *complex of R-modules*), in which case the groups $H^n(C)$ are also R-modules.

Note that if C is a cochain (respectively, chain) complex then C is an exact sequence if and only if all its cohomology (respectively, homology) groups are zero. Thus the n^{th} cohomology (respectively, homology) group measures the failure of exactness of a complex at the n^{th} stage.

Definition. Let $\mathcal{A} = \{A^n\}$ and $\mathcal{B} = \{B^n\}$ be cochain complexes. A *homomorphism of complexes* $\alpha : \mathcal{A} \to \mathcal{B}$ is a set of homomorphisms $\alpha_n : A^n \to B^n$ such that for every n the following diagram commutes:

$$
\begin{tikzcd}
\cdots \arrow[r] & A^n \arrow[r] \arrow[d, "\alpha_n"] & A^{n+1} \arrow[r] \arrow[d, "\alpha_{n+1}"] & \cdots \\
\cdots \arrow[r] & B^n \arrow[r] & B^{n+1} \arrow[r] & \cdots
\end{tikzcd}
\tag{17.4}
$$

Proposition 1. A homomorphism $\alpha : \mathcal{A} \to \mathcal{B}$ of cochain complexes induces group homomorphisms from $H^n(\mathcal{A})$ to $H^n(\mathcal{B})$ for $n \geq 0$ on their respective cohomology groups.

Proof: It is an easy exercise to show that the commutativity of (4) implies that the images and kernels at each stage of the maps in the first row are mapped to the corresponding images and kernels for the maps in the second row, thus giving a well defined map on the respective quotient (cohomology) groups.

Definition. Let $\mathcal{A} = \{A^n\}$, $\mathcal{B} = \{B^n\}$ and $\mathcal{C} = \{C^n\}$ be cochain complexes. A *short exact sequence* of complexes $0 \to \mathcal{A} \xrightarrow{\alpha} \mathcal{B} \xrightarrow{\beta} \mathcal{C} \to 0$ is a sequence of homomorphisms of complexes such that $0 \to A^n \xrightarrow{\alpha_n} B^n \xrightarrow{\beta_n} C^n \to 0$ is short exact for every n.

One of the main features of cochain complexes is that they lead to long exact sequences in cohomology, which is our first main result:

Theorem 2. *(The Long Exact Sequence in Cohomology)* Let $0 \to \mathcal{A} \xrightarrow{\alpha} \mathcal{B} \xrightarrow{\beta} \mathcal{C} \to 0$ be a short exact sequence of cochain complexes. Then there is a long exact sequence of cohomology groups:

$$
\begin{aligned}
0 \to H^0(\mathcal{A}) \to H^0(\mathcal{B}) \to H^0(\mathcal{C}) &\xrightarrow{\delta_0} H^1(\mathcal{A}) \\
&\to H^1(\mathcal{B}) \to H^1(\mathcal{C}) \xrightarrow{\delta_1} H^2(\mathcal{A}) \to \cdots
\end{aligned}
\tag{17.5}
$$

where the maps between cohomology groups at each level are those in Proposition 1. The maps δ_n are called *connecting homomorphisms.*

Proof: The details of this proof are somewhat lengthy. For each n the verification that the sequence $H^n(\mathcal{A}) \to H^n(\mathcal{B}) \to H^n(\mathcal{C})$ is exact is a straightforward check of the definition of exactness of each map, similar to the proof of Theorem 33 in Section 10.5. The construction of a connecting homomorphism δ_n is outlined in Exercise 2. Some work is then needed to show that δ_n is a homomorphism, and that the sequence is exact at δ_n.

One immediate consequence of the existence of the long exact sequence in Theorem 2 is the fact that if any two of the cochain complexes $\mathcal{A}$, $\mathcal{B}$, $\mathcal{C}$ are exact, then so is the third (cf. Exercise 6).

Homomorphisms and the Groups $Ext^n_R(A,B)$

To apply Theorem 2 to analyze the sequence (2), we try to produce a cochain complex whose first few cohomology groups in the long exact sequence (5) agree with the terms in (2). To do this we introduce the notion of a "resolution" of an R-module:

Definition. Let A be any R-module. A *projective resolution* of A is an exact sequence

$$\cdots \longrightarrow P_n \xrightarrow{d_n} P_{n-1} \longrightarrow \cdots \xrightarrow{d_1} P_0 \xrightarrow{\epsilon} A \longrightarrow 0 \qquad (17.6)$$

such that each P_i is a projective R-module.

Every R-module has a projective resolution: Let P_0 be any free (hence projective) R-module on a set of generators of A and define an R-module homomorphism ϵ from P_0 onto A by Theorem 6 in Chapter 10. This begins the resolution $\epsilon : P_0 \to A \to 0$. The surjectivity of ϵ ensures that this sequence is exact. Next let $K_0 = \ker \epsilon$ and let P_1 be any free module mapping onto the submodule K_0 of P_0; this gives the second stage $P_1 \to P_0 \to A$ which, by construction, is also exact. We can continue this way, taking at the n^{th} stage a free R-module P_{n+1} that maps surjectively onto the submodule $\ker d_n$ of P_n, obtaining in fact a *free* resolution of A.

One of the reasons that *projective* modules are used in the resolution of A is that this makes it possible to lift various maps (cf. the proof of Proposition 4 following, for instance).

In general a projective resolution is infinite in length, but if A is itself projective, then it has a very simple projective resolution of finite length, namely $0 \to A \xrightarrow{1} A \to 0$ given by the identity map from A to itself.

Given the projective resolution (6), we may form a related sequence by taking homomorphisms of each of the terms into D, keeping in mind that this reverses the direction of the homomorphisms. This yields the sequence

$$\begin{aligned} 0 \longrightarrow \operatorname{Hom}_R(A, D) &\xrightarrow{\epsilon} \operatorname{Hom}_R(P_0, D) \xrightarrow{d_1} \operatorname{Hom}_R(P_1, D) \xrightarrow{d_2} \cdots \\ \cdots &\xrightarrow{d_{n-1}} \operatorname{Hom}_R(P_{n-1}, D) \xrightarrow{d_n} \operatorname{Hom}_R(P_n, D) \xrightarrow{d_{n+1}} \cdots \end{aligned} \qquad (17.7)$$

where to simplify notation we have denoted the induced maps from $\operatorname{Hom}_R(P_{n-1}, D)$ to $\operatorname{Hom}_R(P_n, D)$ for $n \geq 1$ again by d_n and similarly for the map induced by ϵ (cf. Section 10.5). This sequence is not necessarily exact, however it *is* a cochain complex (this is part of the proof of Theorem 33 in Section 10.5). The corresponding cohomology groups have a special name.

Definition. Let A and D be a R-modules. For any projective resolution of A as in (6) let $d_n : \operatorname{Hom}_R(P_{n-1}, D) \to \operatorname{Hom}_R(P_n, D)$ for all $n \geq 1$ as in (7). Define

$$\operatorname{Ext}^n_R(A, D) = \ker d_{n+1} / \operatorname{image} d_n$$

where $\operatorname{Ext}^0_R(A, D) = \ker d_1$. The group $\operatorname{Ext}^n_R(A, D)$ is called the n^{th} *cohomology group derived from the functor* $\operatorname{Hom}_R(\_\_, D)$. When $R = \mathbb{Z}$ the group $\operatorname{Ext}^n_{\mathbb{Z}}(A, D)$ is also denoted simply $\operatorname{Ext}^n(A, D)$.

Note that the groups $\mathrm{Ext}^n_R(A, D)$ are also the cohomology groups of the cochain complex obtained from (7) by replacing the term $\mathrm{Hom}_R(A, D)$ with zero (which does not effect the cochain property), i.e., they are the cohomology groups of the cochain complex $0 \to \mathrm{Hom}_R(P_0, D) \to \cdots$.

We shall show below that these cohomology groups do not depend on the choice of projective resolution of A. Before doing so we identify the 0^{th} cohomology group and give some examples.

Proposition 3. For any R-module A we have $\mathrm{Ext}^0_R(A, D) \cong \mathrm{Hom}_R(A, D)$.

Proof: Since the sequence $P_1 \xrightarrow{d_1} P_0 \xrightarrow{\epsilon} A \to 0$ is exact, it follows that the corresponding sequence $0 \to \mathrm{Hom}_R(A, D) \xrightarrow{\epsilon} \mathrm{Hom}_R(P_0, D) \xrightarrow{d_1} \mathrm{Hom}_R(P_1, D)$ is also exact by Theorem 33 in Section 10.5 (noting the first comment in the proof). Hence $\mathrm{Ext}^0_R(A, D) = \ker d_1 = \text{image}\, \epsilon \cong \mathrm{Hom}_R(A, D)$, as claimed.

Examples

(1) Let $R = \mathbb{Z}$ and let $A = \mathbb{Z}/m\mathbb{Z}$ for some $m \geq 2$. By the proposition we have $\mathrm{Ext}^0_{\mathbb{Z}}(\mathbb{Z}/m\mathbb{Z}, D) \cong \mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}/m\mathbb{Z}, D)$, and it follows that $\mathrm{Ext}^0_{\mathbb{Z}}(\mathbb{Z}/m\mathbb{Z}, D) \cong {}_mD$, where ${}_mD = \{d \in D \mid md = 0\}$ are the elements of D that have order dividing m. For the higher cohomology groups, we use the simple projective resolution

$$0 \longrightarrow \mathbb{Z} \xrightarrow{m} \mathbb{Z} \longrightarrow \mathbb{Z}/m\mathbb{Z} \longrightarrow 0$$

for A given by multiplication by m on $\mathbb{Z}$. Taking homomorphisms into a fixed $\mathbb{Z}$-module D gives the cochain complex

$$0 \longrightarrow \mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}/m\mathbb{Z}, D) \longrightarrow \mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}, D) \xrightarrow{m} \mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}, D) \longrightarrow 0 \longrightarrow \cdots.$$

We have $D \cong \mathrm{Hom}_{\mathbb{Z}}(\mathbb{Z}, D)$ (cf. Example 4 following Corollary 32 in Section 10.5) and under this isomorphism we have $\mathrm{Ext}^1_{\mathbb{Z}}(\mathbb{Z}/m\mathbb{Z}, D) \cong D/mD$ for any abelian group D. It follows immediately from the definition and the cochain complex above that $\mathrm{Ext}^n_{\mathbb{Z}}(\mathbb{Z}/m\mathbb{Z}, D) = 0$ for all $n \geq 2$ and any abelian group D, which we summarize as

$$\begin{aligned} &\mathrm{Ext}^0_{\mathbb{Z}}(\mathbb{Z}/m\mathbb{Z}, D) \cong {}_mD \\ &\mathrm{Ext}^1_{\mathbb{Z}}(\mathbb{Z}/m\mathbb{Z}, D) \cong D/mD \\ &\mathrm{Ext}^n_{\mathbb{Z}}(\mathbb{Z}/m\mathbb{Z}, D) = 0, \quad \text{for all } n \geq 2. \end{aligned}$$

(2) The same abelian groups may be modules over several different rings R and the Ext_R cohomology groups depend on R. For example, suppose $R = \mathbb{Z}/m\mathbb{Z}$ for some integer $m \geq 1$. An R-module D is the same as an abelian group D with exponent dividing m, i.e., $mD = 0$. In particular, for any divisor d of m, the group $\mathbb{Z}/d\mathbb{Z}$ is an R-module, and

$$\cdots \xrightarrow{m/d} \mathbb{Z}/m\mathbb{Z} \xrightarrow{d} \mathbb{Z}/m\mathbb{Z} \xrightarrow{m/d} \mathbb{Z}/m\mathbb{Z} \xrightarrow{d} \mathbb{Z}/m\mathbb{Z} \longrightarrow \mathbb{Z}/d\mathbb{Z} \longrightarrow 0$$

is a projective (in fact, free) resolution of $\mathbb{Z}/d\mathbb{Z}$ as a $\mathbb{Z}/m\mathbb{Z}$-module, where the final map is the natural projection mapping $x \bmod m$ to $x \bmod d$. Taking homomorphisms into the $\mathbb{Z}/m\mathbb{Z}$-module D, using the isomorphism $\mathrm{Hom}_{\mathbb{Z}/m\mathbb{Z}}(\mathbb{Z}/m\mathbb{Z}, D) \cong D$, and removing the first term gives the cochain complex

$$0 \longrightarrow D \xrightarrow{d} D \xrightarrow{m/d} D \xrightarrow{d} D \xrightarrow{m/d} \cdots.$$

Hence

$$\begin{aligned}
&\operatorname{Ext}^0_{\mathbb{Z}/m\mathbb{Z}}(\mathbb{Z}/d\mathbb{Z}, D) \cong {}_dD,\\
&\operatorname{Ext}^n_{\mathbb{Z}/m\mathbb{Z}}(\mathbb{Z}/d\mathbb{Z}, D) \cong {}_{(m/d)}D/dD, \quad n \text{ odd}, n \geq 1,\\
&\operatorname{Ext}^n_{\mathbb{Z}/m\mathbb{Z}}(\mathbb{Z}/d\mathbb{Z}, D) \cong {}_dD/(m/d)D, \quad n \text{ even}, n \geq 2,
\end{aligned}$$

where ${}_kD = \{d \in D \mid kd = 0\}$ denotes the set of elements of D killed by k. In particular, $\operatorname{Ext}^n_{\mathbb{Z}/p^2\mathbb{Z}}(\mathbb{Z}/p\mathbb{Z}, \mathbb{Z}/p\mathbb{Z}) \cong \mathbb{Z}/p\mathbb{Z}$ for all $n \geq 0$, whereas, for example, $\operatorname{Ext}^n_{\mathbb{Z}}(\mathbb{Z}/p\mathbb{Z}, \mathbb{Z}/p\mathbb{Z}) = 0$ for all $n \geq 2$.

In order to show that the cohomology groups $\operatorname{Ext}^n_R(A, D)$ are independent of the choice of projective resolution of A we shall need to be able to "compare" resolutions. The next proposition shows that an R-module homomorphism from A to B lifts to a homomorphism from a projective resolution of A to a projective resolution of B — this lifting property is one instance where the projectivity of the modules in the resolution is important.

Proposition 4. Let $f : A \to A'$ be any homomorphism of R-modules and take projective resolutions of A and A', respectively. Then for each $n \geq 0$ there is a lift f_n of f such that the following diagram commutes:

$$\begin{tikzcd}
\cdots \arrow[r, "d_2"] & P_1 \arrow[r, "d_1"] \arrow[d, "f_1"'] & P_0 \arrow[r, "\epsilon"] \arrow[d, "f_0"'] & A \arrow[r] \arrow[d, "f"'] & 0 \\
\cdots \arrow[r, "d_2'"] & P_1' \arrow[r, "d_1'"] & P_0' \arrow[r, "\epsilon'"] & A' \arrow[r] & 0
\end{tikzcd} \tag{17.8}$$

where the rows are the projective resolutions of A and A', respectively.

Proof: Given the two rows and map f in (8), then since P_0 is projective we may lift the map $f\epsilon : P_0 \to A'$ to a map $f_0 : P_0 \to P_0'$ in such a way that $\epsilon' f_0 = f\epsilon$ (Proposition 30(2) in Section 10.5). This gives the first lift of f. Proceeding inductively in this fashion, assume f_n has been defined to make the diagram commutative to that point. Thus image $f_n d_{n+1} \subseteq \ker d_n'$. The projectivity of P_{n+1} implies that we may lift the map $f_n d_{n+1} : P_{n+1} \to P_n'$ to a map $f_{n+1} : P_{n+1} \to P_{n+1}'$ to make the diagram commute at the next stage. This completes the proof.

The commutative diagram in Proposition 4 implies that the induced diagram

$$\begin{tikzcd}
0 \arrow[r] & \operatorname{Hom}_R(A, D) \arrow[r] & \operatorname{Hom}_R(P_0, D) \arrow[r] & \operatorname{Hom}_R(P_1, D) \arrow[r] & \cdots \\
0 \arrow[r] & \operatorname{Hom}_R(A, D) \arrow[r] \arrow[u, "f"] & \operatorname{Hom}_R(P_0', D) \arrow[r] \arrow[u, "f_0"] & \operatorname{Hom}_R(P_1', D) \arrow[r] \arrow[u, "f_1"] & \cdots
\end{tikzcd} \tag{17.9}$$

is also commutative. The two rows of this diagram are cochain complexes, and this commutative diagram depicts a homomorphism of these cochain complexes. By Proposition 1 we have an induced map on their cohomology groups:

Proposition 5. Let $f : A \to A'$ be a homomorphism of R-modules and take projective resolutions of A and A' as in Proposition 4. Then for every n there is an induced group homomorphism $\varphi_n : \mathrm{Ext}_R^n(A', D) \to \mathrm{Ext}_R^n(A, D)$ on the cohomology groups obtained via these resolutions, and the maps φ_n depend only on f, not on the choice of lifts f_n in Proposition 4.

Proof: The existence of the map on the cohomology groups Ext_R^n follows from Proposition 1 applied to the homomorphism of cochain complexes (9). The more difficult part is showing these maps do not depend on the choice of lifts f_n in Proposition 4. This is easily seen to be equivalent to showing that if f is the zero map, then the induced maps on cohomology groups are also all zero. Assume then that $f = 0$. By the projectivity of the modules P_i one may inductively define R-module homomorphisms $s_n : P_n \to P'_{n+1}$ with the property that for all n,

$$f_n = d'_{n+1}s_n + s_{n-1}d_n \tag{17.10}$$

so the maps s_n give reverse downward diagonal arrows across the squares in (8). (The collection of maps $\{s_n\}$ is called a *chain homotopy* between the chain homomorphism given by the f_n and the zero chain homomorphism, cf. Exercise 4.) Taking homomorphisms into D gives diagram (9) with additional upward diagonal arrows from the homomorphisms induced by the s_n, and these induced homomorphisms satisfy the relations in (10) (i.e., they form a homotopy between cochain complex homomorphisms). It is now an easy exercise using the diagonal maps added to (9) to see that any element in $\mathrm{Hom}_R(P'_n, D)$ representing a coset in $\mathrm{Ext}_R^n(A', D)$ maps to the zero coset in $\mathrm{Ext}_R^n(A, D)$ (cf. Exercise 4). This completes the argument.

One may also check that the homomorphism $\varphi_0 : \mathrm{Ext}_R^0(A', D) \to \mathrm{Ext}_R^0(A, D)$ in Proposition 5 is the same as the map $f : \mathrm{Hom}_R(A', D) \to \mathrm{Hom}_R(A, D)$ defined in Section 10.5 once the corresponding groups have been identified via the isomorphism in Proposition 3.

Theorem 6. The groups $\mathrm{Ext}_R^n(A, D)$ depend only on A and D, i.e., they are independent of the choice of projective resolution of A.

Proof: In the notation of Proposition 4 let $A' = A$, let $f : A \to A'$ be the identity map and let the two rows of (8) be two projective resolutions of A. For any choice of lifts of the identity map, the resulting homomorphisms on cohomology groups $\varphi_n : \mathrm{Ext}_R^n(A', D) \to \mathrm{Ext}_R^n(A, D)$ are seen to be isomorphisms as follows. Add a third row to the diagram (8) by copying the projective resolution in the top row below the second row. Let g be the identity map from A' to A and lift g to maps $g_n : P'_n \to P_n$ by Proposition 4. Let $\psi_n : \mathrm{Ext}_R^n(A, D) \to \mathrm{Ext}_R^n(A', D)$ be the resulting map on cohomology groups. The maps $g_n \circ f_n : P_n \to P_n$ are now a lift of the identity map $g \circ f$, and they are seen to induce the homomorphisms $\varphi_n \circ \psi_n$ on the cohomology groups. However, since the first and third rows are identical, taking the identity map from P_n to itself for all n is a particular lift of $g \circ f$, and this choice clearly induces the identity map on cohomology groups. The last assertion of Proposition 5 then implies that $\varphi_n \circ \psi_n$ is also the identity on $\mathrm{Ext}_R^n(A, D)$. By a symmetric argument $\psi_n \circ \varphi_n$ is the identity on $\operatorname{Ext}^n_R(A', D)$. This shows the maps φ_n and ψ_n are isomorphisms, as needed to complete the proof.

For a fixed R-module D and fixed integer $n \geq 0$, Proposition 5 and Theorem 6 show that $\operatorname{Ext}^n_R(\_\_, D)$ defines a (contravariant) functor from the category of R-modules to the category of abelian groups.

The next result shows that projective resolutions for a submodule and corresponding quotient module of an R-module M can be fit together to give a projective resolution of M.

Proposition 7. *(Simultaneous Resolution)* Let $0 \to L \to M \to N \to 0$ be a short exact sequence of R-modules, let $L = A$ have a projective resolution as in (6) above, and let N have a similar projective resolution where the projective modules are denoted by $\overline{P}_n$. Then there is a resolution of M by the projective modules $P_n \oplus \overline{P}_n$ such that the following diagram commutes:

$$\begin{tikzcd}
& \vdots \arrow[d] & \vdots \arrow[d] & \vdots \arrow[d] & \\
0 \arrow[r] & P_1 \arrow[r] \arrow[d] & P_1 \oplus \overline{P}_1 \arrow[r] \arrow[d] & \overline{P}_1 \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & P_0 \arrow[r] \arrow[d] & P_0 \oplus \overline{P}_0 \arrow[r] \arrow[d] & \overline{P}_0 \arrow[r] \arrow[d] & 0 \\
0 \arrow[r] & L \arrow[r] \arrow[d] & M \arrow[r] \arrow[d] & N \arrow[r] \arrow[d] & 0 \\
& 0 & 0 & 0 &
\end{tikzcd} \tag{17.11}$$

Moreover, the rows and columns of this diagram are exact and the rows are split.

Proof: The left and right nonzero columns of (11) are exact by hypothesis. The modules in the middle column are projective (cf. Exercise 3, Section 10.5) and the row maps are the obvious ones to make each row a split exact sequence. It remains then to define the vertical maps in the middle column in such a way as to make the diagram commute. This is accomplished in a straightforward manner, working inductively from the bottom upward — the first step in this process is outlined in Exercise 5.

Theorem 2 and Proposition 7 now yield the long exact sequence for Ext_R that extends the exact sequence (2).

Theorem 8. Let $0 \to L \to M \to N \to 0$ be a short exact sequence of R-modules. Then there is a long exact sequence of abelian groups

$$\begin{aligned} 0 \to \operatorname{Hom}_R(N, D) \to \operatorname{Hom}_R(M, D) \to \operatorname{Hom}_R(L, D) \xrightarrow{\delta_0} \operatorname{Ext}^1_R(N, D) \\ \to \operatorname{Ext}^1_R(M, D) \to \operatorname{Ext}^1_R(L, D) \xrightarrow{\delta_1} \operatorname{Ext}^2_R(N, D) \to \cdots \end{aligned} \tag{17.12}$$

where the maps between groups at the same level n are as in Proposition 5 and the connecting homomorphisms δ_n are given by Theorem 2.

Proof: Take a simultaneous projective resolution of the short exact sequence as in Proposition 7 and take homomorphisms into D. To obtain the cohomology groups Ext^n_R from the resulting diagram, as noted in the discussion preceding Proposition 3 we replace the lowest nonzero row in the transformed diagram with a row of zeros to get the following commutative diagram:

$$\begin{tikzcd} & \vdots & \vdots & \vdots & \\ 0 \arrow[r] & \operatorname{Hom}_R(\overline{P}_1, D) \arrow[r] \arrow[u] & \operatorname{Hom}_R(P_1 \oplus \overline{P}_1, D) \arrow[r] \arrow[u] & \operatorname{Hom}_R(P_1, D) \arrow[r] \arrow[u] & 0 \\ 0 \arrow[r] & \operatorname{Hom}_R(\overline{P}_0, D) \arrow[r] \arrow[u] & \operatorname{Hom}_R(P_0 \oplus \overline{P}_0, D) \arrow[r] \arrow[u] & \operatorname{Hom}_R(P_0, D) \arrow[r] \arrow[u] & 0 \\ & 0 \arrow[u] & 0 \arrow[u] & 0 \arrow[u] & \end{tikzcd} \tag{17.13}$$

The columns of (13) are cochain complexes, and the rows are split by Proposition 29(2) of Section 10.5 and the discussion following it. Thus (13) is a short exact sequence of cochain complexes. Theorem 2 then gives a long exact sequence of cohomology groups whose terms are, by definition, the groups $\operatorname{Ext}^n_R(\_\_, D)$, for $n \geq 0$. The 0^{th} order terms are identified by Proposition 3, completing the proof.

Theorem 8 shows how the exact sequence (2) can be extended in a natural way and shows that the group $\operatorname{Ext}^1_R(N, D)$ is the first measure of the failure of (2) to be exact on the right — in fact (2) can be extended to a short exact sequence on the right if and only if the connecting homomorphism δ_0 in (12) is the zero homomorphism. In particular, if $\operatorname{Ext}^1_R(N, D) = 0$ for all R-modules N, then (2) will be exact on the right for *every* exact sequence (1). We have already seen (Corollary 35 in Section 10.5) that this implies the R-module D is injective. Part of the next result shows that the converse is also true and characterizes injective modules in terms of Ext_R groups.

Proposition 9. For an R-module Q the following are equivalent:
(1) Q is injective,
(2) $\operatorname{Ext}^1_R(A, Q) = 0$ for all R-modules A, and
(3) $\operatorname{Ext}^n_R(A, Q) = 0$ for all R-modules A and all $n \geq 1$.

Proof: We showed (2) implies (1) above, and (3) implies (2) is trivial, so it remains to show that if Q is injective then $\mathrm{Ext}_R^n(A, Q) = 0$ for all R-modules A and all $n \geq 1$. Take a projective resolution

$$\cdots \longrightarrow P_n \longrightarrow P_{n-1} \longrightarrow \cdots \longrightarrow P_0 \longrightarrow A \longrightarrow 0$$

for A. Since Q is injective, the sequence

$$0 \to \mathrm{Hom}_R(A, Q) \to \mathrm{Hom}_R(P_0, Q) \to \cdots \to \mathrm{Hom}_R(P_{n-1}, Q) \to \mathrm{Hom}_R(P_n, Q) \to \cdots$$

is still exact (Corollary 35 in Section 10.5), so all of the cohomology groups for this cochain complex are 0. In particular, the groups $\mathrm{Ext}_R^n(A, Q)$ for $n \geq 1$ are all trivial, which is (3).

For a fixed R-module D, the result in Theorem 8 can be viewed as explaining what happens to the short exact sequence $0 \to L \to M \to N \to 0$ on the right after applying the left exact functor $\mathrm{Hom}_R(\_\_, D)$. This is why the (contravariant) functors $\mathrm{Ext}_R^n(\_\_, D)$ are called the *right derived functors* for the functor $\mathrm{Hom}_R(\_\_, D)$.

One can also consider the effect of applying the left exact functor $\mathrm{Hom}_R(D, \_\_)$, i.e., by taking homomorphisms *from* D rather than *into* D. The next theorem shows that in fact the same Ext_R groups define the (covariant) right derived functors for $\mathrm{Hom}_R(D, \_\_)$ as well.

Theorem 10. Let $0 \to L \to M \to N \to 0$ be a short exact sequence of R-modules. Then there is a long exact sequence of abelian groups

$$\begin{aligned} 0 \to \mathrm{Hom}_R(D, L) \to \mathrm{Hom}_R(D, M) \to \mathrm{Hom}_R(D, N) &\xrightarrow{\gamma_0} \mathrm{Ext}_R^1(D, L) \\ \to \mathrm{Ext}_R^1(D, M) \to \mathrm{Ext}_R^1(D, N) &\xrightarrow{\gamma_1} \mathrm{Ext}_R^2(D, L) \to \cdots. \end{aligned} \tag{17.14}$$

Proof: Let $0 \to L \to M \to N \to 0$ be a short exact sequence of R-modules. By taking a projective resolution of D and then applying $\mathrm{Hom}_R(\_\_, L)$, $\mathrm{Hom}_R(\_\_, M)$ and $\mathrm{Hom}_R(\_\_, N)$ to this resolution one obtains the columns in a commutative diagram similar to (13), but with L, M and N in the second positions rather than the first. Applying the Long Exact Sequence Theorem to this array gives (14).

Theorem 10 shows that the group $\mathrm{Ext}_R^1(D, L)$ measures whether the exact sequence

$$0 \longrightarrow \mathrm{Hom}_R(D, L) \longrightarrow \mathrm{Hom}_R(D, M) \longrightarrow \mathrm{Hom}_R(D, N)$$

can be extended to a short exact sequence — it can be extended if and only if γ_0 is the zero homomorphism. In particular, this will always be the case if the module D has the property that $\mathrm{Ext}_R^1(D, B) = 0$ for all R-modules B; in this case it follows by Corollary 32 in Section 10.5 that D is a projective R-module. As in the situation of injective R-modules in Proposition 9, the vanishing of these cohomology groups in fact characterizes projective R-modules:

Proposition 11. For an R-module P the following are equivalent:

(1) P is projective,

(2) $\mathrm{Ext}_R^1(P, B) = 0$ for all R-modules B, and

(3) $\mathrm{Ext}_R^n(P, B) = 0$ for all R-modules B and all $n \geq 1$.

Proof: We proved (2) implies (1) above, and (3) implies (2) is trivial, so it remains to prove that (1) implies (3). If P is a projective R-module, then the simple exact sequence

$$0 \longrightarrow P \xrightarrow{1} P \longrightarrow 0$$

given by the identity map on P is a projective resolution of P. Taking homomorphisms into B gives the simple cochain complex

$$0 \to \mathrm{Hom}_R(P, B) \xrightarrow{1} \mathrm{Hom}_R(P, B) \to 0 \to \cdots \to 0 \to \cdots$$

from which it follows by definition that $\mathrm{Ext}_R^n(P, B) = 0$ for all $n \geq 1$, which gives (3).

Examples

(1) Since $\mathbb{Z}^m$ is a free, hence projective, $\mathbb{Z}$-module, it follows from Proposition 11 that

$$\mathrm{Ext}_{\mathbb{Z}}^n(\mathbb{Z}^m, B) = 0$$

for all abelian groups B, all $m \geq 1$, and all $n \geq 1$.

(2) It is not difficult to show that $\mathrm{Ext}_R^n(A_1 \oplus A_2, B) \cong \mathrm{Ext}_R^n(A_1, B) \oplus \mathrm{Ext}_R^n(A_2, B)$ for all $n \geq 0$ (cf. Exercise 10), so the previous example together with the example following Proposition 3 determines $\mathrm{Ext}_{\mathbb{Z}}^n(A, B)$ for all finitely generated abelian groups A. In particular, $\mathrm{Ext}_{\mathbb{Z}}^n(A, B) = 0$ for all finitely generated groups A, all abelian groups B, and all $n \geq 2$.

We have chosen to define the cohomology group $\mathrm{Ext}_R^n(A, B)$ using a projective resolution of A. There is a parallel development using an *injective resolution* of B:

$$0 \to B \to Q_0 \to Q_1 \to \cdots$$

where each Q_i is injective. In this situation one defines $\mathrm{Ext}_R^n(A, B)$ as the n^{th} cohomology group of the cochain sequence obtained by applying $\mathrm{Hom}_R(A, \_\_)$ to the resolution for B. The theory proceeds in a manner analogous to the development of this section. Ultimately one shows that there is a natural isomorphism between the groups $\mathrm{Ext}_R^n(A, B)$ constructed using both methods.

Examples

(1) Suppose $R = \mathbb{Z}$ and A and B are $\mathbb{Z}$-modules, i.e., are abelian groups. Recall that a $\mathbb{Z}$-module is injective if and only if it is divisible (Proposition 36 in Section 10.5). The group B can be embedded in an injective $\mathbb{Z}$-module Q_0 (Corollary 37 in Section 10.5) and the quotient, Q_1, of Q_0 by the image of B is again injective. Hence we have an injective resolution

$$0 \longrightarrow B \longrightarrow Q_0 \longrightarrow Q_1 \longrightarrow 0$$

of B. Applying $\mathrm{Hom}_{\mathbb{Z}}(A, \_\_)$ to this sequence gives the cochain complex

$$0 \longrightarrow \mathrm{Hom}_{\mathbb{Z}}(A, B) \longrightarrow \mathrm{Hom}_{\mathbb{Z}}(A, Q_0) \longrightarrow \mathrm{Hom}_{\mathbb{Z}}(A, Q_1) \longrightarrow 0 \longrightarrow \cdots$$

from which it follows immediately that

$$\mathrm{Ext}^n_{\mathbb{Z}}(A, B) = 0$$

for all abelian groups A and B and all $n \geq 2$, showing that the result of the previous example holds also when A is not finitely generated.

(2) Suppose A is a torsion abelian group. Then we have $\mathrm{Ext}^0(A, \mathbb{Z}) \cong \mathrm{Hom}(A, \mathbb{Z}) = 0$ since $\mathbb{Z}$ is torsion free. The sequence $0 \to \mathbb{Z} \to \mathbb{Q} \to \mathbb{Q}/\mathbb{Z} \to 0$ gives an injective resolution of $\mathbb{Z}$. Applying $\mathrm{Hom}(A, \_\_)$ gives the cochain complex

$$0 \longrightarrow \mathrm{Hom}(A, \mathbb{Z}) \longrightarrow \mathrm{Hom}(A, \mathbb{Q}) \longrightarrow \mathrm{Hom}(A, \mathbb{Q}/\mathbb{Z}) \longrightarrow 0 \longrightarrow \cdots$$

and since $\mathbb{Q}$ is also torsion free, this shows that

$$\mathrm{Ext}^1_{\mathbb{Z}}(A, \mathbb{Z}) \cong \mathrm{Hom}_{\mathbb{Z}}(A, \mathbb{Q}/\mathbb{Z}).$$

The group $\mathrm{Hom}(A, \mathbb{Q}/\mathbb{Z})$ is called the *Pontriagin dual group* to A. If A is a finite abelian group the Pontriagin dual of A is isomorphic to A (cf. Exercise 14, Section 5.2). In particular, $\mathrm{Ext}^1(A, \mathbb{Z}) \cong A$ is nonzero for all nonzero finite abelian groups A. We have $\mathrm{Ext}^n(A, \mathbb{Z}) = 0$ for all $n \geq 2$ by the previous example.

We record an important property of Ext^1_R, which helps to explain the name for these cohomology groups. Recall that equivalent extensions were defined at the beginning of Section 10.5.

Theorem 12. For any R-modules N and L there is a bijection between $\mathrm{Ext}^1_R(N, L)$ and the set of equivalence classes of extensions of N by L.

Although we shall not prove this result, in Section 4 we establish a similar bijection between equivalence classes of group extensions of G by A and elements of a certain cohomology group, where G is any finite group and A is any $\mathbb{Z}G$-module.

Example

Suppose $R = \mathbb{Z}$ and $A = B = \mathbb{Z}/p\mathbb{Z}$. We showed above that $\mathrm{Ext}^1_R(\mathbb{Z}/p\mathbb{Z}, \mathbb{Z}/p\mathbb{Z}) \cong \mathbb{Z}/p\mathbb{Z}$, so by Theorem 12 there are precisely p equivalence classes of extensions of $\mathbb{Z}/p\mathbb{Z}$ by $\mathbb{Z}/p\mathbb{Z}$. These are given by the direct sum $\mathbb{Z}/p\mathbb{Z} \oplus \mathbb{Z}/p\mathbb{Z}$ (which corresponds to the trivial class in $\mathrm{Ext}^1_R(\mathbb{Z}/p\mathbb{Z}, \mathbb{Z}/p\mathbb{Z})$) and the $p - 1$ extensions

$$0 \longrightarrow \mathbb{Z}/p\mathbb{Z} \longrightarrow \mathbb{Z}/p^2\mathbb{Z} \xrightarrow{\ i\ } \mathbb{Z}/p\mathbb{Z} \longrightarrow 0$$

defined by the map $i(x) = ix \bmod p$ for $i = 1, 2, \ldots, p - 1$. Note that while these are inequivalent as extensions, they all determine the same group $\mathbb{Z}/p^2\mathbb{Z}$.

Tensor Products and the Groups $Tor_n^R(A,B)$

The cohomology groups $\mathrm{Ext}^n_R(A, B)$ determine what happens to short exact sequences on the right after applying the left exact functors $\mathrm{Hom}_R(D, \_\_)$ and $\mathrm{Hom}_R(\_\_, D)$. One may similarly ask for the behavior of short exact sequences on the left after applying the right exact functor $D \otimes_R \_\_$ or the right exact functor $\_\_ \otimes_R D$. This leads to the Tor (homology) groups (whose name derives from their relation to torsion submodules), and we now briefly outline the development of these left derived functors. In some respects this theory is "dual" to the theory for Ext_R. We concentrate on the situation for $D \otimes_R \_\_$ when D is a right R-module. When D is a left R-module there is a completely symmetric theory for $\_\_ \otimes_R D$; when R is commutative and all R-modules have the same left and right R action the homology groups resulting from both developments are isomorphic.

Suppose then that D is a right R-module. Then for every left R-module B the tensor product $D \otimes_R B$ is an abelian group and the functor $D \otimes \_\_$ is covariant and right exact, i.e., for any short exact sequence (1) of left R-modules,

$$D \otimes L \longrightarrow D \otimes M \longrightarrow D \otimes N \longrightarrow 0$$

is an exact sequence of abelian groups. This sequence may be extended at the left end to a long exact sequence as follows. Let

$$\cdots \longrightarrow P_n \xrightarrow{d_n} P_{n-1} \longrightarrow \cdots \xrightarrow{d_1} P_0 \xrightarrow{\epsilon} B \longrightarrow 0$$

be a projective resolution of B, and take tensor products with D to obtain

$$\cdots \longrightarrow D \otimes P_n \xrightarrow{1\otimes d_n} D \otimes P_{n-1} \longrightarrow \cdots \xrightarrow{1\otimes d_1} D \otimes P_0 \xrightarrow{1\otimes \epsilon} D \otimes B \longrightarrow 0. \quad (17.15)$$

It follows from the argument in Theorem 39 of Section 10.5 that (15) is a chain complex — the composition of any two successive maps is zero — so we may form its homology groups.

Definition. Let D be a right R-module and let B be a left R-module. For any projective resolution of B by left R-modules as above let $1 \otimes d_n : D \otimes P_n \to D \otimes P_{n-1}$ for all $n \geq 1$ as in (15). Then

$$\mathrm{Tor}_n^R(D, B) = \ker(1 \otimes d_n)/\,\mathrm{image}(1 \otimes d_{n+1})$$

where $\mathrm{Tor}_0^R(D, B) = (D \otimes P_0)/\,\mathrm{image}(1 \otimes d_1)$. The group $\mathrm{Tor}_n^R(D, B)$ is called the n^{th} *homology group derived from the functor* $D \otimes \_\_$. When $R = \mathbb{Z}$ the group $\mathrm{Tor}_n^{\mathbb{Z}}(D, B)$ is also denoted simply $\mathrm{Tor}_n(D, B)$.

Thus $\mathrm{Tor}_n^R(D, B)$ is the n^{th} homology group of the chain complex obtained from (15) by removing the term $D \otimes B$.

A completely analogous proof to Proposition 3 (but relying on Theorem 39 in Section 10.5) implies the following:

Proposition 13. For any left R-module B we have $\operatorname{Tor}_0^R(D, B) \cong D \otimes B$.

Example

Let $R = \mathbb{Z}$ and let $B = \mathbb{Z}/m\mathbb{Z}$ for some $m \geq 2$. By the proposition, $\operatorname{Tor}_0^{\mathbb{Z}}(D, \mathbb{Z}/m\mathbb{Z})$ is isomorphic to $D \otimes \mathbb{Z}/m\mathbb{Z}$, so we have $\operatorname{Tor}_0^{\mathbb{Z}}(D, \mathbb{Z}/m\mathbb{Z}) \cong D/mD$ (Example 8 following Corollary 12 in Section 10.4). For the higher groups we apply $D \otimes \_\_$ to the projective resolution

$$0 \longrightarrow \mathbb{Z} \xrightarrow{m} \mathbb{Z} \longrightarrow \mathbb{Z}/m\mathbb{Z} \longrightarrow 0$$

of B and use the isomorphisms $D \otimes \mathbb{Z} \cong D$ and $D \otimes \mathbb{Z}/m\mathbb{Z} \cong D/mD$. This gives the chain complex

$$\cdots \longrightarrow 0 \longrightarrow D \xrightarrow{m} D \longrightarrow D/mD \longrightarrow 0.$$

It follows that $\operatorname{Tor}_1^{\mathbb{Z}}(D, \mathbb{Z}/m\mathbb{Z}) \cong {}_mD$ is the subgroup of D annihilated by m and that $\operatorname{Tor}_n^{\mathbb{Z}}(D, \mathbb{Z}/m\mathbb{Z}) = 0$ for all $n \geq 2$, which we summarize as

$$\begin{aligned}
&\operatorname{Tor}_0(D, \mathbb{Z}/m\mathbb{Z}) \cong D/mD,\\
&\operatorname{Tor}_1(D, \mathbb{Z}/m\mathbb{Z}) \cong {}_mD,\\
&\operatorname{Tor}_n(D, \mathbb{Z}/m\mathbb{Z}) = 0, \quad \text{for all } n \geq 2.
\end{aligned}$$

As for Ext, the Tor groups depend on the ring R (cf. Exercise 20).

Following a similar development to that for Ext_R, one shows:

Proposition 14.

(1) The homology groups $\operatorname{Tor}_n^R(D, B)$ are independent of the choice of projective resolution of B, and

(2) for every R-module homomorphism $f : B \to B'$ there are induced maps $\psi_n : \operatorname{Tor}_n^R(D, B) \to \operatorname{Tor}_n^R(D, B')$ on homology groups (depending only on f).

There is a Long Exact Sequence in Homology analogous to Theorem 2, except that all the arrows are reversed, whose proof follows mutatis mutandis from the argument for cohomology. This together with Simultaneous Resolution gives:

Theorem 15. Let $0 \to L \to M \to N \to 0$ be a short exact sequence of left R-modules. Then there is a long exact sequence of abelian groups

$$\begin{aligned}
\cdots \to \operatorname{Tor}_2^R(D, N) \xrightarrow{\delta_1} \operatorname{Tor}_1^R(D, L) \to \operatorname{Tor}_1^R(D, M) \to \\
\operatorname{Tor}_1^R(D, N) \xrightarrow{\delta_0} D \otimes L \to D \otimes M \to D \otimes N \to 0
\end{aligned}$$

where the maps between groups at the same level n are as in Proposition 14 (and the maps δ_n are called connecting homomorphisms).

There is a characterization of flat modules corresponding to Propositions 9 and 11 whose proof is very similar and is left as an exercise.

Proposition 16. For a right R-module D the following are equivalent:

(1) D is a flat R-module,

(2) $\operatorname{Tor}_1^R(D, B) = 0$ for all left R-modules B, and

(3) $\operatorname{Tor}_n^R(D, B) = 0$ for all left R-modules B and all $n \geq 1$.

We have defined $\operatorname{Tor}_n^R(A, B)$ as the homology of the chain complex obtained by tensoring a projective resolution of B on the left with A. The same groups are obtained by taking the homology of the chain complex obtained by tensoring a projective resolution of A on the right by B. Put another way, the $\operatorname{Tor}_n^R(A, B)$ groups define the (covariant) left derived functors for both of the right exact functors $A \otimes_R \_\_$ and $\_\_ \otimes_R B$: if D is a left R-module, then the short exact sequence $0 \to L \to M \to N \to 0$ of right R-modules gives rise to the long exact sequence

$$\cdots \to \operatorname{Tor}_2^R(N, D) \xrightarrow{\gamma_1} \operatorname{Tor}_1^R(L, D) \to \operatorname{Tor}_1^R(M, D) \to$$
$$\operatorname{Tor}_1^R(N, D) \xrightarrow{\gamma_0} L \otimes_R D \to M \otimes_R D \to N \otimes_R D \to 0$$

of abelian groups. In particular, the left R-module D is flat if and only if $\operatorname{Tor}_1^R(A, D) = 0$ for all right R-modules A.

When R is commutative, $A \otimes_R B \cong B \otimes_R A$ (Proposition 20 in Section 10.4) for any two R-modules A and B with the standard R-module structures, and it follows that $\operatorname{Tor}_n^R(A, B) \cong \operatorname{Tor}_n^R(B, A)$ as R-modules. When R is commutative the Tor long exact sequences are exact sequences of R-modules.

Examples

(1) If $R = \mathbb{Z}$, then since $\mathbb{Z}^m$ is free, hence flat (Corollary 42, Section 10.5), we have $\operatorname{Tor}_n(A, \mathbb{Z}^m) = 0$ for all $n \geq 1$ and all abelian groups A.

(2) Since $\operatorname{Tor}_n^R(A, B_1 \oplus B_2) \cong \operatorname{Tor}_n^R(A, B_1) \oplus \operatorname{Tor}_n^R(A, B_2)$ (cf. Exercise 10), the previous two examples together determine $\operatorname{Tor}_n^R(A, B)$ for all abelian groups A and all finitely generated abelian groups B.

(3) As a particular case of the previous example, $\operatorname{Tor}_1(A, B)$ is a torsion group and $\operatorname{Tor}_n(A, B) = 0$ for every abelian group A, every finitely generated abelian group B, and all $n \geq 2$. In fact these results hold without the condition that B be finitely generated.

(4) The exact sequence $0 \to \mathbb{Z} \to \mathbb{Q} \to \mathbb{Q}/\mathbb{Z} \to 0$ gives the long exact sequence

$$\cdots \to \operatorname{Tor}_1(D, \mathbb{Q}) \to \operatorname{Tor}_1(D, \mathbb{Q}/\mathbb{Z}) \to D \otimes \mathbb{Z} \to D \otimes \mathbb{Q} \to D \otimes \mathbb{Q}/\mathbb{Z} \to 0.$$

Since $\mathbb{Q}$ is a flat $\mathbb{Z}$-module (Example 2 following Corollary 42 in Section 10.5), the proposition shows that we have an exact sequence

$$0 \longrightarrow \operatorname{Tor}_1(D, \mathbb{Q}/\mathbb{Z}) \longrightarrow D \longrightarrow D \otimes \mathbb{Q}$$

and so $\operatorname{Tor}_1(D, \mathbb{Q}/\mathbb{Z})$ is isomorphic to the kernel of the natural map from D into $D \otimes \mathbb{Q}$, which is the torsion subgroup of D (cf. Exercise 9 in Section 10.4).

The following results show that, for $R = \mathbb{Z}$, the Tor groups are closely related to torsion subgroups. The Tor groups first arose in applications of torsion abelian groups in topological settings, which helps explain the terminology.

Proposition 17. Let A and B be $\mathbb{Z}$-modules and let $t(A)$ and $t(B)$ denote their respective torsion submodules. Then $\operatorname{Tor}_1(A, B) \cong \operatorname{Tor}_1(t(A), t(B))$.

Proof: In the case where A and B are finitely generated abelian groups this follows by Examples 3 and 4 above. For the general case, cf. Exercise 16.

Corollary 18. If A is an abelian group then A is torsion free if and only if $\operatorname{Tor}_1(A, B) = 0$ for every abelian group B (in which case A is flat as a $\mathbb{Z}$-module).

Proof: By the proposition, if A has no elements of finite order then we have $\operatorname{Tor}_1(A, B) = \operatorname{Tor}_1(t(A), B) = \operatorname{Tor}_1(0, B) = 0$ for every abelian group B. Conversely, if $\operatorname{Tor}_1(A, B) = 0$ for all B, then in particular $\operatorname{Tor}_1(A, \mathbb{Q}/\mathbb{Z}) = 0$, and this group is isomorphic to the torsion subgroup of A by the example above.

The results of Proposition 17 and Corollary 18 hold for any P.I.D. R in place of $\mathbb{Z}$ (cf. Exercise 26 in Section 10.5 and Exercise 16).

Finally, we mention that the cohomology and homology theories we have described may be developed in a vastly more general setting by axiomatizing the essential properties of R-modules and the Hom_R and tensor product functors. This leads to the general notions of *abelian categories* and *additive functors.* In the case of the abelian category of R-modules, any additive functor $\mathcal{F}$ to the category of abelian groups gives rise to a set of *derived functors,* $\mathcal{F}_n$, also from R-modules to abelian groups, for all $n \geq 0$. Then for each short exact sequence $0 \to L \to M \to N \to 0$ of R-modules there is a long exact sequence of (cohomology or homology) groups whose terms are $\mathcal{F}_n(L)$, $\mathcal{F}_n(M)$ and $\mathcal{F}_n(N)$, and these long exact sequences reflect the exactness properties of the functor $\mathcal{F}$. If $\mathcal{F}$ is left or right exact then the 0^{th} derived functor $\mathcal{F}_0$ is naturally equivalent to $\mathcal{F}$ (hence the 0^{th} degree groups $\mathcal{F}_0(X)$ are isomorphic to $\mathcal{F}(X)$), and if $\mathcal{F}$ is an exact functor then $\mathcal{F}_n(X) = 0$ for all $n \geq 1$ and all R-modules X.

EXERCISES

1. Give the details of the proof of Proposition 1.
2. This exercise defines the connecting map δ_n in the Long Exact Sequence of Theorem 2 and proves it is a homomorphism. In the notation of Theorem 2 let $0 \to \mathcal{A} \xrightarrow{\alpha} \mathcal{B} \xrightarrow{\beta} \mathcal{C} \to 0$ be a short exact sequence of cochain complexes, where for simplicity the cochain maps for $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ are all denoted by the same d.
   (a) If $c \in C^n$ represents the class $x \in H^n(\mathcal{C})$ show that there is some $b \in B^n$ with $\beta_n(b) = c$.
   (b) Show that $d_{n+1}(b) \in \ker \beta_{n+1}$ and conclude that there is a unique $a \in A^{n+1}$ such that $\alpha_{n+1}(a) = d_{n+1}(b)$. [Use $c \in \ker d_{n+1}$ and the commutativity of the diagram.]
   (c) Show that $d_{n+2}(a) = 0$ and conclude that a defines a class $\overline{a}$ in the quotient group $H^{n+1}(\mathcal{A})$. [Use the fact that α_{n+2} is injective.]
   (d) Prove that $\overline{a}$ is independent of the choice of b, i.e., if b' is another choice and a' is its unique preimage in A^{n+1} then $\overline{a} = \overline{a'}$, and that $\overline{a}$ is also independent of the choice of c representing the class x.
   (e) Define $\delta_n(x) = \overline{a}$ and prove that δ_n is a group homomorphism from $H^n(\mathcal{C})$ to $H^{n+1}(\mathcal{A})$. [Use the fact that $\delta_n(x)$ is independent of the choices of c and b to compute $\delta_n(x_1 + x_2)$.]

3. Suppose

$$\begin{tikzcd} & A \arrow[r, "\alpha"] \arrow[d, "f"'] & B \arrow[r, "\beta"] \arrow[d, "g"'] & C \arrow[r] \arrow[d, "h"'] & 0 \\ 0 \arrow[r] & A' \arrow[r, "\alpha'"] & B' \arrow[r, "\beta'"] & C' & \end{tikzcd}$$

is a commutative diagram of R-modules with exact rows.

(a) If $c \in \ker h$ and $\beta(b) = c$ prove that $g(b) \in \ker \beta'$ and conclude that $g(b) = \alpha'(a')$ for some $a' \in A'$. [Use the commutativity of the diagram.]

(b) Show that $\delta(c) = a' \bmod \text{image } f$ is a well defined R-module homomorphism from $\ker h$ to the quotient $A' / \text{image } f$.

(c) *(The Snake Lemma)* Prove there is an exact sequence

$$\ker f \longrightarrow \ker g \longrightarrow \ker h \xrightarrow{\delta} \operatorname{coker} f \longrightarrow \operatorname{coker} g \longrightarrow \operatorname{coker} h$$

where $\operatorname{coker} f$ (the *cokernel* of f) is $A'/(\text{image } f)$ and similarly for $\operatorname{coker} g$ and $\operatorname{coker} h$.

(d) Show that if α is injective and β' is surjective (i.e., the two rows in the commutative diagram above can be extended to short exact sequences) then the exact sequence in (c) can be extended to the exact sequence

$$0 \longrightarrow \ker f \longrightarrow \ker g \longrightarrow \ker h \xrightarrow{\delta} \operatorname{coker} f \longrightarrow \operatorname{coker} g \longrightarrow \operatorname{coker} h \longrightarrow 0$$

4. Let $\mathcal{A} = \{A^n\}$ and $\mathcal{B} = \{B^n\}$ be cochain complexes, where the maps $A^n \to A^{n+1}$ and $B^n \to B^{n+1}$ in both complexes are denoted by d_{n+1} for all n. Cochain complex homomorphisms α and β from $\mathcal{A}$ to $\mathcal{B}$ are said to be *homotopic* if for all n there are module homomorphisms $s_n : A^{n+1} \to B^n$ such that the maps $\alpha_n - \beta_n$ from A^n to B^n satisfy

$$\alpha_n - \beta_n = d_n s_{n-1} + s_n d_{n+1}.$$

The collection of maps $\{s_n\}$ is called a *cochain homotopy* from α to β. One may similarly define chain homotopies between chain complexes.

(a) Prove that homotopic maps of cochain complexes induce the same maps on cohomology, i.e., if α and β are homotopic homomorphisms of cochain complexes then the induced group homomorphisms from $H^n(\mathcal{A})$ to $H^n(\mathcal{B})$ are equal for every $n \geq 0$. (Thus "homotopy" gives a sufficient condition for two maps of complexes to induce the same maps on cohomology or homology; this condition is not in general necessary.) [Use the definition of homotopy to show $(\alpha_n - \beta_n)(z) \in \text{image } d_n$ for every $z \in \ker d_{n+1}$.]

(b) Prove that the relation $\alpha \sim \beta$ if α and β are homotopic is an equivalence relation on any set of cochain complex homomorphisms.

5. Establish the first step in the Simultaneous Resolution result of Proposition 7 as follows: assume the first two nonzero rows in diagram (11) are given, except for the map from $P_0 \oplus \overline{P}_0$ to M (where the maps along the row of projective modules are the obvious injection and projection for this split exact sequence). Let $\mu : \overline{P}_0 \to M$ be a lifting to $\overline{P}_0$ of the map $\overline{P}_0 \to N$ (which exists because $\overline{P}_0$ is projective). Let λ be the composition $P_0 \to L \to M$ in the diagram. Define

$$\pi : P_0 \oplus \overline{P}_0 \to M \qquad \text{by} \qquad \pi(x, y) = \lambda(x) + \mu(y).$$

Show that with this definition the first two nonzero rows of (11) form a commutative diagram.

6. Let $0 \to \mathcal{A} \xrightarrow{\alpha} \mathcal{B} \xrightarrow{\beta} \mathcal{C} \to 0$ be a short exact sequence of cochain complexes. Prove that if any two of $\mathcal{A}, \mathcal{B}, \mathcal{C}$ are exact, then so is the third. [Use Theorem 2.]
7. Prove that a finitely generated abelian group A is free if and only if $\operatorname{Ext}^1(A, \mathbb{Z}) = 0$.
8. Prove that if $0 \to L \to M \to N \to 0$ is a split short exact sequence of R-modules, then for every $n \geq 0$ the sequence $0 \to \operatorname{Ext}^n_R(N, D) \to \operatorname{Ext}^n_R(M, D) \to \operatorname{Ext}^n_R(L, D) \to 0$ is also short exact and split. [Use a splitting homomorphism and Proposition 5.]
9. Show that

$$0 \longrightarrow \mathbb{Z}/d\mathbb{Z} \longrightarrow \mathbb{Z}/m\mathbb{Z} \xrightarrow{d} \mathbb{Z}/m\mathbb{Z} \xrightarrow{m/d} \mathbb{Z}/m\mathbb{Z} \xrightarrow{d} \mathbb{Z}/m\mathbb{Z} \xrightarrow{m/d} \cdots$$

is an injective resolution of $\mathbb{Z}/d\mathbb{Z}$ as a $\mathbb{Z}/m\mathbb{Z}$-module. [Use Proposition 36 in Section 10.5.] Use this to compute the groups $\operatorname{Ext}^n_{\mathbb{Z}/m\mathbb{Z}}(A, \mathbb{Z}/d\mathbb{Z})$ in terms of the dual group $\operatorname{Hom}_{\mathbb{Z}/m\mathbb{Z}}(A, \mathbb{Z}/m\mathbb{Z})$. In particular, if $m = p^2$ and $d = p$, give another derivation of the result $\operatorname{Ext}^n_{\mathbb{Z}/p^2\mathbb{Z}}(\mathbb{Z}/p\mathbb{Z}, \mathbb{Z}/p\mathbb{Z}) \cong \mathbb{Z}/p\mathbb{Z}$.
10. (a) Prove that an arbitrary direct sum $\oplus_{i \in I} P_i$ of projective modules P_i is projective and that an arbitrary direct product $\prod_{j \in J} Q_j$ of injective modules Q_j is injective.
    (b) Prove that an arbitrary direct sum of projective resolutions is again projective and use this to show $\operatorname{Ext}^n_R(\oplus_{i \in I} A_i, B) \cong \prod_{i \in I} \operatorname{Ext}^n_R(A_i, B)$ for any collection of R-modules A_i $(i \in I)$. [cf. Exercise 12 in Section 10.5.]
    (c) Prove that an arbitrary direct product of injective resolutions is an injective resolution and use this to show $\operatorname{Ext}^n_R(A, \prod_{j \in J} B_j) \cong \prod_{j \in J} \operatorname{Ext}^n_R(A, B_j)$ for any collection of R-modules B_j $(j \in J)$. [cf. Exercise 12 in Section 10.5.]
    (d) Prove that $\operatorname{Tor}^R_n(A, \oplus_{j \in J} B_j) \cong \oplus_{j \in J} \operatorname{Tor}^R_n(A, B_j)$ for any collection of R-modules B_j $(j \in J)$.
11. *(Bass' Characterization of Noetherian Rings)* Suppose R is a commutative ring.
    (a) If R is Noetherian, and I is any nonzero ideal in R show that the image of any R-module homomorphism $f : I \to \oplus_{j \in \mathcal{J}} Q_j$ from I into a direct sum of injective R-modules Q_j $(j \in \mathcal{J})$ is contained in some finite direct sum of the Q_j.
    (b) If R is Noetherian, prove that an arbitrary direct sum $\oplus_{j \in \mathcal{J}} Q_j$ of injective R-modules is again injective. [Use Baer's Criterion (Proposition 36) and Exercise 4 in Section 10.5 together with (a).]
    (c) Let $I_1 \subseteq I_2 \subseteq \ldots$ be an ascending chain of ideals of R with union I and let $I/I_i \to Q_i$ for $i = 1, 2, \ldots$ be an injection of the quotient I/I_i into an injective R-module Q_i (by Theorem 38 in Section 10.5). Prove that the composition of these injections with the product of the canonical projection maps $I \to I_i$ gives an R-module homomorphism $f : I \to \oplus_{i=1,2,\ldots} Q_i$.
    (d) Prove the converse of (b): if an arbitrary direct sum $\oplus_{j \in \mathcal{J}} Q_j$ of injective R-modules is again injective then R is Noetherian. [If the direct sum in (c) is injective, use Baer's Criterion to lift f to a homomorphism $F : R \to \oplus_{i=1,2,\ldots} Q_i$. If the component of $F(1)$ in Q_i is 0 for $i \geq n$ prove that $I = I_n$ and the ascending chain of ideals is finite.]
12. Prove Proposition 13: $\operatorname{Tor}^R_0(D, A) \cong D \otimes_R A$. [Follow the proof of Proposition 3.]
13. Prove Proposition 16 characterizing flat modules.
14. Suppose $0 \to A \to B \to C \to 0$ is a short exact sequence of R-modules. Prove that if C is a flat R-module, then A is flat if and only if B is also flat. [Use the Tor long exact sequence.] Give an example to show that if A and B are flat then C need not be flat.

15. **(a)** If I is an ideal in R and M is an R-module, prove that $\mathrm{Tor}_1^R(M, R/I)$ is isomorphic to the kernel of the map $M \otimes_R I \to M$ that maps $m \otimes i$ to mi for $i \in I$ and $m \in M$. [Use the Tor long exact sequence associated to $0 \to I \to R \to R/I \to 0$ noting that R is flat.]

(b) *(A Flatness Criterion using* Tor*)* Prove that the R-module M is flat if and only if $\mathrm{Tor}_1^R(M, R/I) = 0$ for every finitely generated ideal I of R. [Use Exercise 25 in Section 10.5.]

16. Suppose R is a P.I.D. and A and B are R-modules. If $t(B)$ denotes the torsion submodule of B show that $\mathrm{Tor}_1^R(A, t(B)) \cong \mathrm{Tor}_1^R(A, B)$ and deduce that $\mathrm{Tor}_1^R(A, B)$ is isomorphic to $\mathrm{Tor}_1^R(t(A), t(B))$. [Use Exercise 26 in Section 10.5 to show that $B/t(B)$ is flat over R, then use the Tor long exact sequence with $D = A$ applied to the short exact sequence $0 \to t(B) \to B \to B/t(B) \to 0$ and the remarks following Proposition 16.]

17. Let $A = \mathbb{Z}/2\mathbb{Z} \oplus \mathbb{Z}/3\mathbb{Z} \oplus \mathbb{Z}/4\mathbb{Z} \oplus \cdots$. Prove that $\mathrm{Ext}^1(A, B) \cong (B/2B) \times (B/3B) \times (B/4B) \times \cdots$ for any abelian group A. [Use Exercise 10.] Prove that $\mathrm{Ext}^1(A, B) = 0$ if and only if B is divisible.

18. Prove that $\mathbb{Z}/2\mathbb{Z}$ is a projective $\mathbb{Z}/6\mathbb{Z}$-module and deduce that $\mathrm{Tor}_1^{\mathbb{Z}/6\mathbb{Z}}(\mathbb{Z}/2\mathbb{Z}, \mathbb{Z}/2\mathbb{Z}) = 0$.

19. Suppose $r \neq 0$ is not a zero divisor in the commutative ring R.

(a) Prove that multiplication by r gives a free resolution $0 \to R \xrightarrow{r} R \to R/rR \to 0$ of the quotient R/rR.

(b) Prove that $\mathrm{Ext}_R^0(R/rR, B) = {}_rB$ is the set of elements $b \in B$ with $rb = 0$, that $\mathrm{Ext}_R^1(R/rR, B) \cong B/rB$, and that $\mathrm{Ext}_R^n(R/rR, B) = 0$ for $n \geq 2$ for every R-module B.

(c) Prove that $\mathrm{Tor}_0^R(A, R/rR) = A/rA$, that $\mathrm{Tor}_1^R(A, R/rR) = {}_rA$ is the set of elements $a \in A$ with $ra = 0$, and that $\mathrm{Tor}_n^R(A, R/rR) = 0$ for $n \geq 2$ for every R-module A.

20. Prove that $\mathrm{Tor}_0^{\mathbb{Z}/m\mathbb{Z}}(A, \mathbb{Z}/d\mathbb{Z}) \cong A/dA$, that $\mathrm{Tor}_n^{\mathbb{Z}/m\mathbb{Z}}(A, \mathbb{Z}/d\mathbb{Z}) \cong {}_dA/(m/d)A$ for n odd, $n \geq 1$, and that $\mathrm{Tor}_n^{\mathbb{Z}/m\mathbb{Z}}(A, \mathbb{Z}/d\mathbb{Z}) \cong {}_{(m/d)}A/dA$ for n even, $n \geq 2$. [Use the projective resolution in Example 2 following Proposition 3.]

21. Let $R = k[x, y]$ where k is a field, and let I be the ideal (x, y) in R.

(a) Let $\alpha : R \to R^2$ be the map $\alpha(r) = (yr, -xr)$ and let $\beta : R^2 \to R$ be the map $\beta((r_1, r_2)) = r_1x + r_2y$. Show that

$$0 \longrightarrow R \xrightarrow{\alpha} R^2 \xrightarrow{\beta} R \longrightarrow k \longrightarrow 0$$

where the map $R \to R/I = k$ is the canonical projection, gives a free resolution of k as an R-module.

(b) Use the resolution in (a) to show that $\mathrm{Tor}_2^R(k, k) \cong k$.

(c) Prove that $\mathrm{Tor}_1^R(k, I) \cong k$. [Use the long exact sequence corresponding to the short exact sequence $0 \to I \to R \to k \to 0$ and (b).]

(d) Conclude from (c) that the torsion free R-module I is not flat (compare to Exercise 26 in Section 10.5).

22. *(Flat Base Change for* Tor*)* Suppose R and S are commutative rings and $f : R \to S$ is a ring homomorphism making S into an R-module as in Example 6 following Corollary 12 in Section 10.4. Prove that if S is flat as an R-module, then $\mathrm{Tor}_n^R(A, B) \cong \mathrm{Tor}_n^S(S \otimes_R A, B)$ for all R-modules A and all S-modules B. [Show that since S is flat, tensoring an R-module projective resolution for A with S gives an S-module projective resolution of $S \otimes_R A$.]

23. *(Localization and* Tor*)* Let $D^{-1}R$ be the localization of the commutative ring R with respect to the multiplicative subset D of R. Prove that localization commutes with Tor, i.e., $D^{-1}\mathrm{Tor}_n^R(A,B) \cong \mathrm{Tor}_n^{D^{-1}R}(D^{-1}A, D^{-1}B)$ for all R-modules A and B and all $n \geq 0$. [Use the previous exercise and the fact that $D^{-1}R$ is flat over R, cf. Proposition 42(6) in Section 15.4.]

24. *(Flatness is local)* Suppose R is a commutative ring. Prove that an R-module M is flat if and only if every localization M_P is a flat R_P-module for every maximal (hence also for every prime) ideal in R. [Use the previous exercise together with the characterization of flatness in terms of Tor.]

25. If R is an integral domain with field of fractions F, prove that $\mathrm{Tor}_1^R(F/R, B) \cong t(B)$ for any R-module B, where $t(B)$ denotes the R-torsion submodule of B.

An R-module M is said to be *finitely presented* if there is an exact sequence

$$R^s \longrightarrow R^t \longrightarrow M \longrightarrow 0$$

of R-modules for some integers s and t. Equivalently, M is finitely generated by t elements and the kernel of the corresponding R-module homomorphism $R^t \to M$ can be generated by s elements.

26. (a) Prove that every finitely generated module over a Noetherian ring R is finitely presented. [Use Exercise 8 in Section 15.1.]
    (b) Prove that an R-module M is finitely presented and projective if and only if M is a direct summand of R^n for some integer $n \geq 1$.

27. Suppose that M is a finitely presented R-module and that $0 \to A \xrightarrow{\alpha} B \xrightarrow{\beta} M \to 0$ is an exact sequence of R-modules. This exercise proves that if B is a finitely generated R-module then A is also a finitely generated R-module.
    (a) Suppose $R^s \xrightarrow{\psi} R^t \xrightarrow{\varphi} M \to 0$ and $e_1, \dots, e_t$ is an R-module basis for R^t. Show that there exist $b_1, \dots, b_t \in B$ so that $\beta(b_i) = \varphi(e_i)$ for $i = 1, \dots, t$.
    (b) If f is the R-module homomorphism from R^t to B defined by $f(e_i) = b_i$ for $i = 1, \dots, t$, show that $f(\psi(R^s)) \subseteq \ker\beta$. [Use $\varphi \circ \psi = 0$.] Conclude that there is a commutative diagram

$$\begin{array}{ccccccccc} & & R^s & \xrightarrow{\psi} & R^t & \xrightarrow{\varphi} & M & \longrightarrow & 0 \\ & & {\scriptstyle g}\downarrow & & {\scriptstyle f}\downarrow & & \| & & \\ 0 & \longrightarrow & A & \xrightarrow{\alpha} & B & \xrightarrow{\beta} & M & \longrightarrow & 0 \end{array}$$

    of R-modules with exact rows.
    (c) Prove that $A/\text{image}\, g \cong B/\text{image}\, f$ and use this to prove that A is finitely generated. [For the isomorphism, use the Snake Lemma in Exercise 3. Then show that image g and $A/\text{image}\, g$ are both finitely generated and apply Exercise 7 of Section 10.3.]
    (d) If I is an ideal of R conclude that R/I is a finitely presented R-module if and only if I is a finitely generated ideal.

28. Suppose R is a local ring with unique maximal ideal $\mathfrak{m}$ and M is a finitely presented R-module. Suppose $m_1, \dots, m_s$ are elements in M whose images in $M/\mathfrak{m}M$ form a basis for $M/\mathfrak{m}M$ as a vector space over the field $R/\mathfrak{m}$.
    (a) Prove that $m_1, \dots, m_s$ generate M as an R-module. [Use Nakayama's Lemma.]
    (b) Conclude from (a) that there is an exact sequence $0 \to \ker\varphi \to R^s \xrightarrow{\varphi} M \to 0$ that maps a set of free generators of R^s to the elements $m_1, \dots, m_s$. Deduce that there is

an exact sequence

$$\operatorname{Tor}_1^R(M, R/\mathfrak{m}) \longrightarrow (\ker\varphi)/\mathfrak{m}(\ker\varphi) \longrightarrow 0.$$

[Use the Tor long exact sequence with respect to tensoring with $R/\mathfrak{m}$, using the fact that $N \otimes R/\mathfrak{m} \cong N/\mathfrak{m}N$ for any R-module N (Example 8 following Corollary 12 in Section 10.4] and the fact that $\varphi : (R/\mathfrak{m})^s \cong M/\mathfrak{m}M$ is an isomorphism by the choice of $m_1, \dots, m_s$.]

(c) Prove that if $\operatorname{Tor}_1^R(M, R/\mathfrak{m}) = 0$ then $m_1, \dots, m_s$ are a set of *free* R-module generators for M. [Use the previous exercise and Nakayama's Lemma to show that $\ker\varphi = 0$.]

29. Suppose R is a local ring with unique maximal ideal $\mathfrak{m}$. This exercise proves that a finitely generated R-module is flat if and only if it is free.

(a) Prove that $M = F/K$ is the quotient of a finitely generated free module F by a submodule K with $K \subseteq \mathfrak{m}F$. [Let F be a free module with $F/\mathfrak{m}F \cong M/\mathfrak{m}M$.]

(b) Suppose $x \in K$ and write $x = a_1e_1 + \cdots + a_ne_n$ where $e_1, \dots, e_n$ are an R-basis for F. Let $I = (a_1, \dots, a_n)$ be the ideal of R generated by $a_1, \dots, a_n$). Prove that if M is flat, then $I = \mathfrak{m}I$ and deduce that $K = 0$, so M is free. [Use Exercise 25(d) of Section 10.5 to see that $x \in IK \subseteq \mathfrak{m}IF$ and conclude that $I \subseteq \mathfrak{m}I$. Then apply Nakayama's Lemma to the finitely generated ideal I.]

30. Suppose R is a local ring with unique maximal ideal $\mathfrak{m}$, M is an R-module, and consider the following statements:

(i) M is a free R-module,
(ii) M is a projective R-module,
(iii) M is a flat R-module, and
(iv) $\operatorname{Tor}_1^R(M, R/\mathfrak{m}) = 0$.

(a) Prove that (i) implies (ii) implies (iii) implies (iv).

(b) Prove that (i) , (ii), and (iii) are equivalent if M is finitely generated. (Exercise 34 below shows (iii) need not imply (i) or (ii) if M is finitely generated but R is not local.) [Use the previous exercise.]

(c) Prove that (i), (ii), (iii), and (iv) are equivalent if M is finitely presented. (Exercise 35 below shows that (iv) need not imply (i), (ii) or (iii) if M is finitely generated but not finitely presented.) [Use Exercise 28.]

Remark: It is a theorem of Kaplansky (cf. *Projective Modules*, Annals of Mathematics, 68(1958), pp. 372-377) that (i) and (ii) are equivalent without the condition that M be finitely generated.

31. *(Localization and* Hom *for Finitely Presented Modules)* Suppose $D^{-1}R$ is the localization of the commutative ring R with respect to the multiplicative subset D of R, and let M be a finitely presented R-module.

(a) For any R-modules A and B prove there is a unique $D^{-1}R$-module homomorphism from $D^{-1}\operatorname{Hom}_R(A, B)$ to $\operatorname{Hom}_{D^{-1}R}(D^{-1}A, D^{-1}B)$ that maps $\varphi \in \operatorname{Hom}_R(A, B)$ to the homomorphism from $D^{-1}A$ to $D^{-1}B$ induced by φ.

(b) For any R-module N and any $m \geq 1$ show that $\operatorname{Hom}_R(R^m, N) \cong N^m$ as R-modules and deduce that $D^{-1}\operatorname{Hom}_R(R^m, N) \cong (D^{-1}N)^m$ as $D^{-1}R$-modules.

(c) Suppose $R^s \longrightarrow R^t \longrightarrow M \longrightarrow 0$ is exact. Prove there is a commutative diagram

$$\begin{tikzcd}
0 \arrow[r] & D^{-1}\operatorname{Hom}_R(M, N) \arrow[r] \arrow[d] & D^{-1}\operatorname{Hom}_R(R^t, N) \arrow[r] \arrow[d] & D^{-1}\operatorname{Hom}_R(R^s, N) \arrow[d] \\
0 \arrow[r] & \operatorname{Hom}_{D^{-1}R}(D^{-1}M, D^{-1}N) \arrow[r] & \operatorname{Hom}_{D^{-1}R}((D^{-1}R)^t, D^{-1}N) \arrow[r] & \operatorname{Hom}_{D^{-1}R}((D^{-1}R)^s, D^{-1}N)
\end{tikzcd}$$

of $D^{-1}R$-modules with exact rows. [For the first row first take R-module homomorphisms from the terms in the presentation for M into N using Theorem 33 of Section 10.5 (noting the first comment in the proof) and then tensor with the flat R-module $D^{-1}R$, cf. Propositions 41 and 42(6) in Section 15.4. For the second row first tensor the presentation with $D^{-1}R$ and then take $D^{-1}R$-module homomorphisms into $D^{-1}N$.]

(d) Use (b) to prove that localization commutes with taking homomorphisms when M is finitely presented, i.e., $D^{-1}\mathrm{Hom}_R(M, N) \cong \mathrm{Hom}_{D^{-1}R}(D^{-1}M, D^{-1}N)$ as $D^{-1}R$-modules. [Show the second two vertical maps in the diagram above are isomorphisms and deduce that the left vertical map is also an isomorphism.] (This result is not true in general if M is not finitely presented.)

32. *(Localization and* Ext *for Finitely Presented Modules)* Suppose $D^{-1}R$ is the localization of the commutative ring R with respect to the multiplicative subset D of R. Prove that if M is a finitely presented R-module then $D^{-1}\mathrm{Ext}^n_R(M, N) \cong \mathrm{Ext}^n_{D^{-1}R}(D^{-1}M, D^{-1}N)$ as $D^{-1}R$-modules for every R-module N and every $n \geq 0$. [Use a projective resolution of N and the previous exercise, noting that tensoring the resolution with $D^{-1}R$ gives a projective resolution for the $D^{-1}R$-module $D^{-1}N$.]

33. Suppose R is a commutative ring and M is a finitely presented R-module (for example a finitely generated module over a Noetherian ring, or a quotient, R/I, of R by a finitely generated ideal I, cf. Exercises 26 and 27). Prove that the following are equivalent:

(a) M is a projective R-module,

(b) M is a flat R-module,

(c) M is locally free, i.e., each localization M_P is a free R_P-module for every maximal (hence also for every prime) ideal P of R.

In particular show that finitely generated projective modules are the same as finitely presented flat modules. [Exercises 24 and 30 show that (b) is equivalent to (c). Use the Ext criterion for projectivity and Exercises 30 and 32 to see that (a) is equivalent to (c).]

34. **(a)** Prove that *every* R-module for the commutative ring R is flat if and only if every finitely generated ideal I of R is a direct summand of R, in which case every finitely generated ideal of R is principal and projective (such a ring is said to be *absolutely flat*). [Use Exercise 15, the previous exercise applied to the finitely presented R-module R/I, and the remarks following Proposition 16.]

(b) Prove that every Boolean ring is absolutely flat. [Use Exercise 24 in Section 7.4, noting that if $I = Rx$ then x is an idempotent so $R = Rx \oplus R(1-x)$.]

(c) Let R be the direct product and I the direct sum of countably many copies of $\mathbb{Z}/2\mathbb{Z}$. Prove that I is an ideal of the Boolean ring R that is not finitely generated and that the cyclic R-module $M = R/I$ is flat but not projective (so finitely generated flat modules need not be projective).

35. Let R be the local ring obtained by localizing the ring of C^∞ functions on the open interval $(-1, 1)$ at the maximal ideal of functions that are 0 at $x = 0$ (cf. Exercise 45 of Section 15.2), let $\mathfrak{m} = (x)$ be the unique maximal ideal of R and let P be the prime ideal $\bigcap_{n\geq 1}\mathfrak{m}^n$. Set $M = R/P$.

(a) Prove that $\mathrm{Tor}_1^R(M, R/\mathfrak{m}) = 0$. [Use Exercise 19 applied with $r = x$, noting that R/P is an integral domain.]

(b) Prove that M is not flat (hence not projective). [Let F be as in Exercise 45 of Section 15.2. Show that the sequence $0 \to R \to R \to R/(F) \to 0$ induced by multiplication by F is exact, but is not exact after tensoring with M.]

17.2 THE COHOMOLOGY OF GROUPS

In this section we consider the application of the general techniques of the previous section in an important special case.

Let G be a group.

Definition. An abelian group A on which G acts (on the left) as automorphisms is called a *G-module.*

Note that a G-module is the same as an abelian group A and a homomorphism $\varphi : G \to \mathrm{Aut}(A)$ of G into the group of automorphisms of A. Since an abelian group is the same as a module over $\mathbb{Z}$, it is also easy to see that a G-module A is the same as a module over the integral group ring,$\mathbb{Z}G$, of G with coefficients in $\mathbb{Z}$. When G is an infinite group the ring $\mathbb{Z}G$ consists of all the finite formal sums of elements of G with coefficients in $\mathbb{Z}$.

As usual we shall often use multiplicative notation and write ga in place of $g \cdot a$ for the action of the element $g \in G$ on the element $a \in A$.

Definition. If A is a G-module, let $A^G = \{a \in A \mid ga = a \text{ for all } g \in G\}$ be the elements of A fixed by all the elements of G.

Examples

(1) If $ga = a$ for all $a \in A$ and $g \in G$ then G is said to act *trivially* on A. In this case $A^G = A$. The abelian group $\mathbb{Z}$ will always be assumed to have trivial G-action for any group G unless otherwise stated.

(2) For any G-module A the fixed points A^G of A under the action of G is clearly a $\mathbb{Z}G$-submodule of A on which G acts trivially.

(3) If V is a vector space over the field F of dimension n and $G = GL_n(F)$ then V is naturally a G-module. In this case $V^G = \{0\}$ since any nonzero element in V can be taken to any other nonzero element in V by some linear transformation.

(4) A semidirect product $E = A \rtimes G$ as in Section 5.5 in the case where A is an abelian normal subgroup gives a G-module A where the action of G is given by the homomorphism $\varphi : G \to \mathrm{Aut}(A)$. The subgroup A^G consists of the elements of A lying in the center of E. More generally, if A is any abelian normal subgroup of a group E, then E acts on A by conjugation and this makes A into a E-module and also an E/A-module. In this case $A^E = A^{E/A}$ also consists of the elements of A lying in the center of E.

(5) If K/F is an extension of fields that is Galois with Galois group G then the additive group K is naturally a G-module, with $K^G = F$. Similarly, the multiplicative group $K^\times$ of nonzero elements in K is a G-module, with fixed points $(K^\times)^G = F^\times$.

The fixed point subgroups in this last example played a central role in Galois Theory in Chapter 14. In general, it is easy to see that a short exact sequence

$$0 \longrightarrow A \longrightarrow B \longrightarrow C \longrightarrow 0$$

of G-modules induces an exact sequence

$$0 \longrightarrow A^G \longrightarrow B^G \longrightarrow C^G \tag{17.15}$$

that in general cannot be extended to a short exact sequence (in general a coset in the quotient C that is fixed by G need not be represented by an *element* in B fixed by G). One way to see that (15) is exact is to observe that A^G can be related to a Hom group:

Lemma 19. Suppose A is a G-module and $\mathrm{Hom}_{\mathbb{Z}G}(\mathbb{Z}, A)$ is the group of all $\mathbb{Z}G$-module homomorphisms from $\mathbb{Z}$ (with trivial G-action) to A. Then $A^G \cong \mathrm{Hom}_{\mathbb{Z}G}(\mathbb{Z}, A)$.

Proof: Any G-module homomorphism α from $\mathbb{Z}$ to A is uniquely determined by its value on 1. Let α_a denote the G-module homomorphism with $\alpha(1) = a$. Since α_a is a G-module homomorphism, $a = \alpha_a(1) = \alpha_a(g \cdot 1) = g \cdot \alpha_a(1) = g \cdot a$ for all $g \in G$, so that a must lie in A^G. Likewise, for any $a \in A^G$ it is easy to check that the map $\alpha_a \mapsto a$ gives an isomorphism from $\mathrm{Hom}_{\mathbb{Z}G}(\mathbb{Z}, A)$ to A^G.

Combined with the results of the previous section, the lemma not only shows that the sequence (15) is exact, it shows that any projective resolution of $\mathbb{Z}$ considered as a $\mathbb{Z}G$-module will give a long exact sequence extending (15). One such projective resolution is the *standard resolution* or *bar resolution* of $\mathbb{Z}$:

$$\cdots \longrightarrow F_n \xrightarrow{d_n} F_{n-1} \rightarrow \cdots \xrightarrow{d_1} F_0 \xrightarrow{\mathrm{aug}} \mathbb{Z} \longrightarrow 0. \tag{17.16}$$

Here $F_n = \mathbb{Z}G \otimes_{\mathbb{Z}} \mathbb{Z}G \otimes_{\mathbb{Z}} \cdots \otimes_{\mathbb{Z}} \mathbb{Z}G$ (where there are $n+1$ factors) for $n \geq 0$, which is a G-module under the action defined on simple tensors by $g \cdot (g_0 \otimes g_1 \otimes \cdots \otimes g_n) = (gg_0) \otimes g_1 \otimes \cdots \otimes g_n$. It is not difficult to see that F_n is a free $\mathbb{Z}G$-module of rank $|G|^n$ with $\mathbb{Z}G$ basis given by the elements $1 \otimes g_1 \otimes g_2 \otimes \cdots \otimes g_n$, where $g_i \in G$. The map $\mathrm{aug} : F_0 \rightarrow \mathbb{Z}$ is the *augmentation map* $\mathrm{aug}(\sum_{g \in G} \alpha_g g) = \sum_{g \in G} \alpha_g$, and the map d_1 is given by $d_1(1 \otimes g) = g - 1$. The maps d_n for $n \geq 2$ are more complicated and their definition, together with a proof that (16) is a projective (in fact, free) resolution can be found in Exercises 1–3.

Applying ($\mathbb{Z}G$-module) homomorphisms from the terms in (16) to the G-module A (replacing the first term by 0) as in the previous section, we obtain the cochain complex

$$0 \longrightarrow \mathrm{Hom}_{\mathbb{Z}G}(F_0, A) \xrightarrow{d_1} \mathrm{Hom}_{\mathbb{Z}G}(F_1, A) \xrightarrow{d_2} \mathrm{Hom}_{\mathbb{Z}G}(F_2, A) \xrightarrow{d_3} \cdots, \tag{17.17}$$

the cohomology groups of which are, by definition, the groups $\mathrm{Ext}^n_{\mathbb{Z}G}(\mathbb{Z}, A)$. Then, as in Theorem 8, the short exact sequence $0 \longrightarrow A \longrightarrow B \longrightarrow C \longrightarrow 0$ of G-modules gives rise to a long exact sequence whose first terms are given by (15) and whose higher terms are the cohomology groups $\mathrm{Ext}^n_{\mathbb{Z}G}(\mathbb{Z}, A)$.

To make this more explicit, we can reinterpret the terms in this cochain complex without explicit reference to the standard resolution of $\mathbb{Z}$, as follows. The elements of $\mathrm{Hom}_{\mathbb{Z}G}(F_n, A)$ are uniquely determined by their values on the $\mathbb{Z}G$ basis elements of F_n, which may be identified with the n-tuples $(g_1, g_2, \ldots, g_n)$ of elements g_i of G. It follows for $n \geq 1$ that the group $\mathrm{Hom}_{\mathbb{Z}G}(F_n, A)$ may be identified with the set of functions from $G \times \cdots \times G$ (n copies) to A. For $n = 0$ we identify $\mathrm{Hom}_{\mathbb{Z}G}(\mathbb{Z}G, A)$ with A.

Definition. If G is a finite group and A is a G-module, define $C^0(G, A) = A$ and for $n \geq 1$ define $C^n(G, A)$ to be the collection of all maps from $G^n = G \times \cdots \times G$ (n copies) to A. The elements of $C^n(G, A)$ are called *n-cochains (of G with values in A)*.

Each $C^n(G, A)$ is an additive abelian group: for $C^0(G, A) = A$ given by the group structure on A; for $n \geq 1$ given by the usual pointwise addition of functions: $(f_1 + f_2)(g_1, g_2, \ldots, g_n) = f_1(g_1, g_2, \ldots, g_n) + f_2(g_1, g_2, \ldots, g_n)$. Under the identification of $\mathrm{Hom}_{\mathbb{Z}G}(F_n, A)$ with $C^n(G, A)$ the cochain maps d_n in (17) can be given very explicitly (cf. also Exercise 3 and the following comment):

Definition. For $n \geq 0$, define the n^{th} *coboundary* homomorphism from $C^n(G, A)$ to $C^{n+1}(G, A)$ by

$$\begin{aligned} d_n(f)(g_1, \ldots, g_{n+1}) = {} & g_1 \cdot f(g_2, \ldots, g_{n+1}) \\ & + \sum_{i=1}^{n} (-1)^i f(g_1, \ldots, g_{i-1}, g_i g_{i+1}, g_{i+2}, \ldots, g_{n+1}) \\ & + (-1)^{n+1} f(g_1, \ldots, g_n) \end{aligned} \tag{17.18}$$

where the product $g_i g_{i+1}$ occupying the i^{th} position of f is taken in the group G.

It is immediate from the definition that the maps d_n are group homomorphisms. It follows from the fact that (17) is a projective resolution that $d_n \circ d_{n-1} = 0$ for $n \geq 1$ (a self contained direct proof just from the definition of d_n above can also be given, but is tedious).

Definition.

(1) Let $Z^n(G, A) = \ker d_n$ for $n \geq 0$. The elements of $Z^n(G, A)$ are called *n-cocycles*.

(2) Let $B^n(G, A) = \mathrm{image}\, d_{n-1}$ for $n \geq 1$ and let $B^0(G, A) = 1$. The elements of $B^n(G, A)$ are called *n-coboundaries*.

Since $d_n \circ d_{n-1} = 0$ for $n \geq 1$ we have $\mathrm{image}\, d_{n-1} \subseteq \ker d_n$, so that $B^n(G, A)$ is always a subgroup of $Z^n(G, A)$.

Definition. For any G-module A the quotient group $Z^n(G, A)/B^n(G, A)$ is called the n^{th} *cohomology group of G with coefficients in A* and is denoted by $H^n(G, A)$, $n \geq 0$.

The definition of the cohomology group $H^n(G, A)$ in terms of cochains will be particularly useful in the following two sections when we examine the low dimensional groups $H^1(G, A)$ and $H^2(G, A)$ and their application in a variety of settings. It should be remembered, however, that $H^n(G, A) \cong \mathrm{Ext}^n(\mathbb{Z}, A)$ for all $n \geq 0$. In particular, these groups can be computed using *any* projective resolution of $\mathbb{Z}$.

Examples

(1) For $f = a \in C^0(G, A)$ we have $d_0(f)(g) = g \cdot a - a$ and so $\ker d_0$ is the set $\{a \in A \mid g \cdot a = a \text{ for all } g \in G\}$, i.e., $Z^0(G, A) = A^G$ and so

$$H^0(G, A) = A^G,$$

for any group G and G-module A.

(2) Suppose $G = 1$ is the trivial group. Then $G^n = \{(1, 1, \dots, 1)\}$ is also the trivial group, so $f \in C^n(G, A)$ is completely determined by $f(1, 1, \dots, 1) = a \in A$. Identifying $f = a$ we obtain $C^n(G, A) = A$ for all $n \geq 0$. Then, if $f = a \in A$,

$$d_n(f)(1, 1, \dots, 1) = a + \sum_{i=1}^{n} (-1)^i a + (-1)^{n+1} a = \begin{cases} 0 & \text{if } n \text{ is even} \\ 1 & \text{if } n \text{ is odd} \end{cases},$$

so $d_n = 0$ if n is even and $d_n = 1$ is the identity if n is odd. Hence

$$H^0(1, A) = A^G = A$$
$$H^n(1, A) = 0 \text{ for all } n \geq 1.$$

Example: (Cohomology of a Finite Cyclic Group)

Suppose G is cyclic of order m with generator σ. Let $N = 1 + \sigma + \sigma^2 + \cdots + \sigma^{m-1} \in \mathbb{Z}G$. Then $N(\sigma - 1) = (\sigma - 1)N = \sigma^m - 1 = 0$, and so we have a particularly simple free resolution

$$\dots \xrightarrow{\sigma-1} \mathbb{Z}G \xrightarrow{N} \mathbb{Z}G \xrightarrow{\sigma-1} \dots \xrightarrow{N} \mathbb{Z}G \xrightarrow{\sigma-1} \mathbb{Z}G \xrightarrow{\text{aug}} \mathbb{Z} \longrightarrow 0$$

where aug denotes the augmentation map (cf. Exercise 8). Taking $\mathbb{Z}G$-module homomorphisms from the terms of this resolution to A (replacing the first term by 0) and using the identification $\mathrm{Hom}_{\mathbb{Z}G}(\mathbb{Z}G, A) = A$ gives the chain complex

$$0 \longrightarrow A \xrightarrow{\sigma-1} A \xrightarrow{N} A \xrightarrow{\sigma-1} A \xrightarrow{N} \dots$$

whose cohomology computes the groups $H^n(G, A)$:

$$H^0(G, A) = A^G, \text{ and } H^n(G, A) = \begin{cases} A^G/NA & \text{if } n \text{ is even, } n \geq 2 \\ {}_NA/(\sigma - 1)A & \text{if } n \text{ is odd, } n \geq 1 \end{cases}$$

where ${}_NA = \{a \in A \mid Na = 0\}$ is the subgroup of A annihilated by N, since the kernel of multiplication by $\sigma - 1$ is A^G.

If in particular $G = \langle \sigma \rangle$ acts trivially on A, then $N \cdot a = ma$, so that in this case $H^0(G, A) = A$, with $H^n(G, A) = A/mA$ for even $n \geq 2$, and $H^n(G, A) = {}_mA$, the elements of A of order dividing m, for odd $n \geq 1$. Specializing even further to $m = 1$ gives Example 2 previously.

Proposition 20. Suppose $mA = 0$ for some integer $m \geq 1$ (i.e., the G-module A has exponent dividing m as an abelian group). Then

$$mZ^n(G, A) = mB^n(G, A) = mH^n(G, A) = 0 \quad \text{for all } n \geq 0.$$

In particular, if A has exponent p for some prime p then the abelian groups $Z^n(G, A)$, $B^n(G, A)$ and $H^n(G, A)$ have exponent dividing p and so these groups are all vector spaces over the finite field $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$.

Proof: If $f \in C^n(G, A)$ is an n-cochain then $f \in A$ (if $n = 0$), in which case $mf = 0$, or f is a function from G^n to A (if $n \geq 1$), in which case mf is a function from G^n to $mA = 0$, so again $mf = 0$. Hence $mZ^n(G, A) = mB^n(G, A) = 0$ since these are subgroups of $C^n(G, A)$. Then $mH^n(G, A) = 0$ since $mZ^n(G, A) = 0$, and the remaining statements in the proposition are immediate.

By Example 1, the long exact sequence in Theorem 10 written in terms of the cohomology groups $H^n(G, A)$ becomes

Theorem 21. *(Long Exact Sequence in Group Cohomology)* Suppose

$$0 \longrightarrow A \longrightarrow B \longrightarrow C \longrightarrow 0$$

is a short exact sequence of G-modules. Then there is a long exact sequence:

$$0 \longrightarrow A^G \longrightarrow B^G \longrightarrow C^G \xrightarrow{\delta_0} H^1(G, A) \longrightarrow H^1(G, B) \longrightarrow H^1(G, C) \xrightarrow{\delta_1} \cdots$$
$$\cdots \xrightarrow{\delta_{n-1}} H^n(G, A) \longrightarrow H^n(G, B) \longrightarrow H^n(G, C) \xrightarrow{\delta_n} H^{n+1}(G, A) \longrightarrow \cdots$$

of abelian groups.

Among many other uses of the long exact sequence in Theorem 21 is a technique called *dimension shifting* which makes it possible to analyze the cohomology group $H^{n+1}(G, A)$ of dimension $n + 1$ for A by instead considering a cohomology group of dimension n for a different G-module. The technique is based on finding a G-module almost all of whose cohomology groups are zero. Such modules are given a name:

Definition. A G-module M is called *cohomologically trivial for G* if $H^n(G, M) = 0$ for all $n \geq 1$.

Corollary 22. *(Dimension Shifting)* Suppose $0 \to A \to M \to C \to 0$ is a short exact sequence of G-modules and that M is cohomologically trivial for G. Then there is an exact sequence

$$0 \longrightarrow A^G \longrightarrow M^G \longrightarrow C^G \longrightarrow H^1(G, A) \longrightarrow 0$$

and

$$H^{n+1}(G, A) \cong H^n(G, C) \text{ for all } n \geq 1.$$

Proof: Since M is cohomologically trivial for G, the portion

$$H^n(G, M) \longrightarrow H^n(G, C) \longrightarrow H^{n+1}(G, A) \longrightarrow H^{n+1}(G, M)$$

of the long exact sequence in Theorem 21 reduces to

$$0 \longrightarrow H^n(G, C) \longrightarrow H^{n+1}(G, A) \longrightarrow 0$$

which shows that $H^n(G, C) \cong H^{n+1}(G, A)$ for $n \geq 1$. Similarly, the first portion of the long exact sequence in Theorem 21 gives the first statement in the corollary.

We now indicate a natural construction that produces a G-module given a module over a subgroup H of G. When $H = 1$ is the trivial group this construction produces a cohomologically trivial module M and an exact sequence as in Corollary 22 for any G-module A.

Definition. If H is a subgroup of G and A is an H-module, define the *induced G-module* $M_H^G(A)$ to be $\mathrm{Hom}_{\mathbb{Z}H}(\mathbb{Z}G, A)$. In other words, $M_H^G(A)$ is the set of maps f from G to A satisfying $f(hx) = hf(x)$ for every $x \in G$ and $h \in H$.

The action of an element $g \in G$ on $f \in M_H^G(A)$ is given by $(g \cdot f)(x) = f(xg)$ for $x \in G$ (cf. Exercise 10 in Section 10.5).

Recall that if H is a subgroup of G and A is an H-module, then the module $\mathbb{Z}G \otimes_{\mathbb{Z}H} A$ obtained by extension of scalars from $\mathbb{Z}H$ to $\mathbb{Z}G$ is a G-module. For a finite group G, or more generally if H has finite index in G, we have $M_H^G(A) \cong \mathbb{Z}G \otimes_{\mathbb{Z}H} A$ (cf. Exercise 10). When G is infinite this need no longer be the case (cf. Exercise 11). The module $\mathbb{Z}G \otimes_{\mathbb{Z}H} A$ is sometimes called the *induced G-module* and the module $M_H^G(A)$ is sometimes referred to as the *coinduced G-module*. For finite groups, associativity of the tensor product shows that $M_H^G(M_K^H(A)) = M_K^G(A)$ for subgroups $K \leq H \leq G$, and the same result holds in general (this follows from the definition using Exercise 7).

Examples

(1) If H is a subgroup of G and $0 \to A \to B \to C \to 0$ is a short exact sequence of H-modules then $0 \to M_H^G(A) \to M_H^G(B) \to M_H^G(C) \to 0$ is a short exact sequence of G-modules, since $M_H^G(A) \cong \mathbb{Z}G \otimes_{\mathbb{Z}H} A$ and $\mathbb{Z}G$ is free, hence flat, over $\mathbb{Z}H$.

(2) When G is finite and A is the trivial H-module $\mathbb{Z}$, the module $M_H^G(\mathbb{Z})$ is a free $\mathbb{Z}$-module of rank $m = |G : H|$. There is a basis $b_1, \dots, b_m$ such that G permutes these basis elements in the same way it permutes the left cosets of H in G by left multiplication, i.e., if we let $b_i \leftrightarrow g_i H$ then $gb_i = b_j$ if and only if $gg_i H = g_j H$. The module $M_H^G(\mathbb{Z})$ is the *permutation module* over $\mathbb{Z}$ for G with stabilizer H. A special case of interest is when $G = S_m$ and $H = S_{m-1}$ where S_m permutes $\{1, 2, \dots, m\}$ as usual. Permutation modules and induced modules over fields are studied in Part VI.

(3) Any abelian group A is an H-module when $H = 1$ is the trivial group. The corresponding induced G-module $M_1^G(A)$ is just the collection of all maps f from G into A. For $g \in G$ the map $g \cdot f \in M_1^G(A)$ satisfies $(g \cdot f)(x) = f(xg)$ for $x \in G$.

(4) Suppose A is a G-module. Then there is a natural map

$$\varphi : A \longrightarrow M_1^G(A)$$

from A into the induced G-module $M_1^G(A)$ in the previous example defined by mapping $a \in A$ to the function f_a with $f_a(x) = xa$ for all $x \in G$. It is clear that φ is a group homomorphism, and $f_{ga}(x) = x(ga) = (xg)a = f_a(xg) = (g \cdot f_a)(x)$ shows that φ is a G-module homomorphism as well. Since $f_a(1) = a$, it follows that f_a is the zero function on G if and only if $a = 0$ in A, so that φ is an injection. Hence we may identify A as a G-submodule of the induced module $M_1^G(A)$.

(5) More generally, if A is a G-module and H is any subgroup of G then the function $f_a(x)$ in the previous example is an element in the subgroup $M_H^G(A)$ since we have $f_a(hx) = (hx)(a) = h(xa) = hf_a(x)$ for all $h \in H$. The associated map from A to $M_H^G(A)$ is an injective G-module homomorphism.

(6) The fixed points $(M_H^G(A))^G$ are maps f from G to A with $gf = f$ for all $g \in G$, i.e., with $(gf)(x) = f(x)$ for all $g, x \in G$. By definition of the G-action on $M_H^G(A)$, this is the equation $f(xg) = f(x)$ for all $g, x \in G$. Taking $x = 1$ shows that f is constant on all of G: $f(g) = f(1) = a \in A$. The constant function $f = a$ is an element of $M_H^G(A)$ if and only if $a = f(hx) = hf(x) = ha$ for all $h \in H$, so $(M_H^G(A))^G \cong A^H$.

An element $f_a(x)$ in the previous example is contained in the subgroup $(M_H^G(A))^G$ if and only if xa is constant for $x \in G$, i.e., if and only if $a \in A^G$.

One of the important properties of the G-module $M_H^G(A)$ induced from the H-module A is that its cohomology with respect to G is the same as the cohomology of A with respect to H:

Proposition 23. *(Shapiro's Lemma)* For any subgroup H of G and any H-module A we have $H^n(G, M_H^G(A)) \cong H^n(H, A)$ for $n \geq 0$.

Proof: Let $\cdots \to P_n \to \cdots \to P_0 \to \mathbb{Z} \to 0$ be a resolution of $\mathbb{Z}$ by projective G-modules (for example, the standard resolution). The cohomology groups $H^n(G, M_H^G(A))$ are computed by taking homomorphisms from this resolution into $M_H^G(A) = \mathrm{Hom}_{\mathbb{Z}H}(\mathbb{Z}G, A)$. Since $\mathbb{Z}G$ is a free $\mathbb{Z}H$-module it follows that this G-module resolution is also a resolution of $\mathbb{Z}$ by projective H-modules, hence by taking homomorphisms into A the same resolution may be used to compute the cohomology groups $H^n(H, A)$. To see that these two collections of cohomology groups are isomorphic, we use the natural isomorphism of abelian groups

$$\Phi : \mathrm{Hom}_{\mathbb{Z}G}(P_n, \mathrm{Hom}_{\mathbb{Z}H}(\mathbb{Z}G, A)) \cong \mathrm{Hom}_{\mathbb{Z}H}(P_n, A)$$

given by $\Phi(f)(p) = f(p)(1)$, for all $f \in \mathrm{Hom}_{\mathbb{Z}G}(P_n, \mathrm{Hom}_{\mathbb{Z}H}(\mathbb{Z}G, A))$ and $p \in P_n$. The inverse isomorphism is defined by taking $\Psi(f')(p)$ to be the map from $\mathbb{Z}G$ to A that takes $g \in G$ to the element $f'(gp)$ in A for all $f' \in \mathrm{Hom}_{\mathbb{Z}H}(P_n, A)$ and $p \in P_n$, i.e., $(\Psi(f')(p))(g) = f'(gp)$. Note this is well defined because P_n is a G-module. (These maps are a special case of an Adjoint Associativity Theorem, cf. Exercise 7.) Since these isomorphisms commute with the cochain maps, they induce isomorphisms on the corresponding cohomology groups, i.e., $H^n(G, M_H^G(A)) \cong H^n(H, A)$, as required.

Corollary 24. For any G-module A the module $M_1^G(A)$ is cohomologically trivial for G, i.e., $H^n(G, M_1^G(A)) = 0$ for all $n \geq 1$.

Proof: This follows immediately from the proposition applied with $H = 1$ together with the computation of the cohomology of the trivial group in Example 2 preceding Proposition 20.

By the corollary, the fourth example above gives us a short exact sequence of G-modules

$$0 \longrightarrow A \xrightarrow{\varphi} M \longrightarrow C \longrightarrow 0$$

where $M = M_1^G(A)$ is cohomologically trivial for G and where C is the quotient of $M_1^G(A)$ by the image of A. The dimension shifting result in Corollary 22 then becomes:

Corollary 25. For any G-module A we have $H^{n+1}(G, A) \cong H^n(G, M_1^G(A)/A)$ for all $n \geq 1$.

We next consider several important maps relating various cohomology groups. Some applications of the use of these homomorphisms appear in the following two sections.

In general, suppose we have two groups G and G' and that A is a G-module and A' is a G'-module. If $\varphi : G' \to G$ is a group homomorphism then A becomes a G'-module by defining $g' \cdot a = \varphi(g')a$ for $g' \in G'$ and $a \in A$. If now $\psi : A \to A'$ is a homomorphism of abelian groups then we consider whether ψ is a G'-module homomorphism:

Definition. Suppose A is a G-module and A' is a G'-module. The group homomorphisms $\varphi : G' \to G$ and $\psi : A \to A'$ are said to be *compatible* if ψ is a G'-module homomorphism when A is made into a G'-module by means of φ, i.e., if $\psi(\varphi(g')a) = g'\psi(a)$ for all $g' \in G'$ and $a \in A$.

The point of compatible homomorphisms is that they induce group homomorphisms on associated cohomology groups, as follows.

If $\varphi : G' \to G$ and $\psi : A \to A'$ are homomorphisms, then φ induces a homomorphism $\varphi^n : (G')^n \to G^n$, and so a homomorphism from $C^n(G, A)$ to $C^n(G', A)$ that maps f to $f \circ \varphi^n$. The map ψ induces a homomorphism from $C^n(G', A)$ to $C^n(G', A')$ that maps f to $\psi \circ f$. Taken together we obtain an induced homomorphism

$$\begin{aligned} \lambda_n : C^n(G, A) &\longrightarrow C^n(G', A') \\ f &\longmapsto \psi \circ f \circ \varphi^n. \end{aligned}$$

If in addition φ and ψ are *compatible* homomorphisms, then it is easy to check that the induced maps λ_n commute with the coboundary operator:

$$\lambda_{n+1} \circ d_n = d_n \circ \lambda_n$$

for all $n \geq 0$. It follows that λ_n maps cocycles to cocycles and coboundaries to coboundaries, hence induces a group homomorphism on cohomology:

$$\lambda_n : H^n(G, A) \longrightarrow H^n(G', A')$$

for $n \geq 0$.

We consider several instances of such maps:

Examples

(1) Suppose $G = G'$ and φ is the identity map. Then to say that the group homomorphism $\psi : A \to A'$ is compatible with φ is simply the statement that ψ is a G-module homomorphism. Hence any G-module homomorphism from A to A' induces a group homomorphism

$$H^n(G, A) \longrightarrow H^n(G, A') \quad \text{for } n \geq 0.$$

In particular, if $0 \to A \to B \to C \to 0$ is a short exact sequence of G-modules we obtain induced homomorphisms from $H^n(G, A)$ to $H^n(G, B)$ and from $H^n(G, B)$ to $H^n(G, C)$ for $n \geq 0$. These are simply the homomorphisms in the long exact sequence of Theorem 21.

(2) *(The Restriction Homomorphism)* If A is a G-module, then A is also an H-module for any subgroup H of G. The inclusion map $\varphi : H \to G$ of H into G and the identity map $\psi : A \to A$ are compatible homomorphisms. The corresponding induced group homomorphism on cohomology is called the *restriction homomorphism*:

$$\text{Res} : H^n(G, A) \longrightarrow H^n(H, A), \quad n \geq 0.$$

The terminology comes from the fact that the map on cochains from $C^n(G, A)$ to $C^n(H, A)$ is simply restricting a map f from G^n to A to the subgroup H^n of G^n.

(3) *(The Inflation Homomorphism)* Suppose H is a normal subgroup of G and A is a G-module. The elements A^H of A that are fixed by H are naturally a module for the quotient group G/H under the action defined by $(gH)\cdot a = g\cdot a$. It is then immediate that the projection $\varphi : G \to G/H$ and the inclusion $\psi : A^H \to A$ are compatible homomorphisms. The corresponding induced group homomorphism on cohomology is called the *inflation homomorphism*:

$$\text{Inf} : H^n(G/H, A^H) \longrightarrow H^n(G, A), \quad n \geq 0.$$

(4) *(The Corestriction Homomorphism)* Suppose that H is a subgroup of G of index m and that A is a G-module. Let $g_1, \dots, g_m$ be representatives for the left cosets of H in G. Define a map

$$\psi : M_H^G(A) \longrightarrow A \qquad \text{by} \qquad f \longmapsto \sum_{i=1}^{m} g_i \cdot f(g_i^{-1}).$$

Note that if we change any coset representative g_i by $g_i h$, then $(g_i h)f((g_i h)^{-1}) = g_i h f(h^{-1} g_i^{-1}) = g_i h h^{-1} f(g_i^{-1}) = g_i f(g_i^{-1})$ so the map ψ is independent of the choice of coset representatives. It is easy to see that ψ is a G-module homomorphism (and even that it is surjective), so we obtain a group homomorphism from $H^n(G, M_H^G(A))$ to $H^n(G, A)$, for all $n \geq 0$. Since A is also an H-module, by Shapiro's Lemma we have an isomorphism $H^n(G, M_H^G(A)) \cong H^n(H, A)$. The composition of these two homomorphisms is called the *corestriction homomorphism:*

$$\text{Cor} : H^n(H, A) \longrightarrow H^n(G, A), \quad n \geq 0.$$

This homomorphism can be computed explicitly by composing the isomorphism Ψ in the proof of Shapiro's Lemma for any resolution of $\mathbb{Z}$ by projective G-modules P_n (note these are G-modules and not simply H-modules) with the map ψ, as follows. For a cocycle $f \in \text{Hom}_{\mathbb{Z}H}(P_n, A)$ representing a cohomology class $c \in H^n(H, A)$, a cocycle $\text{Cor}(f) \in \text{Hom}_{\mathbb{Z}G}(P_n, A)$ representing $\text{Cor}(c) \in H^n(G, A)$ is given by

$$\text{Cor}(f)(p) = \sum_{i=1}^{m} g_i \cdot \Psi(f)(p)(g_i^{-1}) = \sum_{i=1}^{m} g_i f(g_i^{-1} p),$$

for $p \in P_n$. When $n = 0$ this is particularly simple since we can take $P_0 = \mathbb{Z}G$. In this case $f \in \text{Hom}_{\mathbb{Z}H}(\mathbb{Z}G, A) = M_H^G(A)$ is a cocycle if $f = a$ is constant for some $a \in A^H$ and then $\text{Cor}(f)$ is the constant function with value $\sum_{i=1}^{m} g_i \cdot a \in A^G$:

$$\begin{aligned} \text{Cor} : H^0(H, A) = A^H &\longrightarrow A^G = H^0(G, A) \\ a &\longmapsto \sum_{i=1}^{m} g_i \cdot a. \end{aligned}$$

The next result establishes a fundamental relation between the restriction and corestriction homomorphisms.

Proposition 26. Suppose H is a subgroup of G of index m. Then $\mathrm{Cor} \circ \mathrm{Res} = m$, i.e., if c is a cohomology class in $H^n(G, A)$ for some G-module A, then

$$\mathrm{Cor}(\mathrm{Res}(c)) = mc \in H^n(G, A) \quad \text{for all } n \geq 0.$$

Proof: This follows from the explicit formula for corestriction in Example 4 above, as follows. If $f \in \mathrm{Hom}_{\mathbb{Z}H}(P_n, A)$ were in $\mathrm{Hom}_{\mathbb{Z}G}(P_n, A)$, i.e., if f were also a G-module homomorphism, then $g_i f(g_i^{-1} p) = g_i g_i^{-1} f(p) = f(p)$, for $1 \leq i \leq m$. Since restriction is the induced map on cohomology of the natural inclusion of $\mathrm{Hom}_{\mathbb{Z}G}(P_n, A)$ into $\mathrm{Hom}_{\mathbb{Z}H}(P_n, A)$, for such an f we obtain

$$\begin{aligned} \mathrm{Hom}_{\mathbb{Z}G}(P_n, A) \xrightarrow{\mathrm{Res}} \mathrm{Hom}_{\mathbb{Z}H}(P_n, A) \xrightarrow{\mathrm{Cor}} \mathrm{Hom}_{\mathbb{Z}G}(P_n, A) \\ f \longmapsto f \longmapsto mf. \end{aligned}$$

It follows that $\mathrm{Res} \circ \mathrm{Cor}$ is multiplication by m on the cohomology groups as well.

Corollary 27. Suppose the finite group G has order m. Then $mH^n(G, A) = 0$ for all $n \geq 1$ and any G-module A.

Proof: Let $H = 1$, so that $[G : H] = m$, in Proposition 26. Then for any class $c \in H^n(G, A)$ we have $mc = \mathrm{Cor}(\mathrm{Res}(c))$. Since $\mathrm{Res}(c) \in H^n(H, A) = H^n(1, A)$, we have $\mathrm{Res}(c) = 0$ for all $n \geq 1$ by the second example preceding Proposition 20. Hence $mc = 0$ for all $n \geq 1$, which is the corollary.

Corollary 28. If G is a finite group then $H^n(G, A)$ is a torsion abelian group for all $n \geq 1$ and all G-modules A.

Proof: This is immediate from the previous corollary.

Corollary 29. Suppose G is a finite group whose order is relatively prime to the exponent of the G-module A. Then $H^n(G, A) = 0$ for all $n \geq 1$. In particular, if A is a finite abelian group with $(|G|, |A|) = 1$ then $H^n(G, A) = 0$ for all $n \geq 1$.

Proof: This follows since the abelian group $H^n(G, A)$ is annihilated by $|G|$ by the previous corollary and is annihilated by the exponent of A by Proposition 20.

Note that the statements in the preceding corollaries are not in general true for $n = 0$, since then $H^0(G, A) = A^G$, which need not even be torsion.

We mention without proof the following result. Suppose that H is a normal subgroup of G and A is a G-module. The cohomology groups $H^n(H, A)$ can be given the structure of G/H-modules (cf. Exercise 17). It can be shown that there is an exact sequence

$$0 \to H^1(G/H, A^H) \overset{\mathrm{Inf}}{\to} H^1(G, A) \overset{\mathrm{Res}}{\to} H^1(H, A)^{G/H} \overset{\mathrm{Tra}}{\to} H^2(G/H, A^H) \overset{\mathrm{Inf}}{\to} H^2(G, A)$$

where $H^1(H, A)^{G/H}$ denotes the fixed points of $H^1(H, A)$ under the action of G/H and Tra is the so-called *transgression homomorphism.* This exact sequence relates the cohomology groups for G to the cohomology groups for the normal subgroup H and for the quotient group G/H. Put another way, the cohomology for G is related to the cohomology for the factors in the filtration $1 \leq H \leq G$ for G. More generally, one could try to relate the cohomology for G to the cohomology for the factors in a longer filtration for G. This is the theory of *spectral sequences* and is an important tool in homological algebra.

Galois Cohomology and Profinite Groups

One important application of group cohomology occurs when the group G is the Galois group of a field extension K/F. In this case there are many groups of interest on which G acts, for example the additive group of K, the multiplicative group $K^\times$, etc. The Galois group $G = \text{Gal}(K/F)$ is the inverse limit $\varprojlim \text{Gal}(L/F)$ of the Galois groups of the finite extensions L of F contained in K and is a compact topological group with respect to its Krull topology (i.e., the group operations on G are continuous with respect to the topology defined by the subgroups $\text{Gal}(K/L)$ of G of finite index), cf. Section 14.9. In this situation it is useful (and often essential) to take advantage of the additional topological structure of G. For example the subfields of K containing F correspond bijectively with the *closed* subgroups of $G = \text{Gal}(K/F)$, and the example of the composite of the quadratic extensions of $\mathbb{Q}$ discussed in Section 14.9 shows that in general there are many subgroups of G that are not closed. Fortunately, the modifications necessary to define the cohomology groups in this context are relatively minor and apply to arbitrary inverse limits of finite groups (the *profinite* groups). If G is a profinite group then $G = \varprojlim G/N$ where the inverse limit is taken over the open normal subgroups N of G (cf. Exercise 23).

Definition. If G is a profinite group then a *discrete G-module* A is a G-module A with the discrete topology such that the action of G on A is continuous, i.e., the map $G \times A \to A$ mapping (g, a) to $g \cdot a$ is continuous.

Since A is given the discrete topology, every subset of A is open, and in particular every element $a \in A$ is open. The continuity of the action of G on A is then equivalent to the statement that the stabilizer G_a of a in G is an open subgroup of G, hence is of finite index since G is compact (cf. Exercise 22). This in turn is equivalent to the statement that $A = \cup A^H$ where the union is over the open subgroups H of G.

Some care must be taken in defining the cohomology groups $H^n(G, A)$ of a profinite group G acting on a discrete G-module A since there are not enough projectives in this category. For example, when G is infinite, the free G-module $\mathbb{Z}G$ is not a discrete G-module (G does not act continuously, cf. Exercise 25). Nevertheless, the explicit description of $H^n(G, A)$ given in this section (occasionally referred to as the *discrete* cohomology groups) can be easily modified — it is only necessary to require the cochains $C^n(G, A)$ to be *continuous* maps from G^n to A. The definition of the coboundary maps d_n in equation (18) is precisely the same, as is the definition of the groups of cocycles, coboundaries, and the corresponding cohomology groups. It is customary not to introduce a separate notation for these cohomology groups, but to specify which cohomology is meant in the terminology.

Definition. If G is a profinite group and A is a discrete G-module, the cohomology groups $H^n(G, A)$ computed using continuous cochains are called the *profinite* or *continuous* cohomology groups. When $G \doteq \mathrm{Gal}(K/F)$ is the Galois group of a field extension K/F then the *Galois cohomology groups* $H^n(G, A)$ will always mean the cohomology groups computed using continuous cochains.

When G is a finite group, every G-module is a discrete G-module so the discrete and continuous cohomology groups of G are the same. When G is infinite, this need not be the case as shown by the example mentioned previously of the free G-module $\mathbb{Z}G$ when G is an infinite profinite group. All the major results in this section remain valid for the continuous cohomology' groups when "G-module" is replaced by "discrete G-module" and "subgroup" is replaced by "closed subgroup." For example, the Long Exact Sequence in Group Cohomology remains true as stated, the restriction homomorphism requires the subgroup H of G to be a closed subgroup (so that the restriction of a continuous map on G^n to H^n remains continuous), Proposition 26 requires H to be closed, etc.

We can write $G = \varprojlim(G/N)$ and $A = \cup A^N$ where N runs over the open normal subgroups of G (necessarily of finite index in G since G is compact). Then A^N is a discrete G/N-module and it is not difficult to show that

$$H^n(G, A) = \varinjlim_N H^n(G/N, A^N) \tag{17.19}$$

where the cohomology groups are continuous cohomology and the direct limit is taken over the collection of all open normal subgroups N of G (cf. Exercise 24). Since G/N is a finite group, the continuous cohomology groups $H^n(G/N, A^N)$ in this direct limit are just the (discrete) cohomology groups considered earlier in this section. The computation of the continuous cohomology for a profinite group G can therefore always be reduced to the consideration of finite group cohomology where there is no distinction between the continuous and discrete theories.

EXERCISES

1. Let $F_n = \mathbb{Z}G \otimes_{\mathbb{Z}} \mathbb{Z}G \otimes_{\mathbb{Z}} \cdots \otimes_{\mathbb{Z}} \mathbb{Z}G$ ($n+1$ factors) for $n \geq 0$ with G-action defined on simple tensors by $g \cdot (g_0 \otimes g_1 \otimes \cdots \otimes g_n) = (gg_0) \otimes g_1 \otimes \cdots \otimes g_n$.

(a) Prove that F_n is a free $\mathbb{Z}G$-module of rank $|G|^n$ with $\mathbb{Z}G$ basis $1 \otimes g_1 \otimes g_2 \otimes \cdots \otimes g_n$ with $g_i \in G$.

Denote the basis element $1 \otimes g_1 \otimes g_2 \otimes \cdots \otimes g_n$ in (a) by $(g_1, g_2, \ldots, g_n)$ and define the G-module homomorphisms d_n for $n \geq 1$ on these basis elements by $d_1(g_1) = g_1 - 1$ and

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

for $n \geq 2$. Define the $\mathbb{Z}$-module *contracting homomorphisms*

$$\mathbb{Z} \xrightarrow{s_{-1}} F_0 \xrightarrow{s_0} F_1 \xrightarrow{s_1} F_2 \xrightarrow{s_2} \cdots$$

on a $\mathbb{Z}$ basis by $s_{-1}(1) = 1$ and $s_n(g_0 \otimes \cdots \otimes g_n) = 1 \otimes g_0 \otimes \ldots \otimes g_n$.

(b) Prove that

$$\epsilon s_{-1} = 1, \qquad d_1 s_0 + s_{-1}\epsilon = 1, \qquad d_{n+1}s_n + s_{n-1}d_n = 1, \text{ for all } n \geq 1$$

where the map aug : $F_0 \to \mathbb{Z}$ is the augmentation map $\text{aug}(\sum_{g\in G}\alpha_g g) = \sum_{g\in G}\alpha_g$.

(c) Prove that the maps s_n are a chain homotopy (cf. Exercise 4 in Section 1) between the identity (chain) map and the zero (chain) map from the chain

$$\cdots \longrightarrow F_n \xrightarrow{d_n} F_{n-1} \xrightarrow{d_{n-1}} \cdots \xrightarrow{d_1} F_0 \xrightarrow{\text{aug}} \mathbb{Z} \longrightarrow 0 \qquad (*)$$

of $\mathbb{Z}$-modules to itself.

(d) Deduce from (c) that all $\mathbb{Z}$-module homology groups of $(*)$ are zero, i.e., $(*)$ is an exact sequence of $\mathbb{Z}$-modules. Conclude that $(*)$ is a projective G-module resolution of $\mathbb{Z}$.

2. Let P_n denote the free $\mathbb{Z}$-module with basis $(g_0, g_1, g_2, \ldots, g_n)$ with $g_i \in G$ and define an action of G on P_n by $g\cdot(g_0, g_1, \ldots, g_n) = (gg_0, gg_1, \ldots, gg_n)$. For $n \geq 1$ define

$$d_n(g_0, g_1, g_2, \ldots, g_n) = \sum_{i=0}^{n}(-1)^i(g_0, \ldots, \hat{g}_i, \ldots, g_n),$$

where $(g_0, \ldots, \hat{g}_i, \ldots, g_n)$ denotes the term $(g_0, g_1, g_2, \ldots, g_n)$ with g_i deleted.

(a) Prove that P_n is a free $\mathbb{Z}G$-module with basis $(1, g_1, g_2, \ldots, g_n)$ where $g_i \in G$.

(b) Prove that $d_{n-1} \circ d_n = 0$ for $n \geq 1$. [Show that the term $(g_0, \ldots, \hat{g}_j, \ldots, \hat{g}_k, \ldots, g_n)$ missing the entries g_j and g_k occurs twice in $d_{n-1} \circ d_n(g_0, g_1, g_2, \ldots, g_n)$, with opposite signs.]

(c) Prove that $\varphi : P_n \to F_n$ defined by

$$\varphi((g_0, g_1, g_2, \ldots, g_n)) = g_0 \otimes (g_0^{-1}g_1) \otimes (g_1^{-1}g_2) \ldots \otimes (g_{n-1}^{-1}g_n)$$

is a G-module isomorphism with inverse $\psi : P_n \to F_n$ given by

$$\psi(g_0 \otimes g_1 \otimes \ldots \otimes g_n) = (g_0, g_0g_1, g_0g_1g_2, \ldots, g_0g_1g_2\cdots g_n).$$

(d) Prove that if $\epsilon(g_0) = 1$ for all $g_0 \in G$ then

$$\cdots \longrightarrow P_n \xrightarrow{d_n} P_{n-1} \xrightarrow{d_{n-1}} \cdots \xrightarrow{d_1} P_0 \xrightarrow{\epsilon} \mathbb{Z} \longrightarrow 0 \qquad (**)$$

is a free G-module resolution of $\mathbb{Z}$. [Show that the isomorphisms in (c) take the G-module resolutions $(**)$ and $(*)$ of the previous exercise into each other.]

3. Let F_n and P_n be as in the previous two exercises and let A be a G-module.

(a) Prove that $\text{Hom}_{\mathbb{Z}G}(F_n, A)$ can be identified with the collection $C^n(G, A)$ of maps from $G \times G \times \cdots \times G$ (n copies) to A and that under this identification the associated coboundary maps from $C^n(G, A)$ to $C^{n+1}(G, A)$ are given by equation (18).

(b) Prove that $\text{Hom}_{\mathbb{Z}G}(P_n, A)$ can be identified with the collection of maps f from $n+1$ copies $G \times G \times \cdots \times G$ to A that satisfy $f(gg_0, gg_1, \ldots, gg_n) = gf(g_0, g_1, \ldots, g_n)$.

The group $C^n(G, A)$ is sometimes called the group of *inhomogeneous n-cochains of G in A*, and the group in (b) of the previous exercise is called the group of *homogeneous n-cochains of G in A*. The inhomogeneous cochains are easier to describe since there is no restriction on the maps from G^n to A, but the coboundary map d_n on homogeneous cochains is less complicated (and more naturally suggested in topological contexts) than the coboundary map on inhomogeneous cochains. The results of the previous exercises show that the cohomology groups $H^n(G, A)$ defined using either homogeneous or inhomogeneous cochains are the same and indicate the origin of the coboundary maps d_n used in the text. Historically, $H^n(G, A)$ was originally defined using homogeneous cochains.

4. Suppose H is a normal subgroup of the group G and A is a G-module. For every $g \in G$ prove that the map $f(a) = ga$ for $a \in A^H$ defines an automorphism of the subgroup A^H.

5. Suppose the G-module A decomposes as a direct sum $A = A_1 \oplus A_2$ of G-submodules. Prove that for all $n \geq 0$, $H^n(G, A) \cong H^n(G, A_1) \oplus H^n(G, A_2)$.

6. Suppose $0 \to A \to M_1 \to M_2 \to \cdots \to M_k \to C \to 0$ is an exact sequence of G-modules where $M_1, M_2, \ldots, M_k$ are cohomologically trivial. Prove that $H^{n+k}(G, A) \cong H^n(G, C)$ for all $n \geq 1$. [Decompose the exact sequence into a succession of short exact sequences and use Corollary 22. For example, if $0 \to A \xrightarrow{\alpha} M_1 \xrightarrow{\beta} M_2 \xrightarrow{\gamma} C \to 0$ is exact, show that $0 \to A \to M_1 \to B \to 0$ and $0 \to B \to M_2 \to C \to 0$ are both exact, where $B = M_1/\operatorname{image}\alpha = M_1/\ker\beta \cong \operatorname{image}\beta = \ker\gamma$.]

7. *(Adjoint Associativity)* Let R, S and T be rings with 1, let P be a left S-module, let N be a (T, S)-bimodule, and let A be a left T-module. Prove that

$$\Phi : \operatorname{Hom}_S(P, \operatorname{Hom}_T(N, A)) \longrightarrow \operatorname{Hom}_T(N \otimes_S P, A)$$

defined by $\Phi(f)(n \otimes p) = f(p)(n)$ is an isomorphism of abelian groups. (See also Theorem 43 in Section 10.5).

8. Suppose G is cyclic of order m with generator σ and let $N = 1+\sigma+\sigma^2+\cdots+\sigma^{m-1} \in \mathbb{Z}G$.

(a) Prove that the *augmentation* map $\operatorname{aug}(\sum_{i=0}^{m-1} a_i\sigma^i) = \sum_{i=0}^{m-1} a_i$ is a G-module homomorphism from $\mathbb{Z}G$ to $\mathbb{Z}$.

(b) Prove that multiplication by N and by $\sigma - 1$ in $\mathbb{Z}G$ define a free G-module resolution of $\mathbb{Z}$: $\ldots \xrightarrow{\sigma-1} \mathbb{Z}G \xrightarrow{N} \mathbb{Z}G \xrightarrow{\sigma-1} \ldots \xrightarrow{N} \mathbb{Z}G \xrightarrow{\sigma-1} \mathbb{Z}G \xrightarrow{\text{aug}} \mathbb{Z} \longrightarrow 0$.

9. Suppose G is an infinite cyclic group with generator σ.

(a) Prove that multiplication by $\sigma - 1 \in \mathbb{Z}G$ defines a free G-module resolution of $\mathbb{Z}$: $0 \longrightarrow \mathbb{Z}G \xrightarrow{\sigma-1} \mathbb{Z}G \longrightarrow \mathbb{Z} \longrightarrow 0$.

(b) Show that $H^0(G, A) \cong A^G$, that $H^1(G, A) \cong A/(\sigma-1)A$, and that $H^n(G, A) = 0$ for all $n \geq 2$. Deduce that $H^1(G, \mathbb{Z}G) \cong \mathbb{Z}$ (so free modules need not be cohomologically trivial).

10. Suppose H is a subgroup of finite index m in the group G and A is an H-module. Let $x_1, \ldots, x_m$ be a set of left coset representatives for H in G: $G = x_1H \cup \cdots \cup x_mH$.

(a) Prove that $\mathbb{Z}G = \bigoplus_{i=1}^m x_i\mathbb{Z}H = \bigoplus_{i=1}^m \mathbb{Z}Hx_i^{-1}$ and $\mathbb{Z}G \otimes_{\mathbb{Z}H} A = \bigoplus_{i=1}^m (x_i \otimes A)$ as abelian groups.

(b) Let $f_{i,a}$ be the function from $\mathbb{Z}G$ to A defined by

$$f_{i,a}(x) = \begin{cases} ha & \text{if } x = hx_i^{-1} \text{ with } h \in H \\ 0 & \text{otherwise.} \end{cases}$$

Prove that $f_{i,a} \in M_H^G(A) = \operatorname{Hom}_{\mathbb{Z}H}(\mathbb{Z}G, A)$, i.e., $f_{i,a}(h'x) = h'f_{i.a}(x)$ for $h' \in H$.

(c) Prove that the map $\varphi(f) = \sum_{i=1}^m x_i \otimes f(x_i^{-1})$ from $M_H^G(A)$ to $\mathbb{Z}G \otimes_{\mathbb{Z}H} A$ is a G-module homomorphism. [Write $x_i^{-1}g = h_ix_{i'}^{-1}$ for $i = 1, \ldots, m$ and observe that $x_i \otimes f(x_i^{-1}g) = x_i \otimes h_i f(x_{i'}^{-1}) = x_ih_i \otimes f(x_{i'}^{-1}) = gx_{i'} \otimes f(x_{i'}^{-1})$.]

(d) Prove that φ gives a G-module isomorphism $\varphi : M_H^G(A) \cong \mathbb{Z}G \otimes_{\mathbb{Z}H} A$. [For the injectivity observe that an H-module homomorphism is 0 if and only if $f(x_i^{-1}) = 0$ for $i = 1, \ldots, m$. For the surjectivity prove that $\varphi(f_{i,a}) = x_i \otimes a$.]

11. Prove that the isomorphism $M_H^G(A) \cong \mathbb{Z}G \otimes_{\mathbb{Z}H} A$ in (d) of the previous exercise need not hold if H is not of finite index in G. [If G is an infinite cyclic group show that Shapiro's Lemma implies $H^1(G, M_1^G(\mathbb{Z})) = 0$ while $H^1(G, \mathbb{Z}G) \cong \mathbb{Z}$ by Exercise 9.]

12. If H is a subgroup of G and A is an abelian group let $M_{G/H}(A)$ denote the abelian group of all maps from the left cosets gH of H in G to A.

(a) Prove that $M_1^G(A) \cong M_1^H(M_{G/H}(A))$ as H-modules. [If $\{g_i\}_{i\in\mathcal{I}}$ is a choice of left coset representatives of H in G define the correspondence between $f \in M_1^G(A)$ and $F : H \to M_{G/H}(A)$ by $F(h)(g_i H) = f(g_i h)$, and check that this is an isomorphism of H-modules.]

(b) A G-module A such that $H^n(H, A) = 0$ for all $n \geq 1$ and all subgroups H of G is called *cohomologically trivial.* Prove that $M_1^G(A)$ is a cohomologically trivial for any abelian group A.

(c) If G is finite, prove that $\mathbb{Z}G \otimes_{\mathbb{Z}} A$ is cohomologically trivial for all abelian groups A.

13. Suppose A is a G-module and H is a subgroup of G. Prove that the group homomorphism from $H^n(G, A)$ to $H^n(G, M_H^G(A))$ for all $n \geq 0$ induced from the G-module homomorphism from A to $M_H^G(A)$ in Example 3 following Corollary 22 composed with the isomorphism $H^n(G, M_H^G(A)) \cong H^n(H, A)$ of Shapiro's Lemma is the restriction homomorphism from $H^n(G, A)$ to $H^n(H, A)$.

14. Suppose $\varphi : H \to G$ is the inclusion map of the subgroup H of G into G. If A is an H-module and $M_H^G(A)$ the associated induced G-module, define the group homomorphism $\psi : M_H^G(A) \to A$ by mapping f to its value at 1: $\psi(f) = f(1)$.

(a) Prove that φ and ψ are compatible homomorphisms.

(b) Prove that the induced group homomorphism from $H^n(G, M_H^G(A))$ to $H^n(H, A)$ for $n \geq 0$ is the isomorphism in Shapiro's Lemma.

15. Suppose H is a normal subgroup of G and A is a G-module. For fixed $g \in G$, let $\psi(a) = ga$ and $\varphi(h) = g^{-1}hg$ for $h \in H$.

(a) Prove that φ and ψ are compatible homomorphisms.

(b) For each $n \geq 0$, prove that the homomorphism θ_g from $H^n(H, A)$ to $H^n(H, A)$ induced by the compatible homomorphisms φ and ψ is an automorphism of $H^n(H, A)$. [Observe that both φ and ψ have inverses.]

(c) Show that θ_g acting on $H^0(H, A)$ is the automorphism in Exercise 4.

16. Let A be a G-module and for $g \in G$ let θ_g denote the automorphism of $H^n(G, A)$ defined in the previous exercise.

(a) Prove that θ_g acting on $H^0(G, A) = A^G$ is the identity map.

(b) Prove that θ_g acting on $H^n(G, A)$ is the identity map for $n \geq 1$. [By induction on n and dimension shifting. For $n = 1$, use the exact sequence in Corollary 22, together with (a) applied to θ_g on C^G. For $n \geq 2$ use the isomorphism $H^{n+1}(G, A) \cong H^n(G, C)$ in Corollary 22.]

17. Suppose that H is a normal subgroup of G and A is a G-module. For $n \geq 0$ prove that $H^n(H, A)$ is a G/H-module where gH acts by the automorphism θ_g induced by conjugation by g on H and the natural action of g on A as in Exercise 15. [Use the previous exercise to show this action of a coset is well defined.]

18. Suppose that G is cyclic of order m, that H is a subgroup of G of index d, and that $\mathbb{Z}$ is a trivial G-module. Use the projective G-module resolution in Exercise 8 to prove

(a) that Cor : $H^n(H, \mathbb{Z}) \to H^n(G, \mathbb{Z})$ is multiplication by d from $\mathbb{Z}$ to $\mathbb{Z}$ for $n = 0$, from $\mathbb{Z}/(m/d)\mathbb{Z}$ to $\mathbb{Z}/m\mathbb{Z}$ if n is odd, and from 0 to 0 if n is even, $n \geq 2$, and

(b) that Res : $H^n(G, \mathbb{Z}) \to H^n(H, \mathbb{Z})$ is the identity map from $\mathbb{Z}$ to $\mathbb{Z}$ for $n = 0$, and is the natural projection map from $\mathbb{Z}/m\mathbb{Z}$ to $\mathbb{Z}/(m/d)\mathbb{Z}$ or from 0 to 0, depending on the parity of $n \geq 1$.

19. Let p be a prime and let P be a Sylow p-subgroup of the finite group G. Show that for any G-module A and all $n \geq 0$ the map $\text{Res} : H^n(G, A) \to H^n(P, A)$ is injective on the p-primary component of $H^1(G, A)$. Deduce that if $|A| = p^a$ then the restriction map is injective on $H^n(G, A)$. [Use Proposition 26.]

20. Let p be a prime, let $G = \langle \sigma \rangle$ be cyclic of order p^m and let W be a vector space of dimension $d > 0$ over $\mathbb{F}_p$ on which σ acts as a linear transformation. Assume W has a basis such that the matrix of σ is a $d \times d$ elementary Jordan block with eigenvalue 1.

(a) Prove that $d \leq p^m$. [Use facts about the minimal polynomial of an elementary Jordan block.]

(b) Prove that $\dim_{\mathbb{F}_p} W^G = 1$.

(c) Prove that $\dim_{\mathbb{F}_p} (\sigma - 1) W = d - 1$.

(d) If $N = 1 + \sigma + \cdots + \sigma^{p^m - 1}$ is the usual norm element, prove that NW is of dimension 1 if $d = p^m$ (respectively, of dimension 0 if $d < p^m$) and that the dimension of ${}_N W$ is $d - 1$ (respectively, d). [Let R be the group ring $\mathbb{F}_p G$, and show that every nonzero R-submodule of R contains N. Note that W is a cyclic R-module and let $\varphi : R \to W$ be a surjective homomorphism. Conclude that if φ is not an isomorphism then $N \in \ker \varphi$.]

(e) Deduce that if $d = p^m$ then $H^n(G, W) = 0$, and if $d < p^m$ then $H^n(G, W)$ has order p, for all $n \geq 1$ (i.e., these cohomology groups are zero if and only if W is a free $\mathbb{F}_p G$-module).

21. Let p be a prime, let $G = \langle \sigma \rangle$ be cyclic of order p^m and let V be a G-module of exponent p. Let $V = V_1 \oplus V_2 \oplus \cdots \oplus V_k$ be a decomposition of V giving the Jordan Canonical Form of σ, where each V_i is σ-invariant and a matrix of σ on V_i is an $d_i \times d_i$ elementary Jordan block with eigenvalue 1, $d_i \geq 1$ (cf. Section 12.3). Prove that $|V^G| = p^k$ and $|H^n(G, V)| = p^s$ where s is the number of V_i of dimension less than p^m over $\mathbb{F}_p$, for all $n \geq 1$. [Use the preceding exercise and Exercise 5.]

22. Suppose G is a topological group, i.e., there is a topology on G such that the maps $G \times G \to G$ defined by $(g_1, g_2) \mapsto g_1 g_2$ and $G \to G$ defined by $g \mapsto g^{-1}$ are continuous.

(a) If H is an open subgroup of G and $g \in G$, prove that the cosets gH and Hg and the subgroup $g^{-1}Hg$ are also open.

(b) Prove that any open subgroup is also closed. [The complement is the union of cosets as in (a).]

(c) Prove that a closed subgroup of finite index is open.

(d) If G is compact prove that every open subgroup H is of finite index.

23. Suppose G is a compact topological group. Prove the following are equivalent:

(i) G is profinite, i.e., $G = \varprojlim G_i$ is the inverse limit of finite groups G_i.

(ii) There exists a family $\{N_i\}$ $(i \in \mathcal{I})$ of open normal subgroups N_i in G such that $\cap_i N_i = 1$ and in this case $G \cong \varprojlim (G/N_i)$.

(iii) There exists a family $\{H_j\}$ $(j \in \mathcal{J})$ of open subgroups H_j in G such that $\cap_j H_j = 1$.

[To show (iii) implies (ii), let H be open in G and use (d) of the previous exercise to show that $N = \cap_{g \in G} g^{-1}Hg$ is a finite intersection and conclude that $N \subseteq H \subseteq G$ and N is open and normal in G.]

24. Suppose N and N' are open normal subgroups of the profinite group G and $N' \subseteq N$. Prove that the projection homomorphism $\varphi : G/N' \to G/N$ and the injection $\psi : A^N \to A^{N'}$ are compatible homomorphisms and deduce there is an induced homomorphism from $H^n(G/N, A^N)$ to $H^n(G/N', A^{N'})$.

25. If G is an infinite profinite group show that G does not act continuously on $A = \mathbb{Z}G$. [Show that the stabilizer of $a \in A$ is not always of finite index in G.]

17.3 CROSSED HOMOMORPHISMS AND $H^1(G,A)$

In this section we consider in greater detail the cohomology group $H^1(G, A)$ where G is a group and A is a G-module. From the definition of the coboundary map d_1 in equation (18), if $f \in C^1(G, A)$ then

$$d_1(f)(g_1, g_2) = g_1 \cdot f(g_2) - f(g_1 g_2) + f(g_1).$$

Thus any function $f : G \to A$ is a 1-cocycle if and only if it satisfies the identity

$$f(gh) = f(g) + gf(h) \qquad \text{for all } g, h \in G. \tag{17.20}$$

Equivalently, a 1-cocycle is determined by a collection $\{a_g\}_{g \in G}$ of elements in A satisfying $a_{gh} = a_g + ga_h$ for $g, h \in G$ (and then the 1-cocycle f is the function sending g to a_g). Note that if 1 denotes the identity of G, then $f(1) = f(1^2) = f(1) + 1 \cdot f(1) = 2f(1)$, so $f(1) = 0$ is the identity in A. Thus 1-cocycles are necessarily "normalized" at the identity. It then follows from the cocycle condition that $f(g^{-1}) = -g^{-1} f(g)$ for all $g \in G$.

If A is a G-module on which G acts trivially, then the cocycle condition (20) is simply $f(gh) = f(g) + f(h)$, i.e., f is simply a *homomorphism* from the multiplicative group G to the additive group A. Because of this the functions from G to A satisfying (20) are called *crossed homomorphisms*.

A 1-cochain f is a 1-coboundary if there is some $a \in A$ such that

$$f(g) = g \cdot a - a \qquad \text{for all } g \in G, \tag{17.21}$$

(equivalently, $a_g = ga - a$ in the notation above). Note that since $-a \in A$, the coboundary condition in (21) can also be phrased as $f(g) = a - g \cdot a$ for some fixed $a \in A$ and all $g \in G$. The 1-coboundaries are called *principal crossed homomorphisms*. With this terminology the cohomology group $H^1(G, A)$ is the group of crossed homomorphisms modulo the subgroup of principal crossed homomorphisms.

Example: (Hilbert's Theorem 90)

Suppose $G = \text{Gal}(K/F)$ is the Galois group of a finite Galois extension K/F of fields. Then the multiplicative group $K^\times$ is a G-module and $H^1(G, K^\times) = 0$. To see this, let $\{\alpha_\sigma\}$ be the values $f(\sigma)$ of a 1-cocycle f, so that $\alpha_\sigma \in K^\times$ and $\alpha_{\sigma\tau} = \alpha_\sigma \sigma(\alpha_\tau)$ (the cocycle condition written multiplicatively for the group $K^\times$). By the linear independence of automorphisms (Corollary 8 in Section 14.2), there is an element $\gamma \in K$ such that

$$\beta = \sum_{\tau \in G} \alpha_\tau \tau(\gamma)$$

is nonzero, i.e., $\beta \in K^\times$. Then for any $\sigma \in G$ we have

$$\sigma(\beta) = \sum_{\tau \in G} \sigma(\alpha_\tau)\, \sigma\tau(\gamma) = \alpha_\sigma^{-1} \sum_{\tau \in G} \alpha_{\sigma\tau}\, \sigma\tau(\gamma) = \alpha_\sigma^{-1} \beta$$

where the second equality comes from the cocycle condition. Hence $\alpha_\sigma = \beta/\sigma(\beta)$, which is the multiplicative form of the coboundary condition (21) (for the element $a = \beta^{-1}$). Since every 1-cocycle is a 1-coboundary, we have $H^1(G, K^\times) = 0$. The same result holds for infinite Galois extensions by equation (19) in the previous section since $H^1(G, K^\times)$ is the direct limit of trivial groups.

As a special case, suppose K/F is a Galois extension with cyclic Galois group G having generator σ. The cohomology groups for G were computed explicitly in the previous section, and in particular, $H^1(G, A) = {}_NA/(\sigma - 1)A$ for any G-module A (written additively). Since this group is trivial in the present context, we see that an element α in K is in the kernel of the norm map, i.e., $\mathrm{N}_{K/F}(\alpha) = 1$ if and only if $\alpha = \sigma(\beta)/\beta$ for some $\beta \in K$. (For a direct proof of this result in the cyclic case, cf. Exercise 23 in Section 14.2.)

This famous result for cyclic extensions was first proved by Hilbert and appears as "Theorem 90" in his book (known as the "*Zahlbericht*") on number theory in 1897. As a result, the more general result $H^1(G, K^\times) = 0$ is referred to in the literature as "Hilbert's Theorem 90." In general, the higher dimensional cohomology groups $H^n(G, K^\times)$ for $n \geq 2$ can be nontrivial (cf. Exercise 13).

Example

Suppose $G = \mathrm{Gal}(K/F)$ is the Galois group of a finite Galois extension K/F of fields as in the previous example. Then the additive group K is also a G-module and $H^n(G, K) = 0$ for all $n \geq 2$. The proof of this in general uses the fact that there is a *normal basis* for K over F, i.e., there is an element $\alpha \in K$ whose Galois conjugates give a basis for K as a vector space over F, or, equivalently, $K \cong \mathbb{Z}G \otimes_{\mathbb{Z}} F$ as G-modules. The latter isomorphism shows that K is induced as a G-module, and then $H^n(G, K) = 0$ follows from Corollary 24 in Section 2. For a direct proof in the case where G is cyclic, cf. Exercise 26 in Section 14.2.

If G acts trivially on A, then $g \cdot a - a = 0$, so 0 is the only principal crossed homomorphism, i.e., $B^1(G, A) = 0$. This proves the following result:

Proposition 30. If A is a G-module on which G acts trivially then $H^1(G, A) = \mathrm{Hom}(G, A)$, the group of all group homomorphisms from G to H.

If G is a profinite group, then the same result holds for the continuous cohomology group $H^1(G, A)$ provided one takes the group of continuous homomorphisms from G into A.

Examples

(1) If G acts trivially on A then $H^1(G, A) = H^1(G/[G, G], A)$ since any group homomorphism from G to the abelian group A factors through the commutator subgroup $[G, G]$ (cf. Proposition 7(5) in Section 5.4), so computing H^1 for trivial G-action reduces to computing H^1 for some abelian group.

(2) If G is a finite group acting trivially on $\mathbb{Z}$, then $H^1(G, \mathbb{Z}) = 0$ because $\mathbb{Z}$ has no nonzero elements of finite order so there is no nonzero group homomorphism from G to $\mathbb{Z}$.

(3) If A is cyclic of prime order p and G is a p-group then G must act trivially on A (since the automorphism group of A has order $p - 1$), so in this case one always has $H^1(G, A) = \mathrm{Hom}(G, A)$.

(4) If G is a finite group that acts trivially on $\mathbb{Q}/\mathbb{Z}$ then $H^1(G, \mathbb{Q}/\mathbb{Z}) = \mathrm{Hom}(G, \mathbb{Q}/\mathbb{Z}) = \hat{G}$ is the *dual group* of G (cf. Exercise 14 in Section 5.2.). Since $\mathbb{Q}/\mathbb{Z}$ is abelian, any homomorphism of G into $\mathbb{Q}/\mathbb{Z}$ factors through the commutator quotient $G^{\mathrm{ab}} = G/[G, G]$ of G, so $\mathrm{Hom}(G, \mathbb{Q}/\mathbb{Z}) = \mathrm{Hom}(G^{\mathrm{ab}}, \mathbb{Q}/\mathbb{Z})$. It follows that $\mathrm{Hom}(G, \mathbb{Q}/\mathbb{Z}) \cong \widehat{G^{\mathrm{ab}}}$ (which by cf. Exercise 14 again is noncanonically isomorphic to G^{ab}).

If $0 \to A \to B \to C \to 0$ is a short exact sequence of G-modules then the long exact sequence in group cohomology in Theorem 21 of the previous section begins with terms

$$0 \longrightarrow A^G \longrightarrow B^G \longrightarrow C^G \xrightarrow{\delta_0} H^1(G, A) \longrightarrow \cdots.$$

The connecting homomorphism δ_0 is given explicitly as follows: if $c \in C^G$ then there is an element $b \in B$ mapping to c and then $\delta_0(c)$ is the class in $H^1(G, A)$ of the 1-cocycle given by

$$\begin{aligned} \delta_0(c) : G &\longrightarrow A \\ g &\longmapsto g \cdot b - b. \end{aligned}$$

Note that $g \cdot b - b$ is (the image in B of) an element of A for all $g \in G$ since $c \in C^G$. To verify directly that $f = \delta_0(c)$ satisfies the cocycle condition in (20), we compute

$$f(gh) = gh \cdot b - b = (g \cdot b - b) + g \cdot (h \cdot b - b) = f(g) + gf(h).$$

From the explicit expression $f = g \cdot b - b$ it is also clear that $\delta_0(c) \in H^1(G, A)$ maps to 0 in the next term $H^1(G, B)$ of the long exact sequence above since f is the coboundary for the element $b \in B$.

Example: (Kummer Theory)

Suppose that F is a field of characteristic 0 containing the group μ_n of all n^{th} roots of unity for some $n \geq 1$. Let K be an algebraic closure of F and let $G = \mathrm{Gal}(K/F)$. The group G acts trivially on μ_n since $\mu_n \subset F$ by assumption, i.e., $\mu_n \cong \mathbb{Z}/n\mathbb{Z}$ as G-modules. Hence the Galois cohomology group $H^1(G, \mu_n)$ is the group $\mathrm{Hom}_c(G, \mathbb{Z}/n\mathbb{Z})$ of continuous homomorphisms of G into $\mathbb{Z}/n\mathbb{Z}$. If χ is such a continuous homomorphism, then $\ker \chi \subseteq G$ is a closed normal subgroup of G, hence corresponds by Galois theory to a Galois extension L_χ/F. Then $\mathrm{Gal}(L_\chi/F) \cong$ image χ, so L_χ is a cyclic extension of F of degree dividing n. Conversely, every such cyclic extension of F defines an element in $\mathrm{Hom}_c(G, \mathbb{Z}/n\mathbb{Z})$, so there is a bijection between the elements of the Galois cohomology group $H^1(G, \mu_n)$ and the cyclic extensions of F of degree dividing n.

The homomorphism of raising to the n^{th} power is surjective on $K^\times$ (since we can always extract n^{th} roots in K) and has kernel μ_n. Hence the sequence

$$1 \longrightarrow \mu_n \longrightarrow K^\times \xrightarrow{n} K^\times \longrightarrow 1$$

is an exact sequence of discrete G-modules. The associated long exact sequence in Galois cohomology gives

$$1 \longrightarrow \mu_n^G \longrightarrow (K^\times)^G \xrightarrow{n} (K^\times)^G \longrightarrow H^1(G, \mu_n) \longrightarrow H^1(G, K^\times) \longrightarrow \cdots.$$

We have $\mu_n^G = \mu_n$ and $(K^\times)^G = F^\times$ by Galois theory, and $H^1(G, K^\times) = 0$ by Hilbert's Theorem 90, so this exact sequence becomes

$$1 \longrightarrow \mu_n \longrightarrow F^\times \xrightarrow{n} F^\times \longrightarrow H^1(G, \mu_n) \longrightarrow 0,$$

which in turn is equivalent to the isomorphism

$$H^1(G, \mu_n) \cong F^\times / F^{\times n}$$

where $F^{\times n}$ denotes the group of n^{th} powers of elements of $F^{\times}$. This isomorphism is made explicit using the explicit form for the connecting homomorphism given above: for every $\alpha \in F^{\times}$ and $\sigma \in G$, the element $\sqrt[n]{\alpha}$ in $K^{\times}$ maps to α in the exact sequence and

$$\chi(\sigma) = \frac{\sigma(\sqrt[n]{\alpha})}{\sqrt[n]{\alpha}}$$

defines an element in $H^1(G, \mu_n)$ (cf. Exercise 11). The kernel of this homomorphism χ is the field $F(\sqrt[n]{\alpha})$. By the results of the previous paragraph, when F contains the n^{th} roots of unity an extension L/F is Galois with cyclic Galois group of order dividing n if and only if $L = F(\sqrt[n]{\alpha})$ for some $\alpha \in F^{\times}$. Furthermore, the class of α in $F^{\times}/F^{\times n}$ is unique, i.e., α is unique up to an n^{th} power of an element in F. Such an extension is called a *Kummer extension*, cf. Section 14.7 and Exercise 12.

If the characteristic of F is a prime p, the same argument applies when n is not divisible by p, replacing the algebraic closure of F with the separable closure of F (the largest separable algebraic extension of F).

Example: (The Transfer Homomorphism)

Suppose G is a finite group and H is a subgroup. The corestriction defines a homomorphism from $H^1(H, \mathbb{Q}/\mathbb{Z})$ to $H^1(G, \mathbb{Q}/\mathbb{Z})$, which by Example 4 above gives a homomorphism from $\widehat{H^{\text{ab}}}$ to $\widehat{G^{\text{ab}}}$. This gives a homomorphism

$$\text{Ver} : G^{\text{ab}} \longrightarrow H^{\text{ab}}$$

called the *transfer* (or *Verlagerungen*) homomorphism (cf. Exercise 14). To make this homomorphism explicit, consider the exact sequence

$$0 \longrightarrow \mathbb{Q}/\mathbb{Z} \longrightarrow M_1^G(\mathbb{Q}/\mathbb{Z}) \longrightarrow C \longrightarrow 0 \tag{17.22}$$

defined by the homomorphism mapping $a \in \mathbb{Q}/\mathbb{Z}$ to $f_a \in M_1^G(\mathbb{Q}/\mathbb{Z})$ in Example 4 preceding Proposition 23 in the previous section (so $f_a(g) = g \cdot a$ for $g \in G$). This is a short exact sequence of G-modules and hence also of H-modules. The first portions of the associated long exact sequences for the cohomology with respect to H and then G give the rows in the commutative diagram

$$\begin{tikzcd}
\cdots \arrow[r] & C^H \arrow[r, "\delta_0"] \arrow[d, "\text{Cor}"] & H^1(H, \mathbb{Q}/\mathbb{Z}) \arrow[r] \arrow[d, "\text{Cor}"] & 0 \\
\cdots \arrow[r] & C^G \arrow[r, "\delta_0"] & H^1(G, \mathbb{Q}/\mathbb{Z}) \arrow[r] & 0
\end{tikzcd}$$

since $H^1(H, M_1^G(\mathbb{Q}/\mathbb{Z})) = H^1(G, M_1^G(\mathbb{Q}/\mathbb{Z})) = 0$ (cf. Exercise 12 in Section 2). Let $\chi \in H^1(H, \mathbb{Q}/\mathbb{Z})$ and suppose that $c \in C^H$ is an element mapping to χ by the surjective connecting homomorphism δ_0 in the first row of the diagram above. By the commutativity, $\chi' = \text{Cor}(\chi)$ is the image under the connecting homomorphism δ_0 of $c' = \text{Cor}(c) \in C^G$ in the second row of the diagram. By our explicit formula for the coboundary map δ_0, if $F \in M_1^G(\mathbb{Q}/\mathbb{Z})$ is any element mapping to c' in (22) then $g \cdot F - F = f_{a'}$ for a unique $a' \in \mathbb{Q}/\mathbb{Z}$, and we have $\chi'(g) = \delta_0(c')(g) = a'$ for $g \in G$. Since $f_{a'}(x) = x \cdot a' = a'$ for any $x \in G$ because G acts trivially on $\mathbb{Q}/\mathbb{Z}$, the function $g \cdot F - F$ in fact has the constant value a', and so can be evaluated at any $x \in G$ to determine the value of $\chi'(g)$.

Since $c' = \sum_{i=1}^{m} g_i \cdot c \in C^G$ where $g_1, \ldots, g_m$ are representatives of the left cosets of H in G (cf. Example 4 preceding Proposition 26), such an element F is given by

$$F = \sum_{i=1}^{m} g_i \cdot f,$$

where $f \in M_1^G(\mathbb{Q}/\mathbb{Z})$ is any element mapping to c in (22). This f can be used to compute the explicit coboundary of c as before: $h \cdot f - f = f_a$ for a unique $a \in \mathbb{Q}/\mathbb{Z}$ and $\chi(h) = a$ for $h \in H$. As before, the function $h \cdot f - f = f_a$ has the constant value a and so can be evaluated at any element x of G to determine the value of $\chi(h)$.

Computing $g \cdot F - F$ on the element $1 \in G$ it follows that

$$\chi'(g) = \sum_{i=1}^{m} f(gg_i) - \sum_{i=1}^{m} f(g_i).$$

For $i = 1, \ldots, m$, write

$$gg_i = g_j h(g, g_i) \qquad \text{with } h(g, g_i) \in H, \tag{17.23}$$

noting that the resulting set of g_j is some permutation of $\{g_1, \ldots, g_m\}$. Then

$$\sum_{i=1}^{m} f(gg_i) - \sum_{i=1}^{m} f(g_i) = \sum_{i=1}^{m} [f(g_j h(g, g_i)) - f(g_j)] = \sum_{i=1}^{m} \chi(h(g, g_i))$$

since as noted above, $\chi(h) = f(xh) - f(x)$ for any $x \in G$. Hence

$$\chi'(g) = \chi(\prod_{i=1}^{m} h(g, g_i))$$

and so the transfer homomorphism is given by the formula

$$\mathrm{Ver}(g) = \prod_{i=1}^{m} h(g, g_i) \tag{17.24}$$

with the elements $h(g, g_i) \in H$ defined by equation (23). Note that this proves in particular that the map defined in (24) is a homomorphism from G^{ab} to H^{ab} that is independent of the choice of representatives g_i for H in G in (23). Proving that this map is a homomorphism directly is not completely trivial. The same formula also defines the transfer homomorphism when G is infinite and H is a subgroup of finite index in G.

As an example of the transfer, suppose $H = n\mathbb{Z}$ and $G = \mathbb{Z}$ and choose $0, 1, 2, \ldots, n-1$ as coset representatives for H in G. If $g = 1$, then all the elements $h(g, g_i)$ are 0 for $i = 1, 2, \ldots, n-1$ and $h(1, n-1) = n$. Hence the transfer map from $\mathbb{Z}$ to $n\mathbb{Z}$ maps 1 to n, so is simply multiplication by the index. Similarly, the transfer map from any cyclic group G to a subgroup H of index n is the n^{th} power map. See also Exercise 8.

For the cyclic group $\mathbb{F}_p^\times$ for an odd prime p and subgroup $\{\pm 1\}$, it follows that the transfer map is the homomorphism $\mathrm{Ver} : \mathbb{F}_p^\times \to \{\pm 1\}$ given by

$$\mathrm{Ver}(a) = a^{(p-1)/2} = \left(\frac{a}{p}\right) = \begin{cases} +1 & \text{if } a \text{ is a square} \\ -1 & \text{if } a \text{ is not a square} \end{cases}$$

(the symbol $(\frac{a}{p})$ is called the *Legendre symbol* or the *quadratic residue symbol*). If instead we take the elements $1, 2, \ldots, (p-1)/2$ as coset representatives for $\{\pm 1\}$ in $\mathbb{F}_p^\times$ we see that

$$\left(\frac{a}{p}\right) = (-1)^{m(a)}$$

where $m(a)$ is the number of elements among $a, 2a, \ldots, (p-1)a/2$ whose least positive remainder modulo p is greater than $(p-1)/2$ (in which case the element differs by -1 from one of our chosen coset representatives and contributes one factor of -1 to the product in (24)). This result is known as *Gauss' Lemma* in elementary number theory and can be used to prove Gauss' celebrated Quadratic Reciprocity Law (cf. also Exercise 15).

Next we give two important interpretations of $H^1(G, A)$ in terms of semidirect products. If A is a G-module, let E be the semidirect product $E = A \rtimes G$, where A is normal in E and the action of G (viewed as a subgroup of E) on A by conjugation is the same as its G-module action: $gag^{-1} = g \cdot a$. In the notation of Section 5.5, $E = A \rtimes_\varphi G$, where φ is the homomorphism of G into $\mathrm{Aut}(A)$ given by the G-module action. In particular, E will be the direct product of A and G if and only if G acts trivially on A. As in Section 5.5, we shall write the elements of E as (a, g) where $a \in A$ and $g \in G$, with group operation

$$(a_1, g_1)(a_2, g_2) = (a_1 + g_1 \cdot a_2, g_1 g_2).$$

Note that A is written additively, while G and E are written multiplicatively.

Definition. Let X be any group and let Y be a normal subgroup of X. The *stability group* of the series $1 \trianglelefteq Y \trianglelefteq X$ is the group of all automorphisms of X that map Y to itself and act as the identity on both of the factors Y and X/Y, i.e.,

$$\mathrm{Stab}(1 \trianglelefteq Y \trianglelefteq X) = \{\sigma \in \mathrm{Aut}(X) \mid \sigma(y) = y \text{ for all } y \in Y,$$
$$\text{and } \sigma(x) \equiv x \bmod Y \text{ for all } x \in X\}.$$

In the special case where Y is an *abelian* normal subgroup of X, conjugation by elements of Y induce (inner) automorphisms of X that stabilize the series $1 \trianglelefteq Y \trianglelefteq X$, and in this case $Y/C_Y(X)$ is isomorphic to a subgroup of $\mathrm{Stab}(1 \trianglelefteq Y \trianglelefteq X)$ (where $C_Y(X)$ is the elements of Y in the center of X).

Proposition 31. Let A be a G-module and let E be the semidirect product $A \rtimes G$. For each cocycle $f \in Z^1(G, A)$ define $\sigma_f : E \to E$ by

$$\sigma_f((a, g)) = (a + f(g), g).$$

Then the map $f \to \sigma_f$ is a group isomorphism from $Z^1(G, A)$ onto $\mathrm{Stab}(1 \trianglelefteq A \trianglelefteq E)$. Under this isomorphism the subgroup $B^1(G, A)$ of coboundaries maps onto the subgroup $A/C_A(E)$ of the stability group.

Proof: It is an exercise to see that the cocycle condition implies σ_f is an automorphism of E that stabilizes the chain $1 \trianglelefteq A \trianglelefteq E$. Likewise one checks directly that $\sigma_{f_1+f_2} = \sigma_{f_1} \circ \sigma_{f_2}$, so the map $f \mapsto \sigma_f$ is a group homomorphism. By definition of σ_f this map is injective. Conversely, let $\sigma \in \mathrm{Stab}(1 \trianglelefteq A \trianglelefteq E)$. Since σ acts trivially on E/A, each element $(0, g)$ in this semidirect product maps under σ to another element (a, g) in the same coset of A; define $f_\sigma : G \to A$ by letting $f_\sigma(g) = a$. If we identify A with the elements of the form $(a, 1)$ in E, then the group operation in E shows that

$$f_\sigma(g) = \sigma((0, g))(0, g)^{-1}.$$

Because σ is a stability automorphism of E, it is easy to check that f_σ satisfies the cocycle condition. It follows immediately from the definitions that $f_{\sigma_f} = f$, so the map $f \mapsto \sigma_f$ is an isomorphism.

Now f is a coboundary if and only if there is some $x \in A$ such that $f(g) = x - g\cdot x$ for all $g \in G$. Thus f is a coboundary if and only if $\sigma_f((a, g)) = (a + x - g\cdot x, g)$. But conjugation in E by the element $(x, 1)$ maps (a, g) to the same element $(a + x - g\cdot x, g)$, so the automorphism σ_f is conjugation by $(x, 1)$. This proves the remaining assertion of the proposition.

Corollary 32. In the notation of Proposition 31 let φ_a denote the automorphism of E given by conjugation by a for any $a \in A$. Then the cocycles f_1 and f_2 are in the same cohomology class in $H^1(G, A)$ if and only if $\sigma_{f_1} = \varphi_a \circ \sigma_{f_2}$, for some $a \in A$.

The proposition and corollary show that 1-cocycles may be computed by finding automorphisms of E that stabilize the series $1 \trianglelefteq A \trianglelefteq E$, and vice versa. The first cohomology group is then given by taking these automorphisms modulo inner automorphisms, i.e., is the group of "outer stability automorphisms" of this series.

Example

Let $G = Z_2$ act by inversion on $A = \mathbb{Z}/4\mathbb{Z}$. The corresponding semidirect product $E = A \rtimes G$ is the dihedral group of order 8, which has automorphism group isomorphic to D_8; viewing E as a normal (index 2) subgroup of D_{16}, conjugation in the latter group restricted to E exhibits 8 distinct automorphisms of E (cf. Proposition 17 in Section 4.4). The subgroup A of E is characteristic in E, hence every automorphism of E sends A to itself, and therefore also acts on E/A (necessarily trivially since $|E/A| = 2$). Half the automorphisms of E invert A and half centralize A; in fact, the cyclic subgroup of order 8 in D_{16} (which contains A) maps to a cyclic group of order 4 of automorphisms centralizing A. Thus $\mathrm{Stab}(1 \trianglelefteq A \trianglelefteq E) \cong Z_4 \cong Z^1(G, A)$. Since the center of E is a subgroup of A of order 2, $|A/Z(E)| = 2 = |B^1(G, A)|$. This proves $|H^1(G, A)| = 2$.

In the semidirect product E the subgroup G is a complement to A, i.e., $E = AG$ and $A \cap G = 1$; moreover, every E-conjugate of G is also a complement to A. But A may have complements in E that are not conjugate to G in E. Our second interpretation of $H^1(G, A)$ shows that this cohomology group characterizes the E-conjugacy classes of complements of A in E.

Proposition 33. Let A be a G-module and let E be the semidirect product $A \rtimes G$. For each 1-cocycle f let

$$G_f = \{(f(g), g) \mid g \in G\}.$$

Then G_f is a subgroup complement to A in E. The map $f \mapsto G_f$ is a bijection from $Z^1(G, A)$ to the set of complements to A in E. Two complements are conjugate in E if and only if their corresponding 1-cocycles are in the same cohomology class in $H^1(G, A)$, so there is a bijection between $H^1(G, A)$ and the set of E-conjugacy classes of complements to A.

Proof: By the cocycle condition,

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

and it follows that G_f is closed under the group operation in E. As observed earlier, each cocycle necessarily has $f(1) = 0$, so G_f contains the identity $(0, 1)$ of E. The inverse to $(f(g), g)$ in E is $(f(g^{-1}), g^{-1})$, so G_f is closed under inverses. This proves G_f is a subgroup of E. Since the distinct elements of G_f represent the distinct cosets of A in E, G_f is a complement to A in E. Distinct cocycles give different coset representatives, hence they determine different complements.

Conversely, if C is any complement to A in G, then C contains a unique coset representative $a_g g$ of Ag for each $g \in G$. Since C is closed under the group operation the element $(a_g g)(a_h h) = (a_g g a_h g^{-1})gh$ represents the coset Agh, and so a_{gh} is $a_g g a_h g^{-1} = a_g (g \cdot a_h)$ (written additively in A this becomes $a_{gh} = a_g + (g \cdot a_h)$). This shows that the map $f : G \to A$ given by $f(g) = a_g$ is a cocycle, and so $C = G_f$. Hence there is a bijection between 1-cocycles and complements to A in E.

Since $\mathrm{Stab}(1 \trianglelefteq A \trianglelefteq E)$ normalizes A it permutes the complements to A in E. In the notation of Proposition 31, for 1-cocycles f_1 and f_2 it follows immediately from the definition that $\sigma_{f_1}(G_{f_2}) = G_{f_1+f_2}$. This shows that the permutation action of $\mathrm{Stab}(1 \trianglelefteq A \trianglelefteq E)$ on the set of complements to A in E is the (left) regular representation of this group. Furthermore, if $a \in A$ and φ_a is the stability automorphism conjugation by a, then

$$aG_f a^{-1} = \varphi_a(G_f) = G_{f+\beta_a} \tag{17.25}$$

where β_a is the 1-coboundary $\beta_a : g \mapsto a - g \cdot a$. Since G_f is a complement to A, any $e \in E$ may be written as ag for some $a \in A$ and $g \in G_f$. Then $eG_f e^{-1} = aG_f a^{-1}$, i.e., the E-conjugates of G_f are the just the A-conjugates of G_f. Now the complements G_{f_1} and G_{f_2} are conjugate in E if and only if $G_{f_2} = aG_{f_1}a^{-1} = G_{f_1+\beta_a}$ for some $a \in A$ by (25). This shows two complements are conjugate in E if and only if their corresponding cocycles differ by a coboundary, i.e., represent the same cohomology class in $H^1(G, A)$, which completes the proof.

Corollary 34. Under the notation of Proposition 33, all complements to A are conjugate in E if and only if $H^1(G, A) = 0$.

Corollary 35. If A is a finite abelian group whose order is relatively prime to $|G|$ then all complements to A in any semidirect product $E = A \rtimes G$ are conjugate in E.

Examples

(1) Let $A = \langle a \rangle$ and $G = \langle g \rangle$ both be cyclic of order 2. The group G must act trivially on A, hence $A \rtimes G = A \times G$ is a Klein 4-group. Here $A \rtimes G$ is abelian, so every subgroup is conjugate only to itself, and since $H^1(G, A) = \mathrm{Hom}(Z_2, \mathbb{Z}/2\mathbb{Z})$ has order 2, there are precisely two complements to A in E, namely $\langle g \rangle$ and $\langle ag \rangle$.

(2) If $A = \langle a \rangle$ is cyclic of order 2 and $G = \langle x \rangle \times \langle y \rangle$ is a Klein 4-group, then as before G must act trivially on A, so $H^1(G, A) = \mathrm{Hom}(Z_2 \times Z_2, \mathbb{Z}/2\mathbb{Z})$ has order 4. The four complements to A in $A \times G$ are G, $\langle ax, y \rangle$, $\langle x, ay \rangle$ and $\langle ax, ay \rangle$.

(3) Proposition 33 can also be used to compute $H^1(G, A)$. Let $A = \langle r \rangle$ be cyclic of order 4 and let $G = \langle s \rangle$ be cyclic of order 2 acting on A by inversion: $srs^{-1} = r^{-1}$ as in the Example following Corollary 32. Then $A \rtimes G$ is the dihedral group D_8 of order 8. The subgroup A has four complements in D_8, namely the groups generated

by each of the four elements of order 2 not in A: $\langle s \rangle$, $\langle r^2 s \rangle$, $\langle rs \rangle$ and $\langle r^3 s \rangle$. The former pair and the latter pair are conjugate in D_8 (in both cases via r), but $\langle s \rangle$ is not conjugate to $\langle rs \rangle$. Thus A has 2 conjugacy classes of complements in $A \rtimes G$ and hence $H^1(Z_2, \mathbb{Z}/4\mathbb{Z})$ has order 2. This also follows from the computation of the cohomology of cyclic groups in Section 2.

EXERCISES

1. Let G be the cyclic group of order 2 and let A be a G-module. Compute the isomorphism types of $Z^1(G, A)$, $B^1(G, A)$ and $H^1(G, A)$ for each of the following:
   (a) $A = \mathbb{Z}/4\mathbb{Z}$ (trivial action),
   (b) $A = \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ (trivial action),
   (c) $A = \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ (any nontrivial action).
2. Let p be a prime and let P be a p-group.
   (a) Show that $H^1(P, \mathbb{F}_p) \cong P/\Phi(P)$, where $\Phi(P)$ is the Frattini subgroup of P (cf. the exercises in Section 6.1).
   (b) Deduce that the dimension of $H^1(P, \mathbb{F}_p)$ as a vector space over $\mathbb{F}_p$ equals the minimum number of generators of P. [Use Exercise 26(c), Section 6.1.]
3. If G is the cyclic group of order 2 acting by inversion on $\mathbb{Z}$ show that $|H^1(G, \mathbb{Z})| = 2$. [Show that in $E = \mathbb{Z} \rtimes G$ every element of $E - \mathbb{Z}$ has order 2, and there are two conjugacy classes in this coset.]
4. Let A be the Klein 4-group and let $G = \text{Aut}(A) \cong S_3$ act on A in the natural fashion. Prove that $H^1(G, A) = 0$. [Show that in the semidirect product $E = A \rtimes G$, G is the normalizer of a Sylow 3-subgroup of E. Apply Sylow's Theorem to show all complements to A in E are conjugate.]
5. Let G be the cyclic group of order 2 acting on an elementary abelian 2-group A of order 2^n. Show that $H^1(G, A) = 0$ if and only if $n = 2k$ and $|A^G| = 2^k$. [In $E = A \rtimes G$ show that (a, x) is an element of order 2 if and only if $a \in A^G$, where $G = \langle x \rangle$. Then compare the number of complements to A with the number of E-conjugates of x.]
6. *(Thompson Transfer Lemma)* Let G be a finite group of even order, let T be a Sylow 2-subgroup of G, let $M \leq T$ with $|T : M| = 2$, and let x be an element of order 2 in G. Show that if G has no subgroup of index 2 then M contains some G-conjugate of x as follows:
   (a) Let $\text{Ver} : G/[G, G] \to T/[T, T]$ be the transfer homomorphism. Show that
   $$\text{Ver}(x) = \prod_g g^{-1} x g \bmod [T, T]$$
   where the product is over representatives of the cosets gT that are fixed under left multiplication by x.
   (b) Show that under left multiplication x fixes an odd number of left cosets of T in G.
   (c) Show that if G has no subgroup of index 2 then $\text{Ver}(x) \in M/[T, T]$. Deduce that for some $g \in G$ we must have $g^{-1} x g \in M$. [Consider the product $\text{Ver}(x)$ in the group T/M of order 2.]
7. Let H be a subgroup of G and let $x \in G$. The transfer $\text{Ver} : G/[G, G] \to H/[H, H]$ may be computed as follows: let $\mathcal{O}_1, \mathcal{O}_2, \dots, \mathcal{O}_k$ be the distinct orbits of x acting by left multiplication on the left cosets of H in G, let $\mathcal{O}_i$ have length n_i and let $g_i H$ be any representative of $\mathcal{O}_i$.

(a) Show that $\mathcal{O}_i = \{g_i H, xg_i H, x^2 g_i H, \dots, x^{n_i - 1} g_i H\}$ and that $g_i^{-1} x^{n_i} g_i \in H$.
(b) Show that $\mathrm{Ver}(x) = \prod_{i=1}^{k} g_i^{-1} x^{n_i} g_i \bmod [H, H]$.

8. Assume the center, $Z(G)$, of G is of index m. Prove that $\mathrm{Ver}(x) = x^m$, for all $x \in G$, where Ver is the transfer homomorphism from $G/[G, G]$ to $Z(G)$. [Use the preceding exercise.]

9. Let p be a prime, let $n \geq 3$, and let V be an n-dimensional vector space over $\mathbb{F}_p$ with basis $v_1, v_2, \dots, v_n$. Let V be a module for the symmetric group S_n, where each $\pi \in S_n$ permutes the basis in the natural way: $\pi(v_i) = v_{\pi(i)}$.
(a) Show that $|H^1(S_n, V)| = \begin{cases} 0, & \text{if } p \neq 2 \\ 2, & \text{if } p = 2 \end{cases}$. [Use Shapiro's Lemma.]
(b) Show that $H^1(A_n, V) = 0$ for all primes p.

10. Let V be the natural permutation module for S_n over $\mathbb{F}_2$, $n \geq 3$, as described in the preceding exercise, and let $W = \{a_1 v_1 + \cdots + a_n v_n \mid a_1 + \cdots + a_n = 0\}$ (the "trace zero" submodule of V). Show that if n is even then $H^1(A_n, W) \neq 0$. [Show that in the semidirect product $V \rtimes A_n$ the element v_1 induces a nontrivial outer automorphism on $E = W \rtimes A_n$ that stabilizes the series $1 \trianglelefteq W \trianglelefteq E$.]

11. Let F be a field of characteristic not dividing n and let α be any nonzero element in F. Let K be a Galois extension of F containing the splitting field of $x^n - a$, and let $\sqrt[n]{\alpha}$ be a fixed n^{th} root of α in K.
(a) Prove that $\sigma(\sqrt[n]{\alpha})/\sqrt[n]{\alpha}$ is an n^{th} root of unity.
(b) Prove that the function $f(\sigma) = \sigma(\sqrt[n]{\alpha})/\sqrt[n]{\alpha}$ is a 1-cocycle of G with values in the group μ_n of n^{th} roots of unity in K (note μ_n is not assumed to be contained in F).
(c) Prove that the 1-cocycle obtained by a different choice of n^{th} root of α in K differs from the 1-cocycle in (b) by a 1-coboundary.

12. Let F be a field of characteristic not dividing n that contains the n^{th} roots of unity, and suppose L/F is a Galois extension with abelian Galois group of exponent dividing n. Prove that L is the composite of cyclic extensions of F whose degrees are divisors of n and use this to prove that there is a bijection between the subgroups of the multiplicative group $F^\times/F^{\times^n}$ and such extensions L.

13. The Galois group of the extension $\mathbb{C}/\mathbb{R}$ is the cyclic group $G = \langle \tau \rangle$ of order 2 generated by complex conjugation τ. Prove that $H^2(G, \mathbb{C}^\times) \cong \mathbb{R}^\times/\mathbb{R}^+ \cong \mathbb{Z}/2\mathbb{Z}$ where $\mathbb{R}^+$ denotes the positive real numbers.

14. For any group G let $\hat{G} = \mathrm{Hom}(G, \mathbb{Q}/\mathbb{Z})$ denote its dual group.
(a) If $\varphi : G_1 \to G_2$ is a group homomorphism prove that composition with φ induces a homomorphism $\hat{\varphi} : \hat{G}_2 \to \hat{G}_1$ on their dual groups.
(b) For any fixed g in G, show that evaluation at g gives a homomorphism φ_g from $\hat{G}$ to $\mathbb{Q}/\mathbb{Z}$.
(c) Prove that the map taking $g \in G$ to φ_g in (b) defines a homomorphism from G to its *double dual* $\widehat{(\hat{G})}$.
(d) Prove that if G is a finite abelian group then the homomorphism in (c) is an isomorphism of G with its double dual. (By Exercise 14 in Section 5.2 the group G is (noncanonically) isomorphic to its dual $\hat{G}$. This shows that G is *canonically* isomorphic to its double dual — the isomorphism is independent of any choice of generators for G.)
(e) If $\psi : \hat{G}_2 \to \hat{G}_1$ is a homomorphism where G_1 and G_2 are finite abelian groups, then by (a) and (d) there is an induced homomorphism $\varphi : G_1 \to G_2$. Prove that

$\varphi(g_1) = g_2$ if $\chi(g_2) = \chi'(g_1)$ for $\chi' = \psi(\chi)$.

15. Use Gauss' Lemma in the computation of the transfer map for $\mathbb{F}_p^\times$ to $\{\pm 1\}$ to prove that 2 is a square modulo the odd prime p if and only if $p \equiv \pm 1 \bmod 8$. [Count how many elements in $2, 4, \dots, p-1$ are greater than $(p-1)/2$.]

17.4 GROUP EXTENSIONS, FACTOR SETS AND $H^2(G,A)$

If A is a G-module then from the definition of the coboundary map d_2 in equation (18) a function f from $G \times G$ to A is a 2-cocycle if it satisfies the identity

$$f(g,h) + f(gh,k) = g \cdot f(h,k) + f(g,hk) \qquad \text{for all } g,h,k \in G. \qquad (17.26)$$

Equivalently, a 2-cocycle is determined by a collection of elements $\{a_{g,h}\}_{g,h \in G}$ of elements in A satisfying $a_{g,h} + a_{gh,k} = g \cdot a_{h,k} + a_{g,hk}$ for $g, h, k \in G$ (and then the 2-cocycle f is the function sending (g,h) to $a_{g,h}$).

A 2-cochain f is a coboundary if there is a function $f_1 : G \to A$ such that

$$f(g,h) = gf_1(h) - f_1(gh) + f_1(g), \qquad \text{for all } g,h \in G \qquad (17.27)$$

i.e., f is the image under d_1 of the 1-cochain f_1.

One of the main results of this section is to make a connection between the 2-cocycles $Z^2(G,A)$ and the *factor sets* associated to a group extension of G by A, which arise when considering the effect of choosing different coset representatives in defining the multiplication in the extension. In particular, we shall show that there is a bijection between equivalence classes of group extensions of G by A (with the action of G on A fixed) and the elements of $H^2(G,A)$.

We first observe some basic facts about extensions. Let E be any group extension of G by A,

$$1 \longrightarrow A \xrightarrow{\iota} E \xrightarrow{\pi} G \longrightarrow 1. \qquad (17.28)$$

The extension (28) determines an action of G on A, as follows. For each $g \in G$ let e_g be an element of E mapping onto g by π (the choice of such a set of representatives for G in E is called a set-theoretic *section* of π). The element e_g acts by conjugation on the normal subgroup $\iota(A)$ of E, mapping $\iota(a)$ to $e_g\iota(a)e_g^{-1}$. Any other element in E that maps to g is of the form $e_g\iota(a_1)$ for some $a_1 \in A$, and since $\iota(A)$ is abelian, conjugation by this element on $\iota(A)$ is the same as conjugation by e_g, so is independent of the choice of representative for g. Hence G acts on $\iota(A)$, and so also on A since ι is injective. Since conjugation is an automorphism, the extension (28) defines A as a G-module.

Recall from Section 10.5 that two extensions $1 \to A \xrightarrow{\iota_1} E_1 \xrightarrow{\pi_1} G \to 1$ and $1 \to A \xrightarrow{\iota_2} E_2 \xrightarrow{\pi_2} G \to 1$ are *equivalent* if there is a group isomorphism $\beta : E_1 \to E_2$ such that the following diagram commutes:

$$\begin{tikzcd}
1 \arrow[r] & A \arrow[r, "\iota_1"] \arrow[d, "\mathrm{id}"] & E_1 \arrow[r, "\pi_1"] \arrow[d, "\beta"] & G \arrow[r] \arrow[d, "\mathrm{id}"] & 1 \\
1 \arrow[r] & A \arrow[r, "\iota_2"] & E_2 \arrow[r, "\pi_2"] & G \arrow[r] & 1.
\end{tikzcd} \qquad (17.29)$$

In this case we simply say β is the equivalence between the two extensions. As noted in Section 10.5, equivalence of extensions is reflexive, symmetric and transitive. We also observe that

equivalent extensions define the same G-module structure on A.

To see this assume (29) is an equivalence, let g be any element of G and let e_g be any element of E_1 mapping onto g by π_1. The action of g on A given by conjugation in E_1 maps each a to $\iota_1^{-1}(e_g\iota_1(a)e_g^{-1})$. Let $e'_g = \beta(e_g)$. Since the diagram commutes, $\pi_2(e'_g) = g$, so the action of g on A in the second extension is given by conjugation by e'_g. This conjugation maps a to $\iota_2^{-1}(e'_g\iota_2(a)e_g'^{-1})$. Since ι_1, ι_2 and β are injective, the two actions of g on a are equal if and only if they result in the same image in E_2, i.e., $\beta \circ \iota_1(\iota_1^{-1}(e_g\iota_1(a)e_g^{-1})) = e'_g\iota_2(a)e_g'^{-1}$. This equality is now immediate from the definition of e'_g and the commutativity of the diagram.

We next see how an extension as in (28) defines a 2-cocycle in $Z^2(G, A)$. For simplicity we identify A as a subgroup of E via ι and we identify G as E/A via π.

Definition. A map $\mu : G \to E$ with $\pi \circ \mu(g) = g$ and $\mu(1) = 0$, i.e., so that for each $g \in G$, $\mu(g)$ is a representative of the coset Ag of E and the identity of E (which is the zero of A) represents the identity coset, is called a *normalized section* of π.

Fix a section μ of π in (28). Each element of E may be written uniquely in the form $a\mu(g)$, where $a \in A$ and $g \in G$. For $g, h \in G$ the product $\mu(g)\mu(h)$ in E lies in the coset Agh, so there is a unique element $f(g, h)$ in A such that

$$\mu(g)\mu(h) = f(g, h)\mu(gh) \qquad \text{for all } g, h \in G. \tag{17.30}$$

If in addition μ is normalized at the identity we also have

$$f(g, 1) = 0 = f(1, g) \qquad \text{for all } g \in G. \tag{17.31}$$

Definition. The function f defined by equation (30) is called the *factor set* for the extension E associated to the section μ. If f also satisfies (31) then f is called a *normalized* factor set.

We shall see in the examples following that it is possible for different sections μ to give the same factor set f.

We now verify that the factor set f is in fact a 2-cocycle. First note that the group operation in E may be written

$$\begin{aligned}(a_1\mu(g))(a_2\mu(h)) &= (a_1 + \mu(g)a_2\mu(g)^{-1})\mu(g)\mu(h)\\ &= (a_1 + g \cdot a_2)(\mu(g)\mu(h))\\ &= (a_1 + g \cdot a_2 + f(g, h))\mu(gh)\end{aligned} \tag{17.32}$$

where $g \cdot a_2$ denotes the G-module action of g on a_2 given by conjugation in E. Now use (32) and the associative law in E to compute the product $\mu(g)\mu(h)\mu(k)$ in two different ways:

$$\begin{aligned}(\mu(g)\mu(h))\mu(k) &= (f(g, h) + f(gh, k))\mu(ghk)\\ \mu(g)(\mu(h)\mu(k)) &= (gf(h, k) + f(g, hk))\mu(ghk).\end{aligned} \tag{17.33}$$

It follows that the factors in A of the two right hand sides in (33) are equal for every $g, h, k \in G$, and this is precisely the 2-cocycle condition (26) for f. This shows that the factor set associated to the extension E and any choice of section μ is an element in $Z^2(G, A)$.

We next see how the factor set f depends on the choice of section μ. Suppose μ' is another section for the same extension E in (28), and let f' be its associated factor set. Then for all $g \in G$ both $\mu(g)$ and $\mu'(g)$ lie in the same coset Ag, so there is a function $f_1 : G \to A$ such that $\mu'(g) = f_1(g)\mu(g)$ for all g. Then

$$\mu'(g)\mu'(h) = f'(g, h)\mu'(gh) = (f'(g, h) + f_1(gh))\mu(gh).$$

We also have

$$\begin{aligned}\mu'(g)\mu'(h) &= (f_1(g)\mu(g))(f_1(h)\mu(h)) = (f_1(g) + g \cdot f_1(h))(\mu(g)\mu(h)) \\ &= (f_1(g) + g \cdot f_1(h) + f(g, h))\mu(gh).\end{aligned}$$

Equating the factors in A in these two expressions for $\mu'(g)\mu'(h)$ shows that

$$f'(g, h) = f(g, h) + (gf_1(h) - f_1(gh) + f_1(g)) \qquad \text{for all } g, h \in G,$$

in other words f and f' differ by the 2-coboundary of f_1 as in (27).

We have shown that the factor sets associated to the extension E corresponding to different choices of sections give 2-cocycles in $Z^2(G, A)$ that differ by a coboundary in $B^2(G, A)$. Hence associated to the extension E is a well defined cohomology class in $H^2(G, A)$ determined by the factor set in (30) for any choice of section μ.

If the extension E of G by A is a *split* extension (which is to say that $E = A \rtimes G$ is the semidirect product of G by A with the given conjugation action of G on A), then there is a section μ of G that is a *homomorphism* from G to E. In this case the factor set f in (30) is identically 0: $f(g, h) = 0$ for all $g, h \in G$. Hence the cohomology class in $H^2(G, A)$ defined by a split extension is the trivial class.

Suppose now that β is an equivalence between the extension in (28) and an extension E':

$$\begin{array}{ccccccccc} 1 & \longrightarrow & A & \xrightarrow{\iota} & E & \xrightarrow{\pi} & G & \longrightarrow & 1 \\ & & \Big\downarrow{\scriptstyle \mathrm{id}} & & \Big\downarrow{\scriptstyle \beta} & & \Big\downarrow{\scriptstyle \mathrm{id}} & & \\ 1 & \longrightarrow & A & \xrightarrow{\iota'} & E' & \xrightarrow{\pi'} & G & \longrightarrow & 1. \end{array}$$

If μ is a section of π, then $\mu' = \beta \circ \mu$ is a section of π', so what we have just proved can be used to determine the cohomology class in $H^2(G, A)$ corresponding to E'. Applying the homomorphism β to equation (30) gives

$$\beta(\mu(g))\beta(\mu(h)) = \beta(f(g, h))\beta(\mu(gh)) \qquad \text{for all } g, h \in G.$$

Since β restricts to the identity map on A, this is

$$\mu'(g)\mu'(h) = f(g, h)\mu'(gh) \qquad \text{for all } g, h \in G,$$

which shows that the factor set for E' associated to μ' is the same as the factor set for E associated to μ. This proves that equivalent extensions define the same cohomology class in $H^2(G, A)$.

We next show how this procedure may be reversed: Given a class in $H^2(G, A)$ we construct an extension E_f whose corresponding factor set is in the given class in $H^2(G, A)$. The process generalizes the semidirect product construction of Section 5.5 (which is the special case when f is the zero cocycle representing the trivial class).

Note first that any 2-cocycle arising from the factor set of an extension as above where the section μ is normalized satisfies the condition in (31).

Definition. A 2-cocycle f such that $f(g, 1) = 0 = f(1, g)$ for all $g \in G$ is called a *normalized* 2-cocycle.

The construction of E_f is a little simpler when f is a normalized cocycle and for simplicity we indicate the construction in this case (the minor modifications necessary when f is not normalized are indicated in Exercise 4).

We first see that any 2-cocycle f lies in the same cohomology class as a normalized 2-cocycle. Let $d_1 f_1$ be the 2-coboundary of the constant function f_1 on G whose value is $f(1, 1)$. Then $f(1, 1) = d_1 f_1(1, 1)$, and one easily checks from the 2-cocycle condition that $f - d_1 f_1$ is normalized.

We may therefore assume that our cohomology class in $H^2(G, A)$ is represented by the normalized 2-cocycle f. Let E_f be the set $A \times G$, and define a binary operation on E_f by

$$(a_1, g)(a_2, h) = (a_1 + g \cdot a_2 + f(g, h), gh) \tag{17.34}$$

where, as usual, $g \cdot a_2$ denotes the module action of G on A. It is straightforward to check that the group axioms hold: Since f is normalized, the identity element is (0,1) and inverses are given by

$$(a, g)^{-1} = (-g^{-1} \cdot a - f(g^{-1}, g), g^{-1}). \tag{17.35}$$

The cocycle condition implies the associative law by calculations similar to (32) and (33) earlier — the details are left as exercises.

Since f is a normalized 2-cocycle, $A^* = \{(a, 1) \mid a \in A\}$ is a subgroup of E_f, and the map $\iota^* : a \mapsto (a, 1)$ is an isomorphism from A to A^*. Moreover, from (34) and (35) it follows that

$$(0, g)(a, 1)(0, g)^{-1} = (g \cdot a, 1) \qquad \text{for all } g \in G \text{ and all } a \in A. \tag{17.36}$$

Since E_f is generated by A^* together with the set of elements $(0, g)$ for $g \in G$, (36) implies that A^* is a normal subgroup of E_f. Furthermore, it is immediate from (34) that the map $\pi^* : (a, g) \mapsto g$ is a surjective homomorphism from E_f to G with kernel A^*, i.e., $E_f/A^* \cong G$. Thus

$$1 \longrightarrow A \xrightarrow{\iota^*} E_f \xrightarrow{\pi^*} G \longrightarrow 1 \tag{17.37}$$

is a specific extension of G by A, where (36) ensures also that the action of G on A by conjugation in this extension is the module action specified in determining the 2-cocycle f in $H^2(G, A)$. The extension sequence (37) shows that this extension has the normalized section $\mu(g) = (0, g)$ whose corresponding normalized factor set is f. Note that this proves not only that every cohomology class in $H^2(G, A)$ arises from some extension E, but that every normalized 2-cocycle arises as the normalized factor set of some extension.

Finally, suppose f' is another normalized 2-cocycle in the same cohomology class in $H^2(G, A)$ as f and let $E_{f'}$ be the corresponding extension. If f and f' differ by the coboundary of $f_1 : G \to A$ then $f(g, h) - f'(g, h) = gf_1(h) - f_1(gh) + f_1(g)$ for all $g, h \in G$. Setting $g = h = 1$ shows that $f_1(1) = 0$. Define

$$\beta : E_f \longrightarrow E_{f'} \qquad \text{by} \qquad \beta((a, g)) = (a + f_1(g), g).$$

It is immediate that β is a bijection, and

$$\begin{aligned}\beta((a_1, g)(a_2, h)) &= \beta((a_1 + g \cdot a_2 + f(g, h), gh)) \\ &= (a_1 + g \cdot a_2 + f(g, h) + f_1(gh), gh)) \\ &= (a_1 + f_1(g) + g \cdot (a_2 + f_1(h)) + f'(g, h), gh) \\ &= (a_1 + f_1(g), g)(a_2 + f_1(h), h) = \beta((a_1, g))\beta((a_2, h))\end{aligned}$$

shows that β is an isomorphism from E_f to $E_{f'}$.

The restriction of β to A is given by $\beta((a, 1)) = (a + f_1(1), 1) = (a, 1)$, so β is the identity map on A. Similarly β is the identity map on the second component of (a, g), so β induces the identity map on the quotient G. It follows that β defines an equivalence between the extensions E_f and $E_{f'}$. This shows that the equivalence class of the extension E_f depends only on the cohomology class of f in $H^2(G, A)$.

We summarize this discussion in the following theorem.

Theorem 36. Let A be a G-module. Then

(1) A function $f : G \times G \to A$ is a normalized factor set of some extension E of G by A (with conjugation given by the G-module action on A) if and only if f is a normalized 2-cocycle in $Z^2(G, A)$.

(2) There is a bijection between the equivalence classes of extensions E as in (1) and the cohomology classes in $H^2(G, A)$. The bijection takes an extension E into the class of a normalized factor set f for E associated to any normalized section μ of G into E, and takes a cohomology class c in $H^2(G, A)$ to the extension E_f defined by the extension (37) for any normalized cocycle f in the class c.

(3) Under the bijection in (2), split extensions correspond to the trivial cohomology class.

Corollary 37. Every extension of G by the abelian group A splits if and only if $H^2(G, A) = 0$.

Corollary 38. If A is a finite abelian group and $(|A|, |G|) = 1$ then every extension of G by A splits.

Proof: This follows immediately from Corollary 29 in Section 2.

We can use Corollary 38 to prove the same result without the restriction that A be an abelian group.

Theorem 39. (*Schur's Theorem*) If E is any finite group containing a normal subgroup N whose order and index are relatively prime, then N has a complement in E.

Remark: Recall that a subgroup whose order and index are relatively prime is called a *Hall subgroup*, so Schur's Theorem says that every normal Hall subgroup has a complement that splits the group as a semidirect product.

Proof: We use induction on the order of E. Since we may assume $N \neq 1$, let p be a prime dividing $|N|$ and let P be a Sylow p-subgroup of N. Let E_0 be the normalizer in E of P and let $N_0 = N \cap E_0$. By Frattini's Argument (Proposition 6 in Section 6.1) $E = E_0 N$. It follows from the Second Isomorphism Theorem that N_0 is a (normal) Hall subgroup of E_0 and $|E_0 : N_0| = |E : N|$ (cf. Exercise 10 of Section 3.3).

If $E_0 < E$, then by induction applied to N_0 in E_0 we obtain that E_0 contains a complement K to N_0. Since $|K| = |E_0 : N_0|$, K is also a complement to N in E, as needed. Thus we may assume $E_0 = E$, i.e., P is normal in E.

Since the center of P, $Z(P)$, is characteristic in P, it is normal in E (cf. Section 4.4). If $Z(P) = N$, then N is abelian and the theorem follows from Corollary 38. Thus we may assume $Z(P) \neq N$. Let bars denote passage to the quotient group $E/Z(P)$. Then $\overline{N}$ is a normal Hall subgroup of $\overline{E}$. By induction it has a complement $\overline{K}$ in $\overline{E}$. Let E_1 be the complete preimage of $\overline{K}$ in E. Then $|E_1| = |\overline{K}||Z(P)| = |E/N||Z(P)|$, so $Z(P)$ is a normal Hall subgroup of E_1. By induction $Z(P)$ has a complement in E_1 which is seen by order considerations to also be a complement to N in E. This completes the proof.

Examples

(1) If $G = Z_2$ and $A = \mathbb{Z}/2\mathbb{Z}$ then G acts trivially on A and so $H^2(G, A) = A^G/NA = \mathbb{Z}/2\mathbb{Z}$ by the computation of the cohomology of cyclic groups in Section 2, so by Theorem 36 there are precisely two inequivalent extensions of G by A. These are the cyclic group of order 4 and the Klein 4-group, the latter being split and hence corresponding to the trivial class in H^2.

(2) If $G = \langle g \rangle \cong Z_2$ and $A = \langle a \rangle \cong \mathbb{Z}/4\mathbb{Z}$ is a group of order 4 on which G acts trivially, then $H^2(G, A) = A/2A \cong \mathbb{Z}/2\mathbb{Z}$ by the computation of the cohomology of cyclic groups. As in the previous example there are two inequivalent extensions of G by A; evidently these are the groups Z_8 and $Z_4 \times Z_2$, the latter split extension corresponding to the trivial cohomology class.

If $E = \langle r \rangle \times \langle s \rangle$ denotes the split extension of G by A, where $|r| = 4$ and $|s| = 2$, then $\mu_i(g) = r^i s$ for $i = 0, \ldots, 3$ give the four normalized sections of G in E. The sections μ_0, μ_2 both give the zero factor set f. The sections μ_1, μ_3 both give the factor set f' with $f'(g, g) = a^2 \in A$. Both f and f' give normalized 2-cocycles lying in the trivial cohomology class of $H^2(G, A)$. The extension E_f corresponding to the zero 2-cocycle f is the group with the elements $(a, 1)$ and $(1, g)$ as the usual generators (of orders 4 and 2, respectively) for $Z_4 \times Z_2$. In $E_{f'}$, however, $(a, 1)$ has order 4 but so does $(1, g)$ since $(1, g)^2 = (f'(g, g), g^2) = (a^2, 1)$. The 2-cocycles f and f' differ by the coboundary f_1 with $f_1(1) = 1$ and $f_1(g) = r$. The isomorphism $\beta(a, g) = (a + f_1(g), g)$ from E_f to $E_{f'}$ maps the generators $(a, 1)$ and $(1, g)$ of E_f to the generators $(a, 1)$ and (a, g) of $E_{f'}$ and gives the explicit equivalence of these two extensions.

The situation where G acts on A by inversion is handled in Exercise 3.

(3) Suppose $G = Z_2$ and A is the Klein 4-group. If G acts nontrivially on A then G interchanges two of the nonidentity elements, say a and b, of A and fixes the third nonidentity element c. Then $A^G = NA = \{1, c\}$ and so $H^2(G, A) = 0$, and so every extension E of G by A splits. This can be seen directly, as follows. Since the action is nontrivial, such a group must be nonabelian, hence must be D_8. From the lattice of D_8 in Section 2.5 one sees that for each Klein 4-group there is a subgroup of order 2 in D_8 not contained in the 4-group and that subgroup splits the extension.

If G acts trivially on A then $H^2(G, A) = A/2A \cong A$, so there are 4 inequivalent extensions of G by A in this case. These are considered in Exercise 1.

Example: (Groups of Order 8 and $H^2(Z_2 \times Z_2, \mathbb{Z}/2\mathbb{Z})$)

Let $G = \{1, a, b, c\}$ be the Klein 4-group and let $A = \mathbb{Z}/2\mathbb{Z}$. The 2-group G must act trivially on A. The elements of $H^2(G, A)$ classify extensions E of order 8 which has a quotient group by some Z_2 subgroup that is isomorphic to the Klein 4-group. Although there are, up to group isomorphism, only four such groups, we shall see that there are *eight* inequivalent extensions.

Since $G \times G$ has 16 elements, we have $|C^2(G, A)| = 2^{16}$. The cocycle condition (26) here reduces to

$$f(g, h) + f(gh, k) = f(h, k) + f(g, hk) \qquad \text{for all } g, h, k \in G. \tag{17.38}$$

The following relations hold for the subgroup $Z^2(G, A)$ of cocycles:
(1) $f(g, 1) = f(1, g) = f(1, 1)$, for all $g \in G$
(2) $f(g, 1) + f(g, a) + f(g, b) + f(g, c) = 0$, for all $g \in G$
(3) $f(1, h) + f(a, h) + f(b, h) + f(c, h) = 0$, for all $h \in G$.
The first of these come from (38) by setting $h = k = 1$ and by setting $g = h = 1$. The other two relations come from (38) by setting $g = h$ and $h = k$, respectively, using relations (1) and (2). It follows that every 2-cocycle f can be represented by a vector $(\alpha, \beta, \gamma, \delta, \epsilon)$ in $\mathbb{F}_2$ where

$$\alpha = f(1, g) = f(g, 1), \text{ for all } g \in G,$$
$$\beta = f(a, a), \quad \gamma = f(a, b), \quad \delta = f(b, a), \quad \epsilon = f(b, b)$$

because the relations above then determine the remaining values of f:

$$f(a, c) = \alpha + \beta + \gamma \qquad f(b, c) = \alpha + \delta + \epsilon \qquad f(c, a) = \alpha + \beta + \delta$$
$$f(c, b) = \alpha + \gamma + \epsilon \qquad f(c, c) = \alpha + \beta + \gamma + \epsilon.$$

It follows that $|Z^2(G, A)| \leq 2^5$. Although one could eventually show that every function satisfying these relations is a 2-cocycle (hence the order is exactly 32), this will follow from other considerations below.

A cocycle f is a coboundary if there is a function $f_1 : G \to A$ such that

$$f(g, h) = f_1(h) - f_1(gh) + f_1(g), \qquad \text{for all } g, h \in G.$$

This coboundary condition is easily seen to be equivalent to the conditions:
(i) $f(g, 1) = f(1, g) = f(g, g)$ for all $g \in G$, and
(ii) $f(g, h) = f(g', h')$ whenever g, h are distinct nonidentity elements and so are g', h'.

These relations are equivalent to $\alpha = \beta = \epsilon$ and $\gamma = \delta$. Thus $B^2(G, A)$ consists of the vectors $(\alpha, \alpha, \gamma, \gamma, \alpha)$, and so $H^2(G, A)$ has dimension at most 3 (i.e., order at most $2^3 = 8$). It is easy to see that $\{(0, \beta, \gamma, 0, \epsilon)\}$ with β, γ, and ϵ in $\mathbb{F}_2$ gives a set of representatives for $Z^2(G, A)/B^2(G, A)$, and each of these representative cocycles is normalized. We now prove $|H^2(G, A)| = 8$ (and also that $|Z^2(G, A)| = 2^5$) by explicitly exhibiting eight inequivalent group extensions,

Suppose E is an extension of G by A, where for simplicity we assume $A \leq E$. If $\mu : G \to E$ is a section, the factor set for E associated to μ satisfies

$$\mu(g)\mu(h) = f(g, h)\mu(gh).$$

The group E is generated by $\mu(a)$, $\mu(b)$ and A, and A is contained in the center of E since G acts trivially on A. Hence E is abelian if and only if $\mu(a)\mu(b) = \mu(b)\mu(a)$, which by the relation above occurs if and only if $f(a, b) = f(b, a)$. If g is a nonidentity element in G, we also see from the relation above that $\mu(g)$ is an element of order 2 in E if and only if $f(g, g) = 0$. Because A is contained in the center of E, both elements in any nonidentity coset $A\mu(g)$ have the same order (either 2 or 4).

There are four groups of order 8 containing a normal subgroup of order 2 with quotient group isomorphic to the Klein 4-group: $Z_2 \times Z_2 \times Z_2$, $Z_4 \times Z_2$, D_8, and Q_8.

The group $E \cong Z_2 \times Z_2 \times Z_2$ is the split extension of G by A and has $f = 0$ as factor set.

When $E \cong Q_8$, in the usual notation for the quaternion group $A = \langle -1 \rangle$. In this (non-abelian) group every nonidentity coset consists of elements of order 4, and this property is unique to Q_8, so the resulting factor set f satisfies $f(g, g) \neq 0$ for all nonidentity elements in G.

When $E \cong Z_4 \times Z_2 = \langle x \rangle \times \langle y \rangle$ we must have $A = \langle x^2 \rangle$. The cosets Ax and Axy both consist of elements of order 4, and the coset Ay consists of elements of order 2, so exactly one of $\mu(a)$, $\mu(b)$ or $\mu(c)$ is an element of order 2 and the other two must be of order 4. This suggests three homomorphisms from E to G, defined on generators by

$$\begin{array}{ll} \pi_1(y) = a & \pi_1(x) = b \\ \pi_2(y) = b & \pi_2(x) = a\,. \\ \pi_3(y) = c & \pi_3(x) = a \end{array}$$

Each of these homomorphisms maps surjectively onto G, has A as kernel, and has $\mu(a)$ (respectively, $\mu(b)$, $\mu(c)$) an element of order 2 in E. Any isomorphism of E with itself that is the identity on A must take the unique nonidentity coset Ay of A consisting of elements of order 2 to itself. Hence any extension equivalent to the extension E_1 defined by π_1 also maps y to a (since the equivalence is the identity on G). It follows that the three extensions defined by π_1, π_2 and π_3 are inequivalent.

The situation when $E \cong D_8 = \langle r, s \rangle$ is similar. In this case $A = \langle r^2 \rangle$, the cosets As and Asr consist of elements of order 2, and the coset Ar consists of elements of order 4. In this case exactly one of $\mu(a)$, $\mu(b)$ or $\mu(c)$ is an element of order 4 and the other two are of order 2, suggesting the three homomorphisms defined on generators by

$$\begin{array}{ll} \pi_1(r) = a & \pi_1(s) = b \\ \pi_2(r) = b & \pi_2(s) = a\,. \\ \pi_3(r) = c & \pi_3(s) = a \end{array}$$

As before, the corresponding extensions are inequivalent.

The existence of 8 inequivalent extensions of G by A proves that $|H^2(G, A)| = 8$, and hence that these are a complete list of all the inequivalent extensions. In particular, the extension $E_1' \cong Z_4 \times Z_2$ defined by the homomorphism π_1' mapping y to a and x to c must be equivalent to the extension E_1 above (and similarly for the other two extensions isomorphic to $Z_4 \times Z_2$ and the three extensions for D_8). This proves the existence of certain outer automorphisms for these groups, cf. Exercise 9.

Remark: For any prime p the cohomology groups of the elementary abelian group E_{p^m} with coefficients in the finite field $\mathbb{F}_p$ may be determined by relating them to the cohomology groups of the factors in the direct product as mentioned at the end of Section 2. In general, $H^2(E_{p^m}, \mathbb{F}_p)$ is a vector space over $\mathbb{F}_p$ of dimension $\frac{1}{2}m(m+1)$. When $p = 2$ and $m = 2$ this is the result $H^2(Z_2 \times Z_2, \mathbb{Z}/2\mathbb{Z}) \cong (\mathbb{Z}/2\mathbb{Z})^3$ above.

Crossed Product Algebras and the Brauer Group

Suppose F is a field. Recall that an F-algebra B is a ring containing the field F in its center and the identity of B is the identity of F, cf. Section 10.1.

Definition. An F-algebra A is said to be *simple* if A contains no nontrivial proper (two sided) ideals. A *central simple F-algebra* A is a simple F-algebra whose center is F.

Among the easiest central simple F-algebras are the matrix algebras $M_n(F)$ of $n \times n$ matrices with coefficients in F.

If K/F is a finite Galois extension of fields with Galois group $G = \text{Gal}(K/F)$, then we can use the normalized 2-cocycles in $Z^2(G, K^\times)$ to construct certain central simple K-algebras. The construction of these algebras from 2-cocycles and their classification in terms of $H^2(G, K^\times)$ (cf. Theorem 42 below) are important applications of cohomological methods in number theory. Their construction in the case when G is cyclic was one of the precursors leading to the development of abstract cohomology.

Suppose $f = \{a_{\sigma,\tau}\}_{\sigma,\tau \in G}$ is a normalized 2-cocycle in $Z^2(G, K^\times)$. Let B_f be the vector space over L having basis u_σ for $\sigma \in G$:

$$B_f = \left\{ \sum_{\sigma \in G} \alpha_\sigma u_\sigma \mid \alpha_\sigma \in K \right\}. \tag{17.39}$$

Define a multiplication on B_f by

$$u_\sigma \alpha = \sigma(\alpha) u_\sigma \qquad u_\sigma u_\tau = a_{\sigma,\tau} u_{\sigma\tau} \tag{17.40}$$

for $\alpha \in L$ and $\sigma, \tau \in G$. The second equation shows that the $a_{\sigma,\tau}$ give a "factor set" for the elements u_σ in B_f and is one reason this terminology is used. Using this multiplication we find

$$(u_\sigma u_\tau)u_\rho = a_{\sigma,\tau} a_{\sigma\tau,\rho} u_{\sigma\tau\rho} \qquad \text{and} \qquad u_\sigma (u_\tau u_\rho) = \sigma(a_{\tau,\rho}) a_{\sigma,\tau\rho} u_{\sigma\tau\rho}.$$

Since $a_{\sigma,\tau} a_{\sigma\tau,\rho} = \sigma(a_{\tau,\rho}) a_{\sigma,\tau\rho}$ is the multiplicative form of the cocycle condition (26), it follows that the multiplication defined in (40) is associative.

Since the cocycle is normalized we have $a_{1,\sigma} = a_{\sigma,1} = 1$ for all $\sigma \in G$ and it follows from (40) that the element u_1 is an identity in B_f. Identifying K with the elements αu_1 in B_f, we see that B_f is an F-algebra containing the field K and having dimension n^2 over F if $n = [K : F] = |G|$.

Proposition 40. The F-algebra B_f with K-vector space basis u_σ in (39) and multiplication defined by (40) is a central simple F-algebra.

Proof: It remains to show that the center of B_f is F and that B_f contains no nonzero proper ideals. Suppose $x = \sum_{\sigma \in G} \alpha_\sigma u_\sigma$ is an element in the center of B_f. Then $x\beta = \beta x$ for $\beta \in K$ shows that $\sigma(\beta) = \beta$ if $\alpha_\sigma \neq 0$. Since there is an element $\beta \in K$ not fixed by σ for any $\sigma \neq 1$, this shows that $a_\sigma = 0$ for all $\sigma \neq 1$, so $x = \alpha_1 u_1$. Then $xu_\tau = u_\tau x$ if and only if $\tau(\alpha_1) = \alpha_1$, so if this is true for all τ then we must have $\alpha_1 = a \in K$. Hence $x = au_1$ and the center of B_f is F.

To show that B_f is simple, suppose I is a nonzero ideal in B_f and let

$$x = \alpha_{\sigma_1} u_{\sigma_1} + \cdots + \alpha_{\sigma_m} u_{\sigma_m}$$

be a nonzero element of I with the minimal number m of nonzero terms. If $m > 1$ there is an element $\beta \in K^\times$ with $\sigma_m(\beta) \neq \sigma_{m-1}(\beta)$. Then the element $x - \sigma_m(\beta)\, x\, \beta^{-1}$ would be an element of the ideal I with the nonzero element $(1 - \sigma_m(\beta)\, \sigma_{m-1}(\beta)^{-1})\, \alpha_{\sigma_{m-1}}$ as coefficient of $u_{\sigma_{m-1}}$, and would have fewer nonzero terms than x since the coefficient of u_{σ_m} is 0. It follows that $m = 1$ and $x = \alpha\, u_\sigma$ for some $\alpha \in K$ and some σ. This element is a unit, with inverse $\sigma^{-1}(\alpha^{-1})\, u_{\sigma^{-1}}$, so $I = B_f$, completing the proof.

Definition. The central simple F-algebra B_f defined by (39) and (40) is called the *crossed product algebra* for the factor set $\{a_{\sigma,\tau}\}$.

If $f' = a'_{\sigma,\tau}$ is a normalized cocycle in the same cohomology class in $H^2(G, K^\times)$ as $a_{\sigma,\tau}$ then there are elements $b_\sigma \in K^\times$ with

$$a'_{\sigma,\tau} = a_{\sigma,\tau}(\sigma(b_\tau) b_{\sigma\tau}^{-1} b_\sigma)$$

(the multiplicative form of the coboundary condition (27)). If $B_{f'}$ is the F-algebra with K-basis v_σ defined from this cocycle as in (39) and (40), then the K-vector space homomorphism φ defined by mapping u'_σ to $b_\sigma u_\sigma$ satisfies

$$\begin{aligned}\varphi(u'_\sigma u'_\tau) &= \varphi(a'_{\sigma,\tau} u'_{\sigma\tau}) = a'_{\sigma,\tau} b_{\sigma\tau}\, u_{\sigma\tau} = b_\sigma\, \sigma(b_\tau)\, u_\sigma u_\tau \\ &= (b_\sigma u_\sigma)(b_\tau u_\tau) = \varphi(u'_\sigma)\varphi(u'_\tau).\end{aligned}$$

It follows that φ is an F-algebra isomorphism from $B_{f'}$ to B_f.

We have shown that every cohomology class c in $H^2(G, K^\times)$ defines an isomorphism class of central simple F-algebras, namely the isomorphism class of any crossed product algebra for a normalized cocycle $\{a_{\sigma,\tau}\}$ representing the class c. The next result shows that the trivial cohomology class corresponds to the isomorphism class containing $M_n(F)$.

Proposition 41. The crossed product algebra for the trivial cohomology class in $H^2(G, K^\times)$ is isomorphic to the matrix algebra $M_n(F)$ where $n = [K : F]$.

Proof: If $\alpha \in K$ then multiplication by α defines a linear transformation T_α of K viewed as an n-dimensional vector space over F. Similarly, every automorphism $\sigma \in G$ defines an F-linear transformation T_σ of K, and we may view both T_α and T_σ as elements of $M_n(F)$ by choosing a basis for K over F. If B_0 denotes the crossed product algebra for the trivial factor set ($a_{\sigma,\tau} = 1$ for all $\sigma, \tau \in G$), consider the additive map $\varphi : B_0 \to M_n(F)$ defined by $\varphi(\alpha u_\sigma) = T_\alpha T_\sigma$. Since $T_{a\alpha} = aT_\alpha$ for $a \in F$, the map φ is an F-vector space homomorphism. If $x \in K$, we have

$$T_\sigma T_\alpha(x) = T_\sigma(\alpha x) = \sigma(\alpha x) = \sigma(\alpha)\sigma(x) = T_{\sigma(\alpha)}T_\sigma,$$

so $T_\sigma T_\alpha = T_{\sigma(\alpha)}T_\sigma$ as linear transformations on K. It then follows from $u_\sigma u_\tau = u_{\sigma\tau}$ that

$$\begin{aligned}\varphi((\alpha u_\sigma)(\beta u_\tau)) = \varphi(\alpha\sigma(\beta)\,u_{\sigma\tau}) &= T_{\alpha\sigma(\beta)}T_{\sigma\tau} = T_\alpha T_{\sigma(\beta)}T_\sigma T_\tau \\ &= T_\alpha T_\sigma T_\beta T_\tau = \varphi(\alpha u_\sigma)\,\varphi(\beta u_\tau)\end{aligned}$$

which shows that φ is an F-algebra homomorphism from B_0 to $M_n(F)$. Since $\ker\varphi$ is an ideal in B_0 and $\varphi \neq 0$, it follows from Proposition 40 that $\ker\varphi = 0$ and φ is an injection. Since both B_0 and $M_n(F)$ have dimension n^2 as vector spaces over F, it follows that φ is an F-algebra isomorphism, proving the proposition.

Example

If $K = \mathbb{C}$ and $F = \mathbb{R}$, then $G = \mathrm{Gal}(\mathbb{C}/\mathbb{R})$ is of order 2 and generated by complex conjugation τ. We have $|H^2(G, \mathbb{C}^\times)| = 2$. The central simple $\mathbb{R}$-algebra B_0 corresponding to the trivial class is $\mathbb{C}u_1 \oplus \mathbb{C}u_\tau$ with $u_\tau(a+bi) = (a-bi)u_\tau$ and $u_\tau^2 = u_1$. This is isomorphic to the matrix algebra $M_2(\mathbb{R})$ under the map

$$\varphi((a+bi)u_1 + (c+di)u_\tau) = aI + bT_i + cT_\tau + dT_iT_\tau = \begin{pmatrix} a+c & -b+d \\ b+d & a-c \end{pmatrix}.$$

A normalized cocycle f representing the nontrivial cohomology class is defined by the values $a_{1,1} = a_{1,\tau} = a_{\tau,1} = 1$ and $a_{\tau,\tau} = -1$. The corresponding central simple $\mathbb{R}$-algebra B_f is given by $\mathbb{C}v_1 \oplus \mathbb{C}v_\tau$. The element v_1 is the identity of B_f, and we have the relations $v_\tau(a+bi) = (a-bi)v_\tau$ and $v_\tau^2 = -v_1$. Letting $v_1 = 1$ and $v_\tau = j$ we see that B_f is isomorphic as an $\mathbb{R}$-algebra to the real Hamilton Quaternions $\mathbb{R} + \mathbb{R}i + \mathbb{R}j + \mathbb{R}k$.

There is a rich theory of simple algebras and we mention without proof the following results. Let A be a central simple F-algebra of finite dimension over F.

I. If $F \subseteq B \subseteq A$ where B is a simple F-algebra define the *centralizer* B^c of B in A to be the elements of A that commute with all the elements of B. Define the *opposite algebra* B^{opp} to be the set B with opposite multiplication, i.e., the product b_1b_2 in B^{opp} is given by the product b_2b_1 in B. Both B^c and B^{opp} are simple F-algebras and we have

- **a.** $(\dim_F B)(\dim_F B^c) = \dim_F A$
- **b.** $A \otimes_F B^{opp} \cong M_r(B^c)$ as F-algebras, where $r = \dim_F B$
- **c.** $B \otimes_F B^c \cong A$ if B is a central simple F-algebra.

II. If A' is an Artinian (satisfies D.C.C. on left ideals) simple F-algebra, then $A \otimes_F A'$ is an Artinian simple F-algebra with center $(A')^c$.

III. We have $A \cong M_r(\Delta)$ for some division ring Δ whose center is F and some integer $r \geq 1$. The division ring Δ and r are uniquely determined by A. The same statement holds for any Artinian simple F-algebra.

The last result is part of Wedderburn's Theorem described in greater detail in the following chapter.

Definition. If A is a central simple F-algebra then a field L containing F is said to *split* A if $A \otimes_F L \cong M_m(L)$ for some $m \geq 1$.

It follows from (II) that every maximal commutative subalgebra of Δ is a field E with $E = E^c = E^{opp}$; if $[E : F] = m$ we obtain $\dim_F \Delta = m^2$. Applying (II) to $A = \Delta$ and $B = E$ we also see that $\Delta \otimes_F E \cong M_m(E)$. It can also be shown that a maximal subfield E of the central simple F-algebra A also satisfies $E = E^c = E^{opp}$ and so again by (II) it follows that $A \otimes_F E \cong M_r(E)$ ($r^2 = \dim_F A$).

If $A = M_r(\Delta)$ then the field L splits A if and only if L splits Δ, as follows. If $\Delta \otimes_F L \cong M_n(L)$ then

$$A \otimes_F L \cong M_r(\Delta) \otimes_F L \cong M_r(\Delta \otimes_F L) \cong M_r(M_n(L)) \cong M_{rn}(L).$$

Conversely if $A \otimes_F L \cong M_n(L)$ then

$$M_n(L) \cong M_r(\Delta) \otimes_F L \cong M_r(\Delta \otimes_F L).$$

By (II) and (III), $\Delta \otimes_F L \cong M_s(\Delta')$ for some division ring Δ'. Together with the previous isomorphism, the uniqueness statement in (III) shows that $\Delta' \cong L$ and then the isomorphism $\Delta \otimes_F L \cong M_s(L)$ shows that L splits Δ.

We see from the discussion above that a maximal commutative subfield of Δ splits both Δ and $A \cong M_r(\Delta)$ for any $r \geq 1$. It is not too difficult to show from this that every central simple F-algebra of finite dimension over F can be split by a finite Galois extension of F.

Applying (I) by taking A to be the crossed product algebra B_f and taking $B = K$ shows that $K = K^c = K^{opp}$ and $B_f \otimes_F K \cong M_n(K)$. In particular, the crossed product algebras B_f are always split by K.

Example

In the example of the Hamilton Quaternions above we have $B_f \otimes_{\mathbb{R}} \mathbb{C} \cong M_2(\mathbb{C})$. We have $B_f \otimes_{\mathbb{R}} \mathbb{C} = \mathbb{C} + \mathbb{C}i + \mathbb{C}j + \mathbb{C}k$ and an explicit isomorphism φ to $M_2(\mathbb{C})$ is given by

$$\varphi(i) = \begin{pmatrix} \sqrt{-1} & 0 \\ 0 & -\sqrt{-1} \end{pmatrix} \qquad \varphi(j) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

and extending $\mathbb{C}$-linearly.

By (III) every central simple F-algebra A is isomorphic as an F-algebra to $M_r(\Delta)$ for some division ring Δ uniquely determined up to F-isomorphism, called the *division ring part* of A.

Definition. Two central simple F-algebras A and B are *similar* if $A \cong M_r(\Delta)$ and $B \cong M_s(\Delta)$ for the same division ring Δ, i.e., if A and B have the same division ring parts.

Let $[A]$ denote the similarity class of A. By (II), if A and B are central simple F-algebras then $A \otimes_F B$ is again a central simple F-algebra, so we may define a multiplication on similarity classes by $[A][B] = [A \otimes_F B]$. The class $[F]$ is an identity for this multiplication and associativity of the tensor product shows that the multiplication is associative. By (Ib) applied with $B = A$ (so then $B^c = F$ since A is central) we have $[A][A^{opp}] = [F]$, so inverses exist with this multiplication.

Definition. The group of similarity classes of central simple F-algebras with multiplication $[A][B] = [A \otimes_F B]$ is called the *Brauer group* of F and is denoted $Br(F)$.

If L is any extension field of F then by (II) the algebra $A \otimes_F L$ is a central simple L-algebra. It is easy to check that the map $[A] \to [A \otimes_F L]$ is a well defined homomorphism from $Br(F)$ to $Br(L)$. The kernel of this homomorphism consists of the classes of the algebras A with $A \otimes_F L \cong M_m(L)$ for some $m \geq 1$, i.e., the algebras A that are split by L.

Definition. If L/F is a field extension then the *relative Brauer group* $Br(L/F)$ is the group of similarity classes of central simple F-algebras that are split by L. Equivalently, $Br(L/F)$ is the kernel of the homomorphism $[A] \to [A \otimes_F L]$ from $Br(F)$ to $Br(L)$.

The following theorem summarizes some major results in this area and shows the fundamental connection between Brauer groups and the crossed product algebras constructed above.

Theorem 42. Suppose K/F is a Galois extension of degree n with $G = \mathrm{Gal}(K/F)$.

(1) The central simple F-algebra A with $\dim_F A = n^2$ is split by K if and only if $A \otimes_F K \cong M_n(K)$ if and only if A is isomorphic to a crossed product algebra B_f as in (39) and (40).

(2) There is a bijection between the F-isomorphism classes of central simple F-algebras A with $A \otimes_F K \cong M_n(K)$ and the elements of $H^2(G, K^\times)$. Under this bijection the class $c \in H^2(G, K^\times)$ containing the normalized cocycle f corresponds to the isomorphism class of the crossed product algebra B_f defined in (39) and (40), and the trivial cohomology class corresponds to $M_n(F)$.

(3) Every central simple F-algebra of finite dimension over F and split by K is similar to one of dimension n^2 split by K. The bijection in (2) also establishes a bijection between $Br(K/F)$ and $H^2(G, K^\times)$ which is also an isomorphism of groups.

(4) There is a bijection between the collection of F-isomorphism classes of central simple division algebras over F that are split by K and $H^2(G, K^\times)$.

As previously mentioned, every central simple F-algebra of finite dimension over F can be split by some finite Galois extension of F, and it follows that

$$Br(F) = \bigcup_K Br(K/F)$$

where the union is over all finite Galois extensions of F. It follows that there is a bijection between $Br(F)$ and $H^2(\mathrm{Gal}(F^s/F), (F^s)^\times)$ where F^s denotes a separable algebraic closure of F. Here $\mathrm{Gal}(F^s/F)$ is considered as a profinite group and the cohomology group refers to continuous Galois cohomology.

One consequence of this result and Theorem 42 is that a full set of representatives for the F-isomorphism classes of central simple division algebras Δ over F can be obtained from the division algebra parts of the crossed product algebras for finite Galois extensions of F. Those division algebras that are split over K occur for the crossed product algebras for K/F.

Example

Since $H^2(\mathrm{Gal}(\mathbb{F}_{q^d}/\mathbb{F}_q), \mathbb{F}_{q^d}^\times) = 0$ (cf. Exercise 10), we have $Br(\mathbb{F}_{q^d}/\mathbb{F}_q) = 0$ and hence also $Br(\mathbb{F}_q) = 0$. As a consequence, every finite division algebra is a field (cf. Exercise 13 in Section 13.6 for a direct proof), and every finite central simple algebra $\mathbb{F}_q$-algebra is isomorphic to a full matrix ring $M_r(\mathbb{F}_q)$.

EXERCISES

1. Let $A = \{1, a, b, c\}$ be the Klein 4-group and let $G = \langle g \rangle$ be the cyclic group of order 2 acting trivially on A.
   - **(a)** Prove that $|C^2(G, A)| = 2^8$.
   - **(b)** Show that coboundaries are constant functions, and deduce that $|B^2(G, A)| = 4$.
   - **(c)** Use the cocycle condition to show that $|Z^2(G, A)| \leq 2^4$.
   - **(d)** If $E = Z_4 \times Z_2 = \langle x \rangle \times \langle y \rangle$, prove that the extensions $1 \to A \xrightarrow{\iota_i} E \xrightarrow{\pi} G \to 1$ defined by $\pi(x) = g$, $\pi(y) = 1$ and $\iota_1(a) = x^2$, $\iota_1(b) = y$ (respectively, $\iota_2(b) = x^2$, $\iota_2(a) = y$, and $\iota_3(c) = x^2$, $\iota_3(a) = y$), together with the split extension $Z_2 \times Z_2 \times Z_2$ give 4 inequivalent extensions of Z_2 by the Klein 4-group. Deduce that $H^2(G, A)$ has order 4 by explicitly exhibiting the corresponding cocycles.
2. Let $A = \mathbb{Z}/4\mathbb{Z}$ and let G be the cyclic group of order 2 acting trivially on A.
   - **(a)** Prove that $|C^2(G, A)| = 2^8$.
   - **(b)** Use the coboundary condition to show that $|B^2(G, A)| = 2^3$.
   - **(c)** Use the cocycle condition to show that $|Z^2(G, A)| \leq 2^4$.
   - **(d)** Show that $|H^2(G, A)| = 2$ by exhibiting two inequivalent extensions of G by A and their corresponding cocycles.
3. Let $A = \mathbb{Z}/4\mathbb{Z}$ and let G be the cyclic group of order 2 acting by inversion on A.
   - **(a)** Show that there are four coboundaries and that only the zero coboundary is normalized.
   - **(b)** Prove by a direct computation of cocycle and coboundary groups that $|H^2(G, A)| = 2$.
   - **(c)** Exhibit two distinct cohomology classes and their corresponding extension groups.
   - **(d)** Show that for a given extension of G by A with extension group isomorphic to D_8 there are four normalized sections, all of which have the zero 2-cocycle as their factor set.
   - **(e)** Show that for a given extension of G by A with extension group isomorphic to Q_8 there are sixteen sections, four of which are normalized, and all of the latter have the same factor set.
4. For a *non*-normalized 2-cocycle f one defines the extension group E_f on the set $A \times G$ by the same binary operation in equation (34). Verify two of the group axioms in this case by showing that identity is now $(-f(1, 1), 1)$ and inverses are given by
$$(a, x)^{-1} = (-x^{-1} \cdot a - f(x^{-1}, x) - f(1, 1), x^{-1}).$$
(Verification of the associative law is essentially the same as for normalized 2-cocycles.) Prove also that the set $A^{**} = \{(a - f(1, 1), 1) \mid a \in A\}$ is a subgroup of E_f and the map $\iota^{**} : a \mapsto (a - f(1, 1), 1)$ is an isomorphism from A to A^{**}. Show that this extension E_f, with the injection ι^{**} and the usual projection map π^* onto G, is equivalent to an extension derived from a normalized cocycle in the same class as f.
5. Show that the set of equivalences of a given extension $1 \to A \xrightarrow{\iota} E \xrightarrow{\pi} G \to 1$ with itself form a group under composition, and that this group is isomorphic to the stability group

$\mathrm{Stab}(1 \trianglelefteq \iota(A) \trianglelefteq E)$. (Thus Proposition 31 implies $Z^1(G, A)$ is the group of equivalences of the extension with itself).

6. *(Gaschütz's Theorem)* Let p be a prime, let A be an abelian normal p-subgroup of a finite group G, and let P be a Sylow p-subgroup of G. Prove that G is a split extension of G/A by A if and only if P is a split extension of P/A by A. (Note that $A \leq P$ by Exercise 37 in Section 4.5). [Use Sylow's Theorem to show if G splits over A then so too does P. Conversely, show that a normalized 2-cocycle associated to the extension of P/A by A via Theorem 36 is the image of a normalized 2-cocycle in $H^2(G/A, A)$ under the restriction homomorphism $\mathrm{Res} : H^2(G/A, A) \to H^2(P/A, A)$. Then use Proposition 26 and the fact that multiplication by $|G : P|$ is an automorphism of A.]

7. (a) Prove that $H^2(A_4, \mathbb{Z}/2\mathbb{Z}) \neq 0$ by exhibiting a nonsplit extension of A_4 by a cyclic group of order 2. [See Exercise 11, Section 4.5.]
   (b) Prove that $H^2(A_5, \mathbb{Z}/2\mathbb{Z}) \neq 0$ by showing that $SL_2(\mathbb{F}_5)$ is a nonsplit extension of A_5 by a cyclic group of order 2. [Use Propositions 21 and 23 in Section 4.5.]

8. The *Schur multiplier* of a finite group G is defined as the group $H^2(G, \mathbb{C}^\times)$, where the multiplicative group $\mathbb{C}^\times$ of complex numbers is a trivial G-module. Prove that the Schur multiplier is a finite group. [Show that every cohomology class contains a cocycle whose values lie in the n^{th} roots of unity, where $n = |G|$, as follows: If f is any cocycle then by Corollary 27, $f^n \in B^2(G, \mathbb{C}^\times)$. Define $k \in C^2(G, \mathbb{C}^\times)$ by $k(g_1, g_2) = f(g_1, g_2)^{1/n}$ (take any n^{th} roots). Show that $k \in B^2(G, \mathbb{C}^\times)$ and fk^{-1} takes values in the group of n^{th} roots of 1.]

9. Use the classification of the extensions of the Klein 4-group by Z_2 in the example following Theorem 39 to prove the following (in the notation of that example):
   (a) There is an (outer) automorphism of $Z_4 \times Z_2$ which interchanges the cosets Ax and Axy and fixes the coset Ay.
   (b) There is an outer automorphism of D_8 which interchanges the cosets As and Asr and fixes the coset Ar.

10. Suppose $\mathbb{F}_q$ is a finite field with $G = \mathrm{Gal}(\mathbb{F}_{q^d}/\mathbb{F}_q) = \langle \sigma_q \rangle$ where σ_q is the Frobenius automorphism, and let N be the usual norm element for the cyclic group G.
   (a) Use Hilbert's Theorem 90 to prove that $|N(\mathbb{F}_{q^d}^\times)| = (q^d - 1)/(q - 1)$, and deduce that the norm map from $\mathbb{F}_{q^d}$ to $\mathbb{F}_q$ is surjective.
   (b) Prove that $H^n(G, \mathbb{F}_{q^d}^\times) = 0$ for all $n \geq 1$.

Part VI

INTRODUCTION TO THE REPRESENTATION THEORY OF FINITE GROUPS

The final two chapters are an introduction to the representation theory of finite groups together with some applications. We have already seen in Part I how actions of groups on sets, namely permutation representations, are a fundamental tool for unravelling the structure of groups. Cayley's Theorem and Sylow's Theorem as well as many of the results and applications in Sections 6.1 and 6.2 are based on groups acting on sets. The chapter on Galois Theory developed one of the most beautiful correspondences in mathematics where the action of a group as automorphisms of a field gives rise to a correspondence between the lattice of subgroups of the Galois group and the lattice of subfields of a Galois extension of fields. In these final two chapters we study groups acting as linear transformations on vector spaces. We shall be primarily interested in utilizing these linear actions to provide information about the groups themselves.

In Part III we saw that modules are the "representation objects" for rings in the sense that the axioms for an R-module specify a "ring action" of R on some abelian group M which preserves the abelian group structure of M. In the case where M was an $F[x]$-module, x acted as a linear transformation from the vector space M to itself. In Chapter 12 the classification of finitely generated modules over Principal Ideal Domains gave us a great deal of information about these linear transformations of M (e.g., canonical forms). In Chapter 16 we used the ideal structure in Dedekind Domains to generalize the results of Chapter 12 to the classification of finitely generated modules over such domains. In this part we follow a process similar to the study of $F[x]$-modules, replacing the polynomial ring with the group ring FG of G and classifying all finitely generated FG-modules for certain fields F (Wedderburn's Theorem). We then use this classification to derive some results about finite groups such as Burnside's Theorem on the solvability of groups of order $p^a q^b$ in Chapter 19.

CHAPTER 18

Representation Theory and Character Theory

18.1 LINEAR ACTIONS AND MODULES OVER GROUP RINGS

For the remainder of the book the groups we consider will be finite groups, unless explicitly mentioned otherwise. Throughout this section F is a field and G is a finite group. We first introduce the basic terminology. Recall that if V is a vector space over F, then $GL(V)$ is the group of nonsingular linear transformations from V to itself (under composition), and if $n \in \mathbb{Z}^+$, then $GL_n(F)$ is the group of invertible $n \times n$ matrices with entries from F (under matrix multiplication).

Definition. Let G be a finite group, let F be a field and let V be a vector space over F.

(1) A *linear representation* of G is any homomorphism from G into $GL(V)$. The *degree* of the representation is the dimension of V.

(2) Let $n \in \mathbb{Z}^+$. A *matrix representation* of G is any homomorphism from G into $GL_n(F)$.

(3) A linear or matrix representation is *faithful* if it is injective.

(4) The *group ring* of G over F is the set of all formal sums of the form

$$\sum_{g\in G} \alpha_g g, \qquad \alpha_g \in F$$

with componentwise addition and multiplication $(\alpha g)(\beta h) = (\alpha\beta)(gh)$ (where α and β are multiplied in F and gh is the product in G) extended to sums via the distributive law (cf. Section 7.2).

Unless we are specifically discussing permutation representations the term "representation" will always mean "linear representation." When we wish to emphasize the field F we shall say F-representation, or representation of G on V over F.

Recall that if V is a finite dimensional vector space of dimension n, then by fixing a basis of V we obtain an isomorphism $GL(V) \cong GL_n(F)$. In this way any linear representation of G on a finite dimensional vector space gives a matrix representation and vice versa. For the most part our linear representations will be of finite degree and we shall pass freely between linear representations and matrix representations (specifying a basis when we wish to give an explicit correspondence between the two). Furthermore, given a linear representation $\varphi : G \to GL(V)$ of finite degree, a corresponding matrix representation provides numerical invariants (such as the determinant of $\varphi(g)$ for $g \in G$) which are independent of the choice of basis giving the isomorphism between $GL(V)$ and $GL_n(F)$. The exploitation of such invariants will be fundamental to our development.

Before giving examples of representations we recall the group ring FG in greater detail (group rings were introduced in Section 7.2, and some notation and examples were discussed in that section). Suppose the elements of G are $g_1, g_2, \dots, g_n$. Each element of FG is of the form

$$\sum_{i=1}^{n} \alpha_i g_i, \qquad \alpha_i \in F.$$

Two formal sums[1] are equal if and only if all corresponding coefficients of group elements are equal. Addition and multiplication in FG are defined as follows:

$$\sum_{i=1}^{n} \alpha_i g_i \quad + \quad \sum_{i=1}^{n} \beta_i g_i = \sum_{i=1}^{n} (\alpha_i + \beta_i) g_i$$

$$\Big(\sum_{i=1}^{n} \alpha_i g_i\Big)\Big(\sum_{i=1}^{n} \beta_i g_i\Big) = \sum_{k=1}^{n} \Big(\sum_{\substack{i,j \\ g_i g_j = g_k}} \alpha_i \beta_j \Big) g_k$$

where addition and multiplication of the coefficients α_i and β_j is performed in F. Note that by definition of multiplication,

FG is a commutative ring if and only if G is an abelian group.

The group G appears in FG (identifying g_i with $1g_i$) and the field F appears in FG (identifying β with βg_1, where g_1 is the identity of G). Under these identifications

$$\beta\Big(\sum_{i=1}^{n} \alpha_i g_i\Big) = \sum_{i=1}^{n} (\beta\alpha_i) g_i, \qquad \text{for all } \beta \in F.$$

In this way

FG is a vector space over F with the elements of G as a basis.

In particular, FG is a vector space over F of dimension equal to $|G|$. The elements of F commute with all elements of FG, i.e., F is in the *center* of FG. When we wish to emphasize the latter two properties we shall say that FG is an *F-algebra* (in general, an F-algebra is a ring R which contains F in its center, so R is both a ring and an F-vector space).

Note that the operations in FG are similar to those in the F-algebra $F[x]$ (although $F[x]$ is infinite dimensional over F). In some works FG is denoted by $F[G]$, although the latter notation is currently less prevalent.

[1]The formal sum displayed above is a way of writing the function from G to F which takes the value α_i on the group element g_i. This same "formality" was used in the construction of free modules (see Theorem 6 in Section 10.3).

Examples

(1) If $G = \langle g \rangle$ is cyclic of order $n \in \mathbb{Z}^+$, then the elements of FG are of the form

$$\sum_{i=0}^{n-1} \alpha_i g^i .$$

The map $F[x] \to F\langle g \rangle$ which sends x^k to g^k for all $k \geq 0$ extends by F-linearity to a surjective ring homomorphism with kernel equal to the ideal generated by $x^n - 1$. Thus

$$F\langle g \rangle \cong F[x]/(x^n - 1).$$

This is an isomorphism of F-algebras, i.e., is a ring isomorphism which is F-linear.

(2) Under the notation of the preceding example let $r = 1 + g + g^2 + \cdots + g^{n-1}$, so r is a nonzero element of $F\langle g \rangle$. Note that $rg = g + g^2 + \cdots + g^{n-1} + 1 = r$, hence $r(1 - g) = 0$. Thus the ring $F\langle g \rangle$ contains zero divisors (provided $n > 1$). More generally, if G is any group of order > 1, then for any nonidentity element $g \in G$, $F\langle g \rangle$ is a subring of FG, so FG also contains zero divisors.

(3) Let $G = S_3$ and $F = \mathbb{Q}$. The elements $r = 5(1\,2) - 7(1\,2\,3)$ and $s = -4(1\,2\,3) + 12(1\,3\,2)$ are typical members of $\mathbb{Q}S_3$. Their sum and product are seen to be

$$\begin{aligned} r + s &= 5(1\ 2) - 11(1\ 2\ 3) + 12(1\ 3\ 2) \\ rs &= -20(2\ 3) + 28(1\ 3\ 2) + 60(1\ 3) - 84 \end{aligned}$$

(recall that products (compositions) of permutations are computed from right to left). An explicit example of a sum and product of two elements in the group ring $\mathbb{Q}D_8$ appears in Section 7.2.

Before giving specific examples of representations we discuss the correspondence between representations of G and FG-modules (after which we can simultaneously give examples of both). This discussion closely parallels the treatment of $F[x]$-modules in Section 10.1.

Suppose first that $\varphi : G \to GL(V)$ is a representation of G on the vector space V over F. As above, write $G = \{g_1, \ldots, g_n\}$, so for each $i \in \{1, \ldots, n\}$, $\varphi(g_i)$ is a linear transformation from V to itself. Make V into an FG-module by defining the action of a ring element on an element of V as follows:

$$\left(\sum_{i=1}^{n} \alpha_i g_i \right) \cdot v = \sum_{i=1}^{n} \alpha_i \varphi(g_i)(v), \qquad \text{for all } \sum_{i=1}^{n} \alpha_i g_i \in FG,\ v \in V.$$

We verify a special case of axiom 2(b) of a module (see Section 10.1) which shows precisely where the fact that φ is a group homomorphism is needed:

$$\begin{aligned} (g_i g_j) \cdot v &= \varphi(g_i g_j)(v) && \text{(by definition of the action)} \\ &= (\varphi(g_i) \circ \varphi(g_j))(v) && \text{(since } \varphi \text{ is a group homomorphism)} \\ &= \varphi(g_i)(\varphi(g_j)(v)) && \text{(by definition of a composition of linear transformations)} \\ &= g_i \cdot (g_j \cdot v) && \text{(by definition of the action).} \end{aligned}$$

This argument extends by linearity to arbitrary elements of FG to prove that axiom 2(b) of a module holds in general. It is an exercise to check that the remaining module axioms hold.

Note that F is a subring of FG and the action of the field element α on a vector is the same as the action of the ring element $\alpha 1$ on a vector i.e., the FG-module action extends the F action on V.

Suppose now that conversely we are given an FG-module V. We obtain an associated vector space over F and representation of G as follows. Since V is an FG-module, it is an F-module, i.e., it is a vector space over F. Also, for each $g \in G$ we obtain a map from V to V, denoted by $\varphi(g)$, defined by

$$\varphi(g)(v) = g \cdot v \qquad \text{for all } v \in V,$$

where $g \cdot v$ is the given action of the ring element g on the element v of V. Since the elements of F commute with each $g \in G$ it follows by the axioms for a module that for all $v, w \in V$ and all $\alpha, \beta \in F$ we have

$$\begin{aligned}\varphi(g)(\alpha v + \beta w) &= g \cdot (\alpha v + \beta w)\\ &= g \cdot (\alpha v) + g \cdot (\beta w)\\ &= \alpha(g \cdot v) + \beta(g \cdot w)\\ &= \alpha\varphi(g)(v) + \beta\varphi(g)(w),\end{aligned}$$

that is, for each $g \in G$, $\varphi(g)$ is a linear transformation. Furthermore, it follows by axiom 2(b) of a module that

$$\varphi(g_i g_j)(v) = (\varphi(g_i) \circ \varphi(g_j))(v)$$

(this is essentially the calculation above with the steps reversed). This proves that φ is a group homomorphism (in particular, $\varphi(g^{-1}) = \varphi(g)^{-1}$, so every element of G maps to a nonsingular linear transformation, i.e., $\varphi : G \to GL(V)$).

This discussion shows there is a bijection between FG-modules and pairs (V, φ):

$$\left\{ V \text{ an } FG\text{-module} \right\} \longleftrightarrow \left\{ \begin{array}{c} V \text{ a vector space over } F \\ \text{and} \\ \varphi : G \to GL(V) \text{ a representation} \end{array} \right\}.$$

Giving a representation $\varphi : G \to GL(V)$ on a vector space V over F is therefore equivalent to giving an FG-module V. Under this correspondence we shall say that the module V *affords* the representation φ of G.

Recall from Section 10.1 that if a vector space M is made into an $F[x]$-module via the linear transformation T, then the $F[x]$-submodules of M are precisely the T-stable subspaces of M. In the current situation if V is an FG-module affording the representation φ, then a subspace U of V is called *G-invariant* or *G-stable* if $g \cdot u \in U$ for all $g \in G$ and all $u \in U$ (i.e., if $\varphi(g)(u) \in U$ for all $g \in G$ and all $u \in U$). It follows easily that

the FG-submodules of V are precisely the G-stable subspaces of V.

Examples

(1) Let V be a 1-dimensional vector space over F and make V into an FG-module by letting $gv = v$ for all $g \in G$ and $v \in V$. This module affords the representation $\varphi : G \to GL(V)$ defined by $\varphi(g) = I$ = the identity linear transformation, for all $g \in G$. The corresponding matrix representation (with respect to any basis of V) is the homomorphism of G into $GL_1(F)$ which sends every group element to the 1×1 identity matrix. We shall henceforth refer to this as the *trivial representation* of G. The trivial representation has degree 1 and if $|G| > 1$, it is not faithful.

(2) Let $V = FG$ and consider this ring as a left module over itself. Then V affords a representation of G of degree equal to $|G|$. If we take the elements of G as a basis of V, then each $g \in G$ permutes these basis elements under the left regular permutation representation:

$$g \cdot g_i = gg_i.$$

With respect to this basis of V the matrix of the group element g has a 1 in row i and column j if $gg_j = g_i$, and has 0's in all other positions. This (linear or matrix) representation is called the *regular representation* of G. Note that each nonidentity element of G induces a nonidentity permutation on the basis of V so the regular representation is always faithful.

(3) Let $n \in \mathbb{Z}^+$, let $G = S_n$ and let V be an n-dimensional vector space over F with basis $e_1, e_2, \dots, e_n$. Let S_n act on V by defining for each $\sigma \in S_n$

$$\sigma \cdot e_i = e_{\sigma(i)}, \qquad 1 \le i \le n$$

i.e., σ acts by permuting the subscripts of the basis elements. This provides an (injective) homomorphism of S_n into $GL(V)$ (i.e., a faithful representation of S_n of degree n), hence makes V into an FS_n-module. As in the preceding example, the matrix of σ with respect to the basis $e_1, \dots, e_n$ has a 1 in row i and column j if $\sigma \cdot e_j = e_i$ (and has 0 in all other entries). Thus σ has a 1 in row i and column j if $\sigma(j) = i$.

For an example of the ring action, consider the action of FS_3 on the 3-dimensional vector space over F with basis e_1, e_2, e_3. Let σ be the transposition $(1\,2)$, let τ be the 3-cycle $(1\,2\,3)$ and let $r = 2\sigma - 3\tau \in FS_3$. Then

$$\begin{aligned} r \cdot (\alpha e_1 + \beta e_2 + \gamma e_3) &= 2(\alpha e_{\sigma(1)} + \beta e_{\sigma(2)} + \gamma e_{\sigma(3)}) - 3(\alpha e_{\tau(1)} + \beta e_{\tau(2)} + \gamma e_{\tau(3)}) \\ &= 2(\alpha e_2 + \beta e_1 + \gamma e_3) - 3(\alpha e_2 + \beta e_3 + \gamma e_1) \\ &= (2\beta - 3\gamma)e_1 - \alpha e_2 + (2\gamma - 3\beta)e_3. \end{aligned}$$

(4) If $\psi : H \to GL(V)$ is any representation of H and $\varphi : G \to H$ is any group homomorphism, then the composition $\psi \circ \varphi$ is a representation of G. For example, let V be the FS_n-module of dimension n described in the preceding example. If $\pi : G \to S_n$ is any permutation representation of G, the composition of π with the representation above gives a linear representation of G. In other words, V becomes an FG-module under the action

$$g \cdot e_i = e_{\pi(g)(i)}, \qquad \text{for all } g \in G.$$

Note that the regular representation, (2), is just the special case of this where $n = |G|$ and π is the left regular permutation representation of G.

(5) Any homomorphism of G into the multiplicative group $F^\times = GL_1(F)$ is a degree 1 (matrix) representation. For example, suppose $G = \langle g \rangle \cong Z_n$ is the cyclic group of order n and ζ is a fixed n^{th} root of 1 in F. Let $g^i \mapsto \zeta^i$, for all $i \in \mathbb{Z}$. This representation of $\langle g \rangle$ is a faithful representation if and only if ζ is a primitive n^{th} root of 1.

(6) In many situations it is easier to specify an explicit matrix representation of a group G rather than to exhibit an FG-module. For example, recall that the dihedral group D_{2n} has the presentation

$$D_{2n} = \langle r, s \mid r^n = s^2 = 1,\ rs = sr^{-1} \rangle.$$

If R and S are any matrices satisfying the relations $R^n = S^2 = I$ and $RS = SR^{-1}$ then the map $r \mapsto R$ and $s \mapsto S$ extends uniquely to a homomorphism from D_{2n} to the matrix group generated by R and S, hence gives a representation of D_{2n}. An explicit example of matrices $R, S \in M_2(\mathbb{R})$ may be obtained as follows. If a regular n-gon is drawn on the x, y plane centered at the origin with the line $y = x$ as one of its lines of symmetry then the matrix R that rotates the plane through $2\pi/n$ radians and the matrix S that reflects the plane about the line $y = x$ both send this n-gon onto itself. It follows that these matrices act as symmetries of the n-gon and so satisfy the above relations. These matrices are readily computed (cf. Exercise 25, Section 1.6) and so the maps

$$r \mapsto R = \begin{pmatrix} \cos 2\pi/n & -\sin 2\pi/n \\ \sin 2\pi/n & \cos 2\pi/n \end{pmatrix} \qquad \text{and} \qquad s \mapsto S = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

extend uniquely to a (degree 2) representation of D_{2n} into $GL_2(\mathbb{R})$. Since the matrices R and S have orders n and 2 respectively, it follows that they generate a subgroup of $GL_2(\mathbb{R})$ of order $2n$ and hence this representation is faithful.

(7) By using the usual generators and relations for the quaternion group

$$Q_8 = \langle i, j \mid i^4 = j^4 = 1,\ i^2 = j^2,\ i^{-1}ji = j^{-1} \rangle$$

one may similarly obtain (cf. Exercise 26, Section 1.6) a representation φ from Q_8 to $GL_2(\mathbb{C})$ defined by

$$\varphi(i) = \begin{pmatrix} \sqrt{-1} & 0 \\ 0 & -\sqrt{-1} \end{pmatrix} \qquad \text{and} \qquad \varphi(j) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

This representation of Q_8 is faithful.

(8) A 4-dimensional representation of the quaternion group Q_8 may be obtained from the real Hamilton quaternions, $\mathbb{H}$ (cf. Section 7.1). The group Q_8 is a subgroup of the multiplicative group of units of $\mathbb{H}$ and each of the elements of Q_8 acts by left multiplication on the 4-dimensional real vector space $\mathbb{H}$. Since the real numbers are in the center of $\mathbb{H}$ (i.e., since $\mathbb{H}$ is an $\mathbb{R}$-algebra), left multiplication is $\mathbb{R}$-linear. This linear action thus gives a homomorphism from Q_8 into $GL_4(\mathbb{R})$. One can easily write out the explicit matrices of each of the elements of Q_8 with respect to the basis $1, i, j, k$ of $\mathbb{H}$. For example, left multiplication by i acts by $1 \mapsto i$, $i \mapsto -1$, $j \mapsto k$ and $k \mapsto -j$ and left multiplication by j acts by $1 \mapsto j$, $i \mapsto -k$, $j \mapsto -1$ and $k \mapsto i$ so

$$i \mapsto \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \qquad \text{and} \qquad j \mapsto \begin{pmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}.$$

This representation of Q_8 is also faithful.

(9) Suppose that H is a normal subgroup of the group G and suppose that H is an elementary abelian p-group for some prime p. Then $V = H$ is a vector space over $\mathbb{F}_p$, where the scalar a acts on the vector v by $av = v^a$ (see Section 10.1). The action of each element of G by conjugation on V is $\mathbb{F}_p$-linear because $gv^a g^{-1} = (gvg^{-1})^a$ and this action of G on V makes V into an $\mathbb{F}_pG$-module (the automorphisms of elementary abelian p-groups were discussed in Sections 4.4 and 10.1). The kernel of

this representation is the set of elements of G that commute with every element of H, $C_G(H)$ (which always contains the abelian group H itself). Thus the action of a group on subsets of itself often affords linear representations over finite fields. Representations of groups over finite fields are called *modular representations* and these are fundamental to the study of the internal structure of groups.

(10) For an example of an FG-submodule, let $G = S_n$ and let V be the FS_n-module described in Example 3. Let N be the subspace of V consisting of vectors all of whose coordinates are equal, i.e.,

$$N = \{\alpha_1 e_1 + \alpha_2 e_2 + \cdots + \alpha_n e_n \mid \alpha_1 = \alpha_2 = \cdots = \alpha_n\}$$

(this is a 1-dimensional S_n-stable subspace). Each $\sigma \in S_n$ fixes each vector in N so the submodule N affords the trivial representation of S_n. As an exercise, one may show that if $n \geq 3$ then N is the *unique* 1-dimensional subspace of V which is S_n-stable, i.e., N is the unique 1-dimensional FS_n-submodule (N is called the *trace* submodule of FS_n).

Another FS_n-submodule of V is the subspace I of all vectors whose coordinates sum to zero:

$$I = \{\alpha_1 e_1 + \alpha_2 e_2 + \cdots + \alpha_n e_n \mid \alpha_1 + \alpha_2 + \cdots + \alpha_n = 0\}.$$

Again I is an S_n-stable subspace (since each $\sigma \in S_n$ permutes the coordinates of each vector in V, each σ leaves the sum of the coefficients unchanged). Since I is the kernel of the linear transformation from V onto F which sends a vector to the sum of its coefficients (called the augmentation map — cf. Section 7.3), I has dimension $n - 1$.

(11) If $V = FG$ is the regular representation of G described in Example 2 above, then V has FG-submodules of dimensions 1 and $|G| - 1$ as in the preceding example:

$$N = \{\alpha_1 g_1 + \alpha_2 g_2 + \cdots + \alpha_n g_n \mid \alpha_1 = \alpha_2 = \cdots = \alpha_n\}$$
$$I = \{\alpha_1 g_1 + \alpha_2 g_2 + \cdots + \alpha_n g_n \mid \alpha_1 + \alpha_2 + \cdots + \alpha_n = 0\}.$$

In fact N and I are 2-sided ideals of FG (not just left ideals — note that N is in the center of FG). The ideal I is called the *augmentation ideal* of FG and N is called the *trace ideal* of FG.

Recall that in the study of a linear transformation T of a vector space V to itself we made V into an $F[x]$-module (where x acted as T on V); our goal was to decompose V into a direct sum of cyclic submodules. In this way we were able to find a basis of V for which the matrix of T with respect to this basis was in some *canonical* form. Changing the basis of V did not change the module V but changed the matrix representation of T by similarity (i.e., changed the isomorphism between $GL(V)$ and $GL_n(F)$). We introduce the analogous terminology to describe when two FG-modules are the same up to a change of basis.

Definition. Two representations of G are *equivalent* (or *similar*) if the FG-modules affording them are isomorphic modules. Representations which are not equivalent are called *inequivalent*.

Suppose $\varphi : G \to GL(V)$ and $\psi : G \to GL(W)$ are equivalent representations (here V and W must be vector spaces over the same field F). Let $T : V \to W$ be an FG-module isomorphism between them. Since T is, in particular, an F-module isomorphism, T is a vector space isomorphism, so V and W must have the same dimension. Furthermore, for all $g \in G$, $v \in V$ we have $T(g \cdot v) = g \cdot (T(v))$, since T is an isomorphism of FG-modules. By definition of the action of ring elements this means $T(\varphi(g)v) = \psi(g)(T(v))$, that is

$$T \circ \varphi(g) = \psi(g) \circ T \qquad \text{for all } g \in G.$$

In particular, if we identify V and W as vector spaces, then two representations φ and ψ of G on a vector space V are equivalent if and only if there is some $T \in GL(V)$ such that $T \circ \varphi(g) \circ T^{-1} = \psi(g)$ for all $g \in G$. This T is a *simultaneous* change of basis for all $\varphi(g)$, $g \in G$.

In matrix terminology, two representations φ and ψ are equivalent if there is a fixed invertible matrix P such that

$$P\varphi(g)P^{-1} = \psi(g) \qquad \text{for all } g \in G.$$

The linear transformation T or the matrix P above is said to *intertwine* the representations φ and ψ (it gives the "rule" for changing φ into ψ).

In order to study the decomposition of an FG-module into (direct sums of) submodules we shall need some terminology. We state these definitions for arbitrary rings since we shall be discussing direct sum decompositions in greater generality in the next section.

Definition. Let R be a ring and let M be a nonzero R-module.

(1) The module M is said to be *irreducible* (or *simple*) if its only submodules are 0 and M; otherwise M is called *reducible.*

(2) The module M is said to be *indecomposable* if M cannot be written as $M_1 \oplus M_2$ for any nonzero submodules M_1 and M_2; otherwise M is called *decomposable.*

(3) The module M is said to be *completely reducible* if it is a direct sum of irreducible submodules.

(4) A representation is called *irreducible, reducible, indecomposable, decomposable* or *completely reducible* according to whether the FG-module affording it has the corresponding property.

(5) If M is a completely reducible R-module, any direct summand of M is called a *constituent* of M (i.e., N is a constituent of M if there is a submodule N' of M such that $M = N \oplus N'$).

An irreducible module is, by definition, both indecomposable and completely reducible. We shall shortly give examples of indecomposable modules that are not irreducible.

If $R = FG$, an irreducible FG-module V is a nonzero F-vector space with no nontrivial, proper G-invariant subspaces. For example, if $\dim_F V = 1$ then V is necessarily irreducible (its only subspaces are 0 and V).

Suppose V is a finite dimensional FG-module and V is reducible. Let U be a G-invariant subspace. Form a basis of V by taking a basis of U and enlarging it to a basis of V. Then for each $g \in G$ the matrix, $\varphi(g)$, of g acting on V with respect to this basis is of the form

$$\varphi(g) = \begin{pmatrix} \varphi_1(g) & \psi(g) \\ 0 & \varphi_2(g) \end{pmatrix}$$

where $\varphi_1 = \varphi|_U$ (with respect to the chosen basis of U) and φ_2 is the representation of G on V/U (and ψ is not necessarily a homomorphism — $\psi(g)$ need not be a square matrix). So reducible representations are those with a corresponding matrix representation whose matrices are in block upper triangular form.

Assume further that the FG-module V is decomposable, $V = U \oplus U'$. Take for a basis of V the union of a basis of U and a basis of U'. With this choice of basis the matrix for each $g \in G$ is of the form

$$\varphi(g) = \begin{pmatrix} \varphi_1(g) & 0 \\ 0 & \varphi_2(g) \end{pmatrix}$$

(i.e., $\psi(g) = 0$ for all $g \in G$). Thus decomposable representations are those with a corresponding matrix representation whose matrices are in block diagonal form.

Examples

(1) As noted above, all degree 1 representations are irreducible, indecomposable and completely reducible. In particular, this applies to the trivial representation and to the representations described in Example 5 above.

(2) If $|G| > 1$, the regular representation of G is reducible (the augmentation ideal and the trace ideal are proper nonzero submodules). We shall later determine the conditions under which this representation is completely reducible and how it decomposes into a direct sum.

(3) For $n > 1$ the FS_n-module described in Example 10 above is reducible since N and I are proper, nonzero submodules. The module N is irreducible (being 1-dimensional) and if the characteristic of the field F does not divide n, then I is also irreducible.

(4) The degree 2 representation of the dihedral group $D_{2n} = G$ described in Example 6 above is irreducible for $n \geq 3$. There are no G-invariant 1-dimensional subspaces since a rotation by $2\pi/n$ radians sends no line in $\mathbb{R}^2$ to itself. Similarly, the degree 2 complex representation of Q_8 described in Example 7 is irreducible since the given matrix $\varphi(i)$ has exactly two 1-dimensional eigenspaces (corresponding to its distinct eigenvalues $\pm\sqrt{-1}$) and these are not invariant under the matrix $\varphi(j)$. The degree 4 representation $\varphi : Q_8 \to GL_4(\mathbb{R})$ described in Example 8 can also be shown to be irreducible (see the exercises). We shall see, however, that if we view φ as a complex representation $\varphi : Q_8 \to GL_4(\mathbb{C})$ (just by considering the real entries of the matrices to be complex entries) then there is a *complex* matrix P such that $P^{-1}\varphi(g)P$ is a direct sum of 2×2 block matrices for all $g \in Q_8$. Thus an irreducible representation over a field F may become reducible when the field is extended.

(5) Let $G = \langle g \rangle$ be cyclic of order n and assume F contains all the n^{th} roots of 1. As noted in Example 1 in the set of examples of group algebras, $F\langle g \rangle \cong F[x]/(x^n - 1)$. Thus the FG-modules are precisely the $F[x]$-modules annihilated by $x^n - 1$. The latter (finite dimensional) modules are described, up to equivalence, by the Jordan Canonical Form Theorem.

If the minimal polynomial of g acting on an $F\langle g \rangle$-module V has distinct roots in F, there is a basis of V such that g (hence all its powers) is represented by a diagonal matrix (cf. Corollary 25, Section 12.3). In this case, V is a completely reducible $F\langle g\rangle$-module (being a direct sum of 1-dimensional $\langle g\rangle$-invariant subspaces). In general, the minimal polynomial of g acting on V divides $x^n - 1$ so if $x^n - 1$ has distinct roots in F, then V is a completely reducible $F\langle g\rangle$-module. The polynomial $x^n - 1$ has distinct roots in F if and only if the characteristic of F does not divide n. This gives a sufficient condition for every $F\langle g\rangle$-module to be completely reducible.

If the minimal polynomial of g acting on V does *not* have distinct roots (so the characteristic of F does divide n), the Jordan canonical form of g must have an elementary Jordan block of size > 1. Since every linear transformation has a unique Jordan canonical form, g cannot be represented by a diagonal matrix, i.e., V is not completely reducible. It follows from results on cyclic modules in Section 12.3 that the (1-dimensional) eigenspace of g in any Jordan block of size > 1 admits no $\langle g\rangle$-invariant complement, i.e., V is reducible but not completely reducible.

Specifically, let p be a prime, let $F = \mathbb{F}_p$ and let g be of order p. Let V be the 2-dimensional space over $\mathbb{F}_p$ with basis v, w and define an action of g on V by

$$g \cdot v = v \qquad \text{and} \qquad g \cdot w = v + w.$$

This endomorphism of V does have order p (in $GL(V)$) and the matrix of g with respect to this basis is the elementary Jordan block

$$\varphi(g) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Now V is reducible (span$\{v\}$ is a $\langle g\rangle$-invariant subspace) but V is indecomposable (the above 2×2 elementary Jordan matrix is not similar to a diagonal matrix).

The first fundamental result in the representation theory of finite groups shows how Example 5 generalizes to noncyclic groups.

Theorem 1. *(Maschke's Theorem)* Let G be a finite group and let F be a field whose characteristic does not divide $|G|$. If V is any FG-module and U is any submodule of V, then V has a submodule W such that $V = U \oplus W$ (i.e., every submodule is a direct summand).

Remark: The hypothesis of Maschke's Theorem applies to any finite group when F has characteristic 0.

Proof: The idea of the proof of Maschke's Theorem is to produce an FG-module homomorphism

$$\pi : V \to U$$

which is a projection onto U, i.e., which satisfies the following two properties:

(i) $\pi(u) = u$ for all $u \in U$

(ii) $\pi(\pi(v)) = \pi(v)$ for all $v \in V$ (i.e., $\pi^2 = \pi$)

(in fact (ii) is implied by (i) and the fact that $\pi(V) \subseteq U$).

Suppose first that we can produce such an FG-module homomorphism and let $W = \ker \pi$. Since π is a module homomorphism, W is a submodule. We see that W is a direct sum complement to U as follows. If $v \in U \cap W$ then by (i), $v = \pi(v)$ whereas by definition of W, $\pi(v) = 0$. This shows $U \cap W = 0$. To show $V = U + W$ let v be an arbitrary element of V and write $v = \pi(v) + (v - \pi(v))$. By definition, $\pi(v) \in U$. By property (ii) of π,

$$\pi(v - \pi(v)) = \pi(v) - \pi(\pi(v)) = \pi(v) - \pi(v) = 0,$$

i.e., $v - \pi(v) \in W$. This shows $V = U + W$ and hence $V = U \oplus W$. To establish Maschke's Theorem it therefore suffices to find such an FG-module projection π.

Since U is a subspace it has a vector space direct sum complement W_0 in V (take a basis $\mathcal{B}_1$ of U, build it up to a basis $\mathcal{B}$ of V and let W_0 be the span of $\mathcal{B} - \mathcal{B}_1$). Thus $V = U \oplus W_0$ as vector spaces but W_0 need not be G-stable (i.e., need not be an FG-submodule). Let $\pi_0 : V \to U$ be the vector space projection of V onto U associated to this direct sum decomposition, i.e., π_0 is defined by

$$\pi_0(u + w) = u \qquad \text{for all } u \in U,\ w \in W_0.$$

The key idea of the proof is to "average" π_0 over G to form an FG-module projection π. For each $g \in G$ define

$$g\pi_0 g^{-1} : V \to U \qquad \text{by} \qquad g\pi_0 g^{-1}(v) = g \cdot \pi_0(g^{-1} \cdot v), \quad \text{for all } v \in V$$

(here $\cdot$ denotes the action of elements of the ring FG). Since π_0 maps V into U and U is stable under the action of g we have that $g\pi_0 g^{-1}$ maps V into U. Both g and g^{-1} act as F-linear transformations, so $g\pi_0 g^{-1}$ is a linear transformation. Furthermore, if u is in the G-stable space U then so is $g^{-1}u$, and by definition of π_0 we have $\pi_0(g^{-1}u) = g^{-1}u$. From this we obtain that for all $g \in G$,

$$g\pi_0 g^{-1}(u) = u \qquad \text{for all } u \in U$$

(i.e., $g\pi_0 g^{-1}$ is also a vector space projection of V onto U).

Let $n = |G|$ and view n as an element of F ($n = 1 + \cdots + 1$, n times). By hypothesis n is not zero in F and so has an inverse in F. Define

$$\pi = \frac{1}{n} \sum_{g \in G} g\pi_0 g^{-1}.$$

Since π is a scalar multiple of a sum of linear transformations from V to U, it is also a linear transformation from V to U. Furthermore, each term in the sum defining π restricts to the identity map on the subspace U and so $\pi|_U$ is $1/n$ times the sum of n copies of the identity. These observations prove the following:

$$\begin{gathered} \pi : V \to U \text{ is a linear transformation} \\ \pi(u) = u \qquad \text{for all } u \in U \\ \pi^2(v) = \pi(v) \qquad \text{for all } v \in V. \end{gathered}$$

It remains to show that π is an FG-module homomorphism (i.e., is FG-linear). It suffices to prove that for all $h \in G$, $\pi(hv) = h\pi(v)$, for $v \in V$. In this case

$$\begin{aligned}\pi(hv) &= \frac{1}{n}\sum_{g\in G} g\pi_0(g^{-1}hv)\\ &= \frac{1}{n}\sum_{g\in G} h(h^{-1}g)\pi_0((g^{-1}h)v)\\ &= \frac{1}{n}\sum_{\substack{k=h^{-1}g\\ g\in G}} h(k\pi_0(k^{-1}v) = h\pi(v)\end{aligned}$$

(as g runs over all elements of G, so does $k = h^{-1}g$ and the module element h may be brought outside the summation by the distributive law in modules). This establishes the existence of the FG-module projection π and so completes the proof.

The applications of Maschke's Theorem will be to finitely generated FG-modules. Unlike the situation of $F[x]$-modules, however, finitely generated FG-modules are automatically finite dimensional vector spaces (the difference being that FG itself is finite dimensional, whereas $F[x]$ is not). Let V be an FG-module. If V is a finite dimensional vector space over F, then a fortiori V is finitely generated as an FG-module (any F basis gives a set of generators over FG). Conversely, if V is finitely generated as an FG-module, say by $v_1, \dots, v_k$, then one easily sees that V is spanned as a vector space by the finite set $\{g \cdot v_i \mid g \in G,\ 1 \le i \le k\}$. Thus

an FG-module is finitely generated if and only if it is finite dimensional.

Corollary 2. If G is a finite group and F is a field whose characteristic does not divide $|G|$, then every finitely generated FG-module is completely reducible (equivalently, every F-representation of G of finite degree is completely reducible).

Proof: Let V be a finitely generated FG-module. As noted above, V is finite dimensional over F, so we may proceed by induction on its dimension. If V is irreducible, it is completely reducible and the result holds. Suppose therefore that V has a proper, nonzero FG-submodule U. By Maschke's Theorem U has an FG-submodule complement W, i.e., $V = U \oplus W$. By induction, each of U and W are direct sums of irreducible submodules, hence so is V. This completes the induction.

Corollary 3. Let G be a finite group, let F be a field whose characteristic does not divide $|G|$ and let $\varphi : G \to GL(V)$ be a representation of G of finite degree. Then there is a basis of V such that for each $g \in G$ the matrix of $\varphi(g)$ with respect to this basis is block diagonal:

$$\begin{pmatrix}\varphi_1(g) & & & \\ & \varphi_2(g) & & \\ & & \ddots & \\ & & & \varphi_m(g)\end{pmatrix}$$

where φ_i is an irreducible matrix representation of G, $1 \le i \le m$.

Proof: By Corollary 2 we may write $V = U_1 \oplus U_2 \oplus \cdots \oplus U_m$, where U_i is an irreducible FG-submodule of V. Let $\mathcal{B}_i$ be a basis of U_i and let $\mathcal{B}$ be the union of the $\mathcal{B}_i$'s. For each $g \in G$, the matrix of $\varphi(g)$ with respect to the basis $\mathcal{B}$ is of the form in the corollary, where $\varphi_i(g)$ is the matrix of $\varphi(g)|_{U_i}$ with respect to the basis $\mathcal{B}_i$.

The converse of Maschke's Theorem is also true. Namely, if the characteristic of F does divide $|G|$, then G possesses (finitely generated) FG-modules which are not completely reducible. Specifically, the regular representation (i.e., the module FG itself) is not completely reducible.

In Section 18.2 we shall discuss the question of uniqueness of the constituents in direct sum decompositions of FG-modules into irreducible submodules.

EXERCISES

Let F be a field, let G be a finite group and let $n \in \mathbb{Z}^+$.

1. Prove that if $\varphi : G \to GL(V)$ is any representation, then φ gives a faithful representation of $G/\ker\varphi$.

2. Let $\varphi : G \to GL_n(F)$ be a matrix representation. Prove that the map $g \mapsto \det(\varphi(g))$ is a degree 1 representation.

3. Prove that the degree 1 representations of G are in bijective correspondence with the degree 1 representations of the abelian group G/G' (where G' is the commutator subgroup of G).

4. Let V be a (possibly infinite dimensional) FG-module (G is a finite group). Prove that for each $v \in V$ there is an FG-submodule containing v of dimension $\leq |G|$.

5. Prove that if $|G| > 1$ then every irreducible FG-module has dimension $< |G|$.

6. Write out the matrices $\varphi(g)$ for every $g \in G$ for each of the following representations that were described in the second set of examples:
  (a) the representation of S_3 described in Example 3 (let $n = 3$ in that example)
  (b) the representation of D_8 described in Example 6 (i.e., let $n = 4$ in that example and write out the values of all the sines and cosines, for all group elements)
  (c) the representation of Q_8 described in Example 7
  (d) the representation of Q_8 described in Example 8.

7. Let V be the 4-dimensional permutation module for S_4 described in Example 3 of the second set of examples. Let $\pi : D_8 \to S_4$ be the permutation representation of D_8 obtained from the action of D_8 by left multiplication on the set of left cosets of its subgroup $\langle s \rangle$. Make V into an FD_8-module via π as described in Example 4 and write out the 4×4 matrices for r and s given by this representation with respect to the basis $e_1, \ldots, e_4$.

8. Let V be the FS_n-module described in Examples 3 and 10 in the second set of examples.
  (a) Prove that if v is any element of V such that $\sigma \cdot v = v$ for all $\sigma \in S_n$ then v is an F-multiple of $e_1 + e_2 + \cdots + e_n$.
  (b) Prove that if $n \geq 3$, then V has a unique 1-dimensional submodule, namely the submodule N consisting of all F-multiples of $e_1 + e_2 + \cdots + e_n$.

9. Prove that the 4-dimensional representation of Q_8 on $\mathbb{H}$ described in Example 8 in the second set of examples is irreducible. [Show that any Q_8-stable subspace is a left ideal.]

10. Prove that $GL_2(\mathbb{R})$ has no subgroup isomorphic to Q_8. [This may be done by direct computation using generators and relations for Q_8. Simplify these calculations by putting one generator in rational canonical form.]

11. Let $\varphi : S_n \to GL_n(F)$ be the matrix representation given by the permutation module described in Example 3 in the second set of examples, where the matrices are computed with respect to the basis $e_1, \ldots, e_n$. Prove that $\det \varphi(\sigma) = \epsilon(\sigma)$ for all $\sigma \in S_n$, where $\epsilon(\sigma)$ is the sign of the permutation σ. [Check this on transpositions.]

12. Assume the characteristic of F is not 2. Let H be the set of $T \in M_n(F)$ such that T has exactly one nonzero entry in each row and each column and zeros elsewhere, and the nonzero entries are ± 1. Prove that H is a subgroup of $GL_n(F)$ and that H is isomorphic to $E_{2^n} \rtimes S_n$ (semidirect product), where E_{2^n} is the elementary abelian group of order 2^n.

The next few exercises explore an important result known as Schur's Lemma and some of its consequences.

13. Let R be a ring and let M and N be simple (i.e., irreducible) R-modules.
    (a) Prove that every nonzero R-module homomorphism from M to N is an isomorphism. [Consider its kernel and image.]
    (b) Prove Schur's Lemma: if M is a simple R-module then $\mathrm{Hom}_R(M, M)$ is a division ring (recall that $\mathrm{Hom}_R(M, M)$ is the ring of all R-module homomorphisms from M to M, where multiplication in this ring is composition).

14. Let $\varphi : G \to GL(V)$ be a representation of G. The *centralizer* of φ is defined to be the set of all linear transformations, A, from V to itself such that $A\varphi(g) = \varphi(g)A$ for all $g \in G$ (i.e., the linear transformations of V which commute with all $\varphi(g)$'s).
    (a) Prove that a linear transformation A from V to V is in the centralizer of φ if and only if it is an FG-module homomorphism from V to itself (so the centralizer of φ is the same as the *ring* $\mathrm{Hom}_{FG}(V, V)$).
    (b) Show that if z is in the center of G then $\varphi(z)$ is in the centralizer of φ.
    (c) Assume φ is an irreducible representation (so V is a simple FG-module). Prove that if H is any finite *abelian* subgroup of $GL(V)$ such that $A\varphi(g) = \varphi(g)A$ for all $A \in H$ then H is cyclic (in other words, any finite abelian subgroup of the multiplicative group of units in the ring $\mathrm{Hom}_{FG}(V, V)$ is cyclic). [By the preceding exercise, $\mathrm{Hom}_{FG}(V, V)$ is a division ring, so this reduces to proving that a finite abelian subgroup of the multiplicative group of nonzero elements in a division ring is cyclic. Show that the division subring generated by an abelian subgroup of any division ring is a field and use Proposition 18, Section 9.5.]
    (d) Show that if φ is a faithful irreducible representation then the center of G is cyclic.
    (e) Deduce from (d) that if G is abelian and φ is any irreducible representation then $G/\ker\varphi$ is cyclic.

15. Exhibit all 1-dimensional complex representations of a finite cyclic group; make sure to decide which are inequivalent.

16. Exhibit all 1-dimensional complex representations of a finite abelian group. Deduce that the number of inequivalent degree 1 complex representations of a finite abelian group equals the order of the group. [First decompose the abelian group into a direct product of cyclic groups, then use the preceding exercise.]

17. Prove the following variant of Schur's Lemma for complex representations of abelian groups: if G is abelian, any irreducible complex representation, φ, of G is of degree 1 and $G/\ker\varphi$ is cyclic. [This can be done without recourse to Exercise 14 by using the observation that for any $g \in G$ the eigenspaces of $\varphi(g)$ are G-stable. Your proof that φ has degree 1 should also work for infinite abelian groups.]

18. Prove the following general form of Schur's Lemma for complex representations: if $\varphi : G \to GL_n(\mathbb{C})$ is an irreducible matrix representation and A is an $n \times n$ matrix com-

muting with $\varphi(g)$ for all $g \in G$, then A is a scalar matrix. Deduce that if φ is a faithful, irreducible, complex representation then the center of G is cyclic and $\varphi(z)$ is a scalar matrix for all elements z in the center of G. [As in the preceding exercise, the eigenspaces of A are G-stable.]

19. Prove that if G is an abelian group then any finite dimensional complex representation of G is equivalent to a representation into diagonal matrices (i.e., any finite group of commuting matrices over $\mathbb{C}$ can be simultaneously diagonalized). [This can be done without recourse to Maschke's Theorem by looking at eigenspaces.]

20. Prove that the number of degree 1 complex representations of any finite group G equals $|G : G'|$, where G' is the commutator subgroup of G. [Use Exercises 3 and 16.]

21. Let G be a noncyclic abelian group acting by conjugation on an elementary abelian p-group V, where p is a prime not dividing the order of G.

   (a) Prove that if W is an irreducible $\mathbb{F}_pG$-submodule of V then there is some nonidentity element $g \in G$ such that $W \leq C_V(g)$ (here $C_V(g)$ is the subgroup of elements of V that are fixed by g under conjugation).

   (b) Prove that V is generated by the subgroups $C_V(g)$ as g runs over all nonidentity elements of G.

22. Let p be a prime, let P be a p-group and let F be a field of characteristic p. Prove that the only irreducible representation of P over F is the trivial representation. [Do this for a group of order p first using the fact that F contains all p^{th} roots of 1 (namely 1 itself). If P is not of order p, let z be an element of order p in the center of P, prove that z is in the kernel of the irreducible representation and apply induction to $P/\langle z \rangle$.]

23. Let p be a prime, let P be a nontrivial p-group and let F be a field of characteristic p. Prove that the regular representation is not completely reducible. [Use the preceding exercise.]

24. Let p be a prime, let P be a nontrivial p-group and let F be a field of characteristic p. Prove that the regular representation is indecomposable.

18.2 WEDDERBURN'S THEOREM AND SOME CONSEQUENCES

In this section we give a famous classification theorem due to Wedderburn which describes, in particular, the structure of the group algebra FG when the characteristic of F does not divide the order of G. From this classification theorem we shall derive various consequences, including the fact that for each finite group G there are only a finite number of nonisomorphic irreducible FG-modules. This result, together with Maschke's Theorem, in some sense completes the Hölder Program for representation theory of finite groups over such fields. The remainder of the book is concerned with developing techniques for determining and working with the irreducible representations as well as applying this knowledge to obtain group-theoretic information.

Theorem 4. *(Wedderburn's Theorem)* Let R be a nonzero ring with 1 (not necessarily commutative). Then the following are equivalent:

(1) every R-module is projective

(2) every R-module is injective

(3) every R-module is completely reducible

(4) the ring R considered as a left R-module is a direct sum:

$$R = L_1 \oplus L_2 \oplus \cdots \oplus L_n,$$

where each L_i is a simple module (i.e., a simple left ideal) with $L_i = Re_i$, for some $e_i \in R$ with

(i) $e_i e_j = 0$ if $i \neq j$

(ii) $e_i^2 = e_i$ for all i

(iii) $\sum_{i=1}^{n} e_i = 1$

(5) as rings, R is isomorphic to a direct product of matrix rings over division rings, i.e., $R = R_1 \times R_2 \times \cdots \times R_r$ where R_j is a two-sided ideal of R and R_j is isomorphic to the ring of all $n_j \times n_j$ matrices with entries in a division ring Δ_j, $j = 1, 2, \ldots, r$. The integer r, the integers n_j, and the division rings Δ_j (up to isomorphism) are uniquely determined by R.

Proof: A proof of Wedderburn's Theorem is outlined in Exercises 1 to 10

Definition. A ring R satisfying any of the (equivalent) properties in Theorem 4 is called *semisimple with minimum condition.*

Rings R satisfying any of the equivalent conditions of Theorem 4 also satisfy the *minimum condition* or *descending chain condition (D.C.C)* on left ideals:

$$\text{if } I_1 \supseteq I_2 \supseteq \cdots \quad \text{is a descending chain of left ideals of } R$$
$$\text{then there is an } N \in \mathbb{Z}^+ \text{ such that } I_k = I_N \text{ for all } k \geq N$$

(which explains the use of this term in the definition above). The rings we deal with will all have this minimum condition. For example, group algebras always have this property since in any strictly descending chain of ideals the vector space dimensions of the ideals (which are F-subspaces of FG) are strictly decreasing, hence the length of a strictly descending chain is at most the dimension of FG $(= |G|)$. We shall therefore use the term "semisimple" to mean "semisimple with minimum condition." The rings R_i in conclusion (5) of Wedderburn's Theorem are called the *Wedderburn components* of R and the direct product decomposition of R is called its *Wedderburn decomposition*. Note that Wedderburn's Theorem for commutative rings is a consequence of the classification of Artinian rings in Section 16.1. A commutative semisimple ring with minimum condition is an Artinian ring with Jacobson radical equal to zero and so is a direct product of fields (which are its Wedderburn components).

One should note that condition (5) is a two-sided condition which describes the overall structure of R completely (the ring operations in this direct product of rings are componentwise addition and multiplication). In particular it implies that a semisimple ring also has the minimum condition on right ideals. A useful way of thinking of the elements of the direct product $R_1 \times \cdots \times R_r$ in conclusion (5) is as $n \times n$ (block diagonal) matrices of the form

$$\begin{pmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_r \end{pmatrix}$$

where A_i is an arbitrary $n_i \times n_i$ matrix with entries from Δ_i (here $n = \sum_{i=1}^{r} n_i$).

Recall from Section 10.5 that an R-module Q is *injective* if whenever Q is a submodule of any R-module M, then M has a submodule N such that $M = Q \oplus N$. Maschke's Theorem therefore implies:

Corollary 5. If G is a finite group and F is a field whose characteristic does not divide $|G|$, then the group algebra FG is a semisimple ring.

Before obtaining more precise information about how the invariants n, r, Δ_j, etc., relate to invariants in group rings FG for certain fields F, we first study the structure of matrix rings (i.e., the rings described in conclusions (4) and (5) of Wedderburn's Theorem). We introduce some terminology which is used extensively in ring theory. Recall that the *center* of the ring R is the subring of elements commuting with all elements in R; it will be denoted by $Z(R)$ (the center will contain 1 if the ring has a 1).

Definition.

(1) A nonzero element e in a ring R is called an *idempotent* if $e^2 = e$.
(2) Idempotents e_1 and e_2 are said to be *orthogonal* if $e_1e_2 = e_2e_1 = 0$.
(3) An idempotent e is said to be *primitive* if it cannot be written as a sum of two (commuting) orthogonal idempotents.
(4) The idempotent e is called a *primitive central idempotent* if $e \in Z(R)$ and e cannot be written as a sum of two orthogonal idempotents in the ring $Z(R)$.

Proposition 6 describes the ideal structure of a matrix ring and Proposition 8 extends these results to direct products of matrix rings.

Proposition 6. Let Δ be a division ring, let $n \in \mathbb{Z}^+$, let R be the ring of all $n \times n$ matrices with entries from Δ and let I be the identity matrix (= the 1 of R).

(1) The only two-sided ideals of R are 0 and R.
(2) The center of R consists of the scalar matrices αI, where α is in the center of Δ: $Z(R) = \{\alpha I \mid \alpha \in Z(\Delta)\}$, and this is a field isomorphic to $Z(\Delta)$. In particular, if Δ is a field, the center of R is the subring of all scalar matrices. The only central idempotent in R is I (in particular, I is primitive).
(3) Let e_i be the matrix with a 1 in position i, i and zeros elsewhere. Then $e_1, \dots, e_n$ are orthogonal primitive idempotents and $\sum_{i=1}^n e_i = I$.
(4) $L_i = Re_i$ is the left ideal consisting of arbitrary entries in column i and zeros in all other columns. L_i is a simple left R-module. Every simple left R-module is isomorphic to L_1 (in particular, all L_i are isomorphic R-modules) and as a left R-module we have $R = L_1 \oplus \cdots \oplus L_n$.

Before proving this proposition it will be useful to have the following result.

Lemma 7. Let R be an arbitrary nonzero ring.

(1) If M and N are simple R-modules and $\varphi : M \to N$ is a nonzero R-module homomorphism, then φ is an isomorphism.
(2) *(Schur's Lemma)* If M is a simple R-module, then $\operatorname{Hom}_R(M, M)$ is a division ring.

Proof of Lemma 7: To prove (1) note that since φ is nonzero, $\ker\varphi$ is a proper submodule of M. By simplicity of M we have $\ker\varphi = 0$. Similarly, the image of φ is a nonzero submodule of the simple module N, hence $\varphi(M) = N$. This proves φ is bijective, so (1) holds.

By part (1), every nonzero element of the ring $\mathrm{Hom}_R(M, M)$ is an isomorphism, hence has an inverse. This gives (2).

Proof of Proposition 6 Let A be an arbitrary matrix in R whose i, j entry is a_{ij}. Let E_{ij} be the matrix with a 1 in position i, j and zeros elsewhere. The following straightforward computations are left as exercises:

(i) $E_{ij}A$ is the matrix whose i^{th} row equals the j^{th} row of A and all other rows are zero.
(ii) AE_{ij} is the matrix whose j^{th} column equals the i^{th} column of A and all other columns are zero.
(iii) $E_{pq}AE_{rs}$ is the matrix whose p, s entry is a_{qr} and all other entries are zero.

To prove (1) suppose J is any nonzero 2-sided ideal of R and let A be an element of J with a nonzero entry in position q, r. Given any $p, s \in \{1, \dots, n\}$ we obtain from (iii) that

$$E_{ps} = \frac{1}{a_{qr}} E_{pq}AE_{rs} \in J.$$

Since the Δ-linear combinations of $\{E_{ps} \mid 1 \le p \le n,\ 1 \le s \le n\}$ give all of R, it follows that $J = R$. This proves (1).

To prove (2) assume $A \in Z(R)$. Thus for all i, j we have $E_{ij}A = AE_{ij}$. From (i) and (ii) above it follows immediately that all off-diagonal entries of A are zero and all diagonal entries of A are equal. Thus $A = \alpha I$ for some $\alpha \in \Delta$. Furthermore, A must also commute with the set of all scalar matrices βI, $\beta \in \Delta$, i.e., α must commute with all elements of Δ. Finally, since $Z(R)$ is a field, it is immediate that it contains a unique idempotent (namely I). This establishes all parts of (2).

In part (3) it is clear that $e_1, \dots, e_n$ are orthogonal idempotents whose sum is I. We defer proving that they are primitive until we have established (4).

Next we prove (4). From (ii) above it follows that $Re_i = RE_{ii}$ is the set of matrices with arbitrary entries in the i^{th} column and zeros in all other columns. Furthermore, if A is any nonzero element of Re_i, then certainly $RA \subseteq Re_i$. The reverse inclusion holds because if a_{pi} is a nonzero entry of A, then by (i) above

$$e_i = E_{ii} = \frac{1}{a_{pi}} E_{ip}A \in RA.$$

This proves $Re_i = RA$ for any nonzero element $A \in Re_i$, and so Re_i must be a simple R-module.

Let M be any simple R-module. Since $Im = m$ for all $m \in M$ and since $I = \sum_{i=1}^{n} e_i$, there exists some i and some $m \in M$ such that $e_i m \neq 0$. For this i and m the map $re_i \mapsto re_i m$ is a nonzero R-module homomorphism from the simple R-module Re_i to the simple module M. By Lemma 7(1) it is an isomorphism. By (ii), the map $r \mapsto rE_{i1}$ gives $Re_i \cong Re_1$. Finally, every matrix is the direct sum of its columns so $R = L_1 \oplus \cdots \oplus L_n$. This completes the proof of (4).

It remains to prove that the idempotents in part (3) are primitive. If $e_i = a + b$, for some orthogonal idempotents a and b, then we shall see that

$$L_i = Re_i = Ra \oplus Rb.$$

This will contradict the fact that L_i is a simple R-module. To establish the above direct sum note first that since $ab = ba = 0$, we have $ae_i = a \in Re_i$ and $be_i = b \in Re_i$. For all $r \in R$ we have $re_i = ra + rb$, hence $Re_i = Ra + Rb$. Moreover, $Ra \cap Rb = 0$ because if $ra = sb$ for some $r, s \in R$, then $ra = raa = sba = 0$ (recall $a = a^2$ and $ba = 0$). This completes all parts of the proof.

Proposition 8. Let $R = R_1 \times R_2 \times \cdots \times R_r$, where R_i is the ring of $n_i \times n_i$ matrices over the division ring Δ_i, for $i = 1, 2, \dots, r$.

(1) Identify R_i with the i^{th} component of the direct product. Let z_i be the r-tuple with the identity of R_i in position i and zero in all other positions. Then $R_i = z_i R$ and for any $a \in R_i$, $z_i a = a$ and $z_j a = 0$ for all $j \neq i$. The elements $z_1, \dots, z_r$ are all of the primitive central idempotents of R. They are pairwise orthogonal and $\sum_{i=1}^r z_i = 1$.

(2) Let N be any left R-module and let $z_i N = \{z_i x \mid x \in N\}$, $1 \leq i \leq r$. Then $z_i N$ is a left R-submodule of N, each $z_i N$ is an R_i-module on which R_j acts trivially for all $j \neq i$, and

$$N = z_1 N \oplus z_2 N \oplus \cdots \oplus z_r N.$$

(3) The simple R-modules are the simple R_i-modules on which R_j acts trivially for $j \neq i$ in the following sense. Let M_i be the unique simple R_i-module (cf. Proposition 6). We may consider M_i as an R-module by letting R_j act trivially for all $j \neq i$. Then $M_1, \dots, M_r$ are pairwise nonisomorphic simple R-modules and any simple R-module is isomorphic to one of $M_1, \dots, M_r$. Explicitly, the R-module M_i is isomorphic to the simple left ideal $(0, \dots, 0, L^{(i)}, 0, \dots, 0)$ of all elements of R whose i^{th} component, $L^{(i)}$, consists of matrices with arbitrary entries in the first column and zeros elsewhere.

(4) For any R-module N the R-submodule $z_i N$ is a direct sum of simple R-modules, each of which is isomorphic to the module M_i in (3). In particular, if M is a simple R-module, then there is a unique i such that $z_i M = M$ and for this index i we have $M \cong M_i$; for all $j \neq i$, $z_j M = 0$.

(5) If each Δ_i equals the field F, then R is a vector space over F of dimension $\sum_{i=1}^r n_i^2$ and $\dim_F Z(R) = r$.

Proof: In part (1) since multiplication in the direct product of rings is componentwise it is clear that z_i times the element $(a_1, \dots, a_r)$ of R is the r-tuple with a_i in position i and zeros elsewhere. Thus $R_i = z_i R$, z_i is the identity in R_i and $z_i a = 0$ if $a \in R_j$ for any $j \neq i$. It is also clear that $z_1, \dots, z_r$ are pairwise orthogonal central idempotents whose sum is the identity of R. The central idempotents of R are, by definition, the idempotents in $Z(R) = F_1 \times F_2 \times \cdots \times F_r$, where F_i is the center of R_i. By Proposition 6, F_i is the field $Z(\Delta_i)$. If $w = (w_1, \dots, w_r)$ is any central idempotent then $w_i \in F_i$ for all i, and since $w^2 = w$ we have $w_i^2 = w_i$ in the field F_i. Since 0 and 1 are the only solutions to $x^2 = x$ in a field, the only central idempotents in R are r-tuples whose entries are 0's and 1's. Thus $z_1, \dots, z_r$ are primitive central idempotents and since every central idempotent is a sum of these, they are the complete set of primitive central idempotents of R. This proves (1).

To prove (2) let N be any left R-module. First note that for any $z \in Z(R)$ the set $\{zx \mid x \in N\}$ is an R-submodule of N. In particular, $z_i N$ is an R-submodule. Let $z_i x \in z_i N$ and let $a \in R_j$ for some $j \neq i$. By (1) we have that $a = az_j$ and so $az_i x = (az_j)(z_i x) = az_i z_j x = 0$ because $z_i z_j = 0$. Thus the R-submodule $z_i N$ is acted on trivially by R_j for all $j \neq i$. For each $x \in N$ we have by (1) that $x = 1x = z_1 x + \cdots + z_r x$, hence $N = z_1 N + \cdots + z_r N$. Finally, this sum is direct because if, for instance, $x \in z_1 N \cap (z_2 N + \cdots + z_r N)$, then $x = z_1 x$ whereas z_1 times any element of $z_2 N + \cdots + z_r N$ is zero. This proves (2).

In part (3) first note that an R_i-module M becomes an R-module when R_j is defined to act trivially on M for all $j \neq i$. For such a module M the R-submodules are the same as the R_i-submodules. Thus M_i is a simple R-module for each i since it is a simple R_i-module.

Next, let M be a simple R-module. By (2), $M = z_1 M \oplus \cdots \oplus z_r M$. Since M has no nontrivial proper R-submodules, there must be a unique i such that $M = z_i M$ and $z_j M = 0$ for all $j \neq i$. Thus the simple R-module M is annihilated by R_j for all $j \neq i$. This implies that the R-submodules of M are the same as the R_i-submodules of M, so M is therefore a simple R_i-module. By Proposition 6, M is isomorphic as an R_i-module to M_i. Since R_j acts trivially on both M and M_i for all $j \neq i$, it follows that the R_i-module isomorphism may be viewed as an R-module isomorphism as well.

Suppose $i \neq j$ and suppose $\varphi : M_i \to M_j$ is an R-module isomorphism. If $s_i \in M_i$ then $s_i = z_i s_i$ so

$$\varphi(s_i) = \varphi(z_i s_i) = z_i \varphi(s_i) = 0,$$

since $\varphi(s_i) \in M_j$ and z_i acts trivially on M_j. This contradicts the fact that φ is an isomorphism and proves that $M_1, \dots, M_r$ are pairwise nonisomorphic simple R-modules.

Finally, the left ideal of R described in (3) is acted on trivially by R_j for all $j \neq i$ and, by Proposition 6, it is up to isomorphism the unique simple R_i-module. This left ideal is therefore a simple R-module which is isomorphic to M_i. This proves (3).

For part (4) we have already proved that if M is any simple R-module then there is a unique i such that $z_i M = M$ and $z_j M = 0$ for all $j \neq i$. Furthermore, we have shown that for this index i the simple R-module M is isomorphic to M_i. Now let N be any R-module. Then $z_i N$ is a module over R_i which is acted on trivially by R_j for all $j \neq i$. By Wedderburn's Theorem $z_i N$ is a direct sum of simple R-modules. Since each of these simple summands is acted on trivially by R_j for all $j \neq i$, each is isomorphic to M_i. This proves (4).

In part (5) if each Δ_i equals the field F, then as an F-vector space

$$R \cong M_{n_1}(F) \oplus M_{n_2}(F) \oplus \cdots \oplus M_{n_r}(F).$$

Each matrix ring $M_{n_i}(F)$ has dimension n_i^2 over F, hence R has dimension $\sum_{i=1}^r n_i^2$ over F. Furthermore, the center of each $M_{n_i}(F)$ is 1-dimensional (since by Proposition 6(2) it is isomorphic to F), hence $Z(R)$ has dimension r over F. This completes the proof of the proposition.

We now apply Wedderburn's Theorem (and the above ring-theoretic calculations) to the group algebra FG. First of all, in order to apply Wedderburn's Theorem we need the characteristic of F not to divide $|G|$. In fact, since we shall be dealing with numerical data in the sections to come it will be convenient to have the characteristic of F equal to 0. Secondly, it will simplify matters if we force all the division rings which will appear in the Wedderburn decomposition of FG to equal the field F — we shall prove that imposing the condition that F be algebraically closed is sufficient to ensure this. To simplify notation we shall therefore take $F = \mathbb{C}$ for most of the remainder of the text. The reader can easily check that any algebraically closed field of characteristic 0 (e.g., the field of all algebraic numbers) can be used throughout in place of $\mathbb{C}$.

By Corollary 5 the ring $\mathbb{C}G$ is semisimple so by Wedderburn's Theorem

$$\mathbb{C}G \cong R_1 \times R_2 \times \cdots \times R_r$$

where R_i is the ring of $n_i \times n_i$ matrices over some division ring Δ_i. Thinking of the elements of this direct product as $n \times n$ block matrices ($n = \sum_{i=1}^{r} n_i$) where the i^{th} block has entries from Δ_i, the field $\mathbb{C}$ appears in this direct product as scalar matrices and is contained in the center of $\mathbb{C}G$. Note that each Δ_i is a vector space over $\mathbb{C}$ of dimension $\leq n$. The next result shows that this implies each $\Delta_i = \mathbb{C}$.

Proposition 9. If Δ is a division ring that is a finite dimensional vector space over an algebraically closed field F and $F \subseteq Z(\Delta)$, then $\Delta = F$.

Proof: Since $F \subseteq Z(\Delta)$, for each $\alpha \in \Delta$ the division ring generated by α and F is a field. Also, since Δ is finite dimensional over F the field $F(\alpha)$ is a finite extension of F. Because F is algebraically closed it has no nontrivial finite extensions, hence $F(\alpha) = F$ for all $\alpha \in \Delta$, i.e., $\Delta = F$.

This proposition proves that each R_i in the Wedderburn decomposition of $\mathbb{C}G$ is a matrix ring over $\mathbb{C}$:

$$R_i = M_{n_i}(\mathbb{C}).$$

Now Proposition 8(5) implies that

$$\sum_{i=1}^{r} n_i^2 = |G|.$$

The final application in this section is to prove that r (= the number of Wedderburn components in $\mathbb{C}G$) equals the number of conjugacy classes of G. To see this, first note that Proposition 8(5) asserts that $r = \dim_{\mathbb{C}} Z(\mathbb{C}G)$. We compute this dimension in another way.

Let $\mathcal{K}_1, \ldots, \mathcal{K}_s$ be the distinct conjugacy classes of G (recall that these partition G). For each conjugacy class $\mathcal{K}_i$ of G let

$$X_i = \sum_{g \in \mathcal{K}_i} g \quad \in \mathbb{C}G.$$

Note that X_i and X_j have no common terms for $i \neq j$, hence they are linearly independent elements of $\mathbb{C}G$. Furthermore, since conjugation by a group element permutes the elements of each class, $h^{-1}X_ih = X_i$, i.e., X_i commutes with all group elements. This proves that $X_i \in Z(\mathbb{C}G)$.

We show the X_i's form a basis of $Z(\mathbb{C}G)$, which will prove $s = \dim_{\mathbb{C}} Z(\mathbb{C}G) = r$. Since the X_i's are linearly independent it remains to show they span $Z(\mathbb{C}G)$. Let $X = \sum_{g\in G} \alpha_g g$ be an arbitrary element of $Z(\mathbb{C}G)$. Since $h^{-1}Xh = X$,

$$\sum_{g\in G} \alpha_g h^{-1}gh = \sum_{g\in G} \alpha_g g.$$

Since the elements of G form a basis of $\mathbb{C}G$ the coefficients of g in the above two sums are equal:

$$\alpha_{hgh^{-1}} = \alpha_g.$$

Since h was arbitrary, every element in the same conjugacy class of a fixed group element g has the same coefficient in X, hence X can be written as a linear combination of the X_i's.

We summarize these results in the following theorem.

Theorem 10. Let G be a finite group.

(1) $\mathbb{C}G \cong M_{n_1}(\mathbb{C}) \times M_{n_2}(\mathbb{C}) \times \cdots \times M_{n_r}(\mathbb{C})$.

(2) $\mathbb{C}G$ has exactly r distinct isomorphism types of irreducible modules and these have complex dimensions $n_1, n_2, \dots, n_r$ (and so G has exactly r inequivalent irreducible complex representations of the corresponding degrees).

(3) $\sum_{i=1}^{r} n_i^2 = |G|$.

(4) r equals the number of conjugacy classes in G.

Corollary 11.

(1) Let A be a finite abelian group. Every irreducible complex representation of A is 1-dimensional (i.e., is a homomorphism from A into $\mathbb{C}^\times$) and A has $|A|$ inequivalent irreducible complex representations. Furthermore, every finite dimensional complex matrix representation of A is equivalent to a representation into a group of diagonal matrices.

(2) The number of inequivalent (irreducible) degree 1 complex representations of any finite group G equals $|G/G'|$.

Proof: If A is abelian, $\mathbb{C}A$ is a commutative ring. Since a $k \times k$ matrix ring is not commutative whenever $k > 1$ we must have each $n_i = 1$. Thus $r = |A|$ (= the number of conjugacy classes of A). Since every $\mathbb{C}A$-module is a direct sum of irreducible submodules, there is a basis such that the matrices are diagonal with respect to this basis. This establishes the first part of the corollary.

For a general group G, every degree 1 representation, φ, is a homomorphism of G into $\mathbb{C}^\times$. Thus φ factors through G/G'. Conversely, every degree 1 representation of G/G' gives, by composition with the natural projection $G \to G/G'$, a degree 1 representation of G. The degree 1 representations of G are therefore precisely the irreducible representations of the abelian group G/G'. Part (2) is now immediate from (1).

Examples

(1) The irreducible complex representations of a finite abelian group A (i.e., the homomorphisms from A into $\mathbb{C}^\times$) can be explicitly described as follows: decompose A into a direct product of cyclic groups

$$A \cong C_1 \times \cdots \times C_n$$

where $|C_i| = |\langle x_i \rangle| = d_i$. Map each x_i to a (not necessarily primitive) $d_i{}^{\text{th}}$ root of 1 and extend this to all powers of x_i to give a homomorphism. Since there are d_i choices for the image of each x_i, the number of distinct homomorphisms of A into $\mathbb{C}^\times = GL_1(\mathbb{C})$ defined by this process equals $|A|$. By Corollary 11, these are all the irreducible representations of A. Note that it is necessary that the field contain the appropriate roots of 1 in order to realize these representations. An exercise below explores the irreducible representations of cyclic groups over $\mathbb{Q}$.

(2) Let $G = S_3$. By Theorem 10 the number of irreducible complex representations of G is three (= the number of conjugacy classes of S_3). Since the sum of the squares of the degrees is 6, the degrees must be 1, 1 and 2. The two degree 1 representations are immediately evident: the trivial representation and the representation of S_3 into $\{\pm 1\}$ given by mapping a permutation to its sign (i.e., $\sigma \mapsto +1$ if σ is an even permutation and $\sigma \mapsto -1$ if σ is an odd permutation). The degree 2 representation can be found by decomposing the permutation representation on 3 basis vectors (described in Section 1) into irreducibles as follows: let S_3 act on the basis vectors e_1, e_2, e_3 of a vector space V by permuting their indices. The vector $t = e_1 + e_2 + e_3$ is a nonzero fixed vector, so t spans a 1-dimensional G-invariant subspace (which is a copy of the trivial representation). By Maschke's Theorem there is a 2-dimensional G-invariant complement, I. Note that the permutation representation is not a sum of degree 1 representations: otherwise it could be represented by diagonal matrices and the permutations would commute in their action — this is impossible since the representation is faithful and G is non-abelian. Thus I cannot be decomposed further, so I affords *the* irreducible 2-dimensional representation. Indeed, I is the "augmentation" submodule described in Section 1:

$$I = \{w \in V \mid w = \alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_3 \quad \text{with} \quad \alpha_1 + \alpha_2 + \alpha_3 = 0\}.$$

Clearly $e_1 - e_2$ and $e_2 - e_3$ are independent vectors in I, hence they form a basis for this 2-dimensional space. With respect to this basis of I we obtain a matrix representation of S_3 and, for example, this matrix representation on two elements of S_3 is

$$(1\ 2) \mapsto \begin{pmatrix} -1 & 1 \\ 0 & 1 \end{pmatrix} \qquad \text{and} \qquad (1\ 2\ 3) \mapsto \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix}.$$

(3) We decompose the regular representation over $\mathbb{C}$ of an arbitrary finite group. Recall that this is the representation afforded by the left $\mathbb{C}G$-module $\mathbb{C}G$ itself. By Theorem 10, $\mathbb{C}G$ is first of all a direct product of two-sided ideals:

$$\mathbb{C}G \cong M_{n_1}(\mathbb{C}) \times M_{n_2}(\mathbb{C}) \times \cdots \times M_{n_r}(\mathbb{C}).$$

Now by Proposition 6(4) each $M_{n_i}(\mathbb{C})$ decomposes further as a direct sum of n_i isomorphic simple left ideals. These left ideals give a complete set of isomorphism classes of irreducible $\mathbb{C}G$-modules. Thus the regular representation (over $\mathbb{C}$) of G decomposes as *the direct sum of all irreducible representations of G, each appearing with multiplicity equal to the degree of that irreducible representation.*

We record one additional property of $\mathbb{C}G$ which we shall prove in Section 19.2.

Theorem 12. The degree of each complex irreducible representation of a finite group G divides the order of G, i.e., in the notation of Theorem 10, each n_i divides $|G|$ for $i = 1, 2, \dots, r$.

In the next section we shall describe the primitive central idempotents of $\mathbb{C}G$ in terms of the group elements.

EXERCISES

Let G be a finite group and let R be a ring with 1.

1. Prove that conditions (1) and (2) of Wedderburn's Theorem are equivalent.
2. Prove that (3) implies (2) in Wedderburn's Theorem. [Let Q be a submodule of an R-module N. Use Zorn's Lemma to show there is a submodule M maximal with respect to $Q \cap M = 0$. If $Q + M = N$, then (2) holds; otherwise let M_1 be the complete preimage in N of some simple module in N/M not contained in $(Q + M)/M$, and argue that M_1 contradicts the maximality of M.]
3. Prove that (4) implies (3) in Wedderburn's Theorem. [Let N be a nonzero R-module. First show N contains simple submodules by considering a cyclic submodule. Then use Zorn's Lemma applied to the set of direct sums of simple submodules (appropriately ordered) to show that N contains a maximal completely reducible submodule M. If $M \neq N$ let M_1 be the complete preimage in N of a simple module in N/M and contradict the maximality of M.]
4. Prove that (5) implies (4) in Wedderburn's Theorem. [Use the methods in the proofs of Propositions 6 and 8 to decompose each R_i as a left R-module.]

The next six exercises establish some general results about rings and modules that imply the remaining implication of Wedderburn's Theorem: (2) implies (5). In these exercises assume R satisfies (2): every R-module is injective.

5. Show that R has the descending chain condition (D.C.C.) on left ideals. Deduce that R is a finite direct sum of left ideals. [If not, then show that as a left R-module R is a direct sum of an infinite number of nonzero submodules. Derive a contradiction by writing the element 1 in this direct sum.]
6. Show that $R = R_1 \times R_2 \times \cdots \times R_r$ where R_j is a 2-sided ideal and a simple ring (i.e., has no proper, nonzero 2-sided ideals). Show each R_j has an identity and satisfies D.C.C. on left ideals. [Use the preceding exercise to show R has a minimal 2-sided ideal R_1. As a left R-module $R = R_1 \oplus R'$ for some left ideal R'. Show R' is a right ideal and proceed inductively using D.C.C.]
7. Let S be a simple ring with 1 satisfying D.C.C. on left ideals and let L be a minimal left ideal in S. Show that $S \cong L^n$ as left S-modules, where $L^n = L \oplus \cdots \oplus L$ with n factors. [Argue by simplicity that $LS = S$ so $1 = l_1s_1 + \cdots + l_ns_n$ for some $l_i \in L$ and $s_i \in S$ with n minimal. Show that the map $(x_1, \dots, x_n) \mapsto x_1s_1 + \cdots + x_ns_n$ is a surjective homomorphism of left S-modules; use the minimality of L and n to show it is an injection.]
8. Let A be any ring with 1, let L be any left A-module and let L^n be the direct sum of n copies of L with itself.
   (a) Prove the ring isomorphism $\mathrm{Hom}_A(L^n, L^n) \cong M_n(D)$, where $D = \mathrm{Hom}_A(L, L)$ (multiplication in the ring $\mathrm{Hom}_A(X, X)$ is function composition, cf. Proposition 2(4) in Section 10.2).

(b) Deduce that if L is a simple A-module, then $\operatorname{Hom}_A(L^n, L^n)$ is isomorphic to a matrix ring over a division ring. [Use Schur's Lemma and (a).]

(c) Prove the ring isomorphism $\operatorname{Hom}_A(A, A) \cong A^{opp}$, where A^{opp} is the opposite ring to A (the elements and addition are the same as in A but the value of the product $x \cdot y$ in A^{opp} is yx, computed in A), cf. the end of Section 17.4. [Any homomorphism is determined by its value on 1.]

9. Prove that if S is a simple ring with 1 satisfying D.C.C. on left ideals then $S \cong M_n(\Delta)$ for some division ring Δ. (This result together with Exercise 6 completes the existence part of the proof that (2) implies (5) in Wedderburn's Theorem). [Use Exercises 7 and 8 to show $S^{opp} \cong \operatorname{Hom}_S(L^n, L^n) \cong M_n(D)$ for some division ring D. Then show $S \cong M_n(\Delta)$, where Δ is the division ring D^{opp}.]

10. Prove that Δ and n in the isomorphism $S \cong M_n(\Delta)$ of the previous exercise are uniquely determined by S (proving the uniqueness statement in Wedderburn's Theorem), as follows. Suppose $S = M_n(\Delta) \cong M_{n'}(\Delta')$ as rings, where Δ and Δ' are division rings.

(a) Prove that $\Delta \cong \operatorname{Hom}_S(L, L)$ where L is a minimal left ideal in S. Deduce that $\Delta \cong \Delta'$. [Use Proposition 6(4).]

(b) Prove that a finitely generated (left) module over a division ring Δ has a "basis" (a linearly independent generating set), and that any two bases have the same cardinality. Deduce that $n = n'$. [Mimic the proof of Corollary 4(2) of Section 11.1.]

11. Prove that if R is a ring with 1 such that every R-module is free then R is a division ring.

12. Let F be a field, let $f(x) \in F[x]$ and let $R = F[x]/(f(x))$. Find necessary and sufficient conditions on the factorization of $f(x)$ in $F[x]$ so that R is a semisimple ring. When R is semisimple, describe its Wedderburn decomposition. [See Proposition 16 in Section 9.5.]

13. Let G be the cyclic group of order n and let $R = \mathbb{Q}G$. Describe the Wedderburn decomposition of R and find the number and the degrees of the irreducible representations of G over $\mathbb{Q}$. In particular, show that if $n = p$ is a prime then G has exactly one nontrivial irreducible representation over $\mathbb{Q}$ and this representation has degree $p - 1$. [Recall from the first example in Section 1 that $\mathbb{Q}G = \mathbb{Q}[x]/(x^n - 1)$. Use Proposition 16 in Section 9.5 and results from Section 13.6.]

14. Let p be a prime and let $F = \mathbb{F}_p$ be the field of order p. Let G be the cyclic group of order 3 and let $R = FG$. For each of $p = 2$ and $p = 7$ describe the Wedderburn decomposition of R and find the number and the degrees of the irreducible representations of G over F.

15. Prove that if P is a p-group for some prime p, then P has a faithful irreducible complex representation if and only if $Z(P)$ is cyclic. [Use Exercise 18 in Section 1, Theorem 6.1(2) and Example 3.]

16. Prove that if V is an irreducible FG-module and F is an algebraically closed field then $\operatorname{Hom}_{FG}(V, V)$ is isomorphic to F (as a ring).

17. Let F be a field, let $R = M_n(F)$ and let M be the unique irreducible R-module. Prove that $\operatorname{Hom}_R(M, M)$ is isomorphic to F (as a ring).

18. Find all 2-sided ideals of $M_n(\mathbb{Z})$.

18.3 CHARACTER THEORY AND THE ORTHOGONALITY RELATIONS

In general, for groups of large order the representations are difficult to compute and unwieldy if not impossible to write down. For example, a matrix representation of degree 100 involves matrices with 10,000 entries, and a number of 100×100 matrices may be required to describe the representation, even on a set of generators for the group. There are, however, some striking examples where large degree representations have been computed and used effectively. One instance of this is a construction of the simple group J_1 by Z. Janko in 1965 (the existence problem for simple groups was discussed at the end of Section 6.2). Janko was investigating certain properties of simple groups and he found that if any simple group possessed these properties, then it would necessarily have order 175,560 and would be generated by two elements. Furthermore, he proved that a hypothetical simple group with these properties must have a 7-dimensional representation over the field $\mathbb{F}_{11}$ with two generators mapping to the two matrices

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} -3 & 2 & -1 & -1 & -3 & -1 & -3 \\ -2 & 1 & 1 & 3 & 1 & 3 & 3 \\ -1 & -1 & -3 & -1 & -3 & -3 & 2 \\ -1 & -3 & -1 & -3 & -3 & 2 & -1 \\ -3 & -1 & -3 & -3 & 2 & -1 & -1 \\ 1 & 3 & 3 & -2 & 1 & 1 & 3 \\ 3 & 3 & -2 & 1 & 1 & 3 & 1 \end{pmatrix}$$

(note that for any simple group S, every representation of S into $GL_n(F)$ which does not map all group elements to the identity matrix is a faithful representation, so S is isomorphic to its image in $GL_n(F)$). In particular, Janko's calculations showed that the simple group satisfying his properties was unique, if it existed. M. Ward was able to show that these two matrices do generate a subgroup of $GL_7(\mathbb{F}_{11})$ of order 175,560 and it follows that there does exist a simple group satisfying Janko's properties.

In a similar vein, S. Norton, R. Parker and J. Thackray constructed the simple group J_4 of order 86,775,571,046,077,562,880 using a 112-dimensional representation over $\mathbb{F}_2$. This group was shown to be generated by two elements, and explicit matrices in $GL_{112}(\mathbb{F}_2)$ for these two generators were computed in the course of their analysis.

In 1981, R. Griess constructed the largest of the sporadic groups, the so called *Monster*, of order

$$2^{46} \cdot 3^{20} \cdot 5^9 \cdot 7^6 \cdot 11^2 \cdot 13^3 \cdot 17 \cdot 19 \cdot 23 \cdot 29 \cdot 31 \cdot 41 \cdot 47 \cdot 59 \cdot 71.$$

His proof involves calculations of automorphisms of an algebra over $\mathbb{C}$ of dimension 196,884 and leads to a construction of the Monster by means of a representation of this degree.

By analogy, in general it is difficult to write out the explicit permutations associated to a permutation representation $\varphi : G \to S_n$ for large degrees n. There are, however, numerical invariants such as the signs and the cycle types of the permutations $\pi(g)$ and these numerical invariants might be easier to compute than the permutations themselves (i.e., it may be possible to determine the cycle types of elements without actually having to write out the permutations themselves, as in the computation of Galois groups over $\mathbb{Q}$ in Section 14.8). These invariants alone may provide enough information in a given situation to carry out some analysis, such as prove that a given group is not simple (as illustrated in Section 6.2). Furthermore, the invariants just mentioned do not depend on the labelling of the set $\{1, 2, \ldots, n\}$ (i.e., they are independent of a "change of basis" in S_n) and they are the same for elements that are conjugate in G.

In this section we show how to attach numerical invariants to linear representations. These invariants depend only on the equivalence class (isomorphism type) of the representation. In other words, for each representation $\varphi : G \to GL_n(F)$ we shall attach an element of F to each matrix $\varphi(g)$ and we shall see that this number can, in many instances, be computed without knowing the matrix $\varphi(g)$. Moreover, we shall see that these invariants are independent of the similarity class of φ (i.e., are the same for a fixed $g \in G$ if the representation φ is replaced by an equivalent representation) and that they, in some sense, characterize the similarity classes of representations of G.

Throughout this section G is a finite group and, for the moment, F is an arbitrary field. All representations considered are assumed to be finite dimensional.

Definition.

(1) A *class function* is any function from G into F which is constant on the conjugacy classes of G, i.e., $f : G \to F$ such that $f(g^{-1}xg) = f(x)$ for all $g, x \in G$.

(2) If φ is a representation of G afforded by the FG-module V, the *character* of φ is the function

$$\chi : G \to F \qquad \text{defined by} \qquad \chi(g) = \operatorname{tr} \varphi(g),$$

where $\operatorname{tr} \varphi(g)$ is the trace of the matrix of $\varphi(g)$ with respect to some basis of V (i.e., the sum of the diagonal entries of that matrix). The character is called *irreducible* or *reducible* according to whether the representation is irreducible or reducible, respectively. The *degree* of a character is the degree of any representation affording it.

In the notation of the second part of this definition we shall also refer to χ as the character afforded by the FG-module V. In general, a character is *not* a homomorphism from a group into either the additive or multiplicative group of the field.

Examples

(1) The character of the trivial representation is the function $\chi(g) = 1$ for all $g \in G$. This character is called the *principal* character of G.

(2) For degree 1 representations, the character and the representation are usually identified (by identifying a 1×1 matrix with its entry). Thus for abelian groups, irreducible complex representations and their characters are the same (cf. Corollary 11).

(3) Let $\Pi : G \to S_n$ be a permutation representation and let φ be the resulting linear representation on the basis $e_1, \dots, e_n$ of the vector space V:

$$\varphi(g)(e_i) = e_{\Pi(g)(i)}$$

(cf. Example 4 of Section 1). With respect to this basis the matrix of $\varphi(g)$ has a 1 in the diagonal entry i, i if $\Pi(g)$ fixes i; otherwise, the matrix of $\varphi(g)$ has a zero in position i, i. Thus if π is the character of φ then

$$\pi(g) = \text{the number of fixed points of } g \text{ on } \{1, 2, \dots, n\}.$$

In particular, if Π is the permutation representation obtained from left multiplication on the set of left cosets of some subgroup H of G then the resulting character is called the *permutation character* of G on H.

(4) The special case of Example 3 when Π is the regular permutation representation of G is worth recording: if φ is the regular representation of G (afforded by the module FG) and ρ is its character:

$$\rho(g) = \begin{cases} 0 & \text{if } g \neq 1 \\ |G| & \text{if } g = 1. \end{cases}$$

The character of the regular representation of G is called the *regular character* of G. Note that this provides specific examples where a character takes on the value 0 and is not a group homomorphism from G into either F or $F^{\times}$.

(5) Let $\varphi : D_{2n} \to GL_2(\mathbb{R})$ be the explicit matrix representation described in Example 6 in the second set of examples of Section 1. If χ is the character of φ then, by taking traces of the given 2×2 matrices one sees that $\chi(r) = 2\cos(2\pi/n)$ and $\chi(s) = 0$. Since φ takes the identity of D_{2n} to the 2×2 identity matrix, $\chi(1) = 2$.

(6) Let $\varphi : Q_8 \to GL_2(\mathbb{C})$ be the explicit matrix representation described in Example 7 in the second set of examples of Section 1. If χ is the character of φ then, by taking traces of the given 2×2 matrices, $\chi(i) = 0$ and $\chi(j) = 0$. Since the element $-1 \in Q_8$ maps to minus the 2×2 identity matrix, $\chi(-1) = -2$. Since φ takes the identity of Q_8 to the 2×2 identity matrix, $\chi(1) = 2$.

(7) Let $\varphi : Q_8 \to GL_4(\mathbb{R})$ be the matrix representation described in Example 8 in the second set of examples of Section 1. If χ is the character of φ then, by inspection of the matrices exhibited, $\chi(i) = \chi(j) = 0$. Since φ takes the identity of Q_8 to the 4×4 identity matrix, $\chi(1) = 4$.

For $n \times n$ matrices A and B, direct computation shows that $\operatorname{tr} AB = \operatorname{tr} BA$. If A is invertible, this implies that

$$\operatorname{tr} A^{-1}BA = \operatorname{tr} B.$$

Thus the character of a representation is independent of the choice of basis of the vector space affording it, i.e.,

equivalent representations have the same character. (18.1)

Let φ be a representation of G of degree n with character χ. Since $\varphi(g^{-1}xg)$ is $\varphi(g)^{-1}\varphi(x)\varphi(g)$ for all $g, x \in G$, taking traces shows that

the character of a representation is a class function. (18.2)

Since the trace of the $n \times n$ identity matrix is n and φ takes the identity of G to the identity linear transformation (or matrix),

$\chi(1)$ *is the degree of* φ. (18.3)

If V is an FG-module whose corresponding representation has character χ, then each element of the group ring FG acts as a linear transformation from V to V. Thus each $\sum_{g\in G} \alpha_g g \in FG$ has a trace when it is considered as a linear transformation from V to V. The trace of $g \in G$ acting on V is, by definition, $\chi(g)$. Since the trace of any linear combination of matrices is the linear combination of the traces, the trace of $\sum_{g\in G} \alpha_g g$ acting on V is $\sum_{g\in G} \alpha_g \chi(g)$. Note that this trace function on FG is the unique extension of the character χ of G to an F-linear transformation from FG to F. In this way we shall consider characters of G as also being defined on the group ring FG.

Notice in Example 3 above that if the field F has characteristic $p > 0$, the values of the character mod p might be zero even though the number of fixed points is nonzero. In order to circumvent such anomalies and to use the consequences of Wedderburn's Theorem obtained when F is algebraically closed we again specialize the field to be the complex numbers (or any algebraically closed field of characteristic 0). By the results of the previous section

$$\mathbb{C}G \cong M_{n_1}(\mathbb{C}) \times M_{n_2}(\mathbb{C}) \times \cdots \times M_{n_r}(\mathbb{C}). \tag{18.4}$$

For the remainder of this section fix the following notation:

$$\begin{gathered} M_1, M_2, \dots, M_r \text{ are the inequivalent irreducible } \mathbb{C}G\text{-modules,} \\ \chi_i \text{ is the character afforded by } M_i, \quad 1 \le i \le r. \end{gathered} \tag{18.5}$$

Thus r is the number of conjugacy classes of G and we may relabel $M_1, \dots, M_r$ if necessary so that the degree of χ_i is n_i for all i (which is also the dimension of M_i over $\mathbb{C}$).

Now every (finite dimensional) $\mathbb{C}G$-module M is isomorphic (equivalent) to a direct sum of irreducible modules:

$$M \cong a_1 M_1 \oplus a_2 M_2 \oplus \cdots \oplus a_r M_r, \tag{18.6}$$

where a_i is a nonnegative integer indicating the multiplicity of the irreducible module M_i in this direct sum decomposition, i.e.,

$$a_i M_i = \overbrace{M_i \oplus \cdots \oplus M_i}^{a_i \text{ times}}.$$

Note that if the representation φ is afforded by the module M and $M = M_1 \oplus M_2$, then we may choose a basis of M consisting of a basis of M_1 together with a basis of M_2. The matrix representation with respect to this basis is of the form

$$\varphi(g) = \begin{pmatrix} \varphi_1(g) & 0 \\ 0 & \varphi_2(g) \end{pmatrix}$$

where φ_i is the representation afforded by M_i, $i = 1, 2$. One sees immediately that if ψ is the character of φ and ψ_i is the character of φ_i, then $\psi(g) = \psi_1(g) + \psi_2(g)$, i.e., $\psi = \psi_1 + \psi_2$. By induction we obtain:

the character of a representation is the sum of the characters of the constituents appearing in a direct sum decomposition. (18.7)

If ψ is the character afforded by the module M in (6) above, this gives

$$\psi = a_1\chi_1 + a_2\chi_2 + \cdots + a_r\chi_r. \tag{18.8}$$

Thus every (complex) character is a nonnegative integral sum of irreducible (complex) characters. Conversely, by taking direct sums of modules one sees that every such sum of characters is the character of some complex representation of G.

We next prove that the correspondence between characters and equivalence classes of complex representations is *bijective*. Let $z_1, z_2, \dots, z_r$ be the primitive central idempotents of $\mathbb{C}G$ described in the preceding section. Since these are orthogonal (or equivalently, since they are the r-tuples in the decomposition of $\mathbb{C}G$ into a direct product of r

subrings which have a 1 in one position and zeros elsewhere), $z_1, \ldots, z_r$ are $\mathbb{C}$-linearly independent elements of $\mathbb{C}G$. As above, each irreducible character χ_i is a function on $\mathbb{C}G$. By Proposition 8(3) we have

(a) if $j \neq i$ then $z_j M_i = 0$, i.e., z_j acts as the zero matrix on M_j, hence $\chi_j(z_i) = 0$, and

(b) z_i acts as the identity on M_i, hence $\chi_i(z_i) = n_i$.

Thus $\chi_1, \ldots, \chi_r$ are multiples of the dual basis to the independent set $z_1, \ldots, z_r$, hence are linearly independent functions. Now if the $\mathbb{C}G$-module M described in (6) above can be decomposed in a different fashion into irreducibles, say,

$$M \cong b_1 M_1 \oplus b_2 M_2 \oplus \cdots \oplus b_r M_r,$$

then we would obtain a relation

$$a_1\chi_1 + a_2\chi_2 + \cdots + a_r\chi_r = b_1\chi_1 + b_2\chi_2 + \cdots + b_r\chi_r.$$

By linear independence of the irreducible characters, $b_i = a_i$ for all $i \in \{1, \ldots, r\}$. Thus, in any decomposition of M into a direct sum of irreducibles, the multiplicity of the irreducible M_i is the same, $1 \leq i \leq r$. In particular,

two representations are equivalent if and only if they have the same character. (18.9)

This uniqueness can be seen in an alternate way. First, use Proposition 8(2) to decompose an arbitrary finite dimensional $\mathbb{C}G$-module M uniquely as

$$M = z_1 M \oplus z_2 M \oplus \cdots \oplus z_r M.$$

By part (4) of the same proposition, $z_i M$ is a direct sum of simple modules, each of which is isomorphic to M_i. The multiplicity of M_i in a direct sum decomposition of $z_i M$ is, by counting dimensions, equal to $\dfrac{\dim z_i M}{\dim M_i}$. This proves that the multiplicity of M_i in any direct sum decomposition of M into simple submodules is uniquely determined.

Note that, as with decompositions of $F[x]$-modules into cyclic submodules, a $\mathbb{C}G$-module may have many direct sum decompositions into irreducibles — only the multiplicities are unique (see also the exercises). More precisely, comparing with the Jordan canonical form of a single linear transformation, the direct summand $a_i M_i = M_i \oplus \cdots \oplus M_i$ (a_i times) which equals the submodule $z_i M$ is the analogue of the generalized eigenspace corresponding to a single eigenvalue. This submodule of M is unique (as is a generalized eigenspace) and is called the χ_i^{th} *isotypic component* of M. Within the χ_i^{th} isotypic component, the summands M_i are analogous to the 1-dimensional eigenspaces and, just as with the eigenspace of an endomorphism there is no unique basis for the eigenspace. If $G = \langle g \rangle$ is a finite cyclic group, the isotypic components of G are the same as the generalized eigenspaces of g.

Observe that the vector space of all (complex valued) class functions on G has a basis consisting of the functions which are 1 on a given class and zero on all other classes. There are r of these, where r is the number of conjugacy classes of G, so the dimension of the complex vector space of class functions is r. Since the number of (complex) irreducible characters of G equals the number of conjugacy classes and these are linearly independent class functions, we see that

the irreducible characters are a basis for the space of all complex class functions. (18.10)

The next step in the theory of characters is to put an Hermitian inner product structure on the space of class functions and prove that the irreducible characters form an orthonormal basis with respect to this inner product. For class functions θ and ψ define

$$(\theta, \psi) = \frac{1}{|G|} \sum_{g \in G} \theta(g)\overline{\psi(g)}$$

(where the bar denotes complex conjugation). One easily checks that $(\ ,\ )$ is Hermitian: for $\alpha, \beta \in \mathbb{C}$

(a) $(\alpha\theta_1 + \beta\theta_2, \psi) = \alpha(\theta_1, \psi) + \beta(\theta_2, \psi)$,
(b) $(\theta, \alpha\psi_1 + \beta\psi_2) = \overline{\alpha}(\theta, \psi_1) + \overline{\beta}(\theta, \psi_2)$, and
(c) $(\theta, \psi) = \overline{(\psi, \theta)}$.

Our principal aim is to show that the irreducible characters form an orthonormal basis for the space of complex class functions with respect to this Hermitian form (we already know that they are a basis). This fact will follow from the orthogonality of the primitive central idempotents, once we have explicitly determined these in the next proposition.

Proposition 13. Let $z_1, \ldots, z_r$ be the orthogonal primitive central idempotents in $\mathbb{C}G$ labelled in such a way that z_i acts as the identity on the irreducible $\mathbb{C}G$-module M_i, and let χ_i be the character afforded by M_i. Then

$$z_i = \frac{\chi_i(1)}{|G|} \sum_{g \in G} \chi_i(g^{-1})g.$$

Proof: Let $z = z_i$ and write

$$z = \sum_{g \in G} \alpha_g g.$$

Recall from Example 4 in this section that if ρ is the regular character of G then

$$\rho(g) = \begin{cases} 0 & \text{if } g \neq 1 \\ |G| & \text{if } g = 1 \end{cases} \tag{18.11}$$

and recall from the last example in Section 2 that

$$\rho = \sum_{j=1}^{r} \chi_j(1)\chi_j. \tag{18.12}$$

To find the coefficient α_g, apply ρ to zg^{-1} and use linearity of ρ together with equation (11) to obtain

$$\rho(zg^{-1}) = \alpha_g|G|.$$

Computing $\rho(zg^{-1})$ using (12) then gives

$$\sum_{j=1}^{r} \chi_j(1)\chi_j(zg^{-1}) = \alpha_g|G|. \tag{18.13}$$

Let φ_j be the irreducible representation afforded by M_j, $1 \leq j \leq r$. Since we may consider φ_j as an algebra homomorphism from $\mathbb{C}G$ into $\mathrm{End}(M_j)$, we obtain $\varphi_j(zg^{-1}) = \varphi_j(z)\varphi_j(g^{-1})$. Also, we have already observed that $\varphi_j(z)$ is 0 if $j \neq i$ and $\varphi_i(z)$ is the identity endomorphism on M_i. Thus

$$\varphi_j(zg^{-1}) = \begin{cases} 0 & \text{if } j \neq i \\ \varphi_i(g^{-1}) & \text{if } j = i. \end{cases}$$

This proves $\chi_j(zg^{-1}) = \chi_i(g^{-1})\delta_{ij}$, where δ_{ij} is zero if $i \neq j$ and is 1 if $i = j$ (called the Kronecker delta). Substituting this into equation (13) gives $\alpha_g = \dfrac{1}{|G|}\chi_i(1)\chi_i(g^{-1})$. This is the coefficient of g in the statement of the proposition, completing the proof.

The orthonormality of the irreducible characters will follow directly from the orthogonality of the central primitive idempotents via the following calculation:

$$\begin{aligned} z_i\delta_{ij} &= z_i z_j \\ &= \frac{\chi_i(1)}{|G|}\frac{\chi_j(1)}{|G|}\sum_{g,h\in G}\chi_i(g^{-1})\chi_j(h^{-1})gh \\ &= \frac{\chi_i(1)}{|G|}\frac{\chi_j(1)}{|G|}\sum_{y\in G}\Big[\sum_{x\in G}\chi_i(xy^{-1})\chi_j(x^{-1})\Big]y \end{aligned}$$

(to get the latter sum from the former substitute y for gh and x for h). Since the elements of G are a basis of $\mathbb{C}G$ we may equate coefficients with those of z_i found in Proposition 13 to get (the coefficient of g)

$$\delta_{ij}\frac{\chi_i(1)}{|G|}\chi_i(g^{-1}) = \frac{\chi_i(1)\chi_j(1)}{|G|^2}\sum_{x\in G}\chi_i(xg^{-1})\chi_j(x^{-1}).$$

Simplifying (and replacing g by g^{-1}) gives

$$\delta_{ij}\frac{\chi_i(g)}{\chi_j(1)} = \frac{1}{|G|}\sum_{x\in G}\chi_i(xg)\chi_j(x^{-1}) \quad \text{for all } g \in G. \tag{18.14}$$

Taking $g = 1$ in (14) gives

$$\delta_{ij} = \frac{1}{|G|}\sum_{x\in G}\chi_i(x)\chi_j(x^{-1}). \tag{18.15}$$

The sum on the right side would be precisely the inner product (χ_i, χ_j) if $\chi_j(x^{-1})$ were equal to $\overline{\chi_j(x)}$; this is the content of the next proposition.

Proposition 14. If ψ is any character of G then $\psi(x)$ is a sum of roots of 1 in $\mathbb{C}$ and $\psi(x^{-1}) = \overline{\psi(x)}$ for all $x \in G$.

Proof: Let φ be a representation whose character is ψ, fix an element $x \in G$ and let $|x| = k$. Since the minimal polynomial of $\varphi(x)$ divides $X^k - 1$ (hence has distinct roots), there is a basis of the underlying vector space such that the matrix of $\varphi(x)$ with respect to this basis is a diagonal matrix with k^{th} roots of 1 on the diagonal. Since $\psi(x)$ is the sum of the diagonal entries (and does not depend on the choice of basis), $\psi(x)$ is a sum of roots of 1. Moreover, if ϵ is a root of 1, $\epsilon^{-1} = \overline{\epsilon}$. Thus the inverse of a diagonal matrix with roots of 1 on the diagonal is the diagonal matrix with the complex conjugates of those roots of 1 on the diagonal. Since the complex conjugate of a sum is the sum of the complex conjugates, $\psi(x^{-1}) = \operatorname{tr}\varphi(x^{-1}) = \overline{\operatorname{tr}\varphi(x)} = \overline{\psi(x)}$.

Keep in mind that in the proof of Proposition 14 we first fixed a group element x and then chose a basis of the representation space so that $\varphi(x)$ was a diagonal matrix. It is always possible to diagonalize a single element but it is possible to *simultaneously* diagonalize all $\varphi(x)$'s if and only if φ is similar to a sum of degree 1 representations.

Combining the above proposition with equation (15) proves:

Theorem 15. *(The First Orthogonality Relation for Group Characters)* Let G be a finite group and let $\chi_1, \dots, \chi_r$ be the irreducible characters of G over $\mathbb{C}$. Then with respect to the inner product $(\ ,\ )$ above we have

$$(\chi_i, \chi_j) = \delta_{ij}$$

and the irreducible characters are an orthonormal basis for the space of class functions. In particular, if θ is any class function then

$$\theta = \sum_{i=1}^{r} (\theta, \chi_i)\chi_i.$$

Proof: We have just established that the irreducible characters form an orthonormal basis for the space of class functions. If θ is any class function, write $\theta = \sum_{i=1}^{r} a_i \chi_i$, for some $a_i \in \mathbb{C}$. It follows from linearity of the Hermitian product that $a_i = (\theta, \chi_i)$, as stated.

We list without proof the Second Orthogonality Relation; we shall not require it for the applications in this book.

Theorem 16. *(The Second Orthogonality Relation for Group Characters)* Under the notation above, for any $x, y \in G$

$$\sum_{i=1}^{r} \chi_i(x)\overline{\chi_i(y)} = \begin{cases} |C_G(x)| & \text{if } x \text{ and } y \text{ are conjugate in } G \\ 0 & \text{otherwise.} \end{cases}$$

Definition. For θ any class function on G the *norm* of θ is $(\theta, \theta)^{1/2}$ and will be denoted by $\|\theta\|$.

When a class function is written in terms of the irreducible characters, $\theta = \sum \alpha_i \chi_i$, its norm is easily calculated as $\|\theta\| = (\sum \alpha_i^2)^{1/2}$. It follows that

a character has norm 1 if and only if it is irreducible.

Finally, observe that computations of the inner product of characters θ and ψ may be simplified as follows. If $\mathcal{K}_1, \ldots, \mathcal{K}_r$ are the conjugacy classes of G with sizes $d_1, \ldots, d_r$ and representatives $g_1, \ldots, g_r$ respectively, then the value $\theta(g_i)\overline{\psi(g_i)}$ appears d_i times in the sum for (θ, ψ), once for each element of $\mathcal{K}_i$. Collecting these terms gives

$$(\theta, \psi) = \frac{1}{|G|}\sum_{i=1}^{r} d_i\theta(g_i)\overline{\psi(g_i)},$$

a sum only over representatives of the conjugacy classes of G. In particular, the norm of θ is given by

$$\|\theta\|^2 = (\theta, \theta) = \frac{1}{|G|}\sum_{i=1}^{r} d_i|\theta(g_i)|^2.$$

Examples

(1) Let $G = S_3$ and let π be the permutation character of degree 3 described in the examples at the beginning of this section. Recall that $\pi(\sigma)$ equals the number of elements in $\{1, 2, 3\}$ fixed by σ. The conjugacy classes of S_3 are represented by 1, (1 2) and (1 2 3) of sizes 1, 3 and 2 respectively, and $\pi(1) = 3$, $\pi((1\ 2)) = 1$, $\pi((1\ 2\ 3)) = 0$. Hence

$$\begin{aligned}\|\pi\|^2 &= \frac{1}{6}\left[1\,\pi(1)^2 + 3\,\pi((1\ 2))^2 + 2\,\pi((1\ 2\ 3))^2\right]\\ &= \frac{1}{6}(9+3+0) = 2\end{aligned}$$

This implies that π is a sum of two distinct irreducible characters, each appearing with multiplicity 1. Let χ_1 be the principal character of S_3, so that $\chi_1(\sigma) = \overline{\chi_1(\sigma)} = 1$ for all $\sigma \in S_3$. Then

$$\begin{aligned}(\pi, \chi_1) &= \frac{1}{6}\left[1\,\pi(1)\overline{\chi_1(1)} + 3\,\pi((1\ 2))\overline{\chi_1((1\ 2))} + 2\,\pi((1\ 2\ 3))\overline{\chi_1((1\ 2\ 3))}\,\right]\\ &= \frac{1}{6}(3+3+0) = 1\end{aligned}$$

so the principal character appears as a constituent of π with multiplicity 1. This proves $\pi = \chi_1 + \chi_2$ for some irreducible character χ_2 of S_3 of degree 2 (and agrees with our earlier decomposition of this representation). This also shows that the value of χ_2 on $\sigma \in S_3$ is the number of fixed points of σ minus 1.

(2) Let $G = S_4$ and let π be the natural permutation character of degree 4 (so again $\pi(\sigma)$ is the number of fixed points of σ). The conjugacy classes of S_4 are represented by 1, (1 2), (1 2 3), (1 2 3 4) and (1 2)(3 4) of sizes 1, 6, 8, 6 and 3 respectively. Again we compute:

$$\begin{aligned}\|\pi\|^2 &= \frac{1}{24}\left[1\,\pi(1)^2 + 6\,\pi((1\ 2))^2 + 8\,\pi((1\ 2\ 3))^2 + 6\,\pi((1\ 2\ 3\ 4))^2\right.\\ &\qquad\left. + 3\,\pi((1\ 2)(3\ 4))^2\right]\\ &= \frac{1}{24}(16+24+8+0+0) = 2\,.\end{aligned}$$

so π has two distinct irreducible constituents. If χ_1 is the principal character of S_4, then

$$\begin{aligned}(\pi, \chi_1) &= \frac{1}{24}\left[1\,\pi(1) + 6\,\pi((1\ 2)) + 8\,\pi((1\ 2\ 3)) \right.\\ &\qquad \left. + 6\,\pi((1\ 2\ 3\ 4)) + 3\,\pi((1\ 2)(3\ 4))\right]\\ &= \frac{1}{24}(4 + 12 + 8 + 0 + 0) = 1.\end{aligned}$$

This proves that the degree 4 permutation character is the sum of the principal character and an irreducible character of degree 3.

(3) Let $G = D_8$, where

$$D_8 = \langle r, s \mid s^2 = r^4 = 1,\ rs = sr^{-1}\rangle.$$

The conjugacy classes of D_8 are represented by $1, s, r, r^2$ and sr and have sizes 1, 2, 2, 1 and 2, respectively. Let φ be the degree 2 matrix representation of D_8 obtained as in Example 6 in Section 1 from embedding a square in $\mathbb{R}^2$:

$$\varphi(s) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},\ \varphi(r) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},\ \varphi(r^2) = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix},\ \varphi(sr) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Let ψ be the character of this representation (where we consider the real matrices as a subset of the complex matrices). Again, since ψ is real valued one computes

$$\begin{aligned}\|\psi\|^2 &= \frac{1}{8}\left[1\psi(1)^2 + 2\psi(s)^2 + 2\psi(r)^2 + 1\psi(r^2)^2 + 2\psi(sr)^2\right]\\ &= \frac{1}{8}(4 + 0 + 0 + 4 + 0) = 1.\end{aligned}$$

This proves the representation φ is irreducible (even if we allow similarity transformations by complex matrices).

We have seen that the sum of two characters is again a character. Specifically, if ψ_1 and ψ_2 are characters of representations φ_1 and φ_2, then $\psi_1 + \psi_2$ is the character of $\varphi_1 + \varphi_2$.

Proposition 17. If ψ_1 and ψ_2 are characters, then so is their product $\psi_1\psi_2$.

Proof: Let V_1 and V_2 be $\mathbb{C}G$-modules affording characters ψ_1 and ψ_2 and define $W = V_1 \otimes_{\mathbb{C}} V_2$. Since each $g \in G$ acts as a linear transformation on V_1 and V_2, the action of g on simple tensors by $g(v_1 \otimes v_2) = (gv_1) \otimes (gv_2)$ extends by linearity to a well defined linear transformation on W by Proposition 17 in Section 11.2. One easily checks that this action also makes W into a $\mathbb{C}G$-module. By Exercise 38 in Section 11.2 the character afforded by W is $\psi_1\psi_2$.

The next chapter will contain further explicit character computations as well as some applications of group characters to proving theorems about certain classes of groups.

Some Remarks on Fourier Analysis and Group Characters

This brief discussion is intended to indicate some connections of the results above with other areas of mathematics.

The theory of group representations described to this point is a special branch of an area of mathematics called Harmonic Analysis. Readers may already be familiar with the basic theory of Fourier series which also falls into this realm. We make some observations which show how representation theory for finite groups corresponds to "Fourier series" for some infinite groups (in particular, to Fourier series on the circle). To be mathematically precise one needs the Lebesgue integral to ensure completeness of certain (Hilbert) spaces but readers may get the flavor of things by replacing "Lebesgue" by "Riemann."

Let G be the multiplicative group of points on the unit circle in $\mathbb{C}$:

$$G = \{z \in \mathbb{C} \mid |z| = 1\}.$$

We shall usually view G as the interval $[0, 2\pi]$ in $\mathbb{R}$ with the two end points identified, i.e., as the additive group $\mathbb{R}/2\pi\mathbb{Z}$ (the isomorphism is: the real number x corresponds to the complex number e^{ix}). Note that G has a translation invariant measure, namely the Lebesgue measure, and the measure of the circle is 2π. For finite groups, the counting measure is the translation invariant measure (so the measure of a subset H is the number of elements in that subset, $|H|$) and integrals on a finite group with respect to this counting measure are just finite sums.

The space

$$L^2(G) = \{f : G \to \mathbb{C} \mid f \text{ is measurable and } |f|^2 \text{ is integrable over } G\}$$

plays the role of the group algebra of the infinite group G. This space becomes a commutative ring with 1 under the convolution of functions: for $f, g \in L^2(G)$ the product $f * g : G \to \mathbb{C}$ is defined by

$$(f * g)(x) = \frac{1}{2\pi}\int_0^{2\pi} f(x - y)g(y)\, dy \qquad \text{for all } x \in G.$$

(Recall that for a finite group H, the group algebra is also formally the ring of $\mathbb{C}$-valued functions on H under a convolution multiplication and that these functions are written as formal sums – the element $\sum \alpha_g g \in \mathbb{C}G$ denotes the function which sends g to $\alpha_g \in \mathbb{C}$ for all $g \in G$.)

The complete set of continuous homomorphisms of G into $GL_1(\mathbb{C})$ is given by

$$e_n(x) = e^{inx}, \quad x \in [0, 2\pi], \quad n \in \mathbb{Z}.$$

(Recall that for a finite abelian group, all irreducible representations are 1-dimensional and for 1-dimensional representations, characters and representations may be identified.)

The ring $L^2(G)$ admits an Hermitian inner product: for $f, g \in L^2(G)$

$$(f, g) = \frac{1}{2\pi}\int_0^{2\pi} f(t)\overline{g(t)}\, dt.$$

Under this inner product, $\{e_n \mid n \in \mathbb{Z}\}$ is an orthonormal basis (where the term "basis" is used in the analytic sense that these are independent and 0 is the only function orthogonal to all of them). Moreover,

$$L^2(G) = \widehat{\bigoplus_{n\in\mathbb{Z}} E_n}$$

where E_n is the 1-dimensional subspace spanned by e_n, the hat over the direct sum denotes taking the closure of the direct sum in the L^2-topology, and equality indicates equality in the L^2 sense. (Recall that the group algebra of a finite abelian group is the direct sum of the irreducible 1-dimensional submodules, each occurring with multiplicity one.) These facts imply the well known result from Fourier analysis that every square integrable function $f(x)$ on $[0, 2\pi]$ has a Fourier series

$$\sum_{n=-\infty}^{\infty} c_n e^{inx}$$

where the Fourier coefficients, c_n, are given by

$$c_n = (f, e_n) = \frac{1}{2\pi}\int_0^{2\pi} f(t)e^{-int}\, dt.$$

This brief description indicates how the representation theory of finite groups extends to certain infinite groups and the results we have proved may already be familiar in the latter context. In fact, there is a completely analogous theory for arbitrary (not necessarily abelian) compact Lie groups — here the irreducible (complex) representations need not be 1-dimensional but they are all finite dimensional and $L^2(G)$ decomposes as a direct sum of them, each appearing with multiplicity equal to its degree. The emphasis (at least at the introductory level) in this theory is often on the importance of being able to represent functions as (Fourier) series and then using these series to solve other problems (e.g., solve differential equations). The underlying group provides the "symmetry" on which to build this "harmonic analysis," rather than being itself the principal object of study.

EXERCISES

Let G be a finite group. Unless stated otherwise all representations and characters are over $\mathbb{C}$.

1. Prove that $\operatorname{tr} AB = \operatorname{tr} BA$ for $n \times n$ matrices A and B with entries from any commutative ring.

2. In each of (a) to (c) let ψ be the character afforded by the specified representation φ.
- **(a)** Let φ be the degree 2 representation of D_{10} described in Example 6 in the second set of examples in Section 1 (here $n = 5$) and show that $\|\psi\|^2 = 1$ (hence φ is irreducible).
- **(b)** Let φ be the degree 2 representation of Q_8 described in Example 7 in the second set of examples in Section 1 and show that $\|\psi\|^2 = 1$ (hence φ is irreducible).
- **(c)** Let φ be the degree 4 representation of Q_8 described in Example 8 in the second set of examples in Section 1 and show that $\|\psi\|^2 = 4$ (hence even though φ is irreducible over $\mathbb{R}$, φ decomposes over $\mathbb{C}$ as twice an irreducible representation of degree 2).

3. If χ is an irreducible character of G, prove that the χ-isotypic subspace of a $\mathbb{C}G$-module is unique.

4. Prove that if N is any irreducible $\mathbb{C}G$-module and $M = N \oplus N$, then M has infinitely many direct sum decompositions into two copies of N.

5. Prove that a class function is a character if and only if it is a positive integral linear combination of irreducible characters.

6. Let $\varphi : G \to GL(V)$ be a representation with character ψ. Let W be the subspace $\{v \in V \mid \varphi(g)(v) = v \text{ for all } g \in G\}$ of V fixed pointwise by all elements of G. Prove that $\dim W = (\psi, \chi_1)$, where χ_1 is the principal character of G.

7. Assume V is a $\mathbb{C}G$-module on which G acts by permuting the basis $\mathcal{B} = \{e_1, \ldots, e_n\}$. Write $\mathcal{B}$ as a disjoint union of the orbits $\mathcal{B}_1, \ldots, \mathcal{B}_t$ of G on $\mathcal{B}$.
   (a) Prove that V decomposes as a $\mathbb{C}G$-module as $V_1 \oplus \cdots \oplus V_t$, where V_i is the span of $\mathcal{B}_i$.
   (b) Prove that if v_i is the sum of the vectors in $\mathcal{B}_i$ then the 1-dimensional subspace of V_i spanned by v_i is the unique $\mathbb{C}G$-submodule of V_i affording the trivial representation (in other words, any vector in V_i that is fixed under the action of G is a multiple of v_i). [Use the fact that G is transitive on $\mathcal{B}_i$. See also Exercise 8 in Section 1.]
   (c) Let $W = \{v \in V \mid \varphi(g)(v) = v \text{ for all } g \in G\}$ be the subspace of V fixed pointwise by all elements of G. Deduce that $\dim W = t =$ the number of orbits of G on $\mathcal{B}$.

8. Prove the following result (sometimes called Burnside's Lemma although its origin is with Frobenius): let G be a subgroup of S_n and for each $\sigma \in G$ let $\mathrm{Fix}(\sigma)$ denote the number of fixed points of σ on $\{1, \ldots, n\}$. Let t be the number of orbits of G on $\{1, \ldots, n\}$. Then
$$t|G| = \sum_{g \in G} \mathrm{Fix}(g).$$
[Use the preceding two exercises.]

9. Let G be a nontrivial, transitive group of permutations on the finite set Ω and let ψ be the character afforded by the linear representation over $\mathbb{C}$ obtained from Ω (cf. Example 4 in Section 1) so $\psi(\sigma)$ is the number of fixed points of σ on Ω. Now let G act on the set $\Omega \times \Omega$ by $g \cdot (\omega_1, \omega_2) = (g \cdot \omega_1, g \cdot \omega_2)$ and let π be the character afforded by the linear representation obtained from this action.
   (a) Prove that $\pi = \psi^2$.
   (b) Prove that the number of orbits of G on $\Omega \times \Omega$ is given by the inner product (ψ, ψ). [By the preceding exercises, the number of orbits on $\Omega \times \Omega$ is equal to (π, χ_1), where χ_1 is the principal character.]
   (c) Recall that G is said to be *doubly transitive* on Ω if it has precisely 2 orbits in its action on $\Omega \times \Omega$ (it always has at least 2 orbits since the diagonal, $\{(\omega, \omega) \mid \omega \in \Omega\}$, is one orbit). Prove that if G is doubly transitive on Ω then $\psi = \chi_1 + \chi_2$, where χ_1 is the principal character and χ_2 is a nonprincipal irreducible character of G.
   (d) Let $\Omega = \{1, 2, \ldots, n\}$ and let $G = S_n$ act on Ω in the natural fashion. Show that the character of the associated linear representation decomposes as the principal character plus an irreducible character of degree $n - 1$.

10. Let ψ be the character of any 2-dimensional representation of a group G and let x be an element of order 2 in G. Prove that $\psi(x) = 2, 0$ or -2. Generalize this to n-dimensional representations.

11. Let χ be an irreducible character of G. Prove that for every element z in the center of G we have $\chi(z) = \epsilon\chi(1)$, where ϵ is some root of 1 in $\mathbb{C}$. [Use Schur's Lemma.]

12. Let ψ be the character of some representation φ of G. Prove that for $g \in G$ the following hold:
   (a) if $\psi(g) = \psi(1)$ then $g \in \ker\varphi$, and

(b) if $|\psi(g)| = \psi(1)$ and φ is faithful then $g \in Z(G)$ (where $|\psi(g)|$ is the complex absolute value of $\psi(g)$). [Use the method of proof of Proposition 14.]

13. Let $\varphi : G \to GL(V)$ be a representation and let $\chi : G \to \mathbb{C}^\times$ be a degree 1 representation. Prove that $\chi\varphi : G \to GL(V)$ defined by $\chi\varphi(g) = \chi(g)\varphi(g)$ is a representation (note that multiplication of the linear transformation $\varphi(g)$ by the complex number $\chi(g)$ is well defined). Show that $\chi\varphi$ is irreducible if and only if φ is irreducible. Show that if ψ is the character afforded by φ then $\chi\psi$ is the character afforded by $\chi\varphi$. Deduce that the product of any irreducible character with a character of degree 1 is also an irreducible character.

The next few exercises study the notion of *algebraically conjugate* characters. These exercises may be considered as extensions of Proposition 14 and some consequences of these extensions. In particular we obtain a group-theoretic characterization of the conditions under which all irreducible characters of a group take values in $\mathbb{Q}$.

Let F be the subfield of $\mathbb{C}$ of all elements that are algebraic over $\mathbb{Q}$ (the field of algebraic numbers). Thus F is the algebraic closure of $\mathbb{Q}$ contained in $\mathbb{C}$ and all the results established over $\mathbb{C}$ hold without change over F.

14. Note that since $F \subseteq \mathbb{C}$, every representation $\varphi : G \to GL_m(F)$ may also be considered as a complex representation. Prove that if φ is a representation over F that is irreducible over F, then φ is also irreducible when considered over the larger field $\mathbb{C}$ (note that this is not true if F is not algebraically closed — cf. Exercise 2(c) above). Show that the set of irreducible characters of G over F is the same as the set of irreducible characters over $\mathbb{C}$ (i.e., these are exactly the same set of class functions on G). Deduce that every complex representation is equivalent to a representation over F. [Since F is algebraically closed of characteristic 0, the irreducible characters over either F or $\mathbb{C}$ are characterized by the first orthogonality relation.]

Let $\varphi : G \to GL_m(F)$ be any representation with character ψ. Let $\mathbb{Q}(\varphi)$ denote the subfield of F generated by all the entries of the matrices $\varphi(g)$ for all $g \in G$.

15. Prove that $\mathbb{Q}(\varphi)$ is a finite extension of $\mathbb{Q}$.

Now let K be any Galois extension of $\mathbb{Q}$ containing $\mathbb{Q}(\varphi)$ and let $\sigma \in \mathrm{Gal}(K/\mathbb{Q})$. In fact, since every automorphism of K extends to an automorphism of F, we may assume σ is any automorphism of F. The map $\varphi^\sigma : G \to GL_n(F)$ is defined by letting $\varphi^\sigma(g)$ be the $n \times n$ matrix whose entries are obtained by applying the field automorphism σ to the entries of the matrix $\varphi(g)$.

16. Prove that φ^σ is a representation. Prove also that the character of φ^σ is ψ^σ, where $\psi^\sigma(g) = \sigma(\psi(g))$.

17. Prove that φ is irreducible if and only if φ^σ is irreducible.

The representation φ^σ (or character ψ^σ) is called the *algebraic conjugate* of φ by σ (or of ψ, respectively); two representations φ_1 and φ_2 (or characters ψ_1 and ψ_2) are said to be *algebraically conjugate* if there is some automorphism σ of F such that $\varphi_1^\sigma = \varphi_2$ (or $\psi_1^\sigma = \psi_2$, respectively). Some care needs to be taken with this (standard) notation since the exponential notation usually denotes a right action whereas automorphisms of F act on the left on representations: $\varphi^{(\sigma\tau)} = (\varphi^\tau)^\sigma$.

Let $\mathbb{Q}(\psi)$ be the subfield of F generated by the numbers $\psi(g)$ for all $g \in G$. Let $|G| = n$ and let ϵ be a primitive n^{th} root of 1 in F.

18. Prove that $\mathbb{Q}(\psi) \subseteq \mathbb{Q}(\epsilon)$. Deduce that $\mathbb{Q}(\psi)$ is a Galois extension of $\mathbb{Q}$ with abelian Galois group. [See Proposition 14.]

Recall from Section 14.5 that $\mathrm{Gal}(\mathbb{Q}(\epsilon)/\mathbb{Q}) \cong (\mathbb{Z}/n\mathbb{Z})^\times$, where the Galois automorphisms are given on the generator ϵ by $\sigma_a : \epsilon \mapsto \epsilon^a$, where a is an integer relatively prime to n.

19. Prove that if $\sigma_a \in \mathrm{Gal}(\mathbb{Q}(\epsilon)/\mathbb{Q})$ is the field automorphism defined above, then for all $g \in G$ we have $\psi^{\sigma_a}(g) = \psi(g^a)$. [Use the method of Proposition 14.]

20. Prove that if g is an element of G which is conjugate to g^a for all integers a relatively prime to n, then $\psi(g) \in \mathbb{Q}$, for every character ψ of G. [Use the preceding exercise and the fact that $\mathbb{Q}$ is the field fixed by all σ_a's.]

21. Prove that an element $g \in G$ is conjugate to g^a for all integers a relatively prime to $|G|$ if and only if g is conjugate to $g^{a'}$ for all integers a' relatively prime to $|g|$.

22. Show for any positive integer n that every character of the symmetric group S_n is rational valued (i.e., $\psi(g) \in \mathbb{Q}$ for all $g \in S_n$ and all characters ψ of S_n).

The next two exercises establish the converse to Exercise 20.

23. Prove that elements x and y are conjugate in a group G if and only if $\chi(x) = \chi(y)$ for all irreducible characters χ of G.

24. Let $g \in G$ and assume that every irreducible character of G is rational valued on g. Prove that g is conjugate to g^a for every integer a relatively prime to $|G|$. [If g is not conjugate to g^a for some a relatively prime to $|G|$ then by the preceding exercise there is an irreducible character χ such that $\chi(g) \neq \chi(g^a)$. Derive a contradiction from the hypothesis that $\chi(g) \in \mathbb{Q}$.]

25. Describe which irreducible characters of the cyclic group of order n are algebraically conjugate.

26. Prove that every irreducible character of both Q_8 and D_8 is rational valued. Prove that D_{10} has an irreducible character that is not rational valued.

27. Let $G = H \times K$ and let $\varphi : H \to GL(V)$ be an irreducible representation of H with character χ. Then $G \xrightarrow{\pi_H} H \xrightarrow{\varphi} GL(V)$ gives an irreducible representation of G, where π_H is the natural projection; the character, $\tilde{\chi}$, of this representation is $\tilde{\chi}((h,k)) = \chi(h)$. Likewise any irreducible character ψ of K gives an irreducible character $\tilde{\psi}$ of G with $\tilde{\psi}((h,k)) = \psi(k)$.

(a) Prove that the product $\tilde{\chi}\tilde{\psi}$ is an irreducible character of G. [Show it has norm 1.]

(b) Prove that every irreducible character of G is obtained from such products of irreducible characters of the direct factors. [Use Theorem 10, either (3) or (4).]

28. *(Finite subgroups of $GL_2(\mathbb{Q})$)* Let G be a finite subgroup of $GL_2(\mathbb{Q})$.

(a) Show that $GL_2(\mathbb{Q})$ does not contain an element of order n for $n = 5, 7$, or $n \geq 9$. Deduce that $|G| = 2^a 3^b$. [Use rational canonical forms.]

(b) Show that the Klein 4-group is the only noncyclic abelian subgroup of $GL_2(\mathbb{Q})$. Deduce from this and (a) that $|G| \mid 24$.

(c) Show that the only finite subgroups of $GL_2(\mathbb{Q})$ are the cyclic groups of order 1, 2, 3, 4, and 6, the Klein 4-group, and the dihedral groups of order 6, 8, and 12. [Use the classifications of groups of small order in Section 4.5 and Exercise 10 of Section 1 to restrict G to this list. Show conversely that each group listed has a 2-dimensional faithful rational representation.]

CHAPTER 19

Examples and Applications of Character Theory

19.1 CHARACTERS OF GROUPS OF SMALL ORDER

The *character table* of a finite group is the table of character values formatted as follows: list representatives of the r conjugacy classes along the top row and list the irreducible characters down the first column. The entry in the table in row χ_i and column g_j is $\chi_i(g_j)$. The character table of a finite group is unique up to a permutation of its rows and columns. It is customary to make the principal character the first row and the identity the first column and to list the characters in increasing order by degrees. In our examples we shall list the size of the conjugacy classes under each class so the entire table will have $r+2$ rows and $r+1$ columns (although strictly speaking, the character table is the $r \times r$ matrix of character values). This will enable one to easily check the "orthogonality of rows" using the first orthogonality relation: if the classes are represented by $g_1, \dots, g_r$ of sizes $d_1, \dots, d_r$ then

$$(\chi_i, \chi_j) = \frac{1}{|G|}\sum_{k=1}^{r} d_k \chi_i(g_k)\overline{\chi_j(g_k)}.$$

The second orthogonality relation says that the Hermitian product of any two distinct columns of a character table is zero (i.e., it gives an "orthogonality of columns").

A number of character tables are given in the *Atlas of Finite Groups* by Conway, Curtis, Norton, Parker and Wilson, Clarendon Press, 1985. These include the character table of the Monster simple group, M. The group M has 194 irreducible characters. The smallest degree of a nonprincipal irreducible character of M is 196883 and the largest degree is on the order of 2×10^{26}. Nonetheless, it is possible to compute the values of all these characters on all conjugacy classes of M.

For the first example of a character table let $G = \langle x \rangle$ be the cyclic group of order 2. Then G has 2 conjugacy classes and two irreducible characters:

classes:	1	x
sizes:	1	1
χ_1	1	1
χ_2	1	-1

Character Table of Z_2

The characters and representations of this abelian group are the same, and the irreducible representations of any abelian group are described in Example 1 at the end of Section 18.2.

Similarly, if $G = \langle x \rangle$ is cyclic of order 3, and ζ is a fixed primitive cube root of 1 (so $\zeta^2 = \overline{\zeta}$), then the character table of G is the following:

classes:	1	x	x^2
sizes:	1	1	1
χ_1	1	1	1
χ_2	1	ζ	ζ^2
χ_3	1	ζ^2	ζ

Character Table of Z_3

Next we construct the character table of S_3. Recall from Example 2 in Section 18.2 that S_3 has 3 irreducible characters whose values are described in that example and in Example 1 at the end of Section 18.3.

classes:	1	$(1\,2)$	$(1\,2\,3)$
sizes:	1	3	2
χ_1	1	1	1
χ_2	1	-1	1
χ_3	2	0	-1

Character Table of S_3

Next we consider D_8, adopting the notation of Example 3 of Section 18.3. By Corollary 11, D_8 has four characters of degree 1. Also, in Example 3 we constructed an irreducible degree 2 representation. Since the sum of the squares of the degrees of these representations is 8, this accounts for all irreducible representations (or, since there are 5 conjugacy classes, there are 5 irreducible representations). If we let bars denote passage to the commutator quotient group (which is the Klein 4-group), then $\overline{1} = \overline{r^2}$. The degree 1 representations (= their characters) are computed by sending generators $\overline{s}$ and $\overline{r}$ to ± 1 (and the product class is mapped to the product of the values). Matrices for the degree 2 irreducible representation were computed in Example 3 of Section 18.3 and the character of this representation can be read directly from these matrices. The character table of D_8 is therefore the following:

classes:	1	r^2	s	r	sr
sizes:	1	1	2	2	2
χ_1	1	1	1	1	1
χ_2	1	1	-1	1	-1
χ_3	1	1	1	-1	-1
χ_4	1	1	-1	-1	1
χ_5	2	-2	0	0	0

Character Table of D_8

Now we compute the character table of the quaternion group of order 8. We use the usual presentation

$$Q_8 = \langle i, j \mid i^4 = 1,\ i^2 = j^2,\ i^{-1}ji = j^{-1} \rangle$$

and let $k = ij$ and $i^2 = -1$. The conjugacy classes of Q_8 are represented by $1, -1, i$, j and k of sizes $1, 1, 2, 2$ and 2, respectively. Since the commutator quotient of Q_8 is the Klein 4-group, there are four characters of degree 1. The one remaining irreducible character must have degree 2 in order that the sum of the squares of the degrees be 8. Let χ_5 be the degree 2 irreducible character of Q_8. One may check that the representation $\varphi : Q_8 \to GL_2(\mathbb{C})$ described explicitly in Example 7 in the second set of examples of Section 18.1 affords χ_5, but we show how the orthogonality relations give the values of χ_5 without knowing these explicit matrices. If φ is an irreducible representation of degree 2, by Schur's Lemma (cf. Exercise 18 in Section 18.1) $\varphi(-1)$ is a 2×2 scalar matrix and so is $\pm$ the identity matrix since -1 has order 2 in Q_8. Hence $\chi_5(-1) = \pm 2$. Let $\chi_5(i) = a$, $\chi_5(j) = b$ and $\chi_5(k) = c$. The orthogonality relations give

$$1 = (\chi_5, \chi_5) = \frac{1}{8}(2^2 + (\pm 2)^2 + 2a\overline{a} + 2b\overline{b} + 2c\overline{c}).$$

Since $a\overline{a}$, $b\overline{b}$ and $c\overline{c}$ are nonnegative real numbers, they must all be zero. Also, since χ_5 is orthogonal to the principal character we get

$$0 = (\chi_1, \chi_5) = \frac{1}{8}(2 + (\pm 2) + 0 + 0 + 0),$$

hence $\chi_5(-1) = -2$. The complete character table of Q_8 is the following:

classes:	1	-1	i	j	k
sizes:	1	1	2	2	2
χ_1	1	1	1	1	1
χ_2	1	1	-1	1	-1
χ_3	1	1	1	-1	-1
χ_4	1	1	-1	-1	1
χ_5	2	-2	0	0	0

Character Table of Q_8

Observe that D_8 and Q_8 have the same character table, hence

nonisomorphic groups may have the same character table.

Note that the values of the degree 2 representation of Q_8 could also have been easily calculated by applying the second orthogonality relation to each column of the character table. We leave this check as an exercise. Also note that although the degree 2 irreducible characters of D_8 and Q_8 have the same (real number) values the degree 2 representation of D_8 may be realized by real matrices whereas it may be shown that Q_8 has no faithful 2-dimensional representation over $\mathbb{R}$ (cf. Exercise 10 in Section 18.1).

For the next example we construct the character table of S_4. The conjugacy classes of S_4 are represented by 1, $(1\,2)$, $(1\,2\,3)$, $(1\,2\,3\,4)$ and $(1\,2)(3\,4)$ with sizes 1, 6, 8, 6, and 3 respectively. Since $S_4' = A_4$, there are two characters of degree 1: the principal character and the character whose values are the sign of the permutation.

To obtain a degree 2 irreducible character let V be the normal subgroup of order 4 generated by $(1\,2)(3\,4)$ and $(1\,3)(2\,4)$. Any representation φ of $S_4/V \cong S_3$ gives, by composition with the natural projection $S_4 \to S_4/V$, a representation of S_4; if the former is irreducible, so is the latter. Let φ be the composition of the projection with the irreducible 2-dimensional representation of S_3, and let χ_3 be its character. The classes of 1 and $(1\,2)(3\,4)$ map to the identity in the S_3 quotient, $(1\,2)$ and $(1\,2\,3\,4)$ map to transpositions and $(1\,2\,3)$ maps to a 3-cycle. The values of χ_3 can thus be read directly from the values of the character of degree 2 in the table for S_3.

Since S_4 has 5 irreducible characters and the sum of the squares of the degrees is 24, there must be two remaining irreducible characters, each of degree 3. In Example 2 of Section 18.3 one of these was calculated, call it χ_4. Recall that

$$\chi_4(\sigma) = (\text{the number of fixed points of } \sigma) - 1.$$

The remaining irreducible character, χ_5, is $\chi_4\chi_2$. One can either use Proposition 17 in Section 18.3 or Exercise 13 in Section 18.3 to see that this product is indeed a character. The first orthogonality relation verifies that it is irreducible.

classes:	1	$(1\,2)$	$(1\,2\,3)$	$(1\,2\,3\,4)$	$(1\,2)(3\,4)$
sizes:	1	6	8	6	3
χ_1	1	1	1	1	1
χ_2	1	-1	1	-1	1
χ_3	2	0	-1	0	2
χ_4	3	1	0	-1	-1
χ_5	3	-1	0	1	-1

Character Table of S_4

From the character table of S_4 one can easily compute the character table of A_4. Note that A_4 has 4 conjugacy classes. Also $|A_4 : A_4'| = 3$, so A_4 has three characters of degree 1 with $V = A_4'$ in the kernel of each degree 1 representation. The remaining irreducible character must have degree 3. One checks directly from the orthogonality relation applied in A_4 that the character χ_4 of S_4 restricted to A_4 ($= \chi_5|_{A_4}$) is irreducible. This irreducibility check is really necessary since an irreducible representation of a group need not restrict to an irreducible representation of a subgroup (for instance, the irreducible degree 2 representation of S_3 must become reducible when restricted to any proper subgroup, since these are all abelian). The character table of A_4 is the following

classes:	1	$(1\,2)(3\,4)$	$(1\,2\,3)$	$(1\,3\,2)$
sizes:	1	3	4	4
χ_1	1	1	1	1
χ_2	1	1	ζ	ζ^2
χ_3	1	1	ζ^2	ζ
χ_4	3	-1	0	0

Character Table of A_4

where ζ is a primitive cube root of 1 in $\mathbb{C}$.

As a final example we construct the following character table of S_5:

classes:	1	(1 2)	(1 2 3)	(1 2 3 4)	(1 2 3 4 5)	(1 2)(3 4)	(1 2)(3 4 5)
sizes:	1	10	20	30	24	15	20
χ_1	1	1	1	1	1	1	1
χ_2	1	−1	1	−1	1	1	−1
χ_3	4	2	1	0	−1	0	−1
χ_4	4	−2	1	0	−1	0	1
χ_5	5	−1	−1	1	0	1	−1
χ_6	5	1	−1	−1	0	1	1
χ_7	6	0	0	0	1	−2	0

Character Table of S_5

The conjugacy classes and their sizes were computed in Section 4.3. Since $|S_5 : S_5'| = 2$, there are two degree 1 characters: the principal character and the "sign" character.

The natural permutation of S_5 on 5 points gives rise to a permutation character of degree 5. As with S_4 and S_3 the orthogonality relations show that the square of its norm is 2 and it contains the principal character. Thus χ_3 is the permutation character minus the principal character (and, as with the smaller symmetric groups, $\chi_3(\sigma)$ is the number of fixed points of σ minus 1). As argued with S_4, it follows that $\chi_4 = \chi_3\chi_2$ is also an irreducible character.

To obtain χ_5 recall that S_5 has six Sylow 5-subgroups. Its action by conjugation on these gives a faithful permutation representation of degree 6. If ψ is the character of the associated linear representation, then since $\sigma \in S_5$ fixes a Sylow 5-subgroup if and only if it normalizes that subgroup, we have

$$\psi(\sigma) = \text{the number of Sylow 5-subgroups normalized by } \sigma.$$

The normalizer in S_5 of the Sylow 5-subgroup $\langle\, (1\,2\,3\,4\,5)\,\rangle$ is $\langle\ (1\,2\,3\,4\,5)\,,\ (2\,3\,5\,4)\ \rangle$ and all normalizers of Sylow 5-subgroups are conjugate in S_5 to this group. This normalizer contains only the identity, 5-cycles, 4-cycles and products of two disjoint transpositions. No other cycle type normalizes any Sylow 5-subgroup so on any other class, ψ is zero. To compute ψ on the remaining three nonidentity classes note (by inspection in S_6) that in any faithful action on 6 points the following hold: an element of order 5 must be a 5-cycle (hence fixes 1 point); any element of order 4 which fixes one point must be a 4-cycle (hence fixes 2 points); an element of order 2 which is the square of an element of order 4 fixes exactly 2 points also. This gives all the values of ψ. Now direct computation shows that

$$||\,\psi\,||^2 = 2 \qquad \text{and} \qquad (\chi_1, \psi) = 1.$$

Thus $\chi_5 = \psi - \chi_1$ is irreducible of degree 5. By the same theory as for χ_4 one gets that $\chi_6 = \chi_5\chi_2$ is another irreducible character.

Since there are 7 conjugacy classes, there is one remaining irreducible character and its degree is 6. Its values can be obtained immediately from the decomposition of the regular character, ρ (cf. Example 3 in Section 18.2 and Example 4 in Section 18.3):

$$\chi_7 = \frac{\rho - \chi_1 - \chi_2 - 4\chi_3 - 4\chi_4 - 5\chi_5 - 5\chi_6}{6}.$$

A direct calculation by the orthogonality relations checks that χ_7 is irreducible. Note that the values of the character χ_7 were computed without explicitly exhibiting a representation with this character.

EXERCISES

1. Calculate the character tables of $Z_2 \times Z_2$, $Z_2 \times Z_3$ and $Z_2 \times Z_2 \times Z_2$. Explain why the table of $Z_2 \times Z_3$ contains primitive 6^{th} roots of 1.
2. Compute the degrees of the irreducible characters of D_{16}.
3. Compute the degrees of the irreducible characters of A_5. Deduce that the degree 6 irreducible character of S_5 is not irreducible when restricted to A_5. [The conjugacy classes of A_5 are worked out in Section 4.3.]
4. Using the character tables in this section, for each of parts (a) to (d) use the first orthogonality relation to write the specified permutation character (cf. Example 3, Section 18.3) as a sum of irreducible characters:
   (a) the permutation character of the subgroup A_3 of S_3
   (b) the permutation character of the subgroup $\langle (1\ 2\ 3\ 4) \rangle$ of S_4
   (c) the permutation character of the subgroup V_4 of S_4
   (d) the permutation character of the subgroup $\langle (1\ 2\ 3),\ (1\ 2),\ (4\ 5) \rangle$ of S_5 (this subgroup is the normalizer of a Sylow 3-subgroup of S_5).
5. Assume that for any character ψ of a group, ψ^2 is also a character (where $\psi^2(g) = (\psi(g))^2$) — this is a special case of Proposition 17 in Section 18.3. Using the character tables in this section, for each of parts (a) to (e) write out the values of the square, χ^2, of the specified character χ and use the first orthogonality relation to write χ^2 as a sum of irreducible characters:
   (a) $\chi = \chi_3$, the degree 2 character in the table of S_3
   (b) $\chi = \chi_5$, the degree 2 character in the table of Q_8
   (c) $\chi = \chi_5$, the last character in the table of S_4
   (d) $\chi = \chi_4$, the second degree 4 character in the table of S_5
   (e) $\chi = \chi_7$, the last character in the table of S_5.
6. Calculate the character table of A_5.
7. Show that S_6 has an irreducible character of degree 5.
8. Calculate the character table of D_{10}. (This table contains nonreal entries.)
9. Calculate the character table of D_{12}.
10. Calculate the character table of $S_3 \times S_3$.
11. Calculate the character table of $Z_3 \times S_3$.
12. Calculate the character table of $Z_2 \times S_4$.
13. Calculate the character table of $S_3 \times S_4$.
14. Let n be an integer with $n \geq 3$. Show that every irreducible character of D_{2n} has degree 1 or 2 and find the number of irreducible characters of each degree. [The conjugacy classes of D_{2n} were found in Exercises 31 and 32 of Section 4.3 and its commutator subgroup was computed in Section 5.4.]
15. Prove that the character table is an invertible matrix. [Use the orthogonality relations.]
16. For each of A_5 and D_{10} describe which irreducible characters are algebraically conjugate (cf. the exercises in Section 18.3).

17. Let p be any prime and let P be a non-abelian group of order p^3 (up to isomorphism there are two choices for P; for odd p these were constructed when the groups of order p^3 were classified in Section 5.5). This exercise determines the character table of P and shows that both isomorphism types have the same character table (the argument includes the $p = 2$ case worked out in this section).
   (a) Prove that P has p^2 characters of degree 1.
   (b) Prove that P has $p - 1$ irreducible characters of degree p and that these together with the p^2 degree 1 characters are all the irreducible characters of P. [Use Theorem 10(3) and Theorem 12 in Section 18.2.]
   (c) Deduce that (regardless of the isomorphism type) the group P has p^2+p-1 conjugacy classes, p of which are of size 1 (i.e., are central classes) and $p^2 - 1$ of which each have size p. Deduce also that the classes of size p are precisely the nonidentity cosets of the center of P (i.e., if $x \in P - Z(P)$ then the conjugacy class of x is the set of p elements in the coset $xZ(P)$).
   (d) Prove that if χ is an irreducible character of degree p then the representation affording χ is faithful.
   (e) Fix a generator, z, of the center of P and let ϵ be a fixed primitive p^{th} root of 1 in $\mathbb{C}$. Prove that if χ is an irreducible character of degree p then $\chi(z) = p\epsilon^i$ for some $i \in \{1, 2, \ldots, p-1\}$. Prove further that $\chi(x) = 0$ for all $x \in P - Z(P)$. (Note then that the degree p characters are all algebraically conjugate.) [Use the same reasoning as in the construction of the character table of Q_8.]
   (f) Prove that for each $i \in \{1, 2, \ldots, p-1\}$ there is a unique irreducible character χ_i of degree p such that $\chi_i(z) = p\epsilon^i$. Deduce that the character table of P is uniquely determined, and describe it. [Recall from Section 6.1 that regardless of the isomorphism type, $P' = Z(P)$ and $P/P' \cong Z_p \times Z_p$. From this one can write out the degree 1 characters. Part (e) describes the degree p characters.]

19.2 THEOREMS OF BURNSIDE AND HALL

In this section we give a "theoretical" application of character theory: Burnside's $p^a q^b$ Theorem. We also prove Philip Hall's characterization of finite solvable groups, which is a group-theoretic proof relying on Burnside's Theorem as the first step in its induction.

Burnside's Theorem

The following result was proved by Burnside in 1904. Although purely group-theoretic proofs of it were discovered recently (see Theorem 2.8 in *Finite Groups III* by B. Huppert and N. Blackburn, Springer-Verlag, 1982) the original proof by Burnside presented here is very accessible, elegant, and quite brief (given our present knowledge of representation theory).

Theorem 1. (Burnside) For p and q primes, every group of order $p^a q^b$ is solvable.

Before undertaking the proof of Burnside's Theorem itself we establish some results of a general nature. An easy consequence of these preliminary propositions is that the degrees of the irreducible characters of any finite group divide its order. The particular results that lead directly to the proof of Burnside's Theorem appear in Lemmas 6 and 7.

It follows quite easily that a counterexample to Burnside's Theorem of minimal order is a non-abelian simple group, and it is these two character-theoretic lemmas that give the contradiction by proving the existence of a normal subgroup.

We first recall from Section 15.3 the definition of algebraic integers.

Definition. An element $\alpha \in \mathbb{C}$ is called an *algebraic integer* if it is a root of a monic polynomial with coefficients from $\mathbb{Z}$.

The basic results needed for the proof of Burnside's Theorem are:

Proposition 2. Let $\alpha \in \mathbb{C}$.

(1) The following are equivalent:

(i) α is an algebraic integer,

(ii) α is algebraic over $\mathbb{Q}$ and the minimal polynomial of α over $\mathbb{Q}$ has integer coefficients, and

(iii) $\mathbb{Z}[\alpha]$ is a finitely generated $\mathbb{Z}$-module (where $\mathbb{Z}[\alpha]$ is the subring of $\mathbb{C}$ generated by $\mathbb{Z}$ and α, i.e., is the ring of all $\mathbb{Z}$-linear combinations of nonnegative powers of α).

(2) The algebraic integers in $\mathbb{C}$ form a ring and the algebraic integers in $\mathbb{Q}$ are the elements of $\mathbb{Z}$.

Proof: These are established in Section 15.3. (The portion of Section 15.3 consisting of integral extensions and properties of algebraic integers may be read independently from the rest of Chapter 15.)

Corollary 3. For every character ψ of the finite group G, $\psi(x)$ is an algebraic integer for all $x \in G$.

Proof: By Proposition 14 in Section 18.3, $\psi(x)$ is a sum of roots of 1. Each root of 1 is an algebraic integer, so the result follows immediately from Proposition 2(2).

We shall also need some preliminary character-theoretic lemmas before beginning the main proof. Adopt the following notation for the arbitrary finite group G: $\chi_1, \dots, \chi_r$ are the distinct irreducible (complex) characters of G, $\mathcal{K}_1, \dots, \mathcal{K}_r$ are the conjugacy classes of G and φ_i is an irreducible matrix representation whose character is χ_i for each i.

Proposition 4. Define the complex valued function ω_i on $\{\mathcal{K}_1, \dots, \mathcal{K}_r\}$ for each i by

$$\omega_i(\mathcal{K}_j) = \frac{|\mathcal{K}_j|\chi_i(g)}{\chi_i(1)}$$

where g is any element of $\mathcal{K}_j$. Then $\omega_i(\mathcal{K}_j)$ is an algebraic integer for all i and j.

Proof: We first prove that if I is the identity matrix, then

$$\sum_{g \in \mathcal{K}_j} \varphi_i(g) = \omega_i(\mathcal{K}_j) I. \tag{19.1}$$

To see this let X be the left hand side of (1). As we saw in Section 18.2, each $x \in G$ acting by conjugation permutes the elements of $\mathcal{K}_j$ and so X commutes with $\varphi_i(g)$ for all g. By Schur's Lemma (Exercise 18 in Section 18.1) X is a scalar matrix:

$$X = \alpha I \qquad \text{for some } \alpha \in \mathbb{C}.$$

It remains to show that $\alpha = \omega_i(\mathcal{K}_j)$. But

$$\operatorname{tr} X = \sum_{g \in \mathcal{K}_j} \operatorname{tr} \varphi_i(g) = \sum_{g \in \mathcal{K}_j} \chi_i(g) = |\mathcal{K}_j|\chi_i(g).$$

Thus $\alpha\chi_i(1) = \operatorname{tr} X = |\mathcal{K}_j|\chi_i(g)$, as needed to establish (1).

Now let g be a fixed element of $\mathcal{K}_s$ and define a_{ijs} to be the number of ordered pairs g_i, g_j with $g_i \in \mathcal{K}_i$, $g_j \in \mathcal{K}_j$ and $g_i g_j = g$. Notice that a_{ijs} is an integer. It is independent of the choice of g in $\mathcal{K}_s$ because if $x^{-1}gx$ is a conjugate of g, every ordered pair g_i, g_j whose product is g gives rise to an ordered pair $x^{-1}g_i x$, $x^{-1}g_j x$ whose product is $x^{-1}gx$ (and vice versa).

Next we prove that for all $i, j, t \in \{1, \dots, r\}$

$$\omega_t(\mathcal{K}_i)\omega_t(\mathcal{K}_j) = \sum_{s=1}^{r} a_{ijs}\omega_t(\mathcal{K}_s). \tag{19.2}$$

To see this note that by (1), the left hand side of (2) is the diagonal entry of the scalar matrix on the left of the following equation:

$$\begin{aligned}
\Big(\sum_{g \in \mathcal{K}_i} \varphi_t(g)\Big)\Big(\sum_{g \in \mathcal{K}_j} \varphi_t(g)\Big) &= \sum_{g_i \in \mathcal{K}_i}\sum_{g_j \in \mathcal{K}_j} \varphi_t(g_i g_j) \\
&= \sum_{s=1}^{r}\sum_{g \in \mathcal{K}_s} a_{ijs}\varphi_t(g) \\
&= \sum_{s=1}^{r} a_{ijs} \sum_{g \in \mathcal{K}_s} \varphi_t(g) && \text{(since } a_{ijs} \text{ is independent of } g \in \mathcal{K}_s) \\
&= \sum_{s=1}^{r} a_{ijs}\omega_t(\mathcal{K}_s) I && \text{(by (1)).}
\end{aligned}$$

Comparing entries of these scalar matrices gives (2).

Now (2) implies that the subring of $\mathbb{C}$ generated by $\mathbb{Z}$ and $\omega_t(\mathcal{K}_1), \dots, \omega_t(\mathcal{K}_r)$ is a finitely generated $\mathbb{Z}$-module for each $t \in \{1, \dots, r\}$ (it is generated as a $\mathbb{Z}$-module by $1, \omega_t(\mathcal{K}_1), \dots, \omega_t(\mathcal{K}_r)$). Since $\mathbb{Z}$ is a Principal Ideal Domain the submodule $\mathbb{Z}[\omega_t(\mathcal{K}_i)]$ is also a finitely generated $\mathbb{Z}$-module, hence $\omega_t(\mathcal{K}_i)$ is an algebraic integer by Proposition 2. This completes the proof.

Corollary 5. The degree of each complex irreducible representation of a finite group G divides the order of G, i.e., $\chi_i(1) \mid |G|$ for $i = 1, 2, \dots, r$.

Proof: Under the notation of Proposition 4 and with $g_j \in \mathcal{K}_j$ we have

$$\begin{aligned}\frac{|G|}{\chi_i(1)} &= \frac{|G|}{\chi_i(1)}(\chi_i, \chi_i)\\ &= \sum_{j=1}^{r} \frac{|\mathcal{K}_j|\chi_i(g_j)\overline{\chi_i(g_j)}}{\chi_i(1)}\\ &= \sum_{j=1}^{r} \omega_i(\mathcal{K}_j)\overline{\chi_i(g_j)}.\end{aligned}$$

The right hand side is an algebraic integer and the left hand side is rational, hence is an integer. This proves the corollary.

The next two lemmas lead directly to Burnside's Theorem.

Lemma 6. If G is any group that has a conjugacy class $\mathcal{K}$ and an irreducible matrix representation φ with character χ such that $(|\mathcal{K}|, \chi(1)) = 1$, then for $g \in \mathcal{K}$ either $\chi(g) = 0$ or $\varphi(g)$ is a scalar matrix.

Proof: By hypothesis there exist $s, t \in \mathbb{Z}$ such that $s|\mathcal{K}| + t\chi(1) = 1$. Thus

$$s|\mathcal{K}|\chi(g) + t\chi(1)\chi(g) = \chi(g).$$

Divide both sides of this by $\chi(1)$ and note that by Corollary 3 and Proposition 4 both $\chi(g)$ and $\dfrac{|\mathcal{K}|\chi(g)}{\chi(1)}$ are algebraic integers, hence so is $\dfrac{\chi(g)}{\chi(1)}$. Let $a_1 = \dfrac{\chi(g)}{\chi(1)}$ and let $a_1, a_2, \ldots, a_n$ be all its algebraic conjugates over $\mathbb{Q}$ (i.e., the roots of the minimal polynomial of a_1 over $\mathbb{Q}$). Since a_1 is a sum of $\chi(1)$ roots of 1 divided by the integer $\chi(1)$, each a_i is also a sum of $\chi(1)$ roots of 1 divided by $\chi(1)$. Thus a_i has complex absolute value ≤ 1 for all i. Now $b = \prod_{i=1}^{n} a_i \in \mathbb{Q}$ and b is an algebraic integer ($\pm b$ is the constant term of the irreducible polynomial of a_1), hence $b \in \mathbb{Z}$. But

$$|b| = \prod_{i=1}^{n} |a_i| \leq 1,$$

so $b = 0, \pm 1$. Since all a_i's are conjugate, $b = 0 \Leftrightarrow a_1 = 0 \Leftrightarrow \chi(g) = 0$. Also, $b = \pm 1 \Leftrightarrow |a_i| = 1$ for all i. Thus either $\chi(g) = 0$ or $|\chi(g)| = \chi(1)$. In the former situation the lemma is established, so assume $|\chi(g)| = \chi(1)$.

Let φ_1 be a matrix representation equivalent to φ in which $\varphi_1(g)$ is a diagonal matrix:

$$\varphi_1(g) = \begin{pmatrix} \epsilon_1 & & & \\ & \epsilon_2 & & \\ & & \ddots & \\ & & & \epsilon_n \end{pmatrix}.$$

Thus $\chi(g) = \epsilon_1 + \cdots + \epsilon_n$. By the triangle inequality if $\epsilon_i \neq \epsilon_j$ for any i, j, then $|\epsilon_1 + \cdots + \epsilon_n| < n = \chi(1)$. Since this is not the case we must have $\varphi_1(g) = \epsilon I$ (where $\epsilon = \epsilon_i$ for all i). Since scalar matrices are similar only to themselves, $\varphi(g) = \epsilon I$ as well. This completes the proof.

Lemma 7. If $|\mathcal{K}|$ is a power of a prime for some nonidentity conjugacy class $\mathcal{K}$ of G, then G is not a non-abelian simple group.

Proof: Suppose to the contrary that G is a non-abelian simple group and let $|\mathcal{K}| = p^c$. Let $g \in \mathcal{K}$. If $c = 0$ then $g \in Z(G)$, contrary to a non-abelian simple group having a trivial center. As above, let $\chi_1, \dots, \chi_r$ be all the irreducible characters of G with χ_1 the principal character and let ρ be the regular character of G. By decomposing ρ into irreducibles we obtain

$$0 = \rho(g) = 1 + \sum_{i=2}^{r} \chi_i(1)\chi_i(g). \tag{19.3}$$

If $p \mid \chi_j(1)$ for every $j > 1$ with $\chi_j(g) \neq 0$, then write $\chi_j(1) = pd_j$. In this case (3) becomes

$$0 = 1 + p\sum_j d_j\chi_j(g).$$

Thus $\sum_j d_j\chi_j(g) = -1/p$ is an algebraic integer, a contradiction. This proves there is some j such that p does not divide $\chi_j(1)$ and $\chi_j(g) \neq 0$. If φ is a representation whose character is χ_j, then φ is faithful (because G is assumed to be simple) and, by Lemma 6, $\varphi(g)$ is a scalar matrix. Since $\varphi(g)$ commutes with all matrices, $\varphi(g) \in Z(\varphi(G))$. This forces $g \in Z(G)$, contrary to G being a non-abelian simple group. The proof of the lemma is complete.

We now prove Burnside's Theorem. Let G be a group of order $p^a q^b$ for some primes p and q. If $p = q$ or if either exponent is 0 then G is nilpotent hence solvable. Thus we may assume this is not the case. Proceeding by induction let G be a counterexample of minimal order. If G has a proper, nontrivial normal subgroup N, then by induction both N and G/N are solvable, hence so is G (cf. Section 3.4 or Proposition 6.10). Thus we may assume G is a non-abelian simple group. Let $P \in Syl_p(G)$. By Theorem 8 of Chapter 4 there exists $g \in Z(P)$ with $g \neq 1$. Since $P \leq C_G(g)$, the order of the conjugacy class of g (which equals $|G : C_G(g)|$) is prime to p, i.e., is a power of q. This violates Lemma 7 and so completes the proof of Burnside's Theorem.

Philip Hall's Theorem

Recall that a subgroup of a finite group is called a *Hall subgroup* if its order and index are relatively prime. For any subgroup H of a group G a subgroup K such that $G = HK$ and $H \cap K = 1$ is called a *complement* to H in G.

Theorem 8. (P. Hall) Let G be a group of order $p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_t^{\alpha_t}$ where $p_1, \dots, p_t$ are distinct primes. If for each $i \in \{1, \dots, t\}$ there exists a subgroup H_i of G with $|G : H_i| = p_i^{\alpha_i}$, then G is solvable.

Hall's Theorem can also be phrased: if for each $i \in \{1, \dots, t\}$ a Sylow p_i-subgroup of G has a complement, then G is solvable. The converse to Hall's Theorem is also true — this was Exercise 33 in Section 6.1.

We shall first need some elementary lemmas.

Lemma 9. If G is solvable of order > 1, then there exists $P \trianglelefteq G$ with P a nontrivial p-group for some prime p.

Proof: This is a special case of the exercise on minimal normal subgroups of solvable groups at the end of Section 6.1. One can see this easily by letting P be a nontrivial Sylow subgroup of the last nontrivial term, $G^{(n-1)}$, in the derived series of G (where G has solvable length n). In this case $G^{(n-1)}$ is abelian so P is a characteristic subgroup of $G^{(n-1)}$, hence is normal in G.

Lemma 10. Let G be a group of order $p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_t^{\alpha_t}$ where $p_1, \dots, p_t$ are distinct primes. Suppose there are subgroups H and K of G such that for each $i \in \{1, \dots, t\}$, either $p_i^{\alpha_i}$ divides $|H|$ or $p_i^{\alpha_i}$ divides $|K|$. Then $G = HK$ and $|H \cap K| = (|H|, |K|)$.

Proof: Fix some $i \in \{1, \dots, t\}$ and suppose first that $p_i^{\alpha_i}$ divides the order of H. Since HK is a disjoint union of right cosets of H and each of these right cosets has order equal to $|H|$, it follows that $p_i^{\alpha_i}$ divides $|HK|$. Similarly, if $p_i^{\alpha_i}$ divides $|K|$, since HK is a disjoint union of left cosets of K, again $p_i^{\alpha_i}$ divides $|HK|$. Thus $|G| \mid |HK|$ and so $G = HK$. Since

$$|HK| = \frac{|H||K|}{|H \cap K|},$$

it follows that $|H \cap K| = (|H|, |K|)$.

We now begin the proof of Hall's Theorem, proceeding by induction on $|G|$. Note that if $t = 1$ the hypotheses are trivially satisfied for any group ($H_1 = 1$) and if $t = 2$ the hypotheses are again satisfied for any group by Sylow's Theorem (H_1 is a Sylow p_2-subgroup of G and H_2 is a Sylow p_1-subgroup of G). If $t = 1$, G is nilpotent, hence solvable and if $t = 2$, G is solvable by Burnside's Theorem. Assume therefore that $t \geq 3$.

Fix i and note that by the preceding lemma, for all $j \in \{1, \dots, t\} - \{i\}$,

$$|H_i : H_i \cap H_j| = p_j^{\alpha_j}.$$

Thus every Sylow p_j-subgroup of H_i has a complement in H_i: $H_j \cap H_i$. By induction H_i is solvable.

By Lemma 9 we may choose $P \trianglelefteq H_1$ with $|P| = p_i^a > 1$ for some $i > 1$. Since $t \geq 3$ there exists an index $j \in \{1, \dots, t\} - \{1, i\}$. By Lemma 10

$$|H_1 \cap H_j| = p_2^{\alpha_2} \cdots p_{j-1}^{\alpha_{j-1}} p_{j+1}^{\alpha_{j+1}} \cdots p_t^{\alpha_t}.$$

Thus $H_1 \cap H_j$ contains a Sylow p_i-subgroup of H_1. Since P is a normal p_i-subgroup of H_1, P is contained in every Sylow p_i-subgroup of H_1 and so $P \leq H_1 \cap H_j$. By Lemma 10, $G = H_1 H_j$ so each $g \in G$ may be written $g = h_1 h_j$ for some $h_1 \in H_1$ and $h_j \in H_j$. Then

$$gH_jg^{-1} = (h_1h_j)H_j(h_1h_j)^{-1} = h_1H_jh_1^{-1}$$

and so

$$\bigcap_{g \in G} gH_jg^{-1} = \bigcap_{h_1 \in H_1} h_1H_jh_1^{-1}.$$

Now $P \leq H_j$ and $h_1 P h_1^{-1} = P$ for all $h_1 \in H_1$. Thus

$$1 \neq P \leq \bigcap_{h_1 \in H_1} h_1 H_j h_1^{-1}.$$

Thus $N = \bigcap_{g \in G} g H_j g^{-1}$ is a nontrivial, proper normal subgroup of G. It follows that both N and G/N satisfy the hypotheses of the theorem (cf. the exercises in Section 3.3). Both N and G/N are solvable by induction, so G is solvable. This completes the proof of Hall's Theorem.

EXERCISES

1. Show that every character of the symmetric group S_n is integer valued, for all n (i.e., $\psi(g) \in \mathbb{Z}$ for all $g \in S_n$ and all characters ψ of S_n). [See Exercise 22 in Section 18.3.]
2. Let G be a finite group with the property that every maximal subgroup has either prime or prime squared index. Prove that G is solvable. (The simple group $GL_3(\mathbb{F}_2)$ has the property that every maximal subgroup has index either 7 or 8, i.e., either prime or prime cubed index — cf. Section 6.2.). [Let p be the largest prime dividing $|G|$ and let P be a Sylow p-subgroup of G. If $P \trianglelefteq G$, apply induction to G/P. Otherwise let M be a maximal subgroup containing $N_G(P)$. Use Exercise 51 in Section 4.5 to show that $p = 3$ and deduce that $|G| = 2^a 3^b$.]
3. Assume G is a finite group that possesses an abelian subgroup H whose index is a power of a prime. Prove that G is solvable.
4. Repeat the preceding exercise with the word "abelian" replaced by "nilpotent."
5. Use the ideas in the proof of Philip Hall's Theorem to prove Burnside's $p^a q^b$ Theorem in the special case when all Sylow subgroups are abelian (without use of character theory.)

19.3 INTRODUCTION TO THE THEORY OF INDUCED CHARACTERS

Let G be a finite group, let H be a subgroup of G and let φ be a representation of the subgroup H over an arbitrary field F. In this section we show how to obtain a representation of G, called the induced representation, from the representation φ of its subgroup. We also determine a formula for the character of this induced representation, the induced character, in terms of the character of φ and we illustrate this formula by computing some induced characters in specific groups. Finally, we apply the theory of induced characters to prove that there are no simple groups of order $3^3 \cdot 7 \cdot 13 \cdot 409$, a group order which was discussed at the end of Section 6.2 in the context of the existence problem for simple groups. The theory of induced representations and induced characters marks the beginning of more advanced representation theory. This section is intended as an introduction rather than as a comprehensive treatment, and the results we have included were chosen to serve this purpose.

First observe that it may not be possible to extend a representation φ of the subgroup H to a representation Φ of G in such a way that $\Phi|_H = \varphi$. For example, $A_3 \leq S_3$ and A_3 has a faithful representation of degree 1 (cf. Section 1). Since every degree 1 representation of S_3 contains $A_3 = S_3'$ in its kernel, this representation of A_3 cannot be extended to a representation of S_3. For another example of a representation of a subgroup which cannot be extended to the whole group take G to be any simple group and let φ be any representation of H with the property that $\ker\varphi$ is a proper, nontrivial normal subgroup of H. If φ extended to a representation Φ of G then the kernel of Φ would be a proper, nontrivial subgroup of G, contrary to G being a simple group. We shall see that the method of induced characters produces a representation Φ of G from a given representation φ of its subgroup H but that $\Phi|_H \neq \varphi$ in general (indeed, unless $H = G$ the degree of Φ will be greater than the degree of φ).

We saw in Example 5 following Corollary 9 in Section 10.4 that because FH is a subring of FG, the ring FG is an (FG, FH)-bimodule; and so for any left FH-module V, the abelian group $FG \otimes_{FH} V$ is a left FG-module (called the extension of scalars from FH to FG for V). In the representation theory of finite groups this extension is given a special name.

Definition. Let H be a subgroup of the finite group G and let V be an FH-module affording the representation φ of H. The FG-module $FG \otimes_{FH} V$ is called the *induced module* of V and the representation of G it affords is called the *induced representation* of φ. If ψ is the character of φ then the character of the induced representation is called the *induced character* and is denoted by $\mathrm{Ind}_H^G(\psi)$.

Theorem 11. Let H be a subgroup of the finite group G and let $g_1, \dots, g_m$ be representatives for the distinct left cosets of H in G. Let V be an FH-module affording the matrix representation φ of H of degree n. The FG-module $W = FG \otimes_{FH} V$ has dimension nm over F and there is a basis of W such that W affords the matrix representation Φ defined for each $g \in G$ by

$$\Phi(g) = \begin{pmatrix} \varphi(g_1^{-1}gg_1) & \cdots & \varphi(g_1^{-1}gg_m) \\ \vdots & \vdots & \vdots \\ \varphi(g_m^{-1}gg_1) & \cdots & \varphi(g_m^{-1}gg_m) \end{pmatrix}$$

where each $\varphi(g_i^{-1}gg_j)$ is an $n \times n$ block appearing in the i, j block position of $\Phi(g)$, and where $\varphi(g_i^{-1}gg_j)$ is defined to be the zero block whenever $g_i^{-1}gg_j \notin H$.

Proof: First note that FG is a free right FH-module:

$$FG = g_1FH \oplus g_2FH \oplus \cdots \oplus g_mFH.$$

Since tensor products commute with direct sums (Theorem 17, Section 10.4), as abelian groups we have

$$W = FG \otimes_{FH} V \cong (g_1 \otimes V) \oplus (g_2 \otimes V) \oplus \cdots \oplus (g_m \otimes V).$$

Since F is in the center of FG it follows that this is an F-vector space isomorphism as well. Thus if $v_1, v_2, \dots, v_n$ is a basis of V affording the matrix representation φ, then $\{g_i \otimes v_j \mid 1 \le i \le m,\ 1 \le j \le n\}$ is a basis of W. This shows the dimension of W is mn. Order the basis into m sets, each of size n as

$$g_1 \otimes v_1, g_1 \otimes v_2, \dots, g_1 \otimes v_n, g_2 \otimes v_1, \dots, g_2 \otimes v_n, \dots\dots, g_m \otimes v_n.$$

We compute the matrix representation $\Phi(g)$ of each g acting on W with respect to this basis. Fix j and g, and let $gg_j = g_i h$ for some index i and some $h \in H$. Then for every k

$$\begin{aligned} g(g_j \otimes v_k) = (gg_j) \otimes v_k &= g_i \otimes hv_k \\ &= \sum_{t=1}^{n} a_{tk}(h)(g_i \otimes v_t) \end{aligned}$$

where a_{tk} is the t, k coefficient of the matrix of h acting on V with respect to the basis $\{v_1, \ldots, v_n\}$. In other words, the action of g on W maps the j^{th} block of n basis vectors of W to the i^{th} block of basis vectors, and then has the matrix $\varphi(h)$ on that block. Since $h = g_i^{-1} gg_j$, this describes the block matrix $\Phi(g)$ of the theorem, as needed.

Corollary 12. In the notation of Theorem 11

(1) if ψ is the character afforded by V then the induced character is given by

$$\mathrm{Ind}_H^G(\psi)(g) = \sum_{i=1}^{m} \psi(g_i^{-1} gg_i)$$

where $\psi(g_i^{-1}gg_i)$ is defined to be 0 if $g_i^{-1}gg_i \notin H$, and

(2) $\mathrm{Ind}_H^G(\psi)(g) = 0$ if g is not conjugate in G to some element of H. In particular, if H is a normal subgroup of G then $\mathrm{Ind}_H^G(\psi)$ is zero on all elements of $G - H$.

Remark: Since the character ψ of H is constant on the conjugacy classes of H we have $\psi(g) = \psi(h^{-1}gh)$ for all $h \in H$. As h runs over all elements of H, xh runs over all elements of the coset xH. Thus the formula for the induced character may also be written

$$\mathrm{Ind}_H^G(\psi)(g) = \frac{1}{|H|} \sum_{x \in G} \psi(x^{-1}gx)$$

where the elements x in each fixed coset give the same character value $|H|$ times (which accounts for the factor of $1/|H|$), and again $\psi(x^{-1}gx) = 0$ if $x^{-1}gx \notin H$.

Proof: From the matrix of g computed above, the blocks $\varphi(g_i^{-1}gg_i)$ down the diagonal of $\Phi(g)$ are zero except when $g_i^{-1}gg_i \in H$. Thus the trace of the block matrix $\Phi(g)$ is the sum of the traces of the matrices $\varphi(g_i^{-1}gg_i)$ for which $g_i^{-1}gg_i \in H$. Since the trace of $\varphi(g_i^{-1}gg_i)$ is $\psi(g_i^{-1}gg_i)$, part (1) holds.

If $g_i^{-1}gg_i \notin H$ for all coset representatives g_i then each term in the sum for $\mathrm{Ind}_H^G(\psi)(g)$ is zero. In particular, if g is not in the normal subgroup H then neither is any conjugate of g, so $\mathrm{Ind}_H^G(\psi)$ is zero on g.

Examples

(1) Let $G = D_{12} = \langle r, s \mid r^6 = s^2 = 1,\ rs = sr^{-1} \rangle$ be the dihedral group of order 12 and let $H = \{1, s, r^3, sr^3\}$, so that H is isomorphic to the Klein 4-group and $|G : H| = 3$. Following the notation of Theorem 11 we exhibit the matrices for r and s of the induced

representation of a specific representation φ of H. Let the representation of H on a 2-dimensional vector space over $\mathbb{Q}$ with respect to some basis v_1, v_2 be given by

$$\varphi(s) = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} = A, \quad \varphi(r^3) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = B, \quad \varphi(sr^3) = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = C,$$

so $n = 2$, $m = 3$ and the induced representation Φ has degree $nm = 6$. Fix representatives $g_1 = 1$, $g_2 = r$, and $g_3 = r^2$ for the left cosets of H in G, so that $g_k = r^{k-1}$. Then

$$g_i^{-1} r g_j = r^{-(i-1)+1+(j-1)} = r^{j-i+1}, \text{ and}$$
$$g_i^{-1} s g_j = sr^{(i-1)+(j-1)} = sr^{i+j-2}.$$

Thus the 6×6 matrices for the induced representation are seen to be

$$\Phi(r) = \begin{pmatrix} 0 & 0 & B \\ I & 0 & 0 \\ 0 & I & 0 \end{pmatrix} \qquad \Phi(s) = \begin{pmatrix} A & 0 & 0 \\ 0 & 0 & C \\ 0 & C & 0 \end{pmatrix}$$

where the 2×2 matrices A, B and C are given above, I is the 2×2 identity matrix and 0 denotes the 2×2 zero matrix.

(2) If H is any subgroup of G and ψ_1 is the principal character of H, then $\mathrm{Ind}_H^G(\psi_1)(g)$ counts 1 for each coset representative g_i such that $g_i^{-1} g g_i \in H$. Since $g_i^{-1} g g_i \in H$ if and only if g fixes the left coset $g_i H$ under left multiplication, $\mathrm{Ind}_H^G(\psi_1)(g)$ is the number of points fixed by g in the permutation representation of g on the left cosets of H. Thus by Example 3 of Section 18.3 we see that: *if ψ_1 is the principal character of H then $\mathrm{Ind}_H^G(\psi_1)$ is the permutation character on the left cosets of H in G.* In the special case when $H = 1$, this implies *if χ_1 is the principal character of the trivial subgroup $H = 1$ then $\mathrm{Ind}_1^G(\chi_1)$ is the regular character of G.* This also shows that an induced character is not, in general, irreducible even if the character from which it is induced is irreducible.

(3) Let $G = S_3$ and let ψ be a nonprincipal linear character of $A_3 = \langle x \rangle$, so that $\psi(x) = \zeta$, for some primitive cube root of unity ζ (the character tables of $A_3 = Z_3$ and S_3 appear in Section 1). Let $\Psi = \mathrm{Ind}_{A_3}^{S_3}(\psi)$. Thus Ψ has degree $1 \cdot |S_3 : A_3| = 2$ and, by the corollary, Ψ is zero on all transpositions. If y is any transposition then $1, y$ is a set of left coset representatives of A_3 in S_3 and $y^{-1}xy = x^2$. Thus $\Psi(x) = \psi(x) + \psi(x^2)$ equals $\zeta + \zeta^2 = -1$. This shows that if ψ is either of the two nonprincipal irreducible characters of A_3 then the induced character of ψ is the (unique) irreducible character of S_3 of degree 2. In particular, different characters of a subgroup may induce the same character of the whole group.

(4) Let $G = D_8$ have its usual generators and relations and let $H = \langle s \rangle$. Let ψ be the nonprincipal irreducible character of H and let $\Psi = \mathrm{Ind}_H^G(\psi)$. Pick left coset representatives $1, r, r^2, r^3$ for H. By Theorem 11, $\Psi(1) = 4$. Since $\psi(s) = -1$, one computes directly that $\Psi(s) = -2$. By Corollary 12(2) we obtain $\Psi(r) = \Psi(r^2) = \Psi(sr) = 0$. In the notation of the character table of D_8 in Section 1, by the orthogonality relations we obtain $\Psi = \chi_2 + \chi_4 + \chi_5$ (which may be checked by inspection).

For the remainder of this section the field F is taken to be the complex numbers: $F = \mathbb{C}$.

Before concluding with an application of induced characters to simple groups we compute the characters of an important class of groups.

Definition. A finite group G is called a *Frobenius group* with *Frobenius kernel* Q if Q is a proper, nontrivial normal subgroup of G and $C_G(x) \leq Q$ for all nonidentity elements x of Q.

In view of the application to simple groups mentioned at the beginning of this section we shall restrict attention to Frobenius groups G of order $q^a p$, where p and q are distinct primes, such that the Frobenius kernel Q is an elementary abelian q-group of order q^a and the cyclic group G/Q acts irreducibly by conjugation on Q. In other words, we shall assume Q is a direct product of cyclic groups of order q and the only normal subgroups of G that are contained in Q are 1 and Q, i.e., Q is a minimal normal subgroup of G. For example, A_4 is a Frobenius group of this type with Frobenius kernel V_4, its Sylow 2-subgroup. Also, if p and q are distinct primes with $p < q$ and G is a non-abelian group of order pq (one always exists if $p \mid q-1$) then G is a Frobenius group whose Frobenius kernel is its Sylow q-subgroup (which is normal by Sylow's Theorem). We essentially determine the character table of these Frobenius groups. Analogous results on more general Frobenius groups appear in the exercises.

Proposition 13. Let G be a Frobenius group of order $q^a p$, where p and q are distinct primes, such that the Frobenius kernel Q is an elementary abelian q-group of order q^a and the cyclic group G/Q acts irreducibly by conjugation on Q. Then the following hold:

(1) $G = QP$ where P is a Sylow p-subgroup of G. Every nonidentity element of G has order p or q. Every element of order p is conjugate to an element of P and every element of order q belongs to Q. The nonidentity elements of P represent the $p-1$ distinct conjugacy classes of elements of order p and each of these classes has size q^a. There are $(q^a-1)/p$ distinct conjugacy classes of elements of order q and each of these classes has size p.

(2) $G' = Q$ so the number of degree 1 characters of G is p and every degree 1 character contains Q in its kernel.

(3) If ψ is any nonprincipal irreducible character of Q, then $\mathrm{Ind}_Q^G(\psi)$ is an irreducible character of G. Moreover, every irreducible character of G of degree > 1 is equal to $\mathrm{Ind}_Q^G(\psi)$ for some nonprincipal irreducible character ψ of Q. Every irreducible character of G has degree either 1 or p and the number of irreducible characters of degree p is $(q^a-1)/p$.

Proof: Note that QP equals G by order consideration. By definition of a Frobenius group and because Q is abelian, $C_G(h) = Q$ for every nonidentity element h of Q. If x were an element of order pq, then x^p would be an element of order q, hence would lie in the unique Sylow q-subgroup Q of G. But then x would commute with x^p and so x would belong to $C_G(x^p) = Q$, a contradiction. Thus G has no elements of order pq. By Sylow's Theorem every element of order p is conjugate to an element of P and every element of order q lies in Q. No two distinct elements of P are conjugate in G because if $g^{-1}xg = y$ for some $x, y \in P$ then $\overline{g^{-1}xg} = \overline{y}$ in the abelian group $\overline{G} = G/Q$ and so $\overline{x} = \overline{y}$. Then $x = y$ because $\overline{P} \cong P$. Thus there are exactly $p-1$ conjugacy classes of elements of order p and these are represented by the nonidentity elements of P. If x is a nonidentity element of P, then $C_G(x) = P$ and so the conjugacy class of

x consists of $|G : P| = q^a$ elements. Finally, if h is a nonidentity element of Q, then $C_G(h) = Q$ and the conjugacy class of h is $\{h, h^x, \dots, h^{x^{p-1}}\}$, where $P = \langle x \rangle$. This proves all parts of (1).

Since G/Q is abelian, $G' \le Q$. Since G is non-abelian and Q is, by hypothesis, a minimal normal subgroup of G we must have $G' = Q$. Part (2) now follows from Corollary 11 in Section 18.2.

Let ψ be a nonprincipal irreducible character of Q and let $\Psi = \mathrm{Ind}_Q^G(\psi)$. We use the orthogonality relations to show that Ψ is irreducible. Let $1, x, \dots, x^{p-1}$ be coset representatives for Q in G. By Corollary 12, Ψ is zero on $G - Q$ so

$$\begin{aligned}
\|\,\Psi\,\|^2 &= \frac{1}{|G|}\sum_{h\in Q}\Psi(h)\overline{\Psi(h)}\\
&= \frac{1}{|G|}\sum_{h\in Q}\sum_{i=0}^{p-1}\psi(x^i h x^{-i})\overline{\psi(x^i h x^{-i})}\\
&= \frac{p}{|G|}\sum_{h\in Q}\psi(h)\overline{\psi(h)}\\
&= \frac{p|Q|}{|G|} = 1,
\end{aligned}$$

where the second line follows from the definition of the induced character Ψ, the third line follows because each element of Q appears exactly p times in the sum in the second line, and the last line follows from the first orthogonality relation in Q because ψ is an irreducible character of Q. This proves Ψ is an irreducible character of G.

We prove that every irreducible character of G of degree > 1 is the induced character of some nonprincipal degree 1 character of Q by counting the number of distinct irreducible characters of G obtained this way. By parts (1) and (2) the number of irreducible characters of G (= the number of conjugacy classes) is $p + (q^a - 1)/p$ and the number of degree 1 characters is p. Thus the number of irreducible characters of G of degree > 1 is $(q^a - 1)/p$. The group P acts on the set $\mathcal{C}$ of nonprincipal irreducible characters of Q as follows: for each $\psi \in \mathcal{C}$ and each $x \in P$ let ψ^x be defined by

$$\psi^x(h) = \psi(xhx^{-1}) \qquad \text{for all } h \in Q.$$

Since ψ is a nontrivial homomorphism from Q into $\mathbb{C}^\times$ (recall that all irreducible characters of the abelian group Q have degree 1) it follows easily that ψ^x is also a homomorphism. Thus $\psi^x \in \mathcal{C}$ and so P permutes the elements of $\mathcal{C}$. Now let x be a generator for the cyclic group P. Then $1, x, \dots, x^{p-1}$ are representatives for the left cosets of Q in G. By Corollary 12 applied with this set of coset representatives we see that if $\psi \in \mathcal{C}$ then the value of $\mathrm{Ind}_Q^G(\psi)$ on any element h of Q is given by the sum $\psi(h) + \psi^x(h) + \cdots + \psi^{x^{p-1}}(h)$. Thus when the induced character $\mathrm{Ind}_Q^G(\psi)$ is restricted to Q it decomposes into irreducible characters of Q as

$$\mathrm{Ind}_Q^G(\psi)|_Q = \psi + \psi^x + \cdots + \psi^{x^{p-1}}.$$

If ψ_1 and ψ_2 are in different orbits of the action of P on $\mathcal{C}$ then the induced characters $\mathrm{Ind}_Q^G(\psi_1)$ and $\mathrm{Ind}_Q^G(\psi_2)$ restrict to distinct characters of Q (they have no irreducible constituents in common). Thus characters induced from elements of distinct orbits of P on $\mathcal{C}$ are distinct irreducible characters of G. The abelian group Q has $q^a - 1$ nonprincipal irreducible characters (i.e., $|\mathcal{C}| = q^a - 1$) and $|P| = p$ so there are at least $(q^a - 1)/p$ orbits of P on $\mathcal{C}$ and hence at least this number of distinct irreducible characters of G of degree p. Since G has exactly $(q^a - 1)/p$ irreducible characters of degree > 1, every irreducible character of G of degree > 1 must have degree p and must be an induced character from some element of $\mathcal{C}$. The proof is complete.

For the final example we shall require two properties of induced characters. These properties are listed in the next proposition and the proofs are straightforward exercises which follow easily from the formula for induced characters or from the definition of induced modules together with properties of tensor products.

Proposition 14. Let G be a group, let H be a subgroup of G and let ψ and ψ' be characters of H.

(1) *(Induction of characters is additive)* $\mathrm{Ind}_H^G(\psi + \psi') = \mathrm{Ind}_H^G(\psi) + \mathrm{Ind}_H^G(\psi')$.

(2) *(Induction of characters is transitive)* If $H \le K \le G$ then

$$\mathrm{Ind}_K^G(\mathrm{Ind}_H^K(\psi)) = \mathrm{Ind}_H^G(\psi).$$

It follows from part (1) of Proposition 14 that if $\sum_{i=1}^{s} n_i\psi_i$ is any integral linear combination of characters of H with $n_i \ge 0$ for all i then

$$\mathrm{Ind}_H^G\left(\sum_{i=1}^{s} n_i\psi_i\right) = \sum_{i=1}^{s} n_i\,\mathrm{Ind}_H^G(\psi_i). \qquad (*)$$

A class function of H of the form $\sum_{i=1}^{s} n_i\psi_i$, where the coefficients are any integers (not necessarily nonnegative) is called a *generalized character* or *virtual character* of H. For a generalized character of H we define its induced generalized character of G by equation $(*)$, allowing now negative coefficients n_i as well. In this way the function Ind_H^G becomes a group homomorphism from the additive group of generalized characters of H to the additive group of generalized characters of G (which maps characters to characters). This implies that the formula for induced characters in Corollary 12 holds also if ψ is a generalized character of H.

Application to Groups of Order $3^3 \cdot 7 \cdot 13 \cdot 409$

We now conclude with a proof of the following result:

there are no simple groups of order $3^3 \cdot 7 \cdot 13 \cdot 409$.

As mentioned at the beginning of this section, simple groups of this order were discussed at the end of Section 6.2 in the context of the existence problem for simple groups. It is possible to prove that there are no simple groups of this order by arguments involving a permutation representation of degree 819 (cf. the exercises in Section 6.2). We include a character-theoretic proof of this since the methods illustrate some important ideas in the theory of finite groups. The approach is based on M. Suzuki's seminal paper *The nonexistence of a certain type of simple group of odd order*, Proc. Amer. Math. Soc., 8(1957), pp. 686–695, which treats much more general groups. Because we are dealing with a specific group order, our arguments are simpler and numerically more explicit, yet they retain some of the key ideas of Suzuki's work. Moreover, Suzuki's paper and its successor, *Finite groups in which the centralizer of any non-identity element is nilpotent*, by W. Feit, M. Hall and J. Thompson, Math. Zeit., 74(1960), pp. 1–17, are prototypes for the lengthy and difficult Feit–Thompson Theorem (cf. Section 3.4). Our discussion also conveys some of the flavor of these fundamental papers. In particular, each of these papers follows the basic development in which the structure and embedding of the Sylow subgroups is first determined and then character theory (with heavy reliance on induced characters) is applied.

For the remainder of this section we assume G is a simple group of order $3^3 \cdot 7 \cdot 13 \cdot 409$. We list some properties of G which may be verified using the methods stemming from Sylow's Theorem discussed in Section 6.2. The details are left as exercises.

(1) Let $q_1 = 3$, let Q_1 be a Sylow 3-subgroup of G and let $N_1 = N_G(Q_1)$. Then Q_1 is an elementary abelian 3-group of order 3^3 and N_1 is a Frobenius group of order $3^3 \cdot 13$ with Frobenius kernel Q_1 and with N_1/Q_1 acting irreducibly by conjugation on Q_1.

(2) Let $q_2 = 7$, let Q_2 be a Sylow 7-subgroup of G and let $N_2 = N_G(Q_2)$. Then Q_2 is cyclic of order 7 and N_2 is the non-abelian group of order $7 \cdot 3$ (so N_2 is a Frobenius group with Frobenius kernel Q_2).

(3) Let $q_3 = 13$, let Q_3 be a Sylow 11-subgroup of G and let $N_3 = N_G(Q_3)$. Then Q_3 is cyclic of order 13 and N_3 is the non-abelian group of order $13 \cdot 3$ (so N_3 is a Frobenius group with Frobenius kernel Q_3).

(4) Let $q_4 = 409$, let Q_4 be a Sylow 409-subgroup of G and let $N_4 = N_G(Q_4)$. Then Q_4 is cyclic of order 409 and N_4 is the non-abelian group of order $409 \cdot 3$ (so N_4 is a Frobenius group with Frobenius kernel Q_4).

(5) Every nonidentity element of G has prime order and $Q_i \cap Q_i^g = 1$ for every $g \in G - N_i$, for each $i = 1, 2, 3, 4$. The nonidentity conjugacy classes of G are:
   (a) 2 classes of elements of order 3 (each of these classes has size $7 \cdot 13 \cdot 409$)
   (b) 2 classes of elements of order 7 (each of these classes has size $3^3 \cdot 13 \cdot 409$)
   (c) 4 classes of elements of order 13 (each of these classes has size $3^3 \cdot 7 \cdot 409$)
   (d) 136 classes of elements of order 409 (each of these classes has size $3^3 \cdot 7 \cdot 13$),
   and so there are 145 conjugacy classes in G.

Since each of the groups N_i is a Frobenius group satisfying the hypothesis of Proposition 13, the number of characters of N_i of degree > 1 may be read off from that proposition:

(i) N_1 has 2 irreducible characters of degree 13
(ii) N_2 has 2 irreducible characters of degree 3
(iii) N_3 has 4 irreducible characters of degree 3
(iv) N_4 has 136 irreducible characters of degree 3.

From now on, to simplify notation, for any subgroup H of G and any generalized character μ of H let

$$\mu^* = \mathrm{Ind}_H^G(\mu)$$

so a star will always denote induction from a subgroup to the whole group G and the subgroup will be clear from the context.

The following lemma is a key point in the proof. It shows how the vanishing of induced characters described in Corollary 12 (together with the *trivial intersection* property of the Sylow subgroups Q_i, namely the fact that $Q_i \cap Q_i^g = 1$ for all $g \in G - N_G(Q_i)$) may be used to relate inner products of certain generalized characters to the inner products of their induced generalized characters. For these computations it is important that the generalized characters are zero on the identity (which explains why we are considering *differences* of characters of the same degree).

Lemma 15. For any $i \in \{1, 2, 3, 4\}$ let $q = q_i$, let $Q = Q_i$, let $N = N_i$ and let $p = |N : Q|$. Let $\psi_1, \dots, \psi_4$ be any irreducible characters of N of degree p (not necessarily distinct) and let $\alpha = \psi_1 - \psi_2$ and $\beta = \psi_3 - \psi_4$. Then α and β are generalized characters of N which are zero on every element of N of order not equal to q. Furthermore, α^* and β^* are generalized characters of G which are zero on every element of G of order not equal to q and

$$(\alpha^*, \beta^*)_G = (\alpha, \beta)_N$$

(where $(\ ,\ )_H$ denotes the usual Hermitian product of class functions computed in the group H). In other words, induction from N to G is an inner product preserving map on such generalized characters α, β of N.

Proof: By Proposition 13, there are nonprincipal characters $\lambda_1, \dots, \lambda_4$ of Q of degree 1 such that $\psi_j = \mathrm{Ind}_Q^N(\lambda_j)$ for $j = 1, \dots, 4$. By Corollary 12 therefore, each ψ_j vanishes on $N - Q$, hence so do α and β. Note that since $\psi_j(1) = p$ for all j we have $\alpha(1) = \beta(1) = 0$. By the transitivity of induction, $\psi_j^* = \mathrm{Ind}_N^G(\psi_j) = \mathrm{Ind}_Q^G(\lambda_j)$ for all j. Again by Corollary 12 applied to the latter induced character we see that ψ_j^* vanishes on all elements not conjugate in G to some element of Q, hence so do both α^* and β^*. Since the induced characters ψ_j^* all have degree $|G : Q|$, the generalized characters α^* and β^* are zero on the identity. Thus α^* and β^* vanish on all elements of G which are not of order q. Finally, if $g_1, \dots, g_m$ are representatives for the left cosets of N in G with $g_1 = 1$, then because $Q \cap Q^{g_k} = 1$ for all $k > 1$ (by (5) above), it follows immediately from the formula for induced (generalized) characters that $\alpha^*(x) = \alpha(x)$ and $\beta^*(x) = \beta(x)$ for all nonidentity elements $x \in Q$ (i.e., for all elements $x \in N$ of order q). Furthermore, by Sylow's Theorem every element of G of order q lies in a conjugate of Q, hence the collection of G-conjugates of the set $Q - \{1\}$ partition the elements of order q in G into $|G : N|$ disjoint subsets. Since α^* and β^* are class functions on G, the sum of $\alpha^*(x)\overline{\beta^*(x)}$ as x runs over any of these subsets is the same. These facts imply

$$\begin{aligned}
(\alpha^*, \beta^*)_G &= \frac{1}{|G|} \sum_{x \in G} \alpha^*(x)\overline{\beta^*(x)} \\
&= \frac{1}{|G|} \sum_{\substack{x \in G \\ |x| = q}} \alpha^*(x)\overline{\beta^*(x)} \\
&= \frac{1}{|G|} \sum_{\substack{x \in N \\ |x| = q}} |G : N|\alpha^*(x)\overline{\beta^*(x)}
\end{aligned}$$

$$= \frac{1}{|N|}\sum_{x\in N}\alpha(x)\overline{\beta(x)} = (\alpha, \beta)_N.$$

This completes the proof.

The next lemma sets up a correspondence between the irreducible characters of N_i of degree > 1 and some nonprincipal irreducible characters of G.

Lemma 16. For any $i \in \{1, 2, 3, 4\}$ let $q = q_i$, let $Q = Q_i$, let $N = N_i$ and let $p = |N : Q|$. Let $\psi_1, \dots, \psi_k$ be the distinct irreducible characters of N of degree p. Then there are distinct irreducible characters $\chi_1, \dots, \chi_k$ of G, all of which have the same degree, and a fixed sign $\epsilon = \pm 1$ such that $\psi_1^* - \psi_j^* = \epsilon(\chi_1 - \chi_j)$ for all $j = 2, 3, \dots, k$.

Proof: Let $\alpha_j = \psi_1 - \psi_j$ for $j = 2, 3, \dots, k$ so α_j satisfies the hypothesis of Lemma 15. Since $\psi_1 \neq \psi_j$, by Lemma 15

$$2 = ||\alpha_j||^2 = (\alpha_j, \alpha_j)_N = (\alpha_j^*, \alpha_j^*)_G = ||\alpha_j^*||^2$$

for all j. Thus α_j^* must have two distinct irreducible characters of G as its irreducible constituents. Since $\alpha_j^*(1) = 0$ it must be a difference of two distinct irreducible characters, both of which have the same degree. In particular, the lemma holds if $k = 2$ (which is the case for $q = 3$ and $q = 7$). Assume therefore that $k > 2$ and write

$$\alpha_2^* = \psi_1^* - \psi_2^* = \epsilon(\chi - \chi')$$
$$\alpha_3^* = \psi_1^* - \psi_3^* = \epsilon'(\theta - \theta')$$

for some irreducible characters $\chi, \chi', \theta, \theta'$ of G and some signs ϵ, ϵ'. As proved above, $\chi \neq \chi'$ and $\theta \neq \theta'$. Interchanging θ and θ' if necessary, we may assume $\epsilon = \epsilon'$. Thus

$$\alpha_3^* - \alpha_2^* = \psi_2^* - \psi_3^* = \epsilon(\theta - \theta' - \chi + \chi').$$

By Lemma 15, $\psi_2^* - \psi_3^* = (\psi_2 - \psi_3)^*$ also has exactly two distinct irreducible constituents, hence either $\theta = \chi$ or $\theta' = \chi'$. Replacing ϵ by $-\epsilon$ if necessary we may assume that $\theta = \chi$ so that now we have

$$\alpha_2^* = \psi_1^* - \psi_2^* = \epsilon(\chi - \chi')$$
$$\alpha_3^* = \psi_1^* - \psi_3^* = \epsilon(\chi - \theta')$$

where χ, χ' and θ are distinct irreducible characters of G and the sign ϵ is determined. Label $\chi = \chi_1$, $\chi' = \chi_2$ and $\theta = \chi_3$. Now one similarly checks that for each $j \geq 3$ there is an irreducible character χ_j of G such that

$$\alpha_j^* = \psi_1^* - \psi_j^* = \epsilon(\chi_1 - \chi_j)$$

and $\chi_1, \dots, \chi_k$ are distinct. Since all χ_j's have the same degree as χ_1, the proof is complete.

We remark that it need not be the case that $\chi_j = \psi_j^*$ for any j, but only that the differences of irreducible characters of N induce to differences of irreducible characters of G.

The irreducible characters χ_j of G obtained via Lemma 16 are called *exceptional characters* associated to Q.

Lemma 17. The exceptional characters associated to Q_i are all distinct from the exceptional characters associated to Q_j for i and j distinct elements of $\{1, 2, 3, 4\}$.

Proof: Let χ be an exceptional character associated to Q_i and let θ be an exceptional character associated to Q_j. By construction, there are distinct irreducible characters ψ and ψ' of Q_i such that $\psi^* - \psi'^* = \chi - \chi'$ and there are distinct irreducible characters λ and λ' of Q_j such that $\lambda^* - \lambda'^* = \theta - \theta'$. Let $\alpha = \psi - \psi'$ and let $\beta = \lambda - \lambda'$. By Lemma 15, α^* is zero on all elements of G whose order is not equal to q_i (including the identity) and β^* is zero on all elements of G whose order is not equal to q_j. Thus clearly $(\alpha^*, \beta^*) = 0$. It follows easily that the two irreducible constituents of α^* are pairwise orthogonal to those of β^* as well. This establishes the lemma.

It is now easy to show that such a simple group G does not exist. By Lemma 16 and properties (i) to (iv) of G we can count the number of exceptional characters:

- **(i)** there are 2 exceptional characters associated to Q_1
- **(ii)** there are 2 exceptional characters associated to Q_2
- **(iii)** there are 4 exceptional characters associated to Q_3
- **(iv)** there are 136 exceptional characters associated to Q_4 .

Denote the common degree of the exceptional characters associated to Q_i by d_i for $i = 1, \dots, 4$. By Lemma 17, the exceptional characters account for 144 nonprincipal irreducible characters of G hence these, together with the principal character, are all the irreducible characters of G (the number of conjugacy classes of G is 145). The sum of the squares of the degrees of the irreducible characters is the order of G:

$$1 + 2d_1^2 + 2d_2^2 + 4d_3^2 + 136d_4^2 = 1004913.$$

Simplifying this, we obtain

$$d_1^2 + d_2^2 + 2d_3^2 + 68d_4^2 = 502456. \tag{19.4}$$

Finally, since each nonprincipal irreducible representation of the simple group G is faithful and since the smallest degree of a faithful representation of N_1 is 13, each $d_i \geq 13$. Since $d_4 < \sqrt{502456/68} < 86$ and d_4 divides $|G|$, it follows that

$$d_4 \in \{13,\ 21,\ 27,\ 39,\ 63\}.$$

Furthermore, each $d_i \bigm| |G|$ by Corollary 5 and so there are a small number of possibilities for each d_i. One now checks that equation (4) has no solution (this is particularly easy to do by computer). This contradiction completes the proof.

EXERCISES

Throughout the exercises all representations are over the complex numbers.

1. Let $G = S_3$, let $H = A_3$ and let V be the 3-dimensional $\mathbb{C}H$-module which affords the natural permutation representation of A_3. More explicitly, let V have basis e_1, e_2, e_3 and let $\sigma \in A_3$ act on V by $\sigma e_i = e_{\sigma(i)}$. Let 1 and (1 2) be coset representatives for the left cosets of A_3 in S_3 and write out the explicit matrices described in Theorem 11 for the action of S_3 on the induced module W, for each of the elements of S_3.
2. In each of parts (a) to (f) a character ψ of a subgroup H of a particular group G is specified. Compute the values of the induced character $\mathrm{Ind}_H^G(\psi)$ on all the conjugacy classes of G and use the character tables in Section 1 to write $\mathrm{Ind}_H^G(\psi)$ as a sum of irreducible characters:

(a) ψ is the unique nonprincipal degree 1 character of the subgroup $\langle\, (1\ 2)\,\rangle$ of S_3
(b) ψ is the degree 1 character of the subgroup $\langle\, r\,\rangle$ of D_8 defined by $\psi(r) = i$, where $i \in \mathbb{C}$ is a square root of -1
(c) ψ is the degree 1 character of the subgroup $\langle\, r\,\rangle$ of D_8 defined by $\psi(r) = -1$
(d) ψ is any of the nonprincipal degree 1 characters of the subgroup $V_4 = \langle\, (1\ 2),\ (3\ 4)\,\rangle$ of S_4
(e) $\psi = \chi_4$ is the first of the two characters of degree 3 in the character table of $H = S_4$ in Section 1 and H is a subgroup of $G = S_5$
(f) ψ is any of the nonprincipal degree 1 characters of the subgroup $V_4 = \langle\, (1\ 2),\ (3\ 4)\,\rangle$ of S_5.

3. Use Proposition 13 to explicitly write out the character table of each of the following groups:
(a) the dihedral group of order 10
(b) the non-abelian group of order 57
(c) the non-abelian group of order 56 which has a normal, elementary abelian Sylow 2-subgroup.

4. Let H be a subgroup of G, let φ be a representation of H and suppose that N is a normal subgroup of G with $N \leq H$ and N contained in the kernel of φ. Prove that N is also contained in the kernel of the induced representation of φ.

5. Let N be a normal subgroup of G and let ψ_1 be the principal character of N. Let Ψ be the induced character $\mathrm{Ind}_N^G(\psi_1)$ so that by the preceding exercise we may consider Ψ as the character of a representation of G/N. Prove that Ψ is the character of the regular representation of G/N.

6. Let Z be any subgroup of the center of G, let $|G : Z| = m$ and let ψ be a character of Z. Prove that

$$\mathrm{Ind}_Z^G(\psi)(g) = \begin{cases} m\psi(g) & \text{if } g \in Z \\ 0 & \text{if } g \notin Z. \end{cases}$$

7. Let φ be a matrix representation of the subgroup H of G and define matrices $\Phi(g)$ for every $g \in G$ by the displayed formula in the statement of Theorem 11. Prove directly that Φ is a representation by showing that $\Phi(xy) = \Phi(x)\Phi(y)$ for all $x, y \in G$.

8. Let G be a Frobenius group with Frobenius kernel Q. Assume that both Q and G/Q are abelian but G is not abelian (i.e., $G \neq Q$). Let $|Q| = n$ and $|G : Q| = m$.
(a) Prove that G/Q is cyclic and show that $G = QC$ for some cyclic subgroup C of G with $C \cap Q = 1$ (i.e., G is a semidirect product of Q and C and $|C| = m$). [Let q be a prime divisor of n and let G/Q act by conjugation on the elementary abelian q-group $\{h \in Q \mid h^q = 1\}$. Apply Exercise 14(e) of Section 18.1 and the definition of a Frobenius group to an irreducible constituent of this $\mathbb{F}_q G/Q$-module.]
(b) Prove that n and m are relatively prime. [If a prime p divides both the order and index of Q, let P be a Sylow p-subgroup of G. Then $P \cap Q \trianglelefteq P$ and $P \cap Q$ is a Sylow p-subgroup of Q. Consider the centralizer in G of the subgroup $Z(P) \cap Q$ (this intersection is nontrivial by Theorem 1 of Section 6.1).]
(c) Show that G has no elements of order qp, where q is any nontrivial divisor of n and p is any nontrivial divisor of m. [Argue as in Proposition 13.]
(d) Prove that the number of nonidentity conjugacy classes of G contained in Q is $(n-1)/m$ and that each of these classes has size m. [Argue as in Proposition 13.]
(e) Prove that no two distinct elements of C are conjugate in G. Deduce that the nonidentity elements of C are representatives for $m-1$ distinct conjugacy classes of G and that each of these classes has size n. Deduce then that every element of $G - Q$

is conjugate to some element of C and that G has $m + (n-1)/m$ conjugacy classes.

(f) Prove that $G' = Q$ and deduce that G has m distinct characters of degree 1. [To show $Q \leq G'$ let $C = \langle x \rangle$ and argue that the map $h \mapsto [h, x] = x^{-1}h^{-1}xh$ is a homomorphism from Q to Q whose kernel is trivial, hence this map is surjective.]

(g) Show that if ψ is any nonprincipal irreducible character of Q, then $\mathrm{Ind}_Q^G(\psi)$ is an irreducible character of G. Show that every irreducible character of G of degree > 1 is equal to $\mathrm{Ind}_Q^G(\psi)$ for some nonprincipal irreducible character ψ of Q. Deduce that every irreducible character of G has degree either 1 or m and the number of irreducible characters of degree m is $(n-1)/m$. [Check that the proof of Proposition 13(3) establishes this more general result with the appropriate changes to the numbers involved.]

9. Use the preceding exercise to explicitly write out the character table of $\langle (1\,2\,3\,4\,5),\ (2\,3\,5\,4) \rangle$, which is the normalizer in S_5 of a Sylow 5-subgroup (this group is a Frobenius group of order 20).

10. Let N be a normal subgroup of G, let ψ be a character of N and let $g \in G$. Define ψ^g by $\psi^g(h) = \psi(ghg^{-1})$ for all $h \in N$.

(a) Prove that ψ^g is a character of N (ψ and ψ^g are called *G-conjugate* characters of N). Prove that ψ^g is irreducible if and only if ψ is irreducible.

(b) Prove that the map $\psi \mapsto \psi^g$ is a right group action of G on the set of characters of N and N is in the kernel of this action.

(c) Prove that if ψ_1 and ψ_2 are G-conjugate characters of N, then $\mathrm{Ind}_N^G(\psi_1) = \mathrm{Ind}_N^G(\psi_2)$. Prove also that if ψ_1 and ψ_2 are characters of N that are not G-conjugate then $\mathrm{Ind}_N^G(\psi_1) \neq \mathrm{Ind}_N^G(\psi_2)$. [Use the argument in the proof of Proposition 13(3).]

11. Show that if $G = A_4$ and $N = V_4$ is its Sylow 2-subgroup then any two nonprincipal irreducible characters of N are G-conjugate (cf. the preceding exercise).

12. Let $G = D_{2n}$ be presented by its usual generators and relations. Prove that if ψ is any degree 1 character of $H = \langle r \rangle$ such that $\psi \neq \psi^s$, then $\mathrm{Ind}_H^G(\psi)$ is an irreducible character of D_{2n}. Show that every irreducible character of D_{2n} is the induced character of some degree 1 character of $\langle r \rangle$.

13. Prove both parts of Proposition 14.

14. Prove the following result known as *Frobenius Reciprocity*: let $H \leq G$, let ψ be any character of H and let χ be any character of G. Then

$$(\psi\,,\ \chi|_H)_H = (\mathrm{Ind}_H^G(\psi)\,,\ \chi)_G.$$

[Expand the right hand side using the formula for the induced character $\mathrm{Ind}_H^G(\psi)$ or follow the proof of Shapiro's Lemma in Section 17.2.]

15. Assume G were a simple group of order $3^3 \cdot 7 \cdot 13 \cdot 409$ whose Sylow subgroups and their normalizers are described by properties (1) to (5) in this section. Prove that the permutation character of degree 819 obtained from the action of G on the left cosets of the subgroup N_4 decomposes as $\chi_0 + \gamma + \gamma'$, where χ_0 is the principal character of G and γ and γ' are distinct irreducible characters of G of degree 409. [Use Exercise 9 in Section 18.3 to show that this permutation character π has $||\,\pi\,||^2 = 3$.]

APPENDIX I

Cartesian Products and Zorn's Lemma

Section 1 of this appendix contains the definition of the Cartesian product of an arbitrary collection of sets. In the text we shall primarily be interested in products of finitely many (or occasionally countably many) sets. We indicate how the general definition agrees with the familiar "ordered n-tuple" notion of a Cartesian product in these cases. Section 2 contains a discussion of Zorn's Lemma and related topics.

1. CARTESIAN PRODUCTS

A set I is called an *indexing set* or *index set* if the elements of I are used to index some collection of sets. In particular, if A and I are sets, we can form the collection $\{A_i \mid i \in I\}$ by specifying that $A_i = A$ for all $i \in I$. Thus *any* set can be an indexing set; we use this term to emphasize that the elements are used as indices.

Definition.

(1) Let I be an indexing set and let $\{A_i \mid i \in I\}$ be a collection of sets. A *choice function* is any function

$$f : I \to \bigcup_{i \in I} A_i$$

such that $f(i) \in A_i$ for all $i \in I$.

(2) Let I be an indexing set and for all $i \in I$ let A_i be a set. The *Cartesian product* of $\{A_i \mid i \in I\}$ is the set of all choice functions from I to $\cup_{i \in I} A_i$ and is denoted by $\prod_{i \in I} A_i$ (where if either I or any of the sets A_i are empty the Cartesian product is the empty set). The elements of this Cartesian product are written as $\prod_{i \in I} a_i$, where this denotes the choice function f such that $f(i) = a_i$ for each $i \in I$.

(3) For each $j \in I$ the set A_j is called the j^{th} *component* of the Cartesian product $\prod_{i \in I} A_i$ and a_j is the j^{th} *coordinate* of the element $\prod_{i \in I} a_i$.

(4) For $j \in I$ the *projection map* of $\prod_{i \in I} A_i$ onto the j^{th} coordinate, A_j, is defined by $\prod_{i \in I} a_i \mapsto a_j$.

Each choice function f in the Cartesian product $\prod_{i \in I} A_i$ may be thought of as a way of "choosing" an element $f(i)$ from each set A_i.

If $I = \{1, 2, \dots, n\}$ for some $n \in \mathbb{Z}^+$ and if f is a choice function from I to $A_1 \cup \cdots \cup A_n$, where each A_i is nonempty, we can associate to f a unique (ordered) n-tuple:

$$f \to (f(1), f(2), \dots, f(n)).$$

Note that by definition of a choice function, $f(i) \in A_i$ for all i, so the n-tuple above has an element of A_i in the i^{th} position for each i.

Conversely, given an n-tuple $(a_1, a_1, \dots, a_n)$, where $a_i \in A_i$ for all $i \in I$, there is a unique choice function, f, from I to $\cup_{i \in I} A_i$ associated to it, namely

$$f(i) = a_i, \qquad \text{for all } i \in I.$$

It is clear that this map from n-tuples to choice functions is the inverse to the map described in the preceding paragraph. Thus *there is a bijection between ordered n-tuples and elements of $\prod_{i \in I} A_i$*. Henceforth when $I = \{1, 2, \dots, n\}$ we shall write

$$\prod_{i=1}^{n} A_i \quad \text{or} \quad A_1 \times A_2 \times \cdots \times A_n$$

for the Cartesian product and we shall describe the elements as ordered n-tuples.

If $I = \mathbb{Z}^+$, we shall similarly write: $\prod_{i=1}^{\infty} A_i$ or $A_1 \times A_2 \times \cdots$ for the Cartesian product of the A_i's. We shall write the elements as ordered tuples: $(a_1, a_2, \dots)$, i.e., as infinite sequences whose i^{th} terms are in A_i.

Note that when $I = \{1, 2, \dots, n\}$ or $I = \mathbb{Z}^+$ we have used the natural ordering on I to arrange the elements of our Cartesian products into n-tuples. Any other ordering of I (or any ordering on a finite or countable index set) gives a different representation of the elements of the same Cartesian product.

Examples

(1) $A \times B = \{(a, b) \mid a \in A, b \in B\}$.

(2) $\mathbb{R}^n = \mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}$ (n factors) is the usual set of n-tuples with real number entries, Euclidean n-space.

(3) Suppose $I = \mathbb{Z}^+$ and A_i is the same set A, for all $i \in I$. The Cartesian product $\prod_{i \in \mathbb{Z}^+} A$ is the set of all (infinite) sequences $a_1, a_2, a_3 \dots$ of elements of A. In particular, if $A = \mathbb{R}$, then the Cartesian product $\prod_{i \in \mathbb{Z}^+} \mathbb{R}$ is the set of all real sequences.

(4) Suppose I is any indexing set and A_i is the same set A, for all $i \in I$. The Cartesian product $\prod_{i \in I} A$ is just the set of all functions from I to A, where the function $f : I \to A$ corresponds to the element $\prod_{i \in I} f(i)$ in the Cartesian product. This Cartesian product is often (particularly in topology books) denoted by A^I. Note that for each fixed $j \in I$ the projection map onto the j^{th} coordinate sends the function f to $f(j)$, i.e., is evaluation at j.

(5) Let R be a ring and let x be an indeterminate over R. The definition of the ring $R[x]$ of polynomials in x with coefficients from R may be given in terms of Cartesian products rather than in the more intuitive and familiar terms of "formal sums" (in Chapters 7 and 9 we introduced them in the latter form since this is the way we envision and work with them). Let I be the indexing set $\mathbb{Z}^+ \cup \{0\}$ and let $R[x]$ be the subset of the Cartesian product $\prod_{i=0}^{\infty} R$ consisting of elements $(a_0, a_1, a_2, \dots)$ such that only finitely many of the a_i's are nonzero. If $(a_0, a_1, a_2, \dots, a_n, 0, 0, \dots)$ is such a sequence we represent it by the more familiar "formal sum" $\sum_{i=0}^{n} a_i x^i$. Addition and multiplication of these sequences is defined so that the usual rules for addition and multiplication of polynomials hold.

Proposition 1. Let I be a nonempty countable set and for each $i \in I$ let A_i be a set. The cardinality of the Cartesian product is the product of the cardinalities of the sets A_i, i.e.,

$$|\prod_{i \in I} A_i| = \prod_{i \in I} |A_i|,$$

(where if some A_i is an infinite set or if I is infinite and an infinite number of A_i's have cardinality ≥ 2, both sides of this equality are infinity). In particular,

$$|A_1 \times A_2 \times \cdots \times A_n| = |A_1| \times |A_2| \times \cdots \times |A_n|.$$

Proof: In order to count the number of choice functions note that each $i \in I$ may be mapped to any of the $|A_i|$ elements of A_i and for $i \neq j$ the values of choice functions at i and j may be chosen completely independently. Thus the number of choice functions is the product of the cardinalities of the A_i's, as claimed.

For Cartesian products of finitely many sets, $A_1 \times A_2 \times \cdots \times A_n$, one can see this easily from the n-tuple representation: the elements of $A_1 \times A_2 \times \cdots \times A_n$ are n-tuples $(a_1, a_2, \ldots, a_n)$ and each a_i may be chosen as any of the $|A_i|$ elements of A_i. Since these choices are made independently for $i \neq j$, there are $|A_1| \cdot |A_2| \cdots |A_n|$ elements in the Cartesian product.

EXERCISE

1. Let I and J be any two indexing sets and let A be an arbitrary set. For any function $\varphi : J \to I$ define

$$\varphi^* : \prod_{i \in I} A \to \prod_{j \in J} A \quad \text{by} \quad \varphi^*(f) = f \circ \varphi \quad \text{for all choice functions } f \in \prod_{i \in I} A.$$

   (a) Let $I = \{1, 2\}$, let $J = \{1, 2, 3\}$ and let $\varphi : J \to I$ be defined by $\varphi(1) = 2$, $\varphi(2) = 2$ and $\varphi(3) = 1$. Describe explicitly how a 3-tuple in $A \times A \times A$ maps to an ordered pair in $A \times A$ under this φ^*.
   (b) Let $I = J = \{1, 2, \ldots, n\}$ and assume φ is a permutation of I. Describe in terms of n-tuples in $A \times A \times \cdots \times A$ the function φ^*.

2. PARTIALLY ORDERED SETS AND ZORN'S LEMMA

We shall have occasion to use Zorn's Lemma as a form of "infinite induction" in a few places in the text where it is desirable to know the existence of some set which is *maximal* with respect to certain specified properties. For example, Zorn's Lemma is used to show that every vector space has a basis. In this situation a basis of a vector space V is a subset of V which is maximal as a set consisting of linearly independent vectors (the maximality ensures that these vectors span V). For finite dimensional spaces this can be proved by induction; however, for spaces of arbitrary dimension Zorn's Lemma is needed to establish this. By having results which hold in full generality the theory often becomes a little neater in places, although the main results of the text do not require its use.

A specific instance in the text where a maximal object which helps to simplify matters is constructed by Zorn's Lemma is the algebraic closure of a field. An algebraic closure of a field F is an extension of F which is maximal among any collection of algebraic extensions. Such a field contains (up to isomorphism) all elements which are algebraic over F, hence all manipulations involving such algebraic elements can be effected in this one larger field. In any particular situation the use of an algebraic closure can be avoided by adjoining the algebraic elements involved to the base field F, however this becomes tedious (and often obscures matters) in complicated proofs. For the specific fields appearing as examples in this text the use of Zorn's Lemma to construct an algebraic closure can be avoided (for example, the construction of an algebraic closure of any subfield of the complex numbers or of any finite field does not require it).

The first example of the use of Zorn's Lemma appears in the proof of Proposition 11 in Section 7.4.

In order to state Zorn's Lemma we need some terminology.

Definition. A *partial order* on a nonempty set A is a relation $\leq$ on A satisfying

(1) $x \leq x$ for all $x \in A$ (reflexive),

(2) if $x \leq y$ and $y \leq x$ then $x = y$ for all $x, y \in A$ (antisymmetric),

(3) if $x \leq y$ and $y \leq z$ then $x \leq z$ for all $x, y, z \in A$ (transitive).

We shall usually say that A is a partially ordered set under the ordering $\leq$ or that A is partially ordered by $\leq$.

Definition. Let the nonempty set A be partially ordered by $\leq$.

(1) A subset B of A is called a *chain* if for all $x, y \in B$, either $x \leq y$ or $y \leq x$.

(2) An *upper bound* for a subset B of A is an element $u \in A$ such that $b \leq u$, for all $b \in B$.

(3) A *maximal element* of A is an element $m \in A$ such that if $m \leq x$ for any $x \in A$, then $m = x$.

In the literature a chain is also called a *tower* or called a *totally ordered* or *linearly ordered* or *simply ordered* subset.

Some examples below highlight the distinction between upper bounds and maximal elements. Also note that if m is a *maximal* element of A, it is not necessarily the case that $x \leq m$ for all $x \in A$ (i.e., m is not necessarily a *maximum* element).

Examples

(1) Let A be the power set (i.e., set of all subsets) of some set X and $\leq$ be set containment: $\subseteq$. Notice that this is only a *partial* ordering since some subsets of X may not be comparable, e.g. singletons: if $x \neq y$ then $\{x\} \not\subseteq \{y\}$ and $\{y\} \not\subseteq \{x\}$. In this situation an example of a chain is a collection of subsets of X such as

$$X_1 \subseteq X_2 \subseteq X_3 \subseteq \cdots.$$

Any subset B of A has an upper bound, b, namely,

$$b = \bigcup_{x \in B} x.$$

This partially ordered set A has a (unique) maximal element, X.

In many instances the set A consists of some (but not necessarily all) subsets of a set X (i.e., A is a subset of the power set of X) and with the ordering on A again being inclusion. The existence of upper bounds and maximal elements depends on the nature of A.

(2) Let A be the collection of all *proper* subsets of $\mathbb{Z}^+$ ordered under $\subseteq$. In this situation, chains need not have maximal elements, e.g. the chain

$$\{1\} \subseteq \{1, 2\} \subseteq \{1, 2, 3\} \subseteq \cdots$$

does not have an upper bound. The set A does have maximal elements: for example $\mathbb{Z}^+ - \{n\}$ is a maximal element of A for any $n \in \mathbb{Z}^+$.

(3) Let $A = \mathbb{R}$ under the usual $\leq$ relation. In this example every subset of A is a chain (including A itself). The notion of a subset of A having an upper bound is the same as the usual notion of a subset of $\mathbb{R}$ being bounded above by some real number (so some sets, such as intervals of finite length, have upper bounds and others, such as the set of positive reals, do not). The set A does not have a maximal element.

Zorn's Lemma If A is a nonempty partially ordered set in which every chain has an upper bound then A has a maximal element.

It is a nontrivial result that *Zorn's Lemma is independent of the usual (Zermelo–Fraenkel) axioms of set theory*[1] in the sense that if the axioms of set theory are consistent,[2] then so are these axioms together with Zorn's Lemma; and if the axioms of set theory are consistent, then so are these axioms together with the *negation* of Zorn's Lemma. The use of the term "lemma" in Zorn's Lemma is historical.

For the sake of completeness (and to relate Zorn's Lemma to formulations found in other courses) we include two other equivalent formulations of Zorn's Lemma.

The Axiom of Choice The Cartesian product of any nonempty collection of nonempty sets is nonempty. In other words, if I is any nonempty (indexing) set and A_i is a nonempty set for all $i \in I$, then there exists a choice function from I to $\cup_{i \in I} A_i$.

Definition. Let A be a nonempty set. A *well ordering* on A is a total ordering on A such that every nonempty subset of A has a minimum (or smallest) element, i.e., for each nonempty $B \subseteq A$ there is some $s \in B$ such that $s \leq b$, for all $b \in B$.

The Well Ordering Principle Every nonempty set A has a well ordering.

Theorem 2. Assuming the usual (Zermelo–Fraenkel) axioms of set theory, the following are equivalent:

(1) Zorn's Lemma
(2) the Axiom of Choice
(3) the Well Ordering Principle.

Proof: This follows from elementary set theory. We refer the reader to *Real and Abstract Analysis* by Hewitt and Stromberg, Springer-Verlag, 1965, Section 3 for these equivalences and some others.

[1] See P.J. Cohen's papers in: Proc. Nat. Acad. Sci., 50(1963), and 51(1964).

[2] This is not known to be the case!

EXERCISES

1. Let A be the collection of all finite subsets of $\mathbb{R}$ ordered by inclusion. Discuss the existence (or nonexistence) of upper bounds, minimal and maximal elements (where minimal elements are defined analogously to maximal elements). Explain why this is not a well ordering.
2. Let A be the collection of all infinite subsets of $\mathbb{R}$ ordered by inclusion. Discuss the existence (or nonexistence) of upper bounds, minimal and maximal elements. Explain why this is not a well ordering.
3. Show that the following partial orderings on the given sets are not well orderings:
   (a) $\mathbb{R}$ under the usual relation $\leq$.
   (b) $\mathbb{R}^+$ under the usual relation $\leq$.
   (c) $\mathbb{R}^+ \cup \{0\}$ under the usual relation $\leq$.
   (d) $\mathbb{Z}$ under the usual relation $\leq$.
4. Show that $\mathbb{Z}^+$ is well ordered under the usual relation $\leq$.

APPENDIX II

Category Theory

Category theory provides the language and the mathematical foundations for discussing properties of large classes of mathematical objects such as the class of "all sets" or "all groups" while circumventing problems such as Russell's Paradox. In this framework one may explore the commonality across classes of concepts and methods used in the study of each class: homomorphisms, isomorphisms, etc., and one may introduce tools for studying relations between classes: functors, equivalence of categories, etc. One may then formulate precise notions of a "natural" transformation and "natural" isomorphism, both within a given class or between two classes. (In the text we described "natural" as being "coordinate free.") A prototypical example of natural isomorphisms within a class is the isomorphism of an arbitrary finite dimensional vector space with its double dual in Section 11.3. In fact one of the primary motivations for the introduction of categories and functors by S. Eilenberg and S. MacLane in 1945 was to give a precise meaning to the notions of "natural" in cases such as this. Category theory has also played a foundational role for formalizing new concepts such as schemes (cf. Section 15.5) that are fundamental to major areas of contemporary research (e.g., algebraic geometry). Pioneering work of this nature was done by A. Grothendieck, K. Morita and others.

Our treatment of category theory should be viewed more as an introduction to some of the basic language. Since we have not discussed the Zermelo–Fraenkel axioms of set theory or the Gödel–Bernays axioms of classes we make no mention of the foundations of category theory. To remain consistent with the set theory axioms, however, we implicitly assume that there is a *universe* set $\mathbf{U}$ which contains all the sets, groups, rings, etc. that one would encounter in "ordinary" mathematics (so that the category of "all sets" implicitly means "all sets in $\mathbf{U}$"). The reader is referred to books on set theory, logic, or category theory such as *Categories for the Working Mathematician* by S. MacLane, Springer–Verlag, 1971 for further study.

We have organized this appendix so that wherever possible the examples of each new concept use terminology and structures in the order that these appear in the body of the text. For instance, the first example of a functor involves sets and groups, the second example uses rings, etc. In this way the appendix may be read early on in one's study, and a greater appreciation may be gained through rereading the examples as one becomes conversant with a wider variety of mathematical structures.

1. CATEGORIES AND FUNCTORS

We begin with the basic concept of this appendix.

Definition. A *category* $\mathbf{C}$ consists of a class of *objects* and sets of *morphisms* between those objects. For every ordered pair A, B of objects there is a set $\mathrm{Hom}_{\mathbf{C}}(A, B)$ of morphisms from A to B, and for every ordered triple A, B, C of objects there is a *law of composition* of morphisms, i.e., a map

$$\mathrm{Hom}_{\mathbf{C}}(A, B) \times \mathrm{Hom}_{\mathbf{C}}(B, C) \longrightarrow \mathrm{Hom}_{\mathbf{C}}(A, C)$$

where $(f, g) \mapsto gf$, and gf is called the composition of g with f. The objects and morphism satisfy the following axioms: for objects A, B, C and D

(i) if $A \neq B$ or $C \neq D$, then $\mathrm{Hom}_{\mathbf{C}}(A, B)$ and $\mathrm{Hom}_{\mathbf{C}}(C, D)$ are disjoint sets,

(ii) composition of morphisms is associative, i.e., $h(gf) = (hg)f$ for every f in $\mathrm{Hom}_{\mathbf{C}}(A, B)$, g in $\mathrm{Hom}_{\mathbf{C}}(B, C)$ and h in $\mathrm{Hom}_{\mathbf{C}}(C, D)$,

(iii) each object has an identity morphism, i.e., for every object A there is a morphism $1_A \in \mathrm{Hom}_{\mathbf{C}}(A, A)$ such that $f1_A = f$ for every $f \in \mathrm{Hom}_{\mathbf{C}}(A, B)$ and $1_A g = g$ for every $g \in \mathrm{Hom}_{\mathbf{C}}(B, A)$.

Morphisms are also called *arrows*. It is an exercise to see that the identity morphism for each object is unique (by the same argument that the identity of a group is unique). We shall write $\mathrm{Hom}(A, B)$ for $\mathrm{Hom}_{\mathbf{C}}(A, B)$ when the category is clear from the context.

The terminology we use throughout the text is common to all categories: a morphism from A to B will be denoted by $f : A \to B$ or $A \xrightarrow{f} B$. The object A is the *domain* of f and B is the *codomain* of f. A morphism from A to A is an endomorphism of A. A morphism $f : A \to B$ is an isomorphism if there is a morphism $g : B \to A$ such that $gf = 1_A$ and $fg = 1_B$.

There is a natural notion of a *subcategory* category $\mathbf{C}$ of $\mathbf{D}$, i.e., when every object of $\mathbf{C}$ is also an object in $\mathbf{D}$, and for objects A, B in $\mathbf{C}$ we have the containment $\mathrm{Hom}_{\mathbf{C}}(A, B) \subseteq \mathrm{Hom}_{\mathbf{D}}(A, B)$.

Examples

In each of the following examples we leave the details of the verification of the axioms for a category as exercises.

(1) **Set** is the category of all sets. For any two sets A and B, $\mathrm{Hom}(A, B)$ is the set of all functions from A to B. Composition of morphisms is the familiar composition of functions: $gf = g \circ f$. The identity in $\mathrm{Hom}(A, A)$ is the map $1_A(a) = a$, for all $a \in A$. This category contains the category of all finite sets as a subcategory.

(2) **Grp** is the category of all groups, where morphisms are group homomorphisms. Note that the composition of group homomorphisms is again a group homomorphism. A subcategory of **Grp** is **Ab**, the category of all abelian groups. Similarly, **Ring** is the category of all nonzero rings with 1, where morphisms are ring homomorphisms that send 1 to 1. The category **CRing** of all commutative rings with 1 is a subcategory of **Ring**.

(3) For a fixed ring R, the category R–**mod** consists of all left R-modules with morphisms being R-module homomorphisms.

(4) **Top** is the category whose objects are topological spaces and morphisms are continuous maps between topological spaces (cf. Section 15.2). Note that the identity (set) map from a space to itself is continuous in every topology, so $\mathrm{Hom}(A, A)$ always has an identity.

(5) Let **0** be the empty category, with no objects and no morphisms. Let **1** denote the category with one object, A, and one morphism: $\mathrm{Hom}(A, A) = \{1_A\}$. Let **2** be the category with two objects, A_1 and A_2, and only one nonidentity morphism:

$\mathrm{Hom}(A_1, A_2) = \{f\}$ and $\mathrm{Hom}(A_2, A_1) = \emptyset$. Note that the objects A_1 and A_2 and the morphism f are "primitives" in the sense that A_1 and A_2 are not defined to be sets and f is simply an arrow (literally) from A_1 to A_2; it is not defined as a set map on the elements of some set. One can continue this way and define **N** to be the category with N objects $A_1, A_2, \ldots, A_N$ with the only nonidentity morphisms being a unique arrow from A_i to A_j for every $j > i$ (so that composition of arrows is uniquely determined).

(6) If G is a group, form the category **G** as follows. The only object is G and $\mathrm{Hom}(G, G) = G$; the composition of two functions f and g is the product gf in the group G. Note that $\mathrm{Hom}(G, G)$ has an identity morphism: the identity of the group G.

Definition. Let **C** and **D** be categories.

(1) We say $\mathcal{F}$ is a *covariant functor* from **C** to **D** if
  - (a) for every object A in **C**, $\mathcal{F}A$ is an object in **D**, and
  - (b) for every $f \in \mathrm{Hom}_{\mathbf{C}}(A, B)$ we have $\mathcal{F}(f) \in \mathrm{Hom}_{\mathbf{D}}(\mathcal{F}A, \mathcal{F}B)$,

  such that the following axioms are satisfied:
  - (i) if gf is a composition of morphisms in **C**, then $\mathcal{F}(gf) = \mathcal{F}(g)\mathcal{F}(f)$ in **D**, and
  - (ii) $\mathcal{F}(1_A) = 1_{\mathcal{F}A}$.

(2) We say $\mathcal{F}$ is a *contravariant functor* from **C** to **D** if the conditions in (1) hold but property (b) and axiom (i) are replaced by:
  - (b′) for every $f \in \mathrm{Hom}_{\mathbf{C}}(A, B)$, $\mathcal{F}(f) \in \mathrm{Hom}_{\mathbf{D}}(\mathcal{F}B, \mathcal{F}A)$,
  - (i′) if gf is a composition of morphisms in **C**, then $\mathcal{F}(gf) = \mathcal{F}(f)\mathcal{F}(g)$ in **D**

  (i.e., contravariant functors reverse the arrows).

Examples

In each of these examples the verification of the axioms for a functor are left as exercises. Additional examples of functors appear in the exercises at the end of this section.

(1) The identity functor $\mathcal{I}_C$ maps any category **C** to itself by sending objects and morphisms to themselves. More generally, if **C** is a subcategory of **D**, the *inclusion functor* maps **C** into **D** by sending objects and morphisms to themselves.

(2) Let $\mathcal{F}$ be the functor from **Grp** to **Set** that maps any group G to the same set G and any group homomorphism φ to the same set map φ. This functor is called the *forgetful functor* since it "removes" or "forgets" the structure of the groups and the homomorphisms between them. Likewise there are forgetful functors from the categories **Ab**, R–**mod**, **Top**, etc., to **Set**.

(3) The *abelianizing* functor maps **Grp** to **Ab** by sending each group G to the abelian group $G^{\mathrm{ab}} = G/G'$, where G' is the commutator subgroup of G (cf. Section 5.4). Each group homomorphism $\varphi : G \to H$ is mapped to the induced homomorphism on quotient groups:

$$\overline{\varphi} : G^{\mathrm{ab}} \to H^{\mathrm{ab}} \qquad \text{by} \qquad \overline{\varphi}(xG') = \varphi(x)H'.$$

The definition of the commutator subgroup ensures that $\overline{\varphi}$ is well defined and the axioms for a functor are satisfied.

(4) Let R be a ring and let D be a left R-module. For each left R-module N the set $\mathrm{Hom}_R(D, N)$ is an abelian group, and is an R-module if R is commutative (cf. Proposition 2 in Section 10.2). If $\varphi : N_1 \to N_2$ is an R-module homomorphism, then for every $f \in \mathrm{Hom}_R(D, N_1)$ we have $\varphi \circ f \in \mathrm{Hom}_R(D, N_2)$. Thus

$\varphi' : \mathrm{Hom}_R(D, N_1) \to \mathrm{Hom}_R(D, N_2)$ by $\varphi'(f) = \varphi \circ f$. This shows the map

$$\mathcal{H}om(D, \_\_) : N \longrightarrow \mathrm{Hom}_R(D, N)$$
$$\mathcal{H}om(D, \_\_) : \varphi \longrightarrow \varphi'$$

is a covariant functor from R–**Mod** to **Grp**. If R is commutative, it maps R–**Mod** to itself.

(5) In the notation of the preceding example, we observe that if $\varphi : N_1 \to N_2$, then for every $g \in \mathrm{Hom}_R(N_2, D)$ we have $g \circ \varphi \in \mathrm{Hom}_R(N_1, D)$. Thus $\varphi' : \mathrm{Hom}_R(N_2, D) \to \mathrm{Hom}_R(N_1, D)$ by $\varphi'(g) = g \circ \varphi$. In this case the map

$$\mathcal{H}om(\_\_, D) : N \longrightarrow \mathrm{Hom}_R(N, D)$$
$$\mathcal{H}om(\_\_, D) : \varphi \longrightarrow \varphi'$$

defines a *contravariant* functor.

(6) When D is a right R-module the map $D \otimes_R \_\_ : N \to D \otimes_R N$ defines a covariant functor from R–**Mod** to **Ab** (or to R–**Mod** when R is commutative). Here the morphism $\varphi : N_1 \to N_2$ maps to the morphism $1 \otimes \varphi$.

Likewise when D is a left R-module $\_\_ \otimes_R D : N \to N \otimes_R D$ defines a covariant functor from the category of right R-modules to **Ab** (or to R–**Mod** when R is commutative), where the morphism φ maps to the morphism $\varphi \otimes 1$.

(7) Let K be a field and let K–**fdVec** be the category of all finite dimensional vector spaces over K, where morphisms in this category are K-linear transformations. We define the *double dual* functor $\mathcal{D}^2$ from K–**fdVec** to itself. Recall from Section 11.3 that the dual space, V^*, of V is defined as $V^* = \mathrm{Hom}_K(V, K)$; the double dual of V is $V^{**} = \mathrm{Hom}_K(V^*, K)$. Then $\mathcal{D}^2$ is defined on objects by mapping a vector space V to its double dual V^{**}. If $\varphi : V \to W$ is a linear transformation of finite dimensional spaces, then

$$\mathcal{D}^2(\varphi) : V^{**} \to W^{**} \qquad \text{by} \qquad \mathcal{D}^2(\varphi)(E_v) = E_{\varphi(v)},$$

where E_v denotes "evaluation at v" for each $v \in V$. By Theorem 19 in Section 11.3, $E_v \in V^{**}$, and each element of V^{**} is of the form E_v for a unique $v \in V$. Since $\varphi(v) \in W$ we have $E_{\varphi(v)} \in W^{**}$, so $\mathcal{D}^2(\varphi)$ is well defined.

The functor $\mathcal{F}$ from **C** to **D** is called *faithful* (or is called *full*) if for every pair of objects A and B in **C** the map $\mathcal{F} : \mathrm{Hom}(A, B) \to \mathrm{Hom}(\mathcal{F}A, \mathcal{F}B)$ is injective (or surjective, respectively). Thus, for example, the forgetful functor is faithful but not full.

EXERCISES

1. Let N be a group and let **Nor**–N be the collection of all groups that contain N as a normal subgroup. A morphism between objects A and B is any group homomorphism that maps N into N.

(a) Prove that **Nor**–N is a category.

(b) Show how the projection homomorphism $G \mapsto G/N$ may be used to define a functor from **Nor**–N to **Grp**.

2. Let H be a group. Define a map $\mathcal{H}\times$ from **Grp** to itself on objects and morphisms as follows:

$$\mathcal{H}\times : G \to H \times G, \text{ and}$$
$$\text{if } \varphi : G_1 \to G_2 \quad \text{then } \mathcal{H}\times(\varphi) : H \times G_1 \to H \times G_2 \quad \text{by} \quad (h, g) \mapsto (h, \varphi(g)).$$

Prove that $\mathcal{H}\times$ is a functor.

3. Show that the map **Ring** to **Grp** by mapping a ring to its group of units (i.e., $R \mapsto R^{\times}$) defines a functor. Show by explicit examples that this functor is neither faithful nor full.
4. Show that for each $n \geq 1$ the map $\mathcal{GL}_n : R \to GL_n(R)$ defines a functor from **CRing** to **Grp**. [Define $\mathcal{GL}_n$ on morphisms by applying each ring homomorphism to the entries of a matrix.]
5. Supply the details that show the double dual map described in Example 7 satisfies the axioms of a functor.

2. NATURAL TRANSFORMATIONS AND UNIVERSALS

As mentioned in the introduction to this appendix, one of the motivations for the inception of category theory was to give a precise definition of the notion of "natural" isomorphism. We now do so, and see how some natural maps mentioned in the text are instances of the categorical concept. We likewise give the categorical definition of "universal arrows" and view some occurrences of universal properties in the text in this light.

Definition. Let **C** and **D** be categories and let $\mathcal{F}$, $\mathcal{G}$ be covariant functors from **C** to **D**. A *natural transformation* or *morphism of functors* from $\mathcal{F}$ to $\mathcal{G}$ is a map η that assigns to each object A in **C** a morphism η_A in $\mathrm{Hom}_{\mathbf{D}}(\mathcal{F}A, \mathcal{G}A)$ with the following property: for every pair of objects A and B in **C** and every $f \in \mathrm{Hom}_{\mathbf{C}}(A, B)$ we have $\mathcal{G}(f)\eta_A = \eta_B\mathcal{F}(f)$, i.e., the following diagram commutes:

$$\begin{tikzcd} \mathcal{F}A \arrow[r, "\eta_A"] \arrow[d, "\mathcal{F}(f)"'] & \mathcal{G}A \arrow[d, "\mathcal{G}(f)"] \\ \mathcal{F}B \arrow[r, "\eta_B"] & \mathcal{G}B \end{tikzcd}$$

If each η_A is an isomorphism, η is called a *natural isomorphism* of functors.

Consider the special case where $\mathbf{C} = \mathbf{D}$ and **C** is a subcategory of **Set**, and where $\mathcal{F}$ is the identity functor. There is a natural transformation η from the identity functor to $\mathcal{G}$ if whenever $\mathcal{G}$ maps the object A to the object $\mathcal{G}A$ there is a morphism η_A from A to $\mathcal{G}A$, and whenever there is a morphism f from A to B the morphism $\mathcal{G}(f)$ is compatible with f as a map from $\mathcal{G}A$ to $\mathcal{G}B$. In fact $\mathcal{G}(f)$ is uniquely determined by f as a map from the subset $\eta_A(A)$ in $\mathcal{G}A$ to the subset $\eta_B(B)$ of $\mathcal{G}B$. If η is a natural isomorphism, then the value of $\mathcal{G}$ on every morphism is completely determined by η, namely $\mathcal{G}(f) = \eta_B f \eta_A^{-1}$. In this case the functor $\mathcal{G}$ is entirely specified by η. We shall see that some of the examples of functors in the preceding section arise this way.

Examples

(1) For any categories **C** and **D** and any functor $\mathcal{F}$ from **C** to **D** the identity is a natural isomorphism from $\mathcal{F}$ to itself: $\eta_A = 1_{\mathcal{F}A}$ for every object A in **C**.

(2) Let R be a ring and let $\mathcal{F}$ be any functor from R–**Mod** to itself. The zero map is a natural transformation from $\mathcal{F}$ to itself: $\eta_A = 0_A$ for every R-module A, where 0_A is the zero map from A to itself. This is not a natural isomorphism.

(3) Let $\mathcal{F}$ be the identity functor from **Grp** to itself, and let $\mathcal{G}$ be the abelianizing functor (Example 3) considered here as a map from **Grp** to itself. For each group G let $\eta_G : G \to G/G'$ be the usual projection map onto the quotient group. Then η is a natural transformation (but not an isomorphism) with respect to these two functors. (We call the maps η_G the *natural projection* maps.)

(4) Let $\mathcal{G} = \mathcal{D}^2$ be the double dual functor from the category of finite dimensional vector spaces over a field K to itself (Example 7). Then there is a natural isomorphism η from the identity functor to $\mathcal{G}$ given by

$$\eta_V : V \to V^{**} \qquad \text{by} \qquad \eta_V(v) = E_v$$

where E_v is "evaluation at v" for every $v \in V$.

(5) Let $\mathcal{GL}_n$ be the functor from **CRing** to **Grp** defined as follows. Each object (commutative ring) R is mapped by $\mathcal{GL}_n$ to the group $GL_n(R)$ of $n \times n$ invertible matrices with entries from R. For each ring homomorphism $f : R \to S$ let $\mathcal{GL}_n(f)$ be the map of matrices that applies f to each matrix entry. Since f sends 1 to 1 it follows that $\mathcal{GL}_n(f)$ sends invertible matrices to invertible matrices (cf. Exercise 4 in Section 1). Let $\mathcal{G}$ be the functor from **CRing** to **Grp** that maps each ring R to its group of units $R^\times$, and each ring homomorphism f to its restriction to the groups of units (also denoted by f). The *determinant* is a natural transformation from $\mathcal{GL}_n$ to $\mathcal{G}$ because the determinant is defined by the same polynomial for all rings so that the following diagram commutes:

$$\begin{array}{ccc} GL_n(R) & \xrightarrow{\det} & R^\times \\ {\scriptstyle \mathcal{GL}_n(f)}\big\downarrow & & \big\downarrow{\scriptstyle f} \\ GL_n(S) & \xrightarrow{\det} & S^\times \end{array}$$

Let **C**, **D** and **E** be categories, let $\mathcal{F}$ be a functor from **C** to **D**, and let $\mathcal{G}$ be a functor from **D** to **E**. There is an obvious notion of the composition of functors $\mathcal{GF}$ from **C** to **E**. When **E** = **C** the composition $\mathcal{GF}$ maps **C** to itself and $\mathcal{FG}$ maps **D** to itself. We say **C** and **D** are *isomorphic* if for some $\mathcal{F}$ and $\mathcal{G}$ we have $\mathcal{GF}$ is the identity functor $\mathcal{I}_C$, and $\mathcal{FG} = \mathcal{I}_D$. By the discussion in Section 10.1 the categories $\mathbb{Z}$–**Mod** and **Ab** are isomorphic. It also follows from observations in Section 10.1 that the categories of elementary abelian p-groups and vector spaces over $\mathbb{F}_p$ are isomorphic. In practice we tend to identify such isomorphic categories. The following generalization of isomorphism between categories gives a broader and more useful notion of when two categories are "similar."

Definition. Categories **C** and **D** are said to be *equivalent* if there are functors $\mathcal{F}$ from **C** to **D** and $\mathcal{G}$ from **D** to **C** such that the functor $\mathcal{GF}$ is naturally isomorphic to $\mathcal{I}_C$ (the identity functor of **C**) and $\mathcal{FG}$ is naturally isomorphic to the identity functor $\mathcal{I}_D$.

It is an exercise that equivalence of categories is reflexive, symmetric and transitive. The example of Affine k-algebras in Section 15.5 is an equivalence of categories (where one needs to modify the direction of the arrows in the definition of a natural transformation to accommodate the contravariant functors in this example). Another example (which requires some proving) is that for R a commutative ring with 1 the categories of left modules R–**Mod** and $M_{n\times n}(R)$–**Mod** are equivalent.

Finally, we introduce the concepts of universal arrows and universal objects.

Definition.

(1) Let **C** and **D** be categories, let $\mathcal{F}$ be a functor from **C** to **D**, and let X be an object in **D**. A *universal arrow* from X to $\mathcal{F}$ is a pair $(U(X), \iota)$, where $U(X)$ is an object in **C** and $\iota : X \to \mathcal{F}U(X)$ is a morphism in **D** satisfying the following property: for any object A in **C** if φ is any morphism from X to $\mathcal{F}A$ in **D**, then there exists a unique morphism $\Phi : U(X) \to A$ in **C** such that $\mathcal{F}(\Phi)\iota = \varphi$, i.e., the following diagram commutes:

$$\begin{tikzcd} X \arrow[r, "\iota"] \arrow[dr, "\varphi"'] & \mathcal{F}U(X) \arrow[d, "\mathcal{F}(\Phi)"] \\ & \mathcal{F}A \end{tikzcd}$$

(2) Let **C** be a category and let $\mathcal{F}$ be a functor from **C** to the category **Set** of all sets. A *universal element* of the functor $\mathcal{F}$ is a pair (U, ι), where U is an object in **C** and ι is an element of the set $\mathcal{F}U$ satisfying the following property: for any object A in **C** and any element g in the set $\mathcal{F}A$ there is a unique morphism $\varphi : U \to A$ in **C** such that $\mathcal{F}(\varphi)(\iota) = g$.

Examples

(1) (*Universal Arrow: Free Objects*) Let R be a ring with 1. We translate into the language of universal arrows the statement that if $U(X)$ is the free R-module on a set X then any set map from X to an R-module A extends uniquely by R-linearity to an R-module homomorphism from $U(X)$ to A (cf. Theorem 6, Section 10.3): Let $\mathcal{F}$ be the forgetful functor from R–**Mod** to **Set**, so that $\mathcal{F}$ maps an R-module A to the set A, i.e., $A = \mathcal{F}A$ as sets. Let X be any set (i.e., an object in **Set**), let $U(X)$ be the free R-module with basis X, and let $\iota : X \to \mathcal{F}U(X)$ be the set map which sends each $b \in X$ to the basis element b in $U(X)$. Then the universal property of free R-modules is precisely the result that $(U(X), \iota)$ is a universal arrow from X to the forgetful functor $\mathcal{F}$.

Similarly, free groups, vector spaces (which are free modules over a field), polynomial algebras (which are free R-algebras) and the like are all instances of universal arrows.

(2) (*Universal Arrow: Fields of Fractions*) Let $\mathcal{F}$ be the forgetful functor from the category of fields to the category of integral domains, where the morphisms in both categories are *injective* ring homomorphisms. For any integral domain X let $U(X)$ be its field of fractions and let ι be the inclusion of X into $U(X)$. Then $(U(X), \iota)$ is a universal arrow from X to the functor $\mathcal{F}$ (cf. Theorem 15(2) in Section 7.5).

(3) (*Universal Object: Tensor Products*) This example refers to the construction of the tensor product of two modules in Section 10.4. Let **C** $=$ R–**Mod** be the category of R-modules over the commutative ring R, and let M and N be R-modules. For each R-module A let $\text{Bilin}(M, N; A)$ denote the set of all R-bilinear functions from $M \times N$ to A. Define a functor from R–**Mod** to **Set** on objects by

$$\mathcal{F} : A \longrightarrow \text{Bilin}(M, N; A),$$

and if $\varphi : A \to B$ is an R-module homomorphism then

$$\mathcal{F}(\varphi)(h) = \varphi \circ h \qquad \text{for every } h \in \text{Bilin}(M, N; A).$$

Let $U = M \otimes_R N$ and let ι be the bilinear function

$$\iota : M \times N \to M \otimes_R M \qquad \text{by} \qquad \iota(m, n) = m \otimes n,$$

so ι is an element of the set $\text{Bilin}(M, N; M \otimes_R N) = \mathcal{F}U$. Then $(M \otimes_R N, \iota)$ is a universal element of $\mathcal{F}$ because for any R-module A and for any bilinear map $g : M \times N \to A$ (i.e., any element of $\mathcal{F}A$) there is a unique R-module homomorphism $\varphi : M \otimes_R N \to A$ such that $g = \varphi \circ \iota = \mathcal{F}(\varphi)(\iota)$.

EXERCISES

1. Let **Nor**–N be the category described in Exercise 1 of Section 1, and let $\mathcal{F}$ be the inclusion functor from **Nor**–N into **Grp**. Describe a functor $\mathcal{G}$ from **Nor**–N into **Grp** such that the transformation η defined by $\eta_G : G \to G/N$ is a natural transformation from $\mathcal{F}$ to $\mathcal{G}$.

2. Let H and K be groups and let $\mathcal{H}\times$ and $\mathcal{K}\times$ be functors from **Grp** to itself described in Exercise 2 of Section 1. Let $\varphi : H \to K$ be a group homomorphism.

(a) Show that the maps $\eta_A : H \times A \to K \times A$ by $\eta_A(h, a) = (\varphi(h), a)$ determine a natural transformation η from $\mathcal{H}\times$ to $\mathcal{K}\times$.

(b) Show that the transformation η is a natural isomorphism if and only if φ is a group isomorphism.

3. Express the universal property of the commutator quotient group — described in Proposition 7(5) of Section 5.4 — as a universal arrow for some functor $\mathcal{F}$.

Index

A

B

C

D

E

F

G

H

I

J

K

L

M

N

O

P

Q

R

S

T

U

V

W

Z