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

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

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

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

Published by Birkhäuser Boston

Digital Second Edition, not to be sold, no ISBN

Published by the Author

HINTS FOR SOLUTIONS OF PROBLEMS

Chapter I

1. 582.

2. The Euclidean algorithm gives $11 = 1 \cdot 7 + 4$, $7 = 1 \cdot 4 + 3$, $4 = 1 \cdot 3 + 1$, $3 = 3 \cdot 1 + 0$. So the GCD is 1. Reversing the steps gives $1 = 4 - 1 \cdot 3 = (11 - 1 \cdot 7) - 1 \cdot (7 - 1 \cdot 4) = (11 - 1 \cdot 7) - 1 \cdot (7 - 1 \cdot (11 - 1 \cdot 7)) = 2 \cdot 11 - 3 \cdot 7$. So $(x, y) = (2, -3)$ is a solution in (a). For (b), the difference of any two solutions solves $11x + 7y = 0$, and the solutions of this are of the form $(x, y) = n(7, -11)$.

3. Let $d_n = \text{GCD}(a_1, \ldots, a_n)$. The sequence d_n is a monotone decreasing sequence of positive integers, and it must eventually be constant. This eventual constant value is d, and thus $d_n = d$ for suitably large n.

4. These n's divide $x + y - 2$ and the sum of $2x - 3y - 3$ and -2 times the $x + y - 2$, hence $x + y - 2$ and $-5y + 1$. A necessary and sufficient condition for $-5y + 1 = na$ to be solvable for the pair (a, y) is that $\text{GCD}(5, n) = 1$ by Proposition 1.2c. Let us see that the answer to the problem is $\text{GCD}(5, n) = 1$.

The n's we seek must further divide $5(x+y-2) = 5x+5y-10$ and $-5y+1$, hence also the sum $5x-9$, as well as $-5y+1$. If $\text{GCD}(5, n) = 1$, then $5x-9 = nb$ is solvable for (b, x). With our solutions (a, y) and (b, x), we have $5x + 5y - 10 = n(b - a)$. Since 5 divides the left side and $\text{GCD}(5, n) = 1$, 5 divides $b - a$. Write $b - a = 5c$. Then $x + y - 2 = nc$ and $-5y + 1 = na$, and we obtain $2x - 3y - 3 = n(2c + a)$.

5. $Q(x) = (X - 1)P(X) + (X^3 + x^2 + X + 1)$, $P(X) = X(X^3 + x^2 + X + 1)$ $+(X^2+1)$, $X^3+x^2+X+1 = (X+1)(X^2+1)+0$. Hence the GCD is $D(X) = X^2+1$. For (b), we retrace the steps, letting $R(X) = X^3 + X^2 + X + 1$. We have $D(X) = P(X)-XR(X) = P(X)-X(Q(X)-(X-1)P(X)) = (X^2-X+1)P(X)-XQ(X)$. Thus $A(X) = X^2 - X + 1$ and $B(X) = -X$.

6. The computation via the Euclidean algorithm, done within $\mathbb{C}[X]$, retains real numbers as coefficients throughout. By Proposition 1.15a one GCD has real coefficients. By Proposition 1.15c any GCD is a complex multiple of this polynomial with real coefficients.

7. In (a), we may assume, without loss of generality, that P has leading coefficient 1, so that $P(X) = X^n + a_{n-1}X^{n-1} + \cdots + a_0 = \prod_j (X - z_j)^{m_j}$. Define $Q(X) = \prod_j (X - \bar{z}_j)^{m_j}$. Then $Q(\bar{z}) = \prod_j (\bar{z} - \bar{z}_j)^{m_j} = \overline{\prod_j (z - z_j)} = \overline{P(z)}$. Replacing $\bar{z}$ by z gives $Q(z) = \overline{P(\bar{z})} = \overline{\bar{z}^n + a_{n-1}\bar{z}^{n-1} + \cdots + a_0} = z^n + \overline{a_{n-1}}z^{n-1} + \cdots + \overline{a_0}$.

Since P has real coefficients, $Q(z) = P(z)$ for all z. Then $Q - P$ has every z as a root and in particular has more than n roots. Hence it must be the 0 polynomial. So $\prod_j (X - \bar{z}_j)^{m_j} = \prod_j (X - z_j)^{m_j}$, and the result follows from unique factorization (Theorem 1.17).

In (b), the result of (a) shows that we may factor any real polynomial in $\mathbb{C}[X]$ with leading coefficient 1 in the form $\prod_{x_j \text{ real}}(X-x_j)^{m_j} \prod_{z_j \text{ nonreal}} \big((X-z_j)^{n_j}(X-\bar{z}_j)^{n_j}\big)$. The right side equals $\prod_{x_j \text{ real}}(X-x_j)^{m_j} \prod_{z_j \text{ nonreal}} \big(X^2-(z_j+\bar{z}_j)X+z_j\bar{z}_j\big)^{n_j}$. Every factor on the right side is in $\mathbb{R}[X]$, and the only way that the polynomial can be prime in $\mathbb{R}[X]$ is if only one factor is present. Thus the polynomial has degree at most 2.

8. For (a), let $\deg A = d$ and form the equation $A(p/q) = 0$. Multiply through by q^d in order to clear fractions. Every term in the equation except the leading term has q as a factor, and thus q divides the leading term p^d. Since $\text{GCD}(p, q) = 1$, no prime can divide q. Thus $q = \pm 1$, and $n = p/q$ is an integer. Forming the equation $A(n) = 0$, we see that n is a factor of each term except possibly the constant term a_0. Thus n divides a_0.

For (b), we apply (a) to both polynomials. The only possible rational roots of $X^2 - 2$ are ± 1 and ± 2, while the only possible rational roots of $X^3 + X^2 + 1$ are ± 1. Checking directly, we see that none of these possibilities is actually a root. By the Factor Theorem, neither $X^2 - 2$ nor $X^3 + X^2 + 1$ has a first-degree factor in $\mathbb{Q}[X]$. If a polynomial of degree ≤ 3 has a nontrivial factorization, then it has a first-degree factor. We conclude that $X^2 - 2$ and $X^3 + X^2 + 1$ are prime.

9. Computation gives $\text{GCD}(8645, 10465) = 455$. Therefore $8645/10465$ equals $19/23$ in lowest terms.

10. Apart from the identity, the cycle structures are those of (1 2) with 6 representatives, (1 2 3) with 8 representatives, (1 2 3 4) with 6 representatives, and (1 2)(3 4) with 3 representatives. This checks, since there are $4! = 24$ permutations in all.

11. Check that the function $\sigma \mapsto \sigma(1\ 2)$ is one-one from the set of permutations of sign $+1$ onto the set of permutations of sign -1.

12. (a) $x_3 \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$. (b) None. (c) $\begin{pmatrix} -11/3 \\ 10/3 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$.

13. By the definition of "step," an interchange of two rows (type (i)) takes n steps, and a multiplication of a row by a nonzero scalar (type (ii)) takes n steps. Also, replacement of a row by the sum of it and a multiple of another row (type (iii)) takes $2n$ steps. We proceed through the row-reduction algorithm column by column. For each of the n columns, we do possibly one operation of type (i) and then possibly an operation of type (ii). This much requires $\leq 2n$ steps. Then we do at most $n - 1$ operations of type (iii), requiring $\leq 2n(n-1)$ steps. Thus a single column is handled in $\leq 2n(n-1) + 2n = n^2$ steps, and the entire row reduction requires $\leq 2n^3$ steps.

14. $A + B = \begin{pmatrix} -2 & 11 \\ 3 & 8 \end{pmatrix}$, and $AB = \begin{pmatrix} -11 & 25 \\ -21 & 47 \end{pmatrix}$.

15. We induct on n, the result being clear for $n = 1$. Taking into account the fact that B commutes with A, we have $(A + B)^n = (A + B)(A + B)^{n-1} = (A +$

$B)$ $\sum_{k=0}^{n-1}\binom{n-1}{k}A^{n-1-k}B^k = \sum_{k=0}^{n-1}\binom{n-1}{k}A^{n-k}B^k + \sum_{k=0}^{n-1}\binom{n-1}{k}A^{n-1-k}B^{k+1} = \sum_{k=0}^{n-1}\binom{n-1}{k}A^{n-k}B^k + \sum_{k=1}^{n}\binom{n-1}{k-1}A^{n-k}B^k = A^n + \sum_{k=1}^{n-1}\left[\binom{n-1}{k}+\binom{n-1}{k-1}\right]A^{n-k}B^k + B^n$. In turn, the right side equals $\sum_{k=0}^{n}\binom{n}{k}A^{n-k}B^k$ by the Pascal-triangle identity for binomial coefficients.

16. Write $\begin{pmatrix}1&1&0\\0&1&1\\0&0&1\end{pmatrix}$ as $I+B$, where $B=\begin{pmatrix}0&1&0\\0&0&1\\0&0&0\end{pmatrix}$, and apply Problem 15. Since $B^2=\begin{pmatrix}0&0&1\\0&0&0\\0&0&0\end{pmatrix}$ and $B^3=0$, we obtain $(I+B)^n = I+nB+\frac{1}{2}n(n-1)B^2 = \begin{pmatrix}1&n&\frac{1}{2}n(n-1)\\0&1&n\\0&0&1\end{pmatrix}$.

17. $(AD)_{ij}=A_{ij}d_j$ and $(DA)_{ij}=d_iA_{ij}$. Thus $AD=DA$ if and only if $d_i=d_j$ for all (i,j) for which $A_{ij}\neq 0$.

18. $E_{kl}E_{pq}=\delta_{lp}E_{kq}$.

19. Check that $\begin{pmatrix}a&b\\c&d\end{pmatrix}$ times the asserted inverse is the identity. Then the matrix actually is the inverse. Apply the inverse to $\begin{pmatrix}a&b\\c&d\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}=\begin{pmatrix}p\\q\end{pmatrix}$ to obtain the value for $\begin{pmatrix}x\\y\end{pmatrix}$.

20. (a) No inverse. (b) $A^{-1}=\begin{pmatrix}-2/3&-4/3&1\\-2/3&11/3&-2\\1&-2&1\end{pmatrix}$. (c) $A^{-1}=\begin{pmatrix}1&-1&0\\-2&3&-1\\2&-5&4\end{pmatrix}$.

21. No. If the algorithm is followed, then the row of 0's persists throughout the row reduction, at worst moving to a different row at various stages.

22. If $C=(AB)^{-1}$, then $ABC=I$ shows that BC is the inverse of A and $CAB=I$ shows that CA is the inverse of B.

23. $(I+A)(I-A+A^2-A^3+\cdots+(-A)^{k-1}) = I-(-A)^k = I$ shows that $I-A+A^2-A^3+\cdots+(-A)^{k-1}$ is an inverse.

24. Let S be the set of positive integers, and let $f(n)=n+1$. Take $g(n)$ to be $n-1$ for $n>1$ and $g(1)=1$. Then $g\circ f$ is the identity. But f is not onto S, and g is not one-one.

25. Take $A=\begin{pmatrix}1&0\end{pmatrix}$ and $B=\begin{pmatrix}1\\0\end{pmatrix}$. Then $BA=\begin{pmatrix}1&0\\0&0\end{pmatrix}$. More generally, if $A=\begin{pmatrix}a&b\end{pmatrix}$ and $B=\begin{pmatrix}c\\d\end{pmatrix}$, then $BA=\begin{pmatrix}ca&cb\\da&db\end{pmatrix}$. If the upper right entry is 0, then $c=0$ or $b=0$. But then one of the two diagonal entries must be 0, and hence BA cannot be the identity.

26. The set of common multiples is a nonempty set of positive integers because ab is in it. Therefore it has a least element.

27. This is a restatement of Corollary 1.7.

28. Let a and b have prime factorizations $a=p_1^{k_1}\cdots p_r^{k_r}$ and $b=p_1^{l_1}\cdots p_r^{l_r}$. Problem 27 shows that any positive common multiple N of a and b is of the form $p_1^{m_1}\cdots p_r^{m_r}q_1^{n_1}\cdots q_s^{n_s}$ with $m_j\geq k_j$, $m_j\geq l_j$, and $n_j\geq 0$, and certainly any positive

integer of this form is a common multiple. The inequalities for m_j are equivalent with the condition $m_j \geq \max(k_j, l_j)$. The smallest positive integer of this kind has $m_j = \max(k_j, l_j)$ and $n_j = 0$. This proves (a). In combination with the form of N, the formula for $\mathrm{LCM}(a, b)$ proves (b). Conclusion (c) follows from Corollary 1.8 and the identity $k_j + l_j = \min(k_j, l_j) + \max(k_j, l_j)$.

29. If $a_j = p_1^{k_{1,j}} \cdots p_r^{k_{r,j}}$ is a prime factorization of a_j, then $\mathrm{LCM}(a_1, \ldots, a_t) = p_1^{\max_{1 \leq j \leq r}\{k_{1,j}\}} \cdots p_r^{\max_{1 \leq j \leq r}\{k_{r,j}\}}$, just as with Corollary 1.11.

Chapter II

1. The methods at the end of Section 2 lead to the basis $\left\{(\frac{2}{3}, 1, 0), (-\frac{5}{3}, 0, 1)\right\}$ for (a) and to the basis $\left\{(1, -\frac{1}{2}, 2)\right\}$ for (b).

2. For $0 \leq k < n$, the two recursive formulas and one application of associativity give $v_{(k+1)} + v^{(k+2)} = (v_{(k)} + v_{k+1}) + v^{(k+2)} = v_{(k)} + (v_{k+1} + v^{(k+2)}) = v_{(k)} + v^{(k+1)}$, and (a) follows.

For (b), we proceed by induction on n, the cases $n \leq 3$ being handled by associativity. Suppose that the result holds for sums of fewer than n vectors, with $n \geq 4$. In a sum of n vectors, there is some outer plus sign, and the inductive hypothesis means that the sum is of the form $(v_1 + \cdots + v_k) + (v_{k+1} + \cdots + v_n)$, the expressions $v_1 + \cdots + v_k$ and $v_{k+1} + \cdots + v_n$ being unambiguous. The inductive hypothesis means that we have $v_1 + \cdots + v_k = v_{(k)}$ and $v_{k+1} + \cdots + v_n = v^{(k+1)}$, and hence the expression we are studying is of the form $v_{(k)} + v^{(k+1)}$. Part (a) shows that this is independent of k, and hence (b) follows.

3. From Section I.4, σ is a product of transpositions, and hence it is enough to prove the result for a transposition. When $r+1 < s$, iteration of the identity $(r \quad s) = (r \quad r+1)(r+1 \quad s)(r \quad r+1)$ shows that any transposition is a product of transpositions of the form $(r \quad r+1)$, and hence it is enough to prove the formula for $\sigma = (r \quad r+1)$. This case is just the commutative law, and the result follows.

4. (a) $\{(1 \quad 2 \quad -1), (0 \quad 0 \quad 1)\}$; (b) $\left\{\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}\right\}$; (c) 2.

5. If R is a reduced row-echelon form of A, then we know that $R = EA$, where E is a product of invertible elementary matrices. Since A has rank one, R has a single nonzero row r and is of the form $e_1 r$, where e_1 is the first standard basis vector. Then $A = E^{-1}R = (E^{-1}e_1)r$, and we can take $c = E^{-1}e_1$.

6. In (a), let $u_1, \ldots, u_s$ be the rows of R having at least one of the first r entries nonzero, and let $u_{s+1}, \ldots, u_m$ be the other rows. For each i with $1 \leq i \leq s$, the first nonzero entry of u_i corresponds to a corner variable and occurs in the $j(i)^{\text{th}}$ position with $j(i) \leq r$. The most general member of the row space of A is of the form $c_1 u_1 + \cdots + c_m u_m$, and the $j(i)^{\text{th}}$ entry of this is c_i. For this row vector to be in the indicated span, we must have $c_i = 0$ for $i \leq s$.

In (b), let R' be a second reduced row-echelon form, and let its nonzero rows be $v_1, \ldots, v_m$. From part (a), it follows that the linear span of $u_{s+1}, \ldots, u_m$ equals the linear span of $v_{s+1}, \ldots, v_m$ for each s. Moreover, the value of each $j(i)$ has to be the same for u_i as for v_i. Inducting downward, we prove that $u_i = v_i$ for each i. For $i = m$, this follows since the first nonzero entry is 1 for both u_m and v_m. Assuming the result for $s+1$, we write $v_s = c_s u_s + c_{s+1}u_{s+1} + \cdots + c_m u_m$. We have $c_s = 1$ since the first nonzero entry of u_s and v_s is 1, and we have $c_{s+i} = 0$ for $i > 0$ since the $j(s+i)^{\text{th}}$ entry of this equality of row vectors is $0 = c_{s+i}$. Thus $v_s = u_s$, and the induction is complete.

7. Let $E = \{x_1, \ldots, x_N\}$, and let $f_1, \ldots, f_n$ be a basis of U. Form the matrix $A = \begin{pmatrix} f_1(x_1) & \cdots & f_1(x_N) \\ \vdots & \ddots & \vdots \\ f_n(x_1) & \cdots & f_n(x_N) \end{pmatrix}$. By assumption, A has row rank n. Therefore it has column rank n, and there exist n linearly independent columns, say columns $j_1, \ldots, j_n$. Then $D = \{x_{j_1}, \ldots, x_{j_n}\}$.

8. Let the listed basis be Γ, and let Σ be the standard basis. Then $\begin{pmatrix} I \\ \Sigma\Gamma \end{pmatrix} = \begin{pmatrix} 3 & -4 \\ -2 & 3 \end{pmatrix}$, the inverse matrix is $\begin{pmatrix} I \\ \Gamma\Sigma \end{pmatrix} = \begin{pmatrix} 3 & 4 \\ 2 & 3 \end{pmatrix}$, and $\begin{pmatrix} L \\ \Gamma\Gamma \end{pmatrix}$ is the product $\begin{pmatrix} 3 & 4 \\ 2 & 3 \end{pmatrix}\begin{pmatrix} -6 & -12 \\ 6 & 11 \end{pmatrix}\begin{pmatrix} 3 & -4 \\ -2 & 3 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$.

9. One could compute the matrix of $I - D^2$ in an explicit basis, but an easier way is to observe that $D^3 = 0$ and hence $(I - D^2)(I + D^2) = I - D^4 = I$.

10. Since $\text{image}(AB) \subseteq \text{image}\, A$, we have $\text{rank}(AB) = \dim \text{image}(AB) \leq \dim \text{image}\, A = \text{rank}\, A$. Similarly $\text{rank}((AB)^t) = \text{rank}(B^t A^t) \leq \text{rank}\, B^t$. Since a matrix and its transpose have the same rank (by the equality of row rank and column rank), $\text{rank}(AB) \leq \text{rank}\, B$.

11. Since A has n columns, $\text{rank}\, A \leq n$. Applying Problem 10 gives $\text{rank}(AB) \leq \text{rank}\, A \leq n$. Since $n < k = \text{rank}\, I$, we cannot have $AB = I$.

12. Take $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Then $AB = B$ has rank 1 while $BA = 0$ has rank 0.

13. $\{\cosh t, \sinh t\}$.

14. Let $\{v_n \mid n \in \mathbb{Z}\}$ be a countably infinite basis. For each subset S of $\mathbb{Z}$, define v'_S to be the member of V' such that $v'_S(v_n)$ is 1 if n is in S and is 0 if not. Choose by Theorem 2.42 a subset of $\{v'_S\}$ that is a basis for the linear span of all v'_S. Arguing by contradiction, assume that this basis is countable. Number the S's in question as $S_1, S_2, \ldots$. Any v'_S then has a unique expansion as $v'_S = c_1 v'_{S_1} + \cdots + c_k v'_{S_k}$ for some k. Fix k, and let v'_S be expandable for this k. Let $E \subseteq \{1, \ldots, k\}$. Let m and n be such that v_m and v_n are in S_j for j in E and are not in S_j for j in $\{1, \ldots, k\} - E$. Then $v'_{S_j}(v_m) = v'_{S_j}(v_n)$ for $j = 1, \ldots, k$, and hence $v'_S(v_m) = v'_S(v_n)$. Thus with k fixed, the number of S's for which v'_S is expandable is at most 2^k. In particular, it is finite. Taking the union over k, we find that there are only countably many v'_S in

the linear span of $v'_{S_1}, v'_{S_2}, \ldots$. But there are uncountably many subsets S of $\mathbb{Z}$, and we have thus arrived at a contradiction. We conclude that our subset of all v'_S that is a basis for the linear span must have been uncountable.

15. For (a), take L, M, and N to be the three 1-dimensional subspaces of $\mathbb{R}^2$ shown in Figure 2.1. Then $L \cap (M + N) = L$ while $(L \cap M) + (L \cap N) = 0$.

For (b), we always have $\supseteq$ since $L \cap (M+N) \supseteq L \cap M$ and $L \cap (M+N) \supseteq L \cap N$.

For (c), if $l = m + n$ is in $L \cap (M + N)$, then $L \supseteq M$ implies that $n = l - m$ is in L. So $l = m + n$ has $m \in L \cap M$ and $n \in L \cap N$.

16. Take M, N_1, and N_2 to be the three 1-dimensional subspaces of $V = \mathbb{R}^2$ shown in Figure 2.1. Then $M \oplus N_1 = M \oplus N_2 = \mathbb{R}^2$, but $N_1 \neq N_2$.

17. (b) only.

18. In $V_1 \oplus \cdots \oplus V_n$, let p_j pick off the j^{th} coordinate, and let i_j carry v_j to $(0, \ldots, 0, v_j, 0, \ldots, 0)$. Then $p_r i_s$ is I on V_s if $r = s$ and is 0 on V_s if $r \neq s$. Also, $\sum_{k=1}^{n} i_k p_k = I$ on $V_1 \oplus \cdots \oplus V_n$.

19. Corollary 2.15 shows that $\dim \ker T + \dim \text{image } T = n$. Since $\ker T$ and image T have 0 intersection, the union of bases of $\ker T$ and image T is a linearly independent set of n vectors in $\mathbb{R}^n$. This set must be a basis of $\mathbb{R}^n$, and hence $\mathbb{R}^n = \ker T \oplus \text{image } T$. This proves (a).

For (b), let $T^2 = T$ and suppose that v is in $\ker T \cap \text{image } T$. Since v is in image T, we have $v = T(w)$ for some w. Then $v = T(w) = T^2(w) = T(T(w)) = T(v)$, and the right side is 0 since v is in $\ker T$. Consequently $\ker T \cap \text{image } T = 0$.

20. Define $L : V_1' \oplus V_2' \to (V_1 \oplus V_2)'$ by $L(\mu_1, \mu_2)(v_1, v_2) = \mu_1(v_1) + \mu_2(v_2)$.

21. Proposition 2.25 shows that $y \mapsto z$ is onto the subset of z's in V' such that $M \subseteq \ker z$, i.e., is onto Ann M. Since q is onto V/M, $y \mapsto z$ is one-one.

22. The kernel of q is M, and thus the kernel of $q\big|_N$ is $M \cap N$. So $q\big|_N$ is one-one if and only if $M \cap N = 0$.

If $M+N = V$, then any $v \in V$ is of the form $m+n$; so v has $v+M = m+n+M = n + M = q(n)$, and q carries N onto V/M. Conversely if q carries N onto V/M, let $v \in V$ be given, and choose n with $q(n) = v + M$. Then $n + M = v + M$, and hence $v - n$ is in M. This says that $V = M + N$.

Consequently $q\big|_N : N \to V/M$ is an isomorphism if and only if $M \cap N = 0$ and $M + N = V$, and we know from Proposition 2.30 that this pair of conditions is equivalent to the single condition $V = M \oplus N$.

23. If A^{-1} has integer entries, then $\det A$ and $\det A^{-1}$ are integers that are reciprocals, and we conclude that $\det A = \pm 1$. If $\det A = \pm 1$, then Cramer's rule shows that A^{-1} has integer entries.

24. When $r = \text{rank } A$, there exist r linearly independent rows. Say that these are the ones numbered $i_1, \ldots, i_r$. Let A_1 be the r-by-n matrix obtained by deleting the remaining rows. Since A_1 has rank r, it has r linearly independent columns. Say that these are the ones numbered $j_1, \ldots, j_r$. Let A_2 be the r-by-r matrix obtained by

deleting the remaining columns. Then A_2 is a square matrix of rank r, is therefore invertible, and must have nonzero determinant. In the reverse direction if some s-by-s submatrix has nonzero determinant, then the rows of the submatrix are linearly independent, and certainly the corresponding rows of A are linearly independent. Thus $s \le \text{rank } A$.

25. Let the expression in question be $f(t) = \sum_{i=1}^{n} a_i e^{c_i t}$. Put $r_i = e^{c_i}$. The numbers r_i are distinct. The fact that $f(0) = f(1) = \cdots = f(n-1) = 0$ says that the product of the Vandermonde matrix formed from $r_1, \ldots, r_n$ times the column vector $(a_1, \ldots, a_n)$ is the 0 vector. Since the Vandermonde matrix is invertible, it follows that $(a_1, \ldots, a_n)$ is the 0 vector.

26. The characteristic polynomial is $\lambda^2 - 5\lambda + 6 = (\lambda - 2)(\lambda - 3)$. The eigenvectors for $\lambda = 2$ are all nonzero multiples of $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$, and the eigenvectors for $\lambda = 3$ are all nonzero multiples of $\begin{pmatrix} 1 \\ 3 \end{pmatrix}$.

27. $\sum_i (C^{-1}AC)_{ii} = \sum_i \sum_{j,k} (C^{-1})_{ij} A_{jk} C_{ki} = \sum_{j,k} A_{jk} \sum_i C_{ki} (C^{-1})_{ij} = \sum_{j,k} A_{jk} \delta_{kj} = \sum_j A_{jj}$.

28. For $n = 2$, direct computation gives $\lambda^2 - a_1\lambda - a_0$. Similarly we obtain $\lambda^3 - a_2\lambda^2 - a_1\lambda - a_0$ when $n = 3$. We are thus led to the guess in the general case that the determinant is $\lambda^n - a_{n-1}\lambda^{n-1} - \cdots - a_1\lambda - a_0$. This is proved by induction, using expansion in cofactors about the first column. The term from the (1, 1) entry, by the inductive hypothesis, is $\lambda(\lambda^{n-1} - a_{n-1}\lambda^{n-2} - \cdots - a_1)$, and the term from the $(1, n)$ entry is $(-1)^{n+1}(-a_0) \det B$, where B is a lower triangular matrix of size $n-1$ with -1 in every diagonal entry. Then $\det B = (-1)^{n-1}$, and substitution completes the induction.

29. In (a), we have $\det(\lambda I - AB) = \det(A(\lambda A^{-1} - B)) = \det A \det(\lambda A^{-1} - B) = \det(\lambda A^{-1} - B) \det A = \det((\lambda A^{-1} - B)A) = \det(\lambda I - BA)$.

For (b), we know from the fact that the characteristic polynomial of A is a polynomial that there are only finitely many ϵ for which $A + \epsilon I$ fails to be invertible. Thus there is some $\epsilon_0 > 0$ such that $A + \epsilon I$ is invertible when $0 < \epsilon < \epsilon_0$. By (a), these ϵ's have $\det(\lambda I - (A + \epsilon I)B) = \det(\lambda I - B(A + \epsilon I))$. Since det is a polynomial in the entries of the matrix it is applied to, $\det(\lambda I - C)$ is a continuous function of the entries of C. Taking $C = (A + \epsilon I)B$ and then $C = B(A + \epsilon I))$, and letting ϵ tend to 0, we obtain $\det(\lambda I - AB) = \det(\lambda I - BA)$.

30. In $\mathbb{R}^1$, let the n^{th} spanning set consist of $\{(r) \mid 0 < r < 1/n\}$. These each span $\mathbb{R}^1$, but their intersection is empty and the empty set does not span $\mathbb{R}^1$.

31. Let $\{v_\alpha\}$ be a basis of V. For each α, define a member v'_α of V' by saying that $v'_\alpha(v_\beta)$ is 1 for $\beta = \alpha$ and is 0 for $\beta \ne \alpha$. In addition, let w_0 be the member of V' that is 1 on each V_α. Arguing by contradiction, suppose that w_0 is in $\iota(V)$. Then we can write $w_0 = \sum_{\beta \in F} c_\beta \iota(v_\beta)$ for some finite set F, and for each α we have $w_0(v'_\alpha) = \sum_{\beta \in F} c_\beta \iota(v_\beta)(v'_\alpha) = \sum_{\beta \in F} c_\beta v'_\alpha(v_\beta)$. The right side is nonzero only if some $\beta \in F$ has $v'_\alpha(v_\beta) \ne 0$, i.e., only if α is in F. On the other hand, the

left side is 1 for every α. For this equality to happen for all α forces F to be infinite, contradiction.

32. $\operatorname{Ann}(M+N) \subseteq \operatorname{Ann} M$, and $\operatorname{Ann}(M+N) \subseteq \operatorname{Ann} N$; thus $\operatorname{Ann}(M+N) \subseteq \operatorname{Ann} M \cap \operatorname{Ann} N$. If v' is 0 on M and is 0 on N, then it is 0 on $M+N$. Hence $\operatorname{Ann}(M+N) \supseteq \operatorname{Ann} M \cap \operatorname{Ann} N$.

33. $\operatorname{Ann}(M \cap N) \supseteq \operatorname{Ann} M$, and $\operatorname{Ann}(M \cap N) \supseteq \operatorname{Ann} N$; thus $\operatorname{Ann}(M \cap N) \supseteq \operatorname{Ann} M + \operatorname{Ann} N$. Let $\{u_\alpha\}$ be a basis of $M \cap N$, let v_β be vectors added to $\{u_\alpha\}$ to obtain a basis of M, and let w_γ be vectors added to $\{u_\alpha\}$ to obtain a basis of N. Then $\{u_\alpha\} \cup \{v_\beta\} \cup \{w_\gamma\}$ is a basis of $M+N$. Let x_δ be vectors added to this to obtain a basis of V. If v' is given in $\operatorname{Ann}(M \cap N)$, define v_1' to be v' on all the basis vectors but the v_β, where it is to be 0, and define $v_2' = v' - v_1'$. Then $v' = v_1' + v_2'$ with $v_1' \in \operatorname{Ann} M$ and $v_2' \in \operatorname{Ann} N$. So $\operatorname{Ann}(M \cap N) \subseteq \operatorname{Ann} M + \operatorname{Ann} N$.

34. Let v be in M, and let v' be in $\operatorname{Ann} M$. Then $\iota(v)(v') = v'(v) = 0$. This proves (a).

For (b), Propositions 2.19 and 2.20a give $\dim \operatorname{Ann} M = \dim V' - \dim M$ and $\dim \operatorname{Ann}(\operatorname{Ann} M) = \dim V'' - \dim \operatorname{Ann} M = \dim V'' - (\dim V' - \dim M) = \dim M = \dim \iota(M)$. This equality in the presence of the inclusion $\iota(M) \subseteq \operatorname{Ann}(\operatorname{Ann} M)$ implies $\iota(M) = \operatorname{Ann}(\operatorname{Ann} M)$ by Corollary 2.4.

For (c), let V be as in Problem 31, and put $M = V$. Then $\operatorname{Ann}(M) = 0$ and $\operatorname{Ann}(\operatorname{Ann} M) = V'' \neq \iota(V)$.

35. Parts (a) and (b) follow by writing out individual entries of the products as appropriate sums.

36. If A or D is not invertible, then suitable row operations on the matrix on the left side exhibit the matrix on the left as not invertible, and hence both sides are 0. Thus we may assume that A^{-1} and D^{-1} exist. Problem 35c allows us to decompose the given matrix as $\begin{pmatrix} A & B \\ 0 & D \end{pmatrix} = \begin{pmatrix} A & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} I & 0 \\ 0 & D \end{pmatrix} \begin{pmatrix} I & A^{-1}B \\ 0 & I \end{pmatrix}$. The determinant of the product is the product of the determinants. Using the defining formula for det, we see that the first two determinants from the right side are $\det A$ and $\det D$. The third determinant is 1 since the matrix is triangular with 1's on the diagonal.

37. In effect, we do row reduction with blocks, taking advantage of Problem 35c. We have $\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} A & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} I & A^{-1}B \\ C & D \end{pmatrix} = \begin{pmatrix} A & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} I & 0 \\ C & I \end{pmatrix} \begin{pmatrix} I & A^{-1}B \\ 0 & D-CA^{-1}B \end{pmatrix}$. Taking the determinant of both sides and using Problem 36, we obtain $\det \begin{pmatrix} A & B \\ C & D \end{pmatrix} = (\det A) \det(D - CA^{-1}B) = \det(AD - ACA^{-1}B)$, and this equals $\det(AD - CB)$ since $AC = CA$.

38. The matrices $\begin{pmatrix} A \\ 0 \end{pmatrix}$ and $(B \; 0)$ are of size n-by-n, and their products in the two orders are $\begin{pmatrix} AB & 0 \\ 0 & 0 \end{pmatrix}$ and BA. Problem 29 shows that $\det\left(\lambda I_n - \begin{pmatrix} AB & 0 \\ 0 & 0 \end{pmatrix}\right) = \det(\lambda I_n - BA)$. The left side equals $\lambda^{n-k} \det(\lambda I_k - AB)$, and the result follows.

39. Substitute the definitions of the determinants of $A(S)$ and $\widehat{A}(S)$ into the right side, sort out the signs, and verify that the result is the defining expression for $\det A$.

40. Expansion in cofactors about the last row gives $\det A_n = (A_n)_{nn} \det \widehat{(A_n)_{nn}} - (A_n)_{n-1,n} \det \widehat{(A_n)_{n-1,n}} = 2 \det A_{n-1} + \det B$, where B in block form is the square matrix of size $n-1$ given by $B = \begin{pmatrix} A_{n-2} & 0 \\ 0 & -1 \end{pmatrix}$. Expansion by cofactors of $\det B$ about the last row shows that $\det B = -\det A_{n-2}$, and the stated formula results.

41. Inspection gives $\det A_1 = 2$ and $\det A_2 = 3$. The function f with $f(n) = \det A_n - (n+1)$ thus has $f(1) = f(2) = 0$ and $f(n) = 2f(n-1) - f(n-2)$ for $n \geq 3$, and it must be 0 for all $n \geq 1$.

42. The only changes in (a) are notational. For (b), we compute $\det C_2 = \det C_3 = 2$, and the formula $\det C_n = 2$ follows as in Problem 41.

43. For (b), we interchange the first two rows and then interchange the first two columns. The determinant does not change.

44. For (b), we interchange the third and fourth rows and then interchange the third and fourth columns. For (c) we change the list of rows and columns of A_n from 1, 2, 3, 4, 5 to 3, 5, 4, 2, 1.

45. The area of the rectangle is $(a+b)(c+d)$, the two trapezoids have areas $\frac{1}{2}d(a+(a+b))$ and $\frac{1}{2}a(d+(c+d))$, and the two triangles have areas $\frac{1}{2}ac$ and $\frac{1}{2}bd$. The difference is $bc - ad$. The answer is independent of the picture except for a sign. Thus the answer is the absolute value of the determinant.

46. The geometric effect is to leave the left edge where it is and to translate the right edge parallel to itself in the same direction. The area is unchanged because the parallelogram can still be regarded as having base from $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ to $\begin{pmatrix} a \\ c \end{pmatrix}$ and having the same distance between the parallel sides. The algebraic effect is that of the column operation of replacing the second column by the sum of it and s times the first column.

47. Right multiplication by $\begin{pmatrix} 1 & 0 \\ t & 1 \end{pmatrix}$ leaves the bottom edge where it is and translates the top edge parallel to itself in the same direction; algebraically it corresponds to the column operation of replacing the first column by the sum of it and t times the second column. Right multiplication by $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ interchanges the left and bottom sides of the parallelogram and corresponds to interchange of the two columns of the matrix. Right multiplication by $\begin{pmatrix} q & 0 \\ 0 & 1 \end{pmatrix}$ corresponds to stretching the left side by a factor of q if $q > 0$, along with reversing the direction if $q < 0$, and the algebraic effect is the column operation of multiplying the first column by q. The effect of $\begin{pmatrix} 1 & 0 \\ 0 & r \end{pmatrix}$ is similar but affects the bottom edge instead of the left edge.

48. The roles of rows and columns are interchanged by the transpose operation, and the determinant is unaffected by transpose according to Proposition 2.35. In view of Proposition 1.29, A can thus be put in reduced column-echelon form by a sequence of column operations, each of which corresponds to right multiplication by a suitable elementary matrix. The result is an equality saying that the product of A and some elementary matrices is the identity. Using inverses shows that A is

the product of elementary matrices. The product can be applied a step at a time to the cube determined by the standard basis, and each step either preserves volume or multiplies it by a known factor, up to a minus sign. The product of these numerical factors is the determinant, up to a minus sign. Hence the volume of the parallelepiped has to be the product of these factors, with its sign made positive.

Chapter III

1. Since $\operatorname{Tr} B^*A = \sum_{i,j} A_{ij}\overline{B}_{ij}$, the inner product is the usual inner product on the n^2 entries. Then (a) and (b) are immediate. For (c), (a) gives the result of Parseval's equality relative to the orthonormal basis in (b).

For (d), let U be the unitary matrix with columns $u_1, \dots, u_n$, i.e., the matrix $\begin{pmatrix} I \\ \Sigma\Gamma \end{pmatrix}$, where $\Gamma = (u_1, \dots, u_n)$ and Σ is the standard ordered basis. Then $\|A\|_{\mathrm{HS}}^2 = \operatorname{Tr}(A^*A) = \operatorname{Tr}(U^{-1}A^*AU) = \sum_{i,j} |(AU)_{ij}|^2 = \sum_j \|Au_j\|^2$, and this equals $\sum_{i,j} |v_i^* A u_j|^2$ by Parseval's equality.

In (e), $W^\perp$ consists of all matrices that are 0 along the diagonal. It has dimension $n^2 - n$.

2. The system has unknowns $c_0, c_1, \dots, c_n$, where $p_n(x) = c_0 + c_1 x + \cdots + c_n x^n$, and the k^{th} equation, for $0 \le k \le n$, comes from the equality for $f(x) = x^k$, namely $2^{-k} = \sum_{j=0}^n (j + k + 1)^{-1} c_j$.

3. $(LM)^* = M^*L^* = ML$ is equal to LM if and only if $LM = ML$.

4. A vector u is in $\ker L$ if and only if $(L(u), v) = 0$ for all v, if and only if $(u, L^*(v)) = 0$ for all v, if and only if u is in $(\operatorname{image} L^*)^\perp$.

5. There are none. The characteristic polynomial has no real roots, but all roots must be real if A is Hermitian.

6. The map $v_1 \mapsto (L(v_1), v_2)_2$ is a linear functional on V_1 and hence is given by the inner product with a unique member u_1 of V_1, i.e., $((L(v_1), v_2)_2 = (v_1, u_1)_1$, and we define this element u_1 to be $L^*(v_2)$. We readily check that L^* is linear, and (a) is then proved. The proof of (b) proceeds in the same way as in the case that $V_1 = V_2$.

7. In (a), if v is in $S^\perp \cap T^\perp$, then v is in $V^\perp = 0$. Thus $S^\perp + T^\perp$ is a direct sum. We have $\dim V = \dim S + \dim T = (\dim V - \dim S^\perp) + (\dim V - \dim T^\perp) = 2\dim V - \dim S^\perp - \dim T^\perp$. Therefore $\dim V = \dim(S^\perp + T^\perp)$. The inclusion plus the equality of the finite dimensions forces $V = S^\perp + T^\perp$.

In (b), let λ be 0 or 1. Then $E^*u = \lambda u$ if and only if $(E^*u, v) = \lambda(u, v)$ for all v, if and only if $(u, Ev) = \lambda(u, v)$ for all v. When $\lambda = 1$, this says that $E^*u = u$ if and only if $u \perp (I - E)v$ for all v, hence if and only if $u \perp T$, hence if and only if u is in $T^\perp$. When $\lambda = 0$, it says that $E^*u = 0$ if and only if $u \perp Ev$ for all v, hence if and only if $u \perp S$, hence if and only if u is in $S^\perp$.

8. The formulas of the Gram–Schmidt orthogonalization process have $v_j = c_j(ge_j) + \sum_{i<j} a_{ij}v_i$ with $c_j > 0$. Therefore $ge_j = c_j^{-1}v_j + \sum_{i<j} b_{ij}v_i$, and

$$\begin{aligned}(k^{-1}g)_{ij} &= \textstyle\sum_l (k^{-1})_{il}g_{lj} = \sum_l (k^{-1})_{il}(ge_j)_l \\ &= c_j^{-1}\textstyle\sum_l (k^{-1})_{il}(v_j)_l + \sum_l \sum_{m<j} (k^{-1})_{il}b_{mj}(v_m)_l \\ &= c_j^{-1}(k^{-1}v_j)_i + \textstyle\sum_{m<j} b_{mj}(k^{-1}v_m)_i = c_j^{-1}\delta_{ij} + \sum_{m<j} b_{mj}\delta_{im}.\end{aligned}$$

If $i = j$, the right side is c_j^{-1} and is positive. If $i > j$, then every term on the right side is 0. Thus $k^{-1}g$ is upper triangular with positive diagonal entries. Since k carries the standard orthonormal basis to the orthonormal basis $\{v_1, \dots, v_n\}$, k is unitary.

9. For (a), the Spectral Theorem and Corollary 3.22 show that A is similar to a diagonal matrix with positive diagonal entries. Thus $\det A > 0$. In (b), we specialize the inequality $\bar{x}^t Ax > 0$ to x's that are 0 except in the entries numbered $i_1, \dots, i_n$, and we find that the submatrix is positive definite. Then the result follows from Corollary 3.22.

10. Take $g = \sqrt{A}$ in Problem 8 and obtain $\sqrt{A} = kt$ with k unitary and t upper triangular with positive diagonal entries. Then $A = (\sqrt{A})^*(\sqrt{A}) = (kt)^*(kt) = t^*t$.

11. The roots of the characteristic polynomial are $\frac{1}{2}(a+d+s)$ and $\frac{1}{2}(a+d-s)$, where $s = \sqrt{(a-d)^2 + 4|b|^2}$. Let $r = \frac{1}{2}(-a+d+s)$. Then $D = \begin{pmatrix} \frac{1}{2}(a+d+s) & 0 \\ 0 & \frac{1}{2}(a+d-s) \end{pmatrix}$ and $U = (b^2+r^2)^{-1/2}\begin{pmatrix} b & -r \\ r & \bar{b} \end{pmatrix}$.

12. In (a), the conditions $ad - |b|^2 > 0$ and $a + d > 0$ together are necessary and sufficient. In (b), let $\sqrt{D} = \begin{pmatrix} \sqrt{\frac{1}{2}(a+d+s)} & 0 \\ 0 & \sqrt{\frac{1}{2}(a+d-s)} \end{pmatrix}$, and let U be as in the previous problem. Then the positive definite square root of A is $U\sqrt{D}\,U^{-1}$.

13. The Spectral Theorem shows that A has a basis of eigenvectors, each with a real eigenvalue. If v is an eigenvector with eigenvalue λ, then $v^tAv = 0$ says that $\lambda\|v\|^2 = 0$. So every eigenvalue is 0, and A, being similar to a diagonal matrix, has to be 0.

14. Choosing a basis of eigenvectors, we may solve the corresponding problem for diagonal matrices. Thus let A be a diagonal matrix, and assume, without loss of generality, that $A_{11} = \cdots = A_{kk} = 1$ and $A_{jj} \neq 1$ for $j > k$. Then the given equation $(I-A)^2y = (I-A)x$ says that $(1-A_{jj})^2y_j = (1-A_{jj})x_j$ for all j. Thus define y_j to be 0 if $j \le k$, and choose $y_j = (1-A_{jj})^{-1}x_j$ for $j > k$.

15. $LL^* = (UP)(UP)^* = (PU)(PU)^* = PUU^*P = P^2 = PU^*UP = (UP)^*(UP) = L^*L$.

16. The family has a basis of simultaneous eigenvectors, and the matrices are all diagonal in this basis. So the answer is the dimension of the vector space of diagonal matrices, namely n.

17. In (a), $c^t G(v_1, \ldots, v_n)\bar{c} = \sum_{i,j} c_i (v_i, v_j)\bar{c}_j = \left(\sum_i c_i v_i, \sum_j c_j v_j\right) = \|c_1 v_1 + \cdots + c_n v_n\|^2$. Thus Corollary 3.22 shows that $G(v_1, \ldots, v_n)$ is positive semidefinite. Moreover, $\|c_1 v_1 + \cdots + c_n v_n\|^2 = 0$ for some $c \neq 0$ if and only if $v_1, \ldots, v_n$ are linearly dependent. Thus $G(v_1, \ldots, v_n)$ is definite if and only if $v_1, \ldots, v_n$ are linearly independent. We know that a positive semidefinite matrix is definite if and only if it is invertible, and thus $\det G(v_1, \ldots, v_n) > 0$ if and only if $v_1, \ldots, v_n$ are linearly independent; this proves (b). In (c), equality holds in the Schwarz inequality if and only if the two vectors are linearly dependent, i.e., if and only if one of them is a multiple of the other.

18. This is immediate by induction.

19. For (a), the left side is $D^2(X^{n+1}) = (n+1)D(X^n X')$. Comparing with the expected right side, we see that we are to show that

$$nD(X^n X') \stackrel{?}{=} (2n+1)nX''X^n + 4n^2 X^{n-1}.$$

The left side equals nX^{n-1} times $n(X')^2 + XX''$, while the right side equals nX^{n-1} times $(2n+1)X''X + 4n$. Since

$$\begin{aligned} n(X')^2 + XX'' &= 4nx^2 + 2x^2 - 2 \\ &= (4n+2)x^2 - (4n+2) + 4n = (2n+1)X''X + 4n, \end{aligned}$$

(a) is proved.

For (b), the Leibniz rule gives $D^n(X'Y) = X'D^nY + nX''D^{n-1}Y$ for any Y. Meanwhile, application of D^{n-1} to (a) yields

$$D^{n+1}(X^{n+1}) = (2n+1)D^n(X'X^n) - n(2n+1)X''D^{n-1}(X^n) - 4n^2 D^{n-1}(X^{n-1}).$$

Substituting with $Y = (2n+1)X^n$, we obtain (b). The recursion in conclusion (c) follows immediately by multiplying by $(2^{n+1}n!)^{-1}$.

For (d), conclusion (c) and the definition of P_n show that $Q_n = P_n - R_n$ satisfies $Q_0 = Q_1 = 0$ and $(n+1)Q_{n+1}(x) - (2n+1)xQ_n(x) - nQ_{n-1}(x)$. Thus $Q_n(x) = 0$ for every n by induction.

20–21. Write $X = x^2 - 1$. Since $X^n = (x-1)^n f(x)$, the function X^n has all derivatives through order $n-1$ equal to 0 at $x = 1$. The same conclusion applies also at $x = -1$. If $m \leq n$, integration by parts gives

$$\begin{aligned} \textstyle\int_{-1}^1 D^m(X^m)D^n(X^n)\,dx &= [D^m(X^m)D^{n-1}(X^n)]_{-1}^1 - \textstyle\int_{-1}^1 D^{m+1}(X^m)D^{n-1}(X^n)\,dx \\ &= -\textstyle\int_{-1}^1 D^{m+1}(X^m)D^{n-1}(X^n)\,dx \\ &= \cdots = (-1)^k \textstyle\int_{-1}^1 D^{m+k}(X^m)D^{n-k}(X^n)\,dx \end{aligned}$$

for $k \le n$. If $m < n$, then taking $k = m + 1$ gives 0 on the right side because $D^{2m+1}(X^m) = 0$. If $m = n$, then taking $k = n$ gives $(-1)^n \int_{-1}^{1} X^n D^{2n}(X^n)\,dx = (-1)^n(2n)! \int_{-1}^{1} X^n\,dx$ on the right side. Therefore

$$\langle D^n(X^n),\, D^n(X^n)\rangle = (-1)^n(2n)!(-1)^n \tfrac{2(2^n n!))^2}{(2n+1)!} = \tfrac{2(2^n n!)^2}{2n+1},$$

and $\langle P_n,\, P_n\rangle = \frac{2}{2n+1}$.

22. The expansion for (a) is

$$\begin{aligned} &D^{n+1}[(D(X^n))X] \\ &\quad = D^{n+1}(D(X^n))X + (n+1)D^n(D(X^n))X' + \tfrac{1}{2}n(n+1)D^{n-1}(D(X^n))X'' \\ &\quad = XD^2(D^n(X^n)) + (n+1)X'D(D^n(X^n)) + \tfrac{1}{2}n(n+1)X''D^n(X^n), \end{aligned}$$

and the expansion for (b) is

$$\begin{aligned} D^{n+1}[(D(X^n))X] = D^{n+1}(nX^nX') &= D^{n+1}(nX^n)X' + (n+1)D^n(nX^n)X'' \\ &= nD(D^n(X^n))X' + n(n+1)D^n(X^n)X''. \end{aligned}$$

Thus, for (c), we get $(x^2-1)D^2(P_n(x)) + (n+1)2xD(P_n(x)) + \frac{1}{2}n(n+1)2P_n(x) = nD(P_n(x))2x + n(n+1)P_n(x)2$. This simplifies to

$$(x^2-1)P_n'' + 2(n+1)xP_n' + n(n+1)P_n = 2nxP_n' + 2n(n+1)P_n$$

and then to $(1-x^2)P_n'' - 2xP_n' + n(n+1)P_n = 0$.

24. In Problems 24–28, there is no difficulty with addition, and we have to check something only about scalar multiplication. For Problem 24, we need to check in $\overline{V}$ that $(ab)v = a(bv)$, $1v = v$, $a(u+v) = au + av$, and $(a+b)v = av + bv$. These are satisfied in $\overline{V}$ because the identities $\overline{(ab)}v = \bar{a}(\bar{b}v)$, $1v = v$, $\bar{a}(u+v) = \bar{a}u + \bar{a}v$, and $\overline{(a+b)}v = \bar{a}v + \bar{b}v$ hold in V.

25. We are to see that $\overline{L}$ respects scalar multiplication, and the argument is that $\overline{L}(cv) = L(\bar{c}v) = \bar{c}L(v) = c\overline{L}(v)$.

26. We have $(au, bv)_{\overline{V}} = (\bar{b}v, \bar{a}u)_V = a\bar{b}(v, u)_V = a\bar{b}(u, v)_{\overline{V}}$, as required.

27. Let ℓ in V' correspond to v in V, so that $\ell(u) = (u, v)_V = (v, u)_{\overline{V}}$. Then ℓ in V' corresponds to v in $\overline{V}$, while $(c\ell)(u) = c(v, u)_{\overline{V}} = (cv, u)_{\overline{V}}$ shows that $c\ell$ corresponds to cv in $\overline{V}$.

28. Let ℓ in V' correspond to v in $\overline{V}$. Then $L^t(\ell)(u) = \ell(L(u)) = (v, L(u))_V = (v, \overline{L}(u))_{\overline{V}} = ((\overline{L})^*(v), u)_{\overline{V}}$, and this says that $L^t(\ell)$ corresponds to $(\overline{L})^*(v)$, i.e., L^t corresponds to $(\overline{L})^*$.

29. In (a), it is enough to check the result for p and q equal to monomials, and (b) is a direct calculation. In (c), let $p(x) = \sum c_{k_1,\dots,k_n} x_1^{k_1} \cdots x_n^{k_n}$. The bilinearity and (b) show that $\langle p, p\rangle = \sum (c_{k_1,\dots,k_n})^2 \langle x_1^{k_1} \cdots x_n^{k_n},\ x_1^{k_1} \cdots x_n^{k_n}\rangle$, and this is positive unless all the coefficients are 0.

30. The polynomial p is in H_N if and only if $\partial(|x|^2)p = 0$, if and only if $\langle \partial(|x|^2)p, q\rangle = 0$ for all q in V_{N-2}, if and only if $\partial(q)(\partial(|x|^2)p) = 0$ for all q in V_{N-2}, if and only if $\partial(|x|^2 q)p = 0$ for all q in V_{N-2}, if and only if $\langle p, |x|^2 q\rangle = 0$ for all q in V_{N-2}, if and only if p is in $(|x|^2 V_{N-2})^\perp$.

31. Problem 30 gives $V_N = H_N \oplus |x|^2 V_{N-2}$, and we iterate this decomposition.

32. A basis of $|x|^2 V_2$ is $\{|x|^2 x_1^2, |x|^2 x_1 x_2, |x|^2 x_2^2\}$. Apply the Gram–Schmidt orthogonalization procedure to obtain an orthonormal basis $\{|x|^2 u_1, |x|^2 u_2, |x|^2 u_3\}$, and write $x_1^4 + y_1^4 = h_4 + \sum_{j=1}^3 (x_1^4 + y_1^4, |x|^2 u_j)|x|^2 u_j$. Then h_4 is harmonic by Problem 30. A basis of $|x|^2 V_0$ is $|x|^2$, and hence an orthonormal basis consists of the single vector $w = \||x|^2\|^{-1}|x|^2$. Write $u_j = h_{2,j} + (u_j, w)w$ for each j, and substitute. Each $h_{2,j}$ is harmonic. Then we have

$$\begin{aligned} x_1^4 + y_1^4 &= h_4 + \textstyle\sum_{j=1}^3 (x_1^4 + y_1^4, |x|^2 u_j)|x|^2\big(h_{2,j} + (u_j, w)w\big) \\ &= h_4 + |x|^2 \textstyle\sum_{j=1}^3 (x_1^4 + y_1^4, |x|^2 u_j) h_{2,j} \\ &\quad + |x|^4 \textstyle\sum_{j=1}^3 (x_1^4 + y_1^4, |x|^2 u_j)(u_j, w)\||x|^2\|^{-1} \end{aligned}$$

with h_4 in H_4, each $h_{2,j}$ in H_2, and the last sum in H_0.

33. Let P be the positive semidefinite square root of B. Then $AB = APP$, and hence $\det(\lambda I - AB) = \det(\lambda I - PAP)$. Consequently AB has the same eigenvalues as PAP. The latter is positive semidefinite since $(PAPv, v) = (A(Pv), Pv) \geq 0$. Therefore all the eigenvalues of AB are ≥ 0.

34. Since $(P^{-1}ABCP^{-1}v, v) = (ABC(P^{-1}v), P^{-1}v)$, ABC is positive semidefinite if and only if $P^{-1}ABCP^{-1}$ is positive semidefinite, if and only if $P^{-1}ABCP^{-1}$ has all eigenvalues ≥ 0. But $P^{-1}ABCP^{-1}$ has the same eigenvalues as $ABCP^{-1}P^{-1} = AB$, which has all eigenvalues ≥ 0 by the previous problem.

Chapter IV

1. If $a^2 = b^2 = (ab)^2 = 1$, then $a^{-1} = a$, $b^{-1} = b$, and $(ab)^{-1} = ab$. So $ab = (ab)^{-1} = b^{-1}a^{-1} = ba$.

2. Number the vertices counterclockwise as 1, 2, 3, 4. The motions in D_4 are then given by permutations as 1, (1 2)(3 4), (1 4)(2 3), (1 3), (2 4), (1 2 3 4), (1 3)(2 4), (1 4 3 2).

2A. In (c), the result follows from (a) and (b) if $r \neq 0$. If $r = 0$, both sides are 1.

3. Choose integers x and y with $xl + y|G| = 1$. Then $a = a^{xl+y|G|} = (a^l)^x(a^{|G|})^y = (a^l)^x$ since $a^{|G|} = 1$, and this is a power of an element of H.

4. Define $\varphi : G \to G'$ by $\varphi(a) = a$. Then $\varphi(a) \circ \varphi(b) = a \circ b = ba = \varphi(ba) = \varphi(a \circ b)$. From this equality it follows that G' is a group and that φ is an isomorphism.

5. For $n > 0$, $(ab)^n = abab\cdots ab = a^n b^n$; also $(ab)^{-n} = ((ab)^{-1})^n = (b^{-1}a^{-1})^n = (a^{-1}b^{-1})^n = (a^{-1})^n(b^{-1})^n = a^{-n}b^{-n}$. In $\mathfrak{S}_3$, $a \mapsto a^2$ is not a homomorphism since four elements are sent to 1 and since 4 does not divide $|\mathfrak{S}_3| = 6$.

6. Define $\varphi : H \times K \to HK$ by $\varphi(h, k) = hk$. What needs proof is that members of H commute with members of K. If h is in H and k is in K, then $(hkh^{-1})h = hk = k(k^{-1}hk)$. Since H and K are normal, hkh^{-1} is in K and $k^{-1}hk$ is in H. Then $k^{-1}(hkh^{-1}) = (k^{-1}hk)h^{-1}$ and $H \cap K = \{1\}$ together imply $k^{-1}(hkh^{-1}) = 1 = (k^{-1}hk)h^{-1} = 1$. From the first of these, $k = hkh^{-1}$. Therefore $hk = kh$.

7. Since GCD$(1234, 8191) = 1$, there exist x and y with $1234x + 8191y = 1$, and x and y can be found explicitly by the Euclidean algorithm of Section I.1. For this x, $1234x \equiv 1 \bmod 8191$.

8. The members $1, 2, \ldots, p-1$ of $\mathbb{F}_p$ are roots of $X^{p-1} - 1 = 0$. By iterated use of the Factor Theorem, $X^{p-1} - 1 = (X-1)(X-2)\cdots(X-(p-1))Q(X)$, and $Q(X)$ must have degree 0. Checking the coefficient of X^{p-1} on both sides shows that $Q(X) = 1$. Evaluating at $X = 0$ gives $-1 = (-1)(-2)\cdots(-(p-1)) \bmod p$. Since p is odd, this equation reads $(p-1)! = -1 \bmod p$.

9. Corollary 4.39 shows that such a group has to be abelian, and Theorem 4.56 shows that it is the direct sum of cyclic groups. Thus it must be C_{p^2} or $C_p \times C_p$, up to isomorphism.

10. If $y = axa^{-1}$, then $y^n = ax^n a^{-1}$. This proves (a). Also, $ba = a^{-1}(ab)a$ shows that ba and ab are conjugate. This proves (b).

11. There are four classes: $C_1 = \{1\}, C_2 = \{(1\ 2)(3\ 4),\ (1\ 3)(2\ 4),\ (1\ 4)(2\ 3)\}$, $C_3 = \{(1\ 2\ 3),\ (3\ 4\ 1),\ (2\ 1\ 4),\ (4\ 3\ 2)\}, C_4 = \{(1\ 3\ 2),\ (3\ 1\ 4),\ (2\ 4\ 1),\ (4\ 2\ 3)\}$. The centralizer of the first element of each class is $\mathfrak{A}_4$ for C_1, $C_1 \cup C_2$ for C_2, $\{(1\ 2\ 3),\ (1\ 3\ 2)\}$ for C_3 and C_4. Since $\mathfrak{A}_4$ has no element of order 6, it has no subgroup C_6. In a subgroup $\mathfrak{S}_3$, an element of order 3 is conjugate to its square, but no element of order 3 in $\mathfrak{A}_4$ is conjugate to its square.

12. A subgroup of order 30 would have index 2 and would thus be normal, in contradiction to Theorem 4.47.

13. This is a special case of Proposition 4.36.

14. Since H is normal, G acts on H by conjugation. The number of elements in an orbit has to be a divisor of $|G|$, and the smallest divisor of $|G|$ apart from 1 is p, by hypothesis. Since $\{1\}$ is one orbit and there are only $p-1$ other elements in H, each orbit must contain one element. Therefore $ghg^{-1} = h$ for each $g \in G$ and $h \in H$, and each h is in Z_G.

15. Certainly the inner automorphisms are closed under composition and inversion and therefore form a subgroup. If φ is an automorphism and ψ is the inner automorphism $\psi(x) = axa^{-1}$, then $\varphi\circ\psi\circ\varphi^{-1}(x) = \varphi(a\varphi^{-1}(x)a^{-1}) = \varphi(a)x\varphi(a)^{-1}$ shows that $\varphi\circ\psi\circ\varphi^{-1}$ is inner. Hence the subgroup of inner automorphisms is normal. Define a mapping Φ of G into the inner automorphisms by $\Phi(a) = \{x \mapsto axa^{-1}\}$. Then $\Phi(ab) = \Phi(a)\Phi(b)$, and hence Φ is a homomorphism. Certainly Φ is onto the inner automorphisms, and its kernel consists of all elements $a \in G$ with $axa^{-1} = x$ for all x, hence consists of all a in Z_G. Thus Φ exhibits G/Z_G as isomorphic to the group of inner automorphisms.

16. Part (a) is proved in the same way as Lemma 4.45. For (b), choose $m = 8$; then $\mathrm{Aut}\, C_m$ is $C_2 \times C_2$.

17. In (a), each C_k is a conjugacy class, by Proposition 4.42, and it is evident that the C_k's are the only conjugacy classes whose members have order 2. If x and y are in $\mathfrak{S}_n$, then $\tau(xyx^{-1}) = \tau(x)\tau(y)\tau(x)^{-1}$ shows that τ carries any conjugate of y to a conjugate of $\tau(y)$. Therefore conjugacy classes map to conjugacy classes under τ, and $\tau(C_1)$ has to be some C_k.

In (b), the number of ways of selecting $2k$ elements from n is $\binom{n}{2k}$. For each of these, the number of ways of selecting k unordered pairs of elements from $2k$ elements is the multinomial coefficient $\binom{2k}{2,\dots,2} = \frac{(2k)!}{2^k}$. Although the individual pairs are unordered, this enumeration counts one for each different ordering of the k pairs. There are $k!$ orderings, and hence the multinomial coefficient must be divided by $k!$ to discount the enumeration of the pairs. Thus $|C_k|$ is the product of the integer $\binom{n}{2k}$ and the integer $\frac{(2k)!}{2^k k!}$.

In (c), we saw in (b) that $N_k = \frac{(2k)!}{2^k k!}$ is always an integer. Let us bound it below. Canceling every even factor of the numerator by a factor of $k!$ and a factor of 2^k, we see that $N_k = (2k-1)(2k-3)(2k-5)\cdots(3)(1)$. Thus $N_k \geq 2k-1$ with equality only if $2k-1 = 1$, in which case $k = 1$. Also, $N_k \geq (2k-1)(2k-3)$ with equality holding for a value of $k > 1$ only if $2k-3 = 1$, in which case $k = 2$.

Now let us compare $|C_k|$ and $|C_1|$. We have $N_1 = 1$. Also, $|C_k| = \binom{n}{2k}\frac{(2k)!}{2^k k!} = N_k\binom{n}{2k}$ and $|C_1| = N_1\binom{n}{2} = \binom{n}{2}$. The easy comparison is that $|C_k| \geq \binom{n}{2k}$ and this is $> \binom{n}{2} = |C_1|$ unless $k = 1$ or $|n-2k| \leq 2$. Thus $|C_k| > |C_1|$ unless k equals 1 or $\frac{1}{2}n$ or $\frac{1}{2}(n-1)$ or $\frac{1}{2}(n-2)$. We can discard $k = \frac{1}{2}(n-2)$ because in this case $|C_k| = N_k\binom{n}{2} > N_1\binom{n}{2} = |C_1|$ except when $k = 1$.

Consider $k = \frac{1}{2}(n-1)$ with $k > 1$. Then $|C_1| = \frac{1}{2}n(n-1) = nk$ and $|C_k| = N_k\binom{n}{n-1} = nN_k$. From above, the latter is $> n(2k-1) \geq nk = |C_1|$.

Finally consider $k = \frac{1}{2}n$ with $k > 1$. Then $|C_1| = \frac{1}{2}n(n-1) = (n-1)k$ and $|C_k| = N_k\binom{n}{n} = N_k$. From above, the latter for $k > 1$ is $\geq (2k-1)(2k-3) = (n-1)(n-3)$, and this is $> (n-1)k = |C_1|$ unless $k \geq n-3$. When $k \geq n-3$, we obtain $\frac{1}{2}n \geq n-3$ and $n \leq 6$. Since $k = \frac{1}{2}n$, n has to be even with $n \leq 6$. The case $n = 6$ (with $k = 3$) we are allowing, and the case $n = 4$ with $k = 2$ has $|C_2| = 3 \neq 6 = |C_1|$. Thus the only exceptions have $k = 1$ or $n = 6$.

18. In the composition series given for $\mathfrak{S}_4$ in Section 8, take G to be $\mathfrak{A}_4$, N to be the 4-element subgroup in the series, and M to be the 2-element subgroup. For another example, take G to be the dihedral group D_4, N to be the cyclic subgroup of the 4 rotations, and M to be the 2-element subgroup of N.

19. If $\mathrm{GCD}(r, s) = 1$, define a homomorphism $\varphi : \mathbb{Z} \to (\mathbb{Z}/r\mathbb{Z}) \times (\mathbb{Z}/s\mathbb{Z})$ by $\varphi(n) = (n \bmod r, n \bmod s)$. This is 0 for $n = rs$. Thus it descends to a homomorphism $\overline{\varphi} : \mathbb{Z}/rs\mathbb{Z} \to (\mathbb{Z}/r\mathbb{Z}) \times (\mathbb{Z}/s\mathbb{Z})$. The kernel of φ consists of all integers n divisible by r and s. Since r and s are relatively prime, such integers are divisible by rs. Thus $\ker \varphi = rs\mathbb{Z}$, and $\overline{\varphi}$ is one-one. Since the domain and range have the same number of elements, φ is onto.

Conversely if $\mathrm{GCD}(r, s) \neq 1$, then some prime p divides both r and s. The number of elements in C_{rs} of order p is then $p - 1$, while the number of elements in $C_r \times C_s$ of order p is $p(p-1) + (p-1) = p^2 - 1$. So C_{rs} cannot be isomorphic to $C_r \times C_s$.

20. Three, namely C_{27}, $C_9 \times C_3$, and $C_3 \times C_3 \times C_3$.

21. The matrix relating the bases is $C = \begin{pmatrix} 3 & 2 & 5 \\ 0 & 1 & 3 \\ 0 & 1 & 5 \end{pmatrix}$. A row interchange and a column interchange move the entry 1 in the center to the upper left and give $\begin{pmatrix} 1 & 0 & 3 \\ 2 & 3 & 5 \\ 1 & 0 & 5 \end{pmatrix}$. Two row operations and one column operation eliminate the other entries in the first column and first row, yielding $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & -1 \\ 0 & 0 & 2 \end{pmatrix}$. The remaining steps pass from there to

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 3 \\ 0 & 2 & 0 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -3 \\ 0 & 2 & 0 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -3 \\ 0 & 0 & 6 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 6 \end{pmatrix}.$$

Hence $H = \mathbb{Z} \oplus \mathbb{Z} \oplus 6\mathbb{Z}$, and $G/H \cong C_6$.

22. Let the four generators for G be x_1, x_2, x_3, x_4, and let the four generators for H be y_1, y_2, y_3, y_4. Since each is linearly independent over $\mathbb{Q}$, it is linearly independent over $\mathbb{Z}$. The matrix of the y_i's in terms of the x_j's is $C = \begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & -1 \\ 0 & -1 & 0 & 2 \\ 0 & -1 & 2 & 0 \end{pmatrix}$. The reduction procedure on this leads to $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$. Hence $G/H \cong C_2 \times C_2$.

23. Each step of row reduction or column reduction preserves the rank of the matrix as a member of $M_{nn}(\mathbb{Q})$ since row rank equals column rank. Following through the steps of the procedure, we may assume that the matrix is diagonal with diagonal entries $D_{11}, \dots, D_{nn}$ with $D_{jj} \neq 0$ exactly for $1 \leq j \leq r$. Then $H = \bigoplus_{j=1}^{r} D_{jj}\mathbb{Z}$, and we can read off that H has rank r and the $\mathbb{Q}$ rank of the matrix is r.

24. Let G be an abelian group, and let $\widetilde{G} = \bigoplus_{g \in G} \mathbb{Z}$. For each g, form the homomorphism $\varphi_g : \mathbb{Z} \to G$ given in additive notation by $\varphi_g(n) = ng$. Then the universal mapping property of direct sums gives the desired homomorphism of the free abelian group $\widetilde{G}$ onto G.

25. For (a), right translation by any element of $H \cap K$ sends xH to itself and yK to itself, hence sends $xH \cap yK$ to itself. Therefore $xH \cap yK$ is a union of left cosets of $H \cap K$. We are to see that at most one left coset is involved. Thus suppose we have two elements g_1 and g_2 in $xH \cap yK$. Write $g_1 = xh_1 = yk_1$ and $g_2 = xh_2 = yk_2$. Then $g_2^{-1}g_1 = h_2^{-1}h_1 = k_2^{-1}k_1$, and $g_2^{-1}g_1$ is exhibited as in $H \cap K$. So g_1 is in $g_2(H \cap K)$.

For (b), if the sets $x_1H, \dots, x_mH$ exhaust G and the sets $y_1K, \dots, y_nK$ exhaust G, then G is the union of the mn sets $x_iH \cap y_jK$. By (a), G is exhibited as the union of $\le mn$ left cosets of $H \cap K$.

26. Returning to Problem 23, we see that $H = \bigoplus_{j=1}^n D_{jj}\mathbb{Z}$ with each $D_{jj} \neq 0$. Then the index of H in G is $\prod_{j=1}^n D_{jj}$.

27. In (a), take $H_2 = \{(1),\ (1\ 2)(3\ 4),\ (1\ 3)(2\ 4),\ (1\ 4)(2\ 3),\ (1\ 3),\ (2\ 4),\ (1\ 2\ 3\ 4),\ (1\ 4\ 3\ 2)\}$. The number of such subgroups is $2k + 1$ and divides 3. Since H_2 is not normal, the number is > 1. Therefore it is 3.

In (b), take $H_3 = \{(1),\ (1\ 2\ 3),\ (1\ 3\ 2)\}$. The number of such subgroups is $3k + 1$ and divides 8. Since H_3 is not normal, the number is 4.

28. Disproof: In $\mathfrak{S}_3$, take $H = \{(1),\ (1\ 2)\}$. Then $N(H) = H$, and this is not normal.

29. Since $168/7 = 24$, the number of Sylow 7-subgroups is $7k+1$ and divides 24. The group G is assumed simple, and so $k \neq 0$. Then k must be 1, and there are 8 distinct Sylow 7-subgroups. Any two of these intersect only in the identity, and each contains 6 elements of order 7. Hence there are 48 elements of order 7.

30. The number of Sylow q-subgroups is $qk + 1$ and divides p, hence must be 1. So S_q is normal, and the set S_pS_q of products is a subgroup. An argument in the proof of Proposition 4.60 shows that each element of G is uniquely a product of a member of S_p and a member of S_q, and hence G is a semidirect product.

31. Let Γ be the set of subgroups conjugate to H, and form the action $G \times \Gamma \to \Gamma$ by conjugation. The isotropy subgroup at H is $N(H)$, which must have index 1 or index p in G. If it has index 1, then H is normal, and $|\Gamma| = 1$. Otherwise it has index p. Then $N(H) = H$, the orbit of H has $|G|/|H| = p$ elements, and $|\Gamma| = p$.

32. In (a), the subgroup H is a Sylow 2-subgroup, and the number of its conjugates must then be $2k+1$ and divide $24/8 = 3$. Since H is assumed not normal, the number of conjugates has to be 3.

In (b), call the conjugates H, H', and H''. Each member g of G acts on the set $\{H, H', H''\}$ by conjugation of the subgroups, sending H to gHg^{-1}, H' to $gH'g^{-1}$, and H'' to $gH''g^{-1}$. The result is that we obtain a function Φ from G to the permutation group $\mathfrak{S}_3$ on $\{H, H', H''\}$. This function Φ is a group homomorphism.

In (c), the subgroup $\ker \Phi$ is normal, and it is enough to show that this subgroup is neither $\{1\}$ nor G. The image of Φ is not the identity subgroup since some member g of G has $gHg^{-1} = H'$; thus $\ker \Phi \neq G$. Since $24/|\ker \Phi| = |G|/|\ker \Phi| = |\text{image } \Phi| \le 6$, we have $6|\ker \Phi| \ge 24$ and $|\ker \Phi| \ge 4$; thus $\ker \Phi \neq \{1\}$.

33. Let H be a Sylow 3-subgroup, of order 9. If H is normal, then G/H is a subgroup of order 4, necessarily either $C_2 \times C_2$ or C_4. Both of these groups of order 4 are isomorphic to subgroups of $\mathfrak{S}_4$, and thus there is a nontrivial homomorphism of G onto a subgroup of order 4 in $\mathfrak{S}_4$.

If H is not normal, then the number of conjugates of H is $3k + 1$ and divides 4. Then the number of conjugates must be 4. Arguing as in the previous problem we obtain a homomorphism of G into $\mathfrak{S}_4$ by having each element of g map to the corresponding permutation of the conjugates of H. This homomorphism is nontrivial since H can be moved to any of its conjugates by some element of G and since the number of such conjugates is > 1.

34. Let K be a Sylow q-subgroup. The number of conjugates of K is of the form $kq + 1$ and divides $2p$. If $k = 0$, then K is normal. This conclusion disposes of (a) and the first statement of (b) for this case. We come back to the remainder of (b) for this case in a moment.

If $k > 1$, then $kq + 1 \leq 2p$ is impossible since $p < q$. Thus the only other possibility besides $k = 0$ is $k = 1$. Then $q + 1$ divides $2p$. So $q + 1$ equals 1, 2, p, or $2p$. Since $q > p$, the only possibility is $q + 1 = 2p$. This completes the argument for (a).

For the rest we may assume that $q + 1 = 2p$. If either of H or K is normal, then an argument in the proof of Proposition 4.60 shows that HK is a subgroup with pq elements. Since $2p = q + 1$, p divides $q + 1$. If p also divides $q - 1$, then p divides the difference, which is 2, and we obtain a contradiction. So p does not divide $q - 1$, and Proposition 4.60 shows that HK is abelian, hence cyclic.

Thus we are reduced to the situation that $q+1 = 2p$ and K is not normal; we are to prove that H is normal. We have seen in this case that the number of conjugates of K is $q+1$, and hence the number of elements of order q is $(q+1)(q-1) = 2p(q-1) = 2pq-2p$. The number of conjugates of H is of the form $lp+1$ and divides $2q$. If $l = 0$, then H is normal, and we are done. If $l \geq 1$, then the number of elements of order p is $(lp+1)(p-1) \geq (p+1)(p-1) = p^2-1$. Thus the total number of elements of order 1, p, or q is $\geq 1+(p^2-1)+(2pq-2p) = 2pq+(p-1)^2-1 \geq 2pq+2^2-1 > 2pq$, and we have obtained a contradiction.

35. Certainly ψ is one-one and onto. For (h, k) and (h', k') in $H \times_{\varphi_2} K$, we have

$$\psi((h,k)(h',k')) = \psi(hh', ((\varphi_2)_{h'^{-1}}(k))k') = (\varphi(hh'), ((\varphi_2)_{h'^{-1}}(k))k')$$

and

$$\psi(h,k)\psi(h',k') = (\varphi(h),k)(\varphi(h'),k') = (\varphi(hh'), ((\varphi_1)_{\varphi(h')^{-1}}(k))k').$$

The right sides are equal because $(\varphi_2)_{h'^{-1}} = (\varphi_1 \circ \varphi)_{h'^{-1}} = (\varphi_1)_{\varphi(h'^{-1})} = (\varphi_1)_{\varphi(h')^{-1}}$.

36. Again ψ is visibly one-one and onto. The formula for φ_2 in terms of φ_1 is given more concretely as $(\varphi_2)_h(k) = a\big((\varphi_1)_h(a^{-1}(k))\big)$. For (h, k) and (h', k') in

$H \times_{\varphi_1} K$, we then have

$$\psi((h,k)(h',k')) = \psi(hh', ((\varphi_1)_{h'^{-1}}(k))k') \\ = \big(hh', a\big(((\varphi_1)_{h'^{-1}}(k))k'\big)\big) = \big(hh', a\big((\varphi_1)_{h'^{-1}}(k)\big)a(k')\big)$$

and

$$\psi(h,k)\psi(h',k') = (h,a(k))(h',a(k')) = \big(hh', \big((\varphi_2)_{h'^{-1}}(a(k))\big)a(k')\big) \\ = \big(hh', \big(a\big((\varphi_1)_{h'^{-1}}(a^{-1}(a(k)))\big)\big)a(k')\big).$$

The right sides are equal because $a^{-1}(a(k)) = k$.

37. An action of C_p on C_q is a homomorphism of C_p into $\operatorname{Aut} C_q \cong C_{q-1}$. If a is a generator of C_p and b is a generator of C_{q-1}, we may assume that $a \mapsto b^k$ for some k. Since the action is nontrivial, $0 < k < q-1$. Then $1 = a^p$ maps to b^{kp}, and therefore b^{kp} must be 1. This means that kp must be a multiple of $q-1$. So $kp = r(q-1)$. Since $0 < k < q-1$, we see that $p > r$. Therefore p does not divide r and must divide $q-1$.

38. Put $n = (q-1)/p$. Let a be a generator of C_p, and let b be a generator of $\operatorname{Aut} C_q \cong C_{q-1}$. For reference, take $\tau(a) = b^n$. This defines a nontrivial homomorphism of C_p into C_{q-1}. Any other one is of the form $\tau_1(a) = b^{k_1}$ with $0 \le k_1 < q-1$. As in the previous problem, we know that $k_1 p = r(q-1)$. Hence $k_1 = nr$ for some r with $1 \le r \le p-1$. The mapping $\varphi(a^s) = a^{rs}$ is then an automorphism of C_p, and $\tau_1(a) = b^{k_1} = b^{nr} = \tau(a^r) = (\tau \circ \varphi)(a)$. So $\tau_1 = \tau \circ \varphi$. Problem 35 applies and yields the desired isomorphism.

40. For (a), $D_4 \supseteq C_4 \supseteq C_2 \supseteq \{1\}$, where C_4 is the subgroup of rotations. For (b), $H_8 \supseteq C_4 \supseteq C_2 \supseteq \{1\}$, where C_4 is the subgroup $\{\pm 1, \pm \mathbf{i}\}$.

41. For (a), the trivial subgroup, the whole group, and all subgroups of index 2 are automatically normal. The only other possibility is order 2. Since -1 is the only element of order 2, the only subgroup of order 2 is $\{\pm 1\}$. This is the center of H_8 and hence is normal.

For (b), the five conjugacy classes are $\{\pm\mathbf{i}\}$, $\{\pm\mathbf{j}\}$, $\{\pm\mathbf{k}\}$, $\{-1\}$, and $\{1\}$.

For (c), Problem 15 shows that the inner automorphisms form a normal subgroup isomorphic to the quotient of H_8 by its center. The center is $\{\pm 1\}$, and thus the inner automorphisms form a subgroup of the group of all automorphisms isomorphic to $C_2 \times C_2$. The nontrivial inner automorphisms multiply two of $\mathbf{i}, \mathbf{j}, \mathbf{k}$ by -1 and fix the third one. In addition, the cyclic map $\mathbf{i} \mapsto \mathbf{j} \mapsto \mathbf{k} \mapsto \mathbf{i}$ is an automorphism and gives an automorphism of order 3. So is its square. One more automorphism fixes $\mathbf{i}$ and has $\mathbf{j} \mapsto \mathbf{k} \mapsto -\mathbf{j} \mapsto -\mathbf{k}$. Consequently the group of automorphisms G acts transitively on the set of six elements of order 4, and $|G| = 6|H|$, where H is the subgroup fixing $\mathbf{i}$. With $\mathbf{i}$ fixed, an automorphism can carry $\mathbf{j}$ to any of $\pm\mathbf{j}$ and $\pm\mathbf{k}$. Thus $|H| = 4|K|$, where K is the subgroup fixing $\mathbf{i}$ and $\mathbf{j}$. Since $\mathbf{i}$ and $\mathbf{j}$ generate H_8, K is trivial. Hence $|\operatorname{Aut} H_8| = 24$.

42. The only possible orders are the divisors of 8. If it were to have an element of order 8, it would be cyclic, hence abelian. If all elements other than the identity were to have order 2, it would be abelian by Problem 1. Hence it must have an element of order 4.

43. Let C_2 be the subgroup generated by the element of order 2. Proposition 4.44 shows that G is a semidirect product $C_2 \times_\tau K$, and τ has to be nontrivial for G to be nonabelian. By Problem 16a, there is only one possibility for τ. Since D_4 is one such semidirect product, G must be isomorphic to D_4.

44. Let the elements of K be the powers of $\mathbf{i}$. By assumption every element outside K has order 4. Thus $\mathbf{i}^2$ is the only element of order 2. Its conjugacy class therefore contains no other element, and it is central. Let us write -1 for this element. No element other than ± 1 can be central since if the center has order 4, then it commutes with any other element and together they generate an abelian G. So $Z_G = \{\pm 1\}$. Next let $\mathbf{j}$ be an element of order 4 not in K. Define $\mathbf{k} = \mathbf{ij}$. We know that $\mathbf{j}^2 = \mathbf{k}^2 = -1$, and thus the 8 elements are ± 1, $\pm\mathbf{i}$, $\pm\mathbf{j}$, $\pm\mathbf{k}$. From $\mathbf{k} = \mathbf{ij}$, we obtain $\mathbf{kj} = (\mathbf{i})(-1) = -\mathbf{i}$ and similarly $\mathbf{ik} = -\mathbf{j}$. Finally we know that $\mathbf{i}$ and $\mathbf{j}$ do not commute (since G would otherwise be abelian) and that neither $\mathbf{ij}$ nor $\mathbf{ji}$ is a power of $\mathbf{i}$ or $\mathbf{j}$. Thus $\mathbf{ji}$ has to be $\pm\mathbf{k}$ and cannot be $\mathbf{k}$. So $\mathbf{ji} = -\mathbf{k}$, and we then obtain $\mathbf{jk} = \mathbf{i}$ and $\mathbf{ki} = \mathbf{j}$. Thus the multiplication table in G matches that in H_8, and we have an isomorphism.

46. Suppose $K \cong C_4$. If H acts nontrivially on K, then there is a nontrivial homomorphism of $H \cong C_3$ into $\operatorname{Aut} K \cong \operatorname{Aut} C_4 \cong C_2$. Since C_2 has no element of order 3, this is impossible.

If $K \cong C_2 \times C_2$, then $\operatorname{Aut} K \cong \mathfrak{S}_3$, the automorphisms being the permutations of the set $\{(1, 0), (0, 1), (1, 1)\}$. Thus there are two nontrivial homomorphisms of C_3 into $\operatorname{Aut} K$. Since the elements of order 3 in $\mathfrak{S}_3$ are conjugate in $\mathfrak{S}_3$, Problem 36 applies and shows that the two resulting semidirect products are isomorphic. The group $\mathfrak{A}_4$ meets the conditions of this problem, and hence the given G must be isomorphic to $\mathfrak{A}_4$.

47. Certainly one of those conditions holds, and G is abelian if (i) holds. If (ii) holds, then τ has order 2, and τ is determined by its kernel. Let us rewrite the group K as $C_2 \times C_2$ with the second factor as the kernel of τ, so that τ factors through to a homomorphism of the first factor. Then $(C_2 \times C_2) \times_\tau C_3 \cong C_2 \times_{\overline{\tau}} (C_2 \times C_3) \cong C_2 \times_{\overline{\tau}} C_6 \cong D_6$. If (iii) holds, we have a nonnormal subgroup of order 4 in G, and this does not happen in $\mathfrak{A}_4$ or D_6.

48. If (iii) holds, the homomorphism $C_4 \to \operatorname{Aut} C_3$ has to be nontrivial and is then uniquely determined since $\operatorname{Aut} C_3 \cong C_2$. This proves the uniqueness of the group up to isomorphism. The group has 1 element of order 1, 3 elements of order 2, 2 elements of order 3, and 6 elements of order 4.

49. Let H be a Sylow q-subgroup, and let K be a Sylow p-subgroup. The number of conjugates of H is of the form $qk + 1$ and divides p^2. Since p is prime, $qk + 1$

must be 1, p, or p^2. If H is not normal, then $k > 0$ and we cannot have $qk + 1 = p$ since $p < q$; therefore $qk + 1 = p^2$. In this case the number of elements of order q is $(qk+1)(q-1) = p^2(q-1) = p^2q - p^2$, and a Sylow p-subgroup then accounts for all the remaining elements. Consequently H not normal implies K normal.

Now let us analyze what k must be when $qk = p^2 - 1$. Since q is prime, q divides $p+1$ or q divides $p-1$. But the condition q divides $p-1$ is impossible since $p < q$, and thus q divides $p+1$. Since $2q > p+q > p+1$, we must in fact have $q = p+1$. Since all primes but 2 are odd, this says that $p = 2$ and $q = 3$. We conclude that either $p^2q = 12$ or else the condition $qk = p^2 - 1$ is impossible; when $qk = p^2 - 1$ fails, we have seen that H is normal.

50. We form three distinct semidirect products, two with Sylow p-subgroup C_{p^2} and one with Sylow p-subgroup $C_p \times C_p$. For each a Sylow q-subgroup C_q is to be normal. We know from Problem 16a and Corollary 4.27 that the group of automorphisms of the cyclic group C_q is isomorphic with C_{q-1}. We obtain one homomorphism $C_{p^2} \to C_{q-1}$ by mapping a generator of C_{p^2} to an element in C_{q-1} of order p^2 and a second homomorphism by mapping a generator of C_{p^2} to an element in C_{q-1} of order p. The third semidirect product comes by having the first factor C_p of $C_p \times C_p$ act trivially on C_q and having the second factor act with a generator of C_p mapping to an element of order p in C_{q-1}.

51. The second and third groups constructed in the previous problem make sense when p divides $q - 1$.

52. If p does not divide $q-1$, then $p^2q \neq 12$. Problem 49 then shows that a Sylow q-subgroup is normal. Hence the group has to be a semidirect product. The action of a Sylow p-subgroup on C_q corresponds to a homomorphism of C_{p^2} or $C_p \times C_p$ into C_{q-1}, and the condition that p not divide $q - 1$ means that C_{p^2} or $C_p \times C_p$ must map to the identity. Therefore the group is abelian.

53. In (a) and (b), the automorphism group of $\mathbb{Z}/9\mathbb{Z}$ is given by multiplication by the members of $(\mathbb{Z}/9\mathbb{Z})^\times = \{1, 2, 4, 5, 7, 8\}$. The element 4 has square 7 and cube 1 modulo 9, and hence the multiplications by 1, 4, 7 yield a group of automorphisms of order 3 of C_9. Hence C_3 has a nontrivial action by automorphisms on C_9, and there exists a nonabelian semidirect product of C_3 and C_9 with C_9 normal.

In (c), let a be a generator of C_9, let b be a generator of C_3, and let τ_b be the automorphism $a \mapsto a^7$. Then $\tau_{b^{-1}}$ is the automorphism $a^n \mapsto a^{4n}$, and $\tau_{b^{-p}}(a^n) = a^{4^p n}$. Proposition 4.43 says that $(b^m, a^n)(b^p, a^q) = (b^{m+p}, (\tau_{b^{-p}}(a^n))a^q)$, and the right side equals $(b^{m+p}, a^{4^p n+q})$. Taking $m = -1$, $n = 1$, $p = 1$, and $q = 0$, we obtain $(b^{-1}, a)(b, 1) = (1, a^4)$. Abbreviating $(1, a)$ as a and $(b, 1)$ as b, we obtain $a^9 = b^3 = b^{-1}aba^{-4} = 1$.

54. In such a group the subgroup H is normal by Proposition 4.36, and thus the group of order 27 is a semidirect product of C_3 and C_9 with C_9 normal. A nonabelian such semidirect product must have a generator of C_3 mapping into an automorphism of order 3 of C_9. There are two possibilities, and Problem 35 shows that they lead to isomorphic semidirect products.

55. $|\mathrm{GL}(2,\mathbb{F})| = (q^2-1)(q^2-q)$ and $|\mathrm{SL}(2,\mathbb{F})| = (q-1)^{-1}|\mathrm{GL}(2,\mathbb{F})|$ because $|\mathrm{GL}(2,\mathbb{F})| = |\ker\det|\,|\,\mathrm{image}\det|$. This handles (a) and (b). For (c), the scalar matrices of determinant 1 are those for which the scalar has square 1. Since the characteristic is not 2, both ±1 qualify. Since $\mathbb{F}$ is a field, the polynomial X^2-1 can have only two roots. So we factor by a group of order 2, and the number of elements is cut in half. For (d), the order in general is $\frac{(q^2-1)(q^2-q)}{2(q-1)} = \frac{1}{2}(q-1)q(q+1)$. Then $|\mathrm{PSL}(2,\mathbb{F}_7)| = 168$.

56. Regard G as a group of invertible linear mappings that is to be written in the standard basis Σ. Let $\Gamma = (u, v)$. If $A = \begin{pmatrix} M \\ \Gamma\Gamma \end{pmatrix}$, then $A = \begin{pmatrix} 0 & -1 \\ 1 & c \end{pmatrix}$, the upper right entry being -1 because $\det M = 1$. Then $\begin{pmatrix} M \\ \Sigma\Sigma \end{pmatrix} = \begin{pmatrix} M \\ \Sigma\Gamma \end{pmatrix} A \begin{pmatrix} M \\ \Sigma\Gamma \end{pmatrix}^{-1}$. Products AB go into products of such expressions, and conjugates hAh^{-1} by matrices of determinant 1 go into expressions

$$\left(\begin{pmatrix} M \\ \Sigma\Gamma \end{pmatrix} h \begin{pmatrix} M \\ \Sigma\Gamma \end{pmatrix}^{-1}\right)\left(\begin{pmatrix} M \\ \Sigma\Gamma \end{pmatrix} A \begin{pmatrix} M \\ \Sigma\Gamma \end{pmatrix}^{-1}\right)\left(\begin{pmatrix} M \\ \Sigma\Gamma \end{pmatrix} h \begin{pmatrix} M \\ \Sigma\Gamma \end{pmatrix}^{-1}\right)^{-1}$$

that are conjugates of such expressions. Thus if A and such expressions generate $\mathrm{SL}(2,\mathbb{F})$, then the conjugates generate the conjugates, again giving $\mathrm{SL}(2,\mathbb{F})$.

57. In (a), $B^{-1}A^{-1}BA$ is the product of the conjugate $B^{-1}A^{-1}B$ of the inverse of A by A itself and hence is in G. Direct computation shows that the matrix in question is $\begin{pmatrix} a^{-2} & c(a^{-2}-1) \\ 0 & a^2 \end{pmatrix}$. In (b), the diagonal entries are equal if and only if $a^{-2} = a^2$, hence if and only if $a^4 = 1$. In (c), the result of (b) shows that there are at most 4 choices of a to avoid. We must also avoid $a = 0$. Thus if the field has more than 5 elements, a can be chosen nonzero so that $a^4 \neq 1$.

58. As in Problem 57a, the conditions that C is in G and $\det D = 1$ imply that $CDC^{-1}D^{-1}$ is in G. The product in question is $\begin{pmatrix} 1 & x^2-1 \\ 0 & 1 \end{pmatrix}$. Since $x \neq \pm1$, $\lambda = x^2-1$ is not 0.

59. Let Λ be the set of λ such that $E(\lambda) = \begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix}$ is in G. Since $E(\lambda+\lambda') = E(\lambda)E(\lambda')$ and $E(\lambda)^{-1} = E(-\lambda)$, Λ is closed under addition and negation. Since $\begin{pmatrix} \alpha & 0 \\ 0 & \alpha^{-1} \end{pmatrix} E(\lambda) \begin{pmatrix} \alpha & 0 \\ 0 & \alpha^{-1} \end{pmatrix}^{-1} = E(\alpha^2\lambda)$, Λ is closed under multiplication by squares of nonzero elements.

60. The previous problems produce some $\lambda_0 \neq 0$ in Λ, and $-\lambda_0$ is in Λ since Λ is closed under negatives. If $x \neq \pm1$, then $\frac{1}{4}(x+1)^2$ and $\frac{1}{4}(x-1)$ are nonzero squares, and hence $\frac{1}{4}(x+1)^2\lambda_0$ and $\frac{1}{4}(x-1)\lambda_0$ are in λ. Subtracting, we see that $x\lambda_0$ is in Λ. Thus all multiples of λ_0 except possibly for those by $0, +1, -1$ are in Λ. However, we have seen separately that $0, \lambda_0, -\lambda_0$ are in Λ. Hence $\Lambda = \mathbb{F}$.

61. The conjugacy follows from $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 \\ -\lambda & 1 \end{pmatrix}$. Next we have $\begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ b & 1 \end{pmatrix}\begin{pmatrix} 1 & c \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1+ab & c+cab+a \\ b & bc+1 \end{pmatrix}$, and it follows that every member of

$\mathrm{SL}(2,\mathbb{F})$ with lower left entry nonzero is in G. Conjugating by $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, we obtain the same conclusion when the upper right entry is nonzero. Finally $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \alpha & 0 \\ 0 & \alpha^{-1} \end{pmatrix} = \begin{pmatrix} \alpha & \alpha^{-1} \\ 0 & \alpha^{-1} \end{pmatrix}$ says $\begin{pmatrix} \alpha & 0 \\ 0 & \alpha^{-1} \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \alpha & \alpha^{-1} \\ 0 & \alpha^{-1} \end{pmatrix}$ and shows that every matrix $\begin{pmatrix} \alpha & 0 \\ 0 & \alpha^{-1} \end{pmatrix}$ is in G. Hence $G = \mathrm{SL}(2,\mathbb{F})$.

62. Let $\varphi : \mathrm{SL}(2,\mathbb{F}) \to \mathrm{PSL}(2,\mathbb{F})$ be the quotient homomorphism. If H is a normal subgroup $\neq \{1\}$ in $\mathrm{PSL}(2,\mathbb{F})$, then $\varphi^{-1}(H)$ is a normal subgroup of $\mathrm{SL}(2,\mathbb{F})$ containing an element not in the center. By Problem 61, $\varphi^{-1}(H) = \mathrm{SL}(2,\mathbb{F})$. Therefore $H = \varphi(\varphi^{-1}(H)) = \varphi(\mathrm{SL}(2,\mathbb{F})) = \mathrm{PSL}(2,\mathbb{F})$.

63. If a differs from c in a set A of k places and if b differs from c in a set B of l places, then a differs from b at most in the places of $A \cup B$, hence in at most $k + l$ places. Therefore $d(a, b) \leq d(a, c) + d(c, b)$.

If $d(w, a) \leq (D-1)/2$ and $d(w, b) \leq (D-1)/2$ with a and b distinct in C, then it follows that $d(a, b) \leq (D-1)$ and hence that $\delta(C) = \min_{x \neq y \text{ in } C} d(x, y) \leq d(a, b) \leq (D-1) < D$.

64. Since C is linear, 0 is in C. Then $\delta(C) \leq d(0, c)$ for every c in C, and we obtain $\delta(C) \leq \min_{c \in C} d(0, c)$. On the other hand, we certainly have $d(a, b) = d(0, a - b)$ for all a, b in $\mathbb{F}^n$. If a and b are in C, then the linearity of C forces $a - b$ to be in C, and hence $d(a, b) = d(0, a - b) \geq \min_{c \in C} d(0, c)$. Taking the minimum over all a and b, we obtain $\delta(C) \geq \min_{c \in C} d(0, c)$. Hence equality holds.

65. $n + 1$ and 0, 1 and n, n and 1, 2 and $n - 1$.

66. In (a), a basis vector c is 1 in one of the entries corresponding to the corner variables, and it is 0 in the other entries corresponding to corner variables. At worst it could be 1 in every entry corresponding to an independent variable. The number of independent variables is n minus the rank, i.e., n minus $\dim C$. Thus $\mathrm{wt}(c) \leq 1 + n - \dim C$. Since $\delta(C) \leq \mathrm{wt}(c)$, $\dim C + \delta(C) \leq n + 1$.

For (b), one can take the parity-check code.

For (c), the alternative would be $\dim C + \delta(C) = n + 1$. Then $\dim C + \mathrm{wt}(c) \geq n + 1$ for every c in C. Consequently every basis vector of C must have a 1 in every position corresponding to an independent variable. Since $\dim C \geq 2$, there are at least two such basis vectors. Their sum gets a contribution of 2 to its weight from the corner variables and can have a 0 in at most 1 position corresponding to an independent variable. But their sum is 0 in every position corresponding to an independent variable. Hence there is at most one such position, and we conclude that $n - \dim C = 1$, in contradiction to the hypothesis $\dim C \leq n - 2$.

67. A direct check of all seven nonzero elements of C shows that each has weight 3. Therefore $\delta(C) = 3$.

68. In (a), the basis vectors each have one 1 in positions 3, 5, 6, 7, and at least two of the parity bits in positions 1, 2, 4 are 1 since none of 3, 5, 6, 7 is a power of 2. Any sum of two distinct basis vectors has two 1's in positions 3, 5, 6, 7, and the parity bits cannot all be 0 since the parity bits for each of the basis vectors identify the

basis vector and since the two basis vectors in question are distinct. Finally the sum of three or more basis vectors has 1 in three or more positions 3, 5, 6, 7 and hence has weight ≥ 3. Thus all code words have weight ≥ 3, and therefore $\delta(C_7) \geq 3$. Since the first basis vector has weight 3, $\delta(C_7) = 3$.

In (b), each word in C_8 is a word of C_7 plus a parity bit. The part from C_7 has weight ≥ 3, by (a), and the parity bit means that the weight has to be even. Thus the weight of every word in C_8 is ≥ 4.

In (c) for C_{2^r-1}, we distinguish between the r bits whose indices are a power of two and the other $2^r - 1 - r$ bits. The first are the check bits, and the others are the message bits. The message bits are allowed to be arbitrary, and the check bits will depend on them. Thus $\dim C_8 = 2^r - r - 1$. For a given pattern of message bits, the check bit in position 2^j counts, modulo 2, the number of 1's in message bits that occur is positions requiring 2^j in their binary expansions. Then C_{2^r} is obtained by adjoining a parity bit to each word of C_{2^r-1}.

The first conclusion of (d) was proved in the course of answering (c), and the other two conclusions follow by the same argument that was given for $r = 3$ in (a) and (b).

69. In (a), the dimension of the null space of H is the number of columns minus the rank, hence is $7 - 3 = 4$. Since C_7 lies in the null space and $\dim C_7 = 4$, the null space equals C_7.

In (b), let c be in C_7. If e_i denotes the usual i^{th} basis vector, then $H(c + e_i) = Hc + He_i = He_i$, and this is the i^{th} column of H.

70. Take a basis of C, write it as the rows of a matrix, row reduce the matrix, and permute the variables so that all the corner variables precede all the independent variables. The resulting matrix in block form is $(I \;\; A)$ for some matrix A with $\dim C$ rows and $n - \dim C$ columns. Since each basis vector has weight ≥ 3, each row of A has at least two 1's. Since each sum of two distinct basis vectors has weight ≥ 3, the sum of two distinct rows of A cannot be 0. Thus the rows of A must be distinct.

Arguing by contradiction, suppose that $\dim C > n - r$, so that A has $\leq r - 1$ columns. The number of possible rows in A with at least two 1's is then $\leq 2^{r-1} - 1 - (r - 1) = 2^{r-1} - r$. Hence $n - r < \dim C \leq 2^{r-1} - r$, and $n < 2^{r-1}$, contradiction.

71. For (a), the answers are X^n, $(X + Y)^n$, $X^n + Y^n$, $\frac{1}{2}((X + Y)^n + \frac{1}{2}(X - Y)^n)$, $X^6 + 7X^3Y^3$, $X^7 + 7X^4Y^3 + 7X^3Y^4 + Y^7$, and $X^8 + 14X^4Y^4 + Y^8$. The last three are by a direct count of the number of code words of each weight.

In (b), the 0 word is the unique code word of weight 0, and it is present in every linear code.

In (c), the expression $X^{n-\text{wt}(c)}Y^{\text{wt}(c)}$ makes a contribution of 0 to the coefficient $N_k(C)$ of $X^{n-k}Y^k$ if $\text{wt}(c) \neq k$ and makes a contribution of 1 to the coefficient if $\text{wt}(c) = k$. Summing on c yields $\sum_{k=0}^{n} N_k(C)X^{n-k}Y^k = \sum_{c \in C} X^{n-\text{wt}(c)}Y^{\text{wt}(c)}$.

72. The equality $(1+X+X^2+X^4)(1+X+X^3) = 1+X^7$ produces a member of C with weight 2. Therefore $\delta(C) \leq 2$. On the other hand, the product of $1+X+X^2+X^4$ with a polynomial can never be a monomial, and therefore no code word has weight 1. Thus $\delta(C) > 1$.

73. In essence we use the method suggested by the solution to Problem 70, except that we put coefficients corresponding to low degrees on the left and we row reduce the matrix into the form $(A\ \ I)$. Let $8 \le n \le 19$. Form the images of as many of the following polynomials as have degree $\le n$:

$$1,\ X,\ X^2,\ X^3,\ X^4,\ X^5,\ X^6+1,\ X^7+X+1,\ X^k(X^8+X^2+X+1) \text{ for } k \ge 0.$$

The list stops with $k = n - 16$. Assemble the coefficients of the image polynomials as the rows of a matrix as in Problem 70. The images form a basis of C. They all have weight 4, and thus every member of C has even weight. Since the image of 1 has weight 4, $\delta(C)$ must be 2 or 4.

Imagine doing a row reduction as in the solution of Problem 70. We want to rule out $\delta(C) = 2$, and it is enough to show that the basis vectors and all sums of two distinct basis vectors have weight > 2. To handle the basis vectors, it is enough to show that the A part of the reduced matrix $(A\ \ I)$ never has just one 1 in a row. To handle the sums of two distinct basis vectors, it is enough to show that the sum of two rows of A is never 0, i.e., that the rows of A are distinct.

The matrix A will have 8 columns, corresponding to powers X^l with $l \le 7$. The rows of $(A\ \ I)$ are thus to correspond to polynomials of the form $X^m +$"lower," where each expression "lower" has degree at most 7 and m takes on the values $8, 9, \ldots, n$. The polynomials whose images correspond to the rows of the reduced matrix are

$$1,\ X, \ldots,\ X^5,\ X^6+1,\ X^7+X+1,$$
$$X^8+X^2+X+1,\ X(X^8+X^2+X+1), \ldots,\ X^3(X^8+X^2+X+1),$$

and the left part A of the reduced matrix is

$$A = \begin{pmatrix}
1&1&1&0&0&0&0&0\\
0&1&1&1&0&0&0&0\\
0&0&1&1&1&0&0&0\\
0&0&0&1&1&1&0&0\\
0&0&0&0&1&1&1&0\\
0&0&0&0&0&1&1&1\\
1&1&1&0&0&0&1&1\\
1&0&0&1&0&0&0&1\\
1&0&1&0&1&0&0&0\\
0&1&0&1&0&1&0&0\\
0&0&1&0&1&0&1&0\\
0&0&0&1&0&1&0&1
\end{pmatrix}.$$

No row of A is 0, and no two distinct rows are equal. This completes the proof.

74. Suppose that $\{X_s\}_{s\in S}$ is an object in $\mathcal{C}^S$ and that $f_s : X_s \to A$ for each s is a function, A being a particular set. The disjoint union of the X_s's consists of all ordered pairs (x_s, s) with $s \in S$ and $x_s \in X_s$, and we define $i_s(x_s) = (x_s, s)$. To define a function f from the disjoint union of the X_s's into A such that $fi_s = f_s$ for all s, we let $f(x_s, s) = f_s(x_s)$. Then $fi_s(x_s) = f(x_s, s) = f_s(x_s)$. Thus f exists. On the other hand, the condition that $fi_s = f_s$ forces $f(x_s, s)$ to be $f_s(x_s)$, and hence the f in the universal mapping property is unique, as it is required to be.

76. Peeking ahead to Problem 80, we take the category to be $\mathcal{C}^{\text{opp}}$, where $\mathcal{C}$ is the category defined in Section 11 after Example 4 of products. The category $\mathcal{C}$ has no product functor when S has two elements.

77. The existence of the identity and associativity are part of the definition. The existence of inverses is given in the hypothesis. The answer to the question is "yes"; if a group G is given, define a category with one object, namely the set G, define $\text{Morph}(G, G)$ to be the set G, and let the law of composition be the group law.

78. To see that $\circ^{\text{opp}}$ is well defined, let f be in $\text{Morph}_{\mathcal{C}^{\text{opp}}}(A, B)$, and let g be in $\text{Morph}_{\mathcal{C}^{\text{opp}}}(B, C)$. The definition is $g \circ^{\text{opp}} f = f \circ g$, and this is meaningful since g is in $\text{Morph}_{\mathcal{C}}(C, B)$ and f is in $\text{Morph}_{\mathcal{C}}(B, A)$. The associativity and the existence of the identity are straightforward to check. It is clear from the definition that $(\mathcal{C}^{\text{opp}})^{\text{opp}} = \mathcal{C}$.

In a diagram the vertices stay where they are, and so do the morphisms, since the objects and the sets of morphisms do not change. However, the direction of each arrow is reversed since "domain" and "range" are interchanged in passing from $\mathcal{C}$ to $\mathcal{C}^{\text{opp}}$. Thus diagrams map to diagrams with the arrows reversed.

Compositions correspond because of the definition of $\circ^{\text{opp}}$, and it follows that commutative diagrams map to commutative diagrams.

79. Let A and B be sets such that A has three elements and B has one element. The number of functions from A to B is then one, and the number of functions from B to A is three. Since $\text{Morph}_{\mathcal{C}^{\text{opp}}}(A, B) = \text{Morph}_{\mathcal{C}}(B, A)$, $\text{Morph}_{\mathcal{C}^{\text{opp}}}(A, B)$ has three elements and cannot be accounted for by functions from A to B.

80. For (a), if $(X, \{p_s\}_{s\in S})$ is a product of $\{X_s\}_{s\in S}$, we set up the diagram of the universal mapping property of the product. Passing to $\mathcal{C}^{\text{opp}}$ and using Problem 78, we obtain the same diagram in $\mathcal{C}^{\text{opp}}$ but with the arrows reversed. Then it follows that $(X, \{p_s\}_{s\in S})$, when interpreted in $\mathcal{C}^{\text{opp}}$, satisfies the condition of being a coproduct. The other half proceeds in the same way.

For (b), we start with two coproducts in $\mathcal{C}$ and pass to $\mathcal{C}^{\text{opp}}$, where they become products, according to (a). Proposition 4.63 shows that the two products are canonically isomorphic in $\mathcal{C}^{\text{opp}}$. This isomorphism, when reinterpreted in $\mathcal{C}$, is still an isomorphism, and the result is that the two coproducts in $\mathcal{C}$ are canonically isomorphic.

Chapter V

1. For (a), we have $((g_1, h_1)((g_2, h_2)x)) = (g_1, h_1)(g_2 x h_2^{-1}) = g_1 g_2 x h_2^{-1} h_1^{-1} = (g_1 g_2)x(h_1 h_2)^{-1} = (g_1 g_2, h_1 h_2)x$ and $(1, 1)x = 1x1^{-1} = x$.

For (b), left multiplications by $\text{GL}(m, \mathbb{C})$ preserve the row space, hence the rank, and right multiplications by $\text{GL}(n, \mathbb{C})$ preserve the column space, hence the rank. Hence all members of an orbit have the same rank.

Row operations, which correspond to left multiplications by elementary matrices, can be used to bring the matrix into reduced row-echelon form, and then column

operations, which correspond to right multiplications by elementary matrices, can be used to bring the result into reduced column-echelon form. If $r = \min(m, n)$, then the resulting matrix is 1 in entries $(1, 1)$, $(2, 2)$, $\ldots$, (l, l) for some $l \leq r$ and 0 elsewhere. This has rank l and answers (c) and the remainder of (b).

2. If A has minimal polynomial $X^k + c_{k-1}X^{k-1} + \cdots + c_1X + c_0$, with $c_0 \neq 0$, then $I = A(-c_0^{-1}(A^{k-1} + c_{k-1}A^{k-2} + \cdots + c_1I))$, and A is invertible. Conversely if $c_0 = 0$, then X is a factor of the minimal polynomial and must be a factor of the characteristic polynomial, by Corollary 5.10. Then 0 is an eigenvalue, and the null space is nonzero. Hence A is not invertible.

3. Proposition 5.12 shows that $l_j \geq \max(r_j, s_j)$. For u in U, we know that $P_1(L)^{r_1} \cdots P_k(L)^{r_k}(u) = 0$. For w in W, we know that $P_1(L)^{s_1} \cdots P_k(L)^{s_k}(w) = 0$. Thus any v in U or W has $P_1(L)^{\max(r_1,s_1)} \cdots P_k(L)^{\max(r_k,s_k)}(v) = 0$. Forming sums, we see that $P_1(L)^{\max(r_1,s_1)} \cdots P_k(L)^{\max(r_k,s_k)}(v) = 0$ for all v in V. Thus the minimal polynomial divides $P_1(X)^{\max(r_1,s_1)} \cdots P_k(X)^{\max(r_k,s_k)}$, and we must have $l_j \leq \max(r_j, s_j)$.

4. For any monomial $P(X) = X^j$, the monomial $Q(X) = XP(X) = X^{j+1}$ has $Q(BA) = BA(BA)^j = B(AB)^jA = BP(AB)A$. Taking suitable linear combinations of this result as j varies, we obtain (a).

For (b), let $M_{AB}(X)$ and $M_{BA}(X)$ be the minimal polynomials of AB and BA. Part (a) implies that $M_{BA}(X)$ divides $XM_{AB}(X)$. Reversing the roles of A and B, we see that $M_{AB}(X)$ divides $XM_{BA}(X)$. By unique factorization all the prime powers in the prime factorizations of $M_{AB}(X)$ and $M_{BA}(X)$ are the same except for the power of X. The powers of X in the factorizations of $M_{AB}(X)$ and $M_{BA}(X)$ differ at most by 1.

5. Theorem 5.14 allows us to write $\mathbb{K}^n = U_1 \oplus \cdots \oplus U_k$ and $\mathbb{K}^n = W_1 \oplus \cdots \oplus W_l$, where the U_j are the eigenspaces for the distinct eigenvalues of D and the W_j are the eigenspaces for the distinct eigenvalues of D'. These decompositions are the primary decompositions as in Theorem 5.19, and (e) of that theorem shows that $W_j = (W_j \cap U_1) \oplus \cdots \oplus (W_j \cap U_k)$ for $1 \leq j \leq l$. Summing on j, we see that $\mathbb{K}^n$ is the direct sum of all $U_i \cap W_j$. Each of D and D' is scalar on $U_i \cap W_j$, and (a) follows by translating this result into a statement about matrices.

The matrices $N = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ and $N' = \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}$ commute, and both have N uniquely as Jordan form. If C were to exist with $C^{-1}NC$ and $C^{-1}N'C$ both in Jordan form, we would have $C^{-1}NC = C^{-1}N'C$ and $N = N'$, contradiction. This answers (b).

6. If E is the projection of V on U along W, then each member of U is an eigenvector with eigenvalue 1, and each member of W is an eigenvector with eigenvalue 0. The union of bases of U and W is then a basis of eigenvectors for E, and (a) follows from Theorem 5.14. In view of Proposition 5.15, two projections are given by similar matrices if and only if they have the same rank.

7. For (a), $EF = F$ implies image $F \subseteq$ image E, which implies $EF = F$. Reversing the roles of E and F, we see that $FE = E$ if and only if image $E \subseteq$

image F.

For (b), $EF = E$ implies $\ker F \subseteq \ker E$, while $FE = F$ implies $\ker E \subseteq \ker F$. So $EF = E$ and $FE = F$ implies $\ker E = \ker F$. Conversely if $\ker F \subseteq \ker E$, then $EF = E$ on $\ker F$ and $EF = E$ on image F; so $EF = E$. Reversing the roles of E and F, we see that $\ker E \subseteq \ker F$ implies $FE = F$.

8. If $EF = FE$, then $(EF)^2 = EFEF = E(FE)F = E(EF)F = E^2F^2 = EF$. So EF is a projection. This proves (a).

For (b), let $E = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ and $F = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$. Each is a projection, and $EF = F$, so that EF is a projection. However, $FE = E$. Since $E \neq F$, $EF \neq FE$.

9. If E is a projection, then $U = 2E - I$ has $U^2 = 4E^2 - 4E + I = 4E - 4E + I = I$; so U is an involution. If U is an involution, then $E = \frac{1}{2}(U + I)$ has $E^2 = \frac{1}{4}(U^2 + 2U + I) = \frac{1}{4}(I + 2U + I) = \frac{1}{2}(U + I) = E$. So E is a projection. The two formulas $U = 2E - I$ and $E = \frac{1}{2}(U + I)$ are inverse to each other.

10. Apply Theorem 5.19, and take U to be the primary subspace for the prime polynomial X and W to be the sum of the remaining primary subspaces. Then (i), (ii), and (iii) are immediate from the theorem. For (iv), let U_j be the primary subspace for some other prime polynomial $P(X)$. The theorem shows that $L\big|_{U_j}$ has a power of $P(X)$ as minimal polynomial. Since X does not divide $P(X)$, Problem 2 shows that $L\big|_{U_j}$ is invertible. Hence $L\big|_{U_j}$ is invertible on the direct sum of the U_j's other than the one for the polynomial X.

11. Let $V = U_1 \oplus \cdots \oplus U_k$ be the primary decomposition, with U_1 corresponding to the prime X. By (ii) and Theorem 5.19e, $U = (U_1 \cap U) \oplus \cdots \oplus (U_k \cap U)$ and similarly for W. Then $U_j \cap U = 0$ for $j \geq 2$ by (iii), and hence $U \subseteq U_1$. By (iv), $U_1 \cap W = 0$, so that $W \subseteq U_2 \oplus \cdots \oplus U_k$. By (i), $U = U_1$ and $W = U_2 \oplus \cdots \oplus U_k$.

12. Part (a) is immediate, and a basis for (b) consists of the union of bases for the individual U_j's. Part (f) is evident.

For (d) and (e), since D is a linear combination of the E_j's and each E_j is a polynomial in L, D is a polynomial in L, say $D = P(L)$. Then $N = L - P(L)$ commutes with L, and this is (d). Applying the division algorithm to P, we have $P = AM + R$ with $R = 0$ or $\deg R < \deg M$. Evaluating at L gives $D = P(L) = A(L)M(L) + R(L) = R(L)$ since $M(L) = 0$. Thus R will serve in place of P if $\deg P \geq \deg M$. This proves the existence in (e) of the polynomial for D. Since $N = L - D$, N is a polynomial in L, and again we can take this polynomial to be 0 or to have degree $< \deg M$. This proves the existence in (e) of the polynomial for N. For uniqueness if P_1 is a second polynomial that yields D, then $0 = D - D = P(L) - P_1(L)$ shows that $P - P_1$ is a multiple of M, and the condition on the degrees of P and P_1 forces $P - P_1 = 0$. So P is unique. Similarly the polynomial representing N is unique. This completes the proof of uniqueness in (e).

If $Q_j(X)$ is the polynomial $(X - \lambda_0)^{l_j}$, then $N^{l_j} = (L - D)^{l_j} = Q_j(L)$ on U_j, and Theorem 5.19f shows that $Q_j(L)$ is 0 on U_j. Therefore a power of N is 0 on each U_j, and N is nilpotent. This proves (c). Part (g) now follows.

13. Each eigenvector of D must lie in some U_j by Theorem 5.19e. If V_i is the eigenspace of D with eigenvalue c_i, it follows that $V_i \subseteq U_{j(i)}$ for some $j = j(i)$. Thus each U_j is the sum of full eigenspaces of D. Property (d) forces N to carry V_i into itself. By (c), $(L - D)^n$ is 0 on V_i for $n = \dim V$; hence $(L - c_i I)^n$ is 0 on V_i. Since $V_i \subseteq U_j$, $(L - \lambda_j I)^n$ is 0 on U_j. Application of Problem 10 to $L - c_i I$ shows that $L - \lambda_j I$ is nonsingular on V_i if $c_i \neq \lambda_j$, in contradiction to the fact that $(L - \lambda_j I)^n$ is 0 on U_j, and therefore $c_i = \lambda_j$. The conclusion is that $V_i = U_{j(i)}$, and the desired uniqueness follows.

A slightly shorter argument is available if one takes the constructive proof of existence of a decomposition $L = D + N$ as known, so that Problem 12 is available for that decomposition. If there is a second decomposition $L = D' + N'$ satisfying (a) through (d), then D' and N' commute with L and hence with all polynomials in L. Thus they commute with D and N. The equality $L = D + N = D' + N'$ implies $D - D' = N' - N$. Problem 5a shows that $D - D'$ has a basis of eigenvectors, and $N' - N$ is nilpotent because the commutativity of N and N' shows that the the Binomial Theorem applies, in view of Problem 15 in Chapter I. Thus $D - D' = N' - N = 0$.

14. In (a), Lemma 5.22 says that $\det(XI - N') = X^{n'}$. Consequently $\det(XI - (N' + cI)) = \det((X - c)I - N') = (X - c)^{n'}$.

In (b), form the primary decomposition of L as in Theorem 5.19, and let notation be as in Problem 12. On the subspace U_j, which is carried to itself by L, $L = D + N$ acts as $\lambda_j I + N$, and the characteristic polynomial on that subspace is $(X - \lambda_j)^{n_j}$, by (a). On the whole space V, the characteristic polynomial of L is the product of the contributions from each U_j, since as a consequence of Proposition 5.11, the determinant of a block diagonal matrix is the product of the determinants of the blocks. Therefore L has characteristic polynomial $\prod_{j=1}^{n} (X - \lambda_j)^{n_j}$, and this matches the characteristic polynomial of D.

15. The characteristic polynomial is $X^2 - 2X + 1 = (X - 1)^2$. Since $A - I \neq 0$, the minimal polynomial is $(X - 1)^2$ rather than $X - 1$. Thus the Jordan form is $J = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Solving shows that $\ker(A - I)$ consists of the multiples of $\begin{pmatrix} 3/2 \\ 1 \end{pmatrix}$. Use $\begin{pmatrix} 3 \\ 2 \end{pmatrix}$ as the first column of C, and solve $(A - I)X = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$ to get $X = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ as one answer for the second column. Then $C = \begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix}$, $C^{-1} = \begin{pmatrix} 1 & -1 \\ -2 & 3 \end{pmatrix}$, and one readily checks that $C^{-1}AC = J$.

16. The characteristic polynomial is $P(X) = \det(XI - A) = X^3$. Thus A is nilpotent, and in fact $A^2 = 0$. Then $J = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, and the computation proceeds as in Example 1 in Section 7, yielding $C = \begin{pmatrix} 4 & 1 & -1 \\ -8 & 0 & 4 \\ 8 & 0 & 0 \end{pmatrix}$ and $C^{-1} = \begin{pmatrix} 0 & 0 & \frac{1}{8} \\ 1 & \frac{1}{4} & -\frac{1}{4} \\ 0 & \frac{1}{4} & \frac{1}{4} \end{pmatrix}$.

17. The characteristic polynomial is $(X - 2)^6(X - 3)$ by inspection. Thus there is a primary subspace for $X - 2$ with dimension 6 and a primary subspace for $X - 3$

with dimension 1. For the Jordan form let $K_j = \ker(A - 2I)^j$. By raising $A - 2I$ to powers and row reducing, we see that $\dim K_3 = 6$, $\dim K_2 = 5$, and $\dim K_1 = 3$. We do not have to proceed beyond K_3 since we have reached the full dimension 6 of the primary subspace for $X - 2$. Therefore the number of Jordan blocks for $X - 2$ of size ≥ 3 is $6 - 5 = 1$, of size ≥ 2 is $5 - 3 = 2$, and of size ≥ 1 is 3. Hence there is one block of each size 1, 2, and 3, and

$$J = \begin{pmatrix} 2 & 1 & 0 & & & & \\ & 2 & 1 & & & & \\ & & 2 & & & & \\ & & & 2 & 1 & & \\ & & & & 2 & & \\ & & & & & 2 & \\ & & & & & & 3 \end{pmatrix}.$$

Solving $(A - 3I)X = 0$, we find that the eigenvectors for eigenvalue 3 are the multiples of $(5, 2, 2, 3, 2, 1, 1)$. Thus this vector can be taken to be the last column of C.

The next step is to express K_1, K_2, and K_3 explicitly in terms of parameters by using the standard solution procedure for systems of homogeneous linear equations. The result is that

$$K_1 = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\}, \qquad K_2 = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ 0 \\ 0 \end{pmatrix} \right\}, \qquad K_3 = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ 0 \end{pmatrix} \right\}.$$

Following the method of Example 1 in Section 7, we choose W_2 such that $K_3 = K_2 \oplus W_2$, and then we form $U_1 = (A - 2I)(W_2)$:

$$W_2 = \left\{ \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ x_6 \\ 0 \end{pmatrix} \right\} \qquad \text{and} \qquad U_1 = \left\{ \begin{pmatrix} 0 \\ x_6 \\ 0 \\ x_6 \\ x_6 \\ \\ 0 \end{pmatrix} \right\}.$$

We choose W_1 such that $K_2 = K_1 \oplus U_1 \oplus W_1$, and we form $U_0 = (A - 2I)(U_1 + W_1)$:

$$W_1 = \left\{ \begin{pmatrix} 0 \\ 0 \\ 0 \\ x_4 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\} \qquad \text{and} \qquad U_0 = \left\{ \begin{pmatrix} x_4 + 2x_6 \\ 0 \\ x_6 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\}.$$

Finally we choose W_0 such that $K_1 = K_0 \oplus U_0 \oplus W_0$. Here $K_0 = 0$, and we can take $W_0 = \{(0, x_2, 0, 0, 0, 0, 0)\}$.

To form C we take a basis of each W_j, apply powers of $A-2I$ in turn to its members, and line up the resulting columns, along with the eigenvector for eigenvalue 3, as C:

$$C = \begin{pmatrix} 2&0&0&1&0&0&5 \\ 0&1&0&0&0&1&2 \\ 1&0&0&0&0&0&2 \\ 0&1&0&0&1&0&3 \\ 0&1&0&0&0&0&2 \\ 0&0&1&0&0&0&1 \\ 0&0&0&0&0&0&1 \end{pmatrix}.$$

18. In (a), if every prime-power factor of the minimal polynomial is of degree 1, then the matrix is similar to a diagonal matrix, and the multiplicities of the eigenvalues can be seen from the characteristic polynomial. If the minimal polynomial is $(X-c)^2$, then the matrix has to be similar to $\begin{pmatrix} c&1&0 \\ 0&c&0 \\ 0&0&c \end{pmatrix}$. If the minimal polynomial instead is $(X-c)^2(X-d)$, then the matrix has to be similar to $\begin{pmatrix} c&1&0 \\ 0&c&0 \\ 0&0&d \end{pmatrix}$. If the minimal polynomial is $(X-c)^3$, then the matrix has to be similar to $\begin{pmatrix} c&1&c \\ 0&c&1 \\ 0&0&c \end{pmatrix}$. There are no other possibilities.

For (b), $\begin{pmatrix} 0&1&0&0 \\ 0&0&0&0 \\ 0&0&0&1 \\ 0&0&0&0 \end{pmatrix}$ and $\begin{pmatrix} 0&1&0&0 \\ 0&0&0&0 \\ 0&0&0&0 \\ 0&0&0&0 \end{pmatrix}$ both have minimal polynomial X^2 and characteristic polynomial X^4, but they are not similar because their ranks are unequal.

19. If the diagonal entries are c and N denotes the strictly upper-triangular part, then $J^k = (cI+N)^k = \sum_{j=0}^{k} \binom{k}{j} c^{k-j} N^j$. The term from $j=1$ is not canceled by any other term, and hence J^k is not diagonal.

20. Choose J in Jordan form and C invertible with $J = C^{-1}AC$. Then $J^n = CA^nC^{-1} = CC^{-1} = I$. By Problem 19, every Jordan block in J is of size 1-by-1. Thus A is similar to a diagonal matrix D, and each diagonal entry of D must be an n^{th} root of unity. Any n-tuple of n^{th} roots of unity can form the diagonal entries, and the corresponding matrices are similar if and only if one is a permutation of the other.

21. The minimal polynomial has to divide $X(X^2-1) = X(X+1)(X-1)$. Hence there is a basis of eigenvectors, the allowable eigenvalues being 1, -1, and 0. A similarity class is therefore given by an unordered triple of elements from the set $\{1,-1,0\}$. There are three possibilities for a single eigenvalue, six possibilities for one eigenvalue of multiplicity 2 and one of multiplicity 1, and one possibility with all three eigenvalues present. So the answer is ten.

22. If $A^2 = N$ and $N^n = 0$, then $A^{2n} = 0$. So A is nilpotent and $A^n = 0$. Since $N^{n-1} \neq 0$, $A^{2n-2} \neq 0$. Therefore $n > 2n-2$, and $n = 1$.

23. If J is of size n, then the matrix C with $C_{i,n+1-i} = 1$ for $1 \leq i \leq n$ and $C_{ij} = 0$ otherwise has $C^{-1}JC = J^t$.

24. Choose C with $C^{-1}AC = J$ in Jordan form. Problem 23 shows that there is a block-diagonal matrix B with $B^{-1}JB = J^t$. Then $B^{-1}C^{-1}ACB = J^t$ and $C^tA^t(C^{-1})^t = J^t$. So $B^{-1}C^{-1}ACB = C^tA^t(C^{-1})^t$, and the result follows.

25. The matrices A and B have $A^2 = B^2 = 0$ and hence are nilpotent. Since each of A and B has rank 2, $\dim\ker A = \dim\ker B = 2$. The numbers $\dim\ker A^k$ and $\dim\ker B^k$ being equal for all k, the two matrices have the same Jordan form and are therefore similar.

26. If $M(X)$ is the minimal polynomial of L, then $M(L)v = 0$. Hence $M(X)$ is in $\mathcal{I}_v$. Then Proposition 5.8 shows that $M_v(X)$ exists.

27. The polynomial $M_v(X)$ has to divide the minimal polynomial of $L\big|_{\mathcal{P}(v)}$, and the latter has degree $\le \dim\mathcal{P}(v)$. Hence $\deg M_v(X) \le \dim\mathcal{P}(v)$. If $v, L(v), \dots, L^{\deg M_v - 1}(v)$ are linearly dependent, then there is a nonzero polynomial $Q(X)$ of degree $\le \deg M_v - 1$ with $Q(L)(v) = 0$, and that fact contradicts the minimality of the degree of $M_v(X)$. Hence they are independent, and $\deg M_v(X) \ge \dim\mathcal{P}(v)$. Thus equality holds, and the linearly independent set is a basis. This proves (a) and (b).

Since $M_v(X)$ divides the minimal polynomial of $L\big|_{\mathcal{P}(v)}$, which divides the characteristic polynomial of $L\big|_{\mathcal{P}(v)}$, and since the end polynomials have degree $\dim\mathcal{P}(v)$, these three polynomials are all equal. This proves (c).

28. Use the ordered basis $(L^{d-1}(v), L^{d-2}(v), \dots, L(v), v)$.

29. Since $P(X)$ is prime and does not divide $Q(X)$, there exist polynomials $A(X)$ and $B(X)$ with $A(X)P(X) + B(X)Q(X) = 1$. Using the substitution that sends X to L and applying both sides to v, we obtain $B(L)Q(L)(v) = v$. Hence $\mathcal{P}(Q(L)(v)) \supseteq \mathcal{P}(v)$. Since the reverse inclusion is clear, the result follows.

30. In (a), the base case of the induction is that $\dim V = \deg P(X)$, and then the result follows from Problem 27. For the inductive step, the same problem shows that there must be a nontrivial invariant subspace U. Proposition 5.12 shows that the minimal polynomial for U and V/U is $P(X)$, and induction shows that the characteristic polynomial for U and V/U is a power of $P(X)$. Proposition 5.11 then shows that the characteristic polynomial for V is a power of $P(X)$.

For (b), we induct on l, using (a) to handle the case $l = 1$. For general l, form the invariant subspace $U = \ker P(X)^{l-1}$, for which the minimal polynomial is some $P(X)^r$ with $r < l$. The minimal polynomial of V/U is certainly $P(X)$. By induction, U and V/U have characteristic polynomials equal to powers of $P(X)$, and Proposition 5.11 shows that the same thing is true for V.

In (c), (b) says that the characteristic polynomial is of the form $P(X)^r$ for some r. Then the degree of the characteristic polynomial is rd, where $d = \deg P(X)$.

32–34. These are proved word-for-word in the same way as Lemmas 5.23 through 5.25 except that n is to be replaced by l and N is to be replaced by $P(L)$.

35. If $Q(X)$ is in $\mathbb{K}[X]$, we successively apply the division algorithm to write

$$\begin{aligned} Q &= A_0P + B_0 \quad \text{with } \deg B_0 < \deg P, \\ A_0 &= A_1P + B_1 \quad \text{with } \deg B_1 < \deg P, \\ A_1 &= A_2P + B_2 \quad \text{with } \deg B_2 < \deg P, \end{aligned}$$

etc., and then we substitute and find that

$$\begin{aligned} Q &= A_0P + B_0 = A_1P^2 + B_1P + B_0 = A_2P^3 + B_2P^2 + B_1P + B_0 \\ &= \cdots = A_jP^{j+1} + B_jP^j + \cdots + B_2P^2 + B_1P + B_0 \end{aligned}$$

with each B_i equal to 0 or of degree $< \deg P$. The fact that $W_j \subseteq K_{j+1}$ implies that $P^{j+1}(L)(v) = 0$. Consequently

$$\mathcal{P}(v) = \{(B_jP^j + \cdots + B_1P + B_0)(L)(v) \mid B_i = 0 \text{ or } \deg B_i < d \text{ for } 0 \le i \le j\},$$

and the given set spans $\mathcal{P}(v)$.

For the linear independence suppose that some such expression is 0 with not all $B_i(X)$ equal to 0. Fix i as small as possible with $B_i(X) \neq 0$. Since $P(L)^{j+1}(v) = 0$, $B_r(L)P(L)^r(v)$ is annihilated by $P(L)^{j-i}$ if $r > i$. Application of $P(L)^{j-i}$ to the dependence relation yields

$$P(L)^{j-i}(B_j(L)P(L)^j(v) + \cdots + B_{i+1}(L)P(L)^{i+1} + B_i(L)P(L)^i)(v) = 0$$

and therefore also $B_i(L)P(L)^j(v) = 0$. Since $\deg B_i < \deg P$, Problem 29 shows that $P(L)^j(v) = 0$. Therefore v is in K_j. Since $W_j \cap K_j$, we conclude $v = 0$, contradiction.

36. We show at the same time that it is possible to arrange for each U_j and W_j to be such that $K_j + U_j$ and $K_j + W_j$ are invariant under L. We proceed by induction downward on j. The construction begins with $U_{l-1} = 0$ and W_{l-1} chosen such that $K_l = K_{l-1} \oplus W_{l-1}$. Then we have $L(W_{l-1}) \subseteq W_{l-1} + K_{l-1}$ and $L(U_{l-1}) \subseteq U_{l-1} + K_{l-1}$. Select some $v_1^{(l-1)} \neq 0$ in W_{l-1}. If there is a polynomial $B(X) \neq 0$ with $\deg B < \deg P$ such that $B(L)(v_1^{(l-1)})$ is in K_{l-1}, then it follows from Problem 29 and the invariance of K_{l-1} under L that $v_1^{(l-1)}$ is in K_{l-1}, contradiction. So there is no such polynomial, and the vectors $v_1^{(l-1)}, L(v_1^{(l-1)}), \ldots, L^{d-1}(v_1^{(l-1)})$ are linearly independent with span $T_1^{(l-1)}$ such that $K_{l-1} + T_1^{(l-1)}$ is a direct sum.

If $K_{l-1} + T_1^{(l-1)} \neq K_l$, then we form $v_2^{(l-1)}$ and $T_2^{(l-1)}$ in the same way. If there is a polynomial $B(X) \neq 0$ with $\deg B < \deg P$ such that $B(L)(v_2^{(l-1)})$ is in $K_{l-1} + T_1^{(l-1)}$, then Problem 29 shows that $v_2^{(l-1)}$ is in $K_{l-1} + T_1^{(l-1)}$, contradiction. We conclude that $K_{l-1} + T_1^{(l-1)} + T_2^{(l-1)}$ is a direct sum. Continuing in this way,

we obtain enough linearly independent vectors to have a basis for a complement $W_{l-1} = T_1^{(l-1)} + T_2^{(l-1)} + \cdots$ to K_{l-1}.

Now suppose inductively in the construction of U_j and W_j that $j \le l-2$ and that $U_{j+1} + K_{j+1}$ and $W_{j+1} + K_{j+1}$ are invariant under L. We define $U_j = P(L)(U_{j+1} \oplus W_{j+1})$, and the assumed invariance implies that $U_j + K_j$ is invariant under L. We now construct W_j in the same way that we constructed W_{l-1}, insisting that $(U_j + K_j) \cap W_j = 0$. If we choose $v_1^{(j)}$ in K_{j+1} but not $U_j + K_j$, then the invariance of $U_j + K_j$ under L implies that the vectors $v_1^{(j)}, L(v_1^{(j)}), \dots, L^{d-1}(v_1^{(j)})$ are linearly independent and their linear span $T_1^{(j)}$ is such that $U_j + K_j + T_1^{(j)}$ is a direct sum. Continuing in this way, we obtain the required basis of a complement W_j to $K_j \oplus U_j$.

37. Problem 36 arranges that the vectors $L^r(v_{i_j}^{(j)})$ for $0 \le r \le d-1$ and all i_j form a basis of W_j. We show by induction downward for $j \le l-1$ that the vectors $L^r P(L)^k(v_{i_{j+k}}^{(j+k)})$ for $0 \le r \le d-1$, $k > 0$, and all i_{j+k} form a basis of U_j. This holds for $j = l-1$ since $U_{l-1} = 0$. If it is true for $j+1$, then $U_{j+1} \oplus W_{j+1}$ has a basis consisting of all $L^r P(L)^k(v_{i_{j+1+k}}^{(j+1+k)})$ for $0 \le r \le d-1$, $k \ge 0$, and all i_{j+1+k}. Since Problem 33 shows that $P(L)$ is one-one from $U_{j+1} \oplus W_{j+1}$ onto U_j, U_j has a basis consisting of all $L^r P(L)^{k+1}(v_{i_{j+1+k}}^{(j+1+k)})$ for $0 \le r \le d-1$, $k \ge 0$, and all i_{j+1+k}, i.e., all $L^r P(L)^k(v_{i_{j+k}}^{(j+k)})$ for $0 \le r \le d-1$, $k > 0$, and all i_{j+k}. This completes the induction.

38. Problem 35 gives a basis for the cyclic subspace generated by $v_{i_j}^{(j)}$, Problem 37 shows that the members within $U_i \oplus W_i$ of the union of these bases, as j and i_j vary, form a basis of $U_i \oplus W_i$, and Problem 34 allows us to conclude that as i varies, we obtain a basis of V.

39. Because of the linear independence proved in Problem 38, the left side of the formula in question equals the number of vectors $v_{i_k}^{(k)}$ in any W_k with $k \ge j$, which equals $\sum_{k \ge j} (\dim W_k)/d$. Iterated application of Problem 33 gives

$$\begin{aligned} \dim K_{j+1} - \dim K_j &= \dim U_j + \dim W_j = \dim U_{j+1} + \dim W_{j+1} + \dim W_j \\ &= \cdots = \textstyle\sum_{k \ge j} \dim W_k, \end{aligned}$$

and the result follows.

40. The minimal polynomial for any cyclic subspace must divide the minimal polynomial for V and hence must be a power of $P(X)$. Problem 28 shows that the restrictions of L to any two cyclic subspaces with the same minimal polynomial are isomorphic. Hence the decomposition into cyclic subspaces will be unique up to isomorphism as soon as it is proved that the number of cyclic direct summands with minimal polynomial of the form $P(X)^k$ with $k \ge j+1$ equals $(\dim K_{j+1} - \dim K_j)/d$.

Suppose that V is the direct sum of cyclic subspaces C_i, with v_i as the generator of C_i. Since each C_i is invariant under L, each K_r is the direct sum of the subspaces

$K_r \cap C_i$. Thus

$$\dim K_{j+1} - \dim K_j = \sum_i \big(\dim(K_{j+1} \cap C_i) - \dim(K_j \cap C_i)\big).$$

If $P(X)^k$ is the minimal polynomial of C_i, it is enough to show that the right side of this displayed formula equals d if $k \geq j+1$ and equals 0 if $k \leq j$. By Problem 35, C_i has a basis consisting of all vectors $L^r P(L)^s(v_i)$ with $0 \leq r \leq d-1$ and $0 \leq s \leq k-1$. The nonzero vectors among the $L^r P(L)^{s+j+1}(v_i)$ are still linearly independent; these are the ones with $s+j+1 < k$, i.e., $s < k-j-1$. The vectors $L^r P(L)^s(v_i)$ that are not sent to 0 by $P(L)^{j+1}$ are a basis of $K_{j+1} \cap C_i$. These are the ones with $s \geq k-j-1$. This is the full basis of C_i if $j+1 > k$, and there are $d(j+1)$ such vectors if $j+1 \leq k$. Thus

$$\dim K_{j+1} \cap C_i = \begin{cases} dk & \text{if } j+1 > k, \\ d(j+1) & \text{if } j+1 \leq k. \end{cases}$$

Similarly

$$\dim K_j \cap C_i = \begin{cases} dk & \text{if } j > k, \\ dj & \text{if } j \leq k. \end{cases}$$

Subtracting and taking the cases into account, we see that

$$\dim(K_{j+1} \cap C_i) - \dim(K_j \cap C_i) = \begin{cases} d & \text{if } j+1 \leq k, \\ 0 & \text{otherwise.} \end{cases}$$

41. (a) $\begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix}$, (b) $\begin{pmatrix} \cosh t & \sinh t \\ \sinh t & \cosh t \end{pmatrix}$, (c) the diagonal matrix with diagonal entries $e^{d_1}, \ldots, e^{d_n}$.

42. Suppose that J has diagonal entry c. Let N be the strictly upper-triangular part of J. Then $e^{tJ} = e^{tcI+tN} = e^{tc}e^{tN}$. Here $e^{tN} = I + tN + \frac{1}{2!}t^2N^2 + \cdots + \frac{1}{(n-1)!}t^{n-1}N^{n-1}$ since $N^n = 0$. The powers of N were observed to have the diagonal of 1's move one step at a time up and to the right.

43. $\dfrac{d}{dt}(e^{tA}v) = (Ae^{tA})v = A(e^{tA}v)$.

44. Suppose that $y(t)$ is a solution. The product rule for derivatives is valid in this situation by the usual derivation. Hence $\frac{d}{dt}(e^{-tA}y(t)) = \frac{d}{dt}(e^{-tA})y(t) + e^{-tA}y'(t) = -e^{-tA}Ay(t) + e^{-tA}y'(t) = e^{-tA}(-Ay(t) + y'(t))$. The right side is 0 since $y(t)$ solves the differential equation. Since $\frac{d}{dt}(e^{-tA}y(t)) = 0$, each component of $e^{-tA}y(t)$ is constant. Thus for a suitable vector v of complex constants, $e^{-tA}y(t) = v$, and the conclusion is that $y(t) = e^{tA}v$.

45. The first formula follows by making a term-by-term calculation with the defining series. Multiplication of C has to be interchanged with the infinite sum, and

similarly for C^{-1}, but these operations are simply the operations of taking certain linear combinations of limits.

Suppose that $z(t)$ satisfies $\frac{d}{dt}z(t) = (C^{-1}AC)z(t)$ and $z(0) = u$. Multiplying by C gives $\frac{d}{dt}Cz(t) = ACz(t)$. Thus $y(t) = Cz(t)$ satisfies $\frac{d}{dt}y(t) = Ay(t)$ and $y(0) = Cz(0) = Cu$. We can invert the correspondence by using C^{-1}.

46. Example 3 in Section 7 says that $C^{-1}AC = J$ holds for $J = \begin{pmatrix} 3 & 1 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{pmatrix}$ and $C = \begin{pmatrix} -1 & 1 & 0 \\ -1 & 0 & 0 \\ -1 & 0 & 1 \end{pmatrix}$. Define $u = C^{-1}\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 0 & -1 & 0 \\ 1 & -1 & 0 \\ 0 & -1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} -2 \\ -1 \\ 1 \end{pmatrix}$. Problems 42–43 show that the unique solution of $\frac{d}{dt}z(t) = Jz(t)$ with $z(0) = u$ is $z(t) = e^{tJ}u$. Problem 45 shows that the unique solution to $\frac{d}{dt}y(t) = Ay(t)$ with $y(0) = Cu = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ is $y(t) = Cz(t) = Ce^{tJ}u$. By Problem 42, this is

$$\begin{aligned} y(t) &= \begin{pmatrix} -1 & 1 & 0 \\ -1 & 0 & 0 \\ -1 & 0 & 1 \end{pmatrix}\begin{pmatrix} e^{3t} & 0 & 0 \\ 0 & e^{3t} & 0 \\ 0 & 0 & e^{2t} \end{pmatrix}\begin{pmatrix} 1 & t & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} -2 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 & 1 & 0 \\ -1 & 0 & 0 \\ -1 & 0 & 1 \end{pmatrix}\begin{pmatrix} e^{3t} & te^{3t} & 0 \\ 0 & e^{3t} & 0 \\ 0 & 0 & e^{2t} \end{pmatrix}\begin{pmatrix} -2 \\ -1 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} -e^{3t} & -te^{3t}+e^{3t} & 0 \\ -e^{3t} & -te^{3t} & 0 \\ -e^{3t} & -te^{3t} & e^{2t} \end{pmatrix}\begin{pmatrix} -2 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} e^{3t}+te^{3t} \\ 2e^{3t}+te^{3t} \\ 2e^{3t}+te^{3t}+e^{2t} \end{pmatrix}. \end{aligned}$$

Chapter VI

1. In (a), the linear function $\varphi : V \to V'$ given by $\varphi(v) = \langle v, \cdot \rangle$ has kernel equal to the left radical of the bilinear form, hence 0. Therefore φ is one-one, and $\dim \text{image}\,\varphi = \dim V = \dim V'$. Since $\dim V' < \infty$, φ is onto V'. In (b), $v \mapsto (v, \cdot)$ is a linear functional and by (a) is of the form $(v, u) = \langle w, u \rangle$ for some unique w depending on v. Set $w = L(v)$. The uniqueness shows that $L(v_1 + v_2) = L(v_1) + L(v_1)$ and $L(cv) = cL(v)$. Hence L is linear.

2. Since M^tAM would have to be nonsingular, the only possibility would be M^tAM equal to the identity. Writing M^{-1} as $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$, we obtain the conditions $a + c = b + d = 0$ and $ab + cd = 1$. A check of cases shows that these have no solution.

3. Take $M = \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix}$.

5. Define $(a + bi)w = aw + bJ(w)$ for a and b real. The crucial property to show in order to obtain a complex vector space is that $((a + bi)(c + di))(w) = (a + bi)((c + di)w)$; expansion of both sides shows that both sides are equal to $(ac - bd)w + (bc + ad)J(w)$ since $J^2 = -I$. Thus $W = V_{\mathbb{R}}$ for a suitable V.

Next define $(v, w) = \langle J(v), w \rangle + i\langle v, w \rangle$. This is bilinear over $\mathbb{R}$. It is complex linear in the first variable because $(J(v), w) = \langle J^2(v), w \rangle + i\langle J(v), w \rangle = -\langle v, w \rangle +$

$i\langle J(v), w\rangle = i(v, w)$. It is Hermitian because $\overline{(w, v)} = \langle J(w), v\rangle - i\langle v, w\rangle = \langle J^2(w), J(v)\rangle - i\langle w, v\rangle = -\langle w, J(v)\rangle - i\langle w, v\rangle = \langle J(v), w\rangle + i\langle v, w\rangle = (v, w)$.

6. For (a), U isotropic implies $U^\perp \supseteq U$. If v is a vector in $U^\perp$ but not U, then $U \oplus \mathbb{K}v$ is isotropic. Maximality thus implies that $U^\perp = U$. Proposition 6.3 says that $\dim V = \dim U + \dim U^\perp$, and we conclude that $\dim V = 2\dim U$. So $\dim U = n$.

The proof of (b) goes by induction on the dimension, the base case being dimension 2, where there is no problem. Assuming the result for spaces of dimension less than $\dim V$, let S_1 be maximal isotropic in V, so that $\dim S_1 = \frac{1}{2}\dim V$ by (a). Fix a basis $\{v_1, \ldots, v_n\}$ of S_1. Choose u_1 with $\langle v_1, u_1\rangle = 1$; this exists by nondegeneracy. Put $U = \mathbb{K}v_1 \oplus \mathbb{K}u_1$. Then $\langle\,\cdot\,,\,\cdot\,\rangle\big|_{U\times U}$ is evidently nondegenerate, and Corollary 6.4 shows that $V = U \oplus U^\perp$. Certainly $S_1 \cap U^\perp$ is an isotropic subspace of $U^\perp$. It contains the $n-1$ linearly independent elements $v_j - \langle v_j, u_1\rangle v_1$ for $2 \le j \le n$ and hence has dimension $\ge n-1$. Therefore it is maximal isotropic. By induction, there is a maximal isotropic subspace T of $U^\perp$ with $(S_1 \cap U^\perp) \cap T = 0$. Put $S_2 = T \oplus \mathbb{K}u_1$. Since $\langle u_1, U^\perp\rangle = 0$, $\langle u_1, T\rangle = 0$. Therefore S_2 is isotropic, hence maximal isotropic in V. Suppose that the element $t + cu_1$ of S_2 lies in S_1. From $\langle v_1, t + cu_1\rangle = 0$, $v_1 \in U$, $t \in U^\perp$, and $\langle v_1, u_1\rangle = 1$, we obtain $c = 0$. Then $t + cv_1$ lies in $(S_1 \cap U^\perp) \cap T$, which is 0. We conclude that $S_1 \cap S_2 = 0$.

For (c), if $\langle\,\cdot\,, s_2\rangle$ is the 0 function on S_1, then the fact that S_1 is maximal isotropic implies that $s_2 = 0$. Therefore the mapping $s_2 \mapsto \langle\,\cdot\,, s_2\rangle\big|_{S_1}$ is one-one. A count of dimensions shows that it is onto S_1'.

In (d), choose any basis $\{p_1, \ldots, p_n\}$ of S_1, and let $\{q_1, \ldots, q_n\}$ be the dual basis of S_1', which has been identified with S_2 by (c).

7. In (a), first suppose that $h : \bigoplus_s U_s \to V$ is given. Then hi_s is in $\mathrm{Hom}_{\mathbb{K}}(U_s, V)$, and the map from left to right may be taken to be $h \mapsto \{hi_s\}_{s\in S}$. Next suppose that $h_s : U_s \to V$ is given for each s. Then the universal mapping property of $\bigoplus_s U_s$ supplies $h : \bigoplus_s U_s \to V$ with $hi_s = h_s$ for all s. The map from right to left may be taken as $\{h_s\}_{s\in S} \mapsto h$. These two maps invert each other.

In (b), first suppose that $h_s : U \to V_s$ is given for each s. Then the universal mapping property of the direct product produces $h : U \to \prod_s V_s$. The map from right to left may be taken as $\{h_s\}_{s\in S} \mapsto h$. Next suppose that $h : U \to \prod_s V_s$ is given. Then $p_s h$ is in $\mathrm{Hom}_{\mathbb{K}}(U, V_s)$ for each $s \in S$. Consequently the S-tuple $\{h_s\}_{s\in S}$ is in $\prod_s \mathrm{Hom}_{\mathbb{K}}(U, V_s)$. Then the map from left to right can be taken as $h \mapsto \{p_s h\}_{s\in S}$. These two maps invert each other.

For (c), we treat (a) and (b) separately. In the case of (a), take S countably infinite with each $U_s = \mathbb{K}$ and with $V = \mathbb{K}$. Then $\mathrm{Hom}_{\mathbb{K}}(\bigoplus_{s\in S} U_s, V)$ has uncountable dimension and $\bigoplus_{s\in S} \mathrm{Hom}_{\mathbb{K}}(U_s, V)$ has countable dimension.

In the case of (b), take S to be countably infinite with each $V_s = \mathbb{K}$ and with $U = \bigoplus_{s\in S} V_s$. Each member of $\mathrm{Hom}_{\mathbb{K}}(U, V_{s_0})$ has its values in V_{s_0}, and hence each member of $\bigoplus_s \mathrm{Hom}_{\mathbb{K}}(U, V_s)$ has its values in finitely many V_s. On the other hand, the identity function from U into $\bigoplus_s V_s$ is in $\mathrm{Hom}_{\mathbb{K}}(U, \bigoplus V_s)$ and takes values in all V_s's.

8. For (a), we have $g_1(g_2(x)) = g_1(g_2 x g_2^t) = g_1 g_2 x g_2^t g_1^t = (g_1 g_2) x (g_1 g_2)^t = (g_1 g_2)(x)$. If x is alternating, then $(gxg^t)^t = gx^t g^t = -gxg^t$, and $(gxg^t)_{ii} = \sum_{j,k} g_{ij} x_{jk} g_{ik} = \sum_{j<k} g_{ij} x_{jk} g_{ik} + \sum_{j>k} g_{ij} x_{jk} g_{ik} = \sum_{j<k} g_{ij}(x_{jk} - x_{jk}) g_{ik} = 0$; hence gxg^t is alternating. If x is symmetric, then $(gxg^t)^t = gx^t g^t = gxg^t$, and gxg^t is symmetric.

For (b), certainly x and gxg^t have the same rank if g is nonsingular. Theorem 6.7 shows that an alternating matrix x can be transformed by some nonsingular g to a matrix gxg^t that is block diagonal with k blocks of the form $\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$, where $2k$ is the rank, followed by 0's down the diagonal. This proves that any two alternating matrices of the same rank lie in the same orbit. It also gives an example of a matrix in each orbit.

For (c), certainly x and gxg^t have the same rank if g is nonsingular. The Principal Axis Theorem (Theorem 6.5) shows that any symmetric matrix over $\mathbb{C}$ can be transformed by some nonsingular g to a matrix gxg^t that is diagonal, say with diagonal entries $d_1, \dots, d_n$. We may assume that $d_1, \dots, d_k$ are nonzero and the others are 0. Taking h to be the diagonal matrix with diagonal entries $(d_1^{-1/2}, \dots, d_k^{-1/2}, 0, \dots, 0)$ and forming $h(gxg^t)h^t$, we obtain a diagonal matrix in the same orbit whose first k diagonal entries are 1 and whose other diagonal entries are 0. As k varies, these matrices have different ranks and hence lie in different orbits. They provide examples of matrices in each orbit.

9. In (a), the formula is $T_{UV}\big(\sum_i (u_i' \otimes v_i)\big)(u) = \sum_i u_i'(u) v_i$, and we may assume that $\{v_i\}$ is linearly independent. If this is 0 for all u, then the linear independence of the v_i's implies that $u_i'(u) = 0$ for all i and all u. Then all u_i' are 0, and hence $\sum_i (u_i' \otimes v_i) = 0$. Thus T_{UV} is one-one.

In (b), Problem 7a shows that it is enough to handle $U = \mathbb{K}$. Thus we are to show that $\mathbb{K}' \otimes_{\mathbb{K}} V$ maps onto $\mathrm{Hom}_{\mathbb{K}}(\mathbb{K}, V) \cong V$. One member of $\mathbb{K}'$ is the identity function $1'$ on $\mathbb{K}$, and $1' \otimes V$ certainly maps onto V.

For (c), if $U = V$ and if $\dim U$ is infinite, every member of the image of T_{UU} has finite rank, but $\mathrm{Hom}_{\mathbb{K}}(U, U)$ contains the identity function, which has infinite rank.

In (d), let $L : U_1 \to U$ and $M : V \to V_1$ be given, so that $F(L, M)$ carrying $(U' \otimes_{\mathbb{K}} V)$ to $(U_1' \otimes_{\mathbb{K}} V_1)$ is given by $F(L, M)(u' \otimes v) = L^t(u') \otimes M(v)$ and $G(L, M)$ carrying $\mathrm{Hom}_{\mathbb{K}}(U, V)$ to $\mathrm{Hom}_{\mathbb{K}}(U_1, V_1)$ has $(G(L, M)(\varphi))(u_1) = M(\varphi(L(u_1))$ Then

$$\begin{aligned} T_{U_1 V_1} F(L, M)(u' \otimes v)(u_1) &= T_{U_1 V_1}\big(L^t(u') \otimes M(v)\big)(u_1) \\ &= L^t(u')(u_1) M(v) = u'(L(u_1)) M(v), \\ G(L, M) T_{UV}(u' \otimes v)(u_1) &= M((T_{UV}(u' \otimes v))(L(u_1))) \\ &= M(u'(L(u_1))v) = u'(L(u_1)) M(v). \end{aligned}$$

The right sides are equal, and hence $\{T_{UV}\}$ is a natural transformation.

In (e), the answer is no because the maps T_{UV} need not be isomorphisms, according to (c).

10. To see that $\Psi(E)$ is a vector space, one has to verify that $(l+l')\varphi = l\varphi + l'\varphi$, $l(\varphi+\varphi') = l\varphi + l\varphi'$, and $(ll')\varphi = l(l'\varphi)$, and these are all routine. If μ is in $\mathrm{Hom}_{\mathbb{K}}(E, F)$, then $\Psi(\mu) : \mathrm{Hom}_{\mathbb{K}}(\mathbb{L}, E) \to \mathrm{Hom}_{\mathbb{K}}(\mathbb{L}, F)$ has to be given by left-by-μ, and the key step is to show that $\Psi(\mu)$ is $\mathbb{L}$ linear, not merely $\mathbb{K}$ linear. For φ in $\mathrm{Hom}_{\mathbb{K}}(\mathbb{L}, E)$ and l, l' in $\mathbb{L}$, we have $(\Psi(\mu)(l\varphi))(l') = \mu((l\varphi)(l')) = \mu(\varphi(ll')) = (\Psi(\mu)\varphi)(ll') = (l(\Psi(\mu)\varphi))(l')$. Hence $\Psi(\mu)(l\varphi) = l(\Psi(\mu)\varphi)$ as required. It is routine to check that $\Psi(1) = 1$ and that $\mu \to \Psi(\mu)$ respects compositions, and hence Ψ is a functor.

11. Let $\Gamma = (v_1, \dots, v_n)$ be an ordered basis of E, $\Delta = (w_1, \dots, w_m)$ be an ordered basis of F, and $A = [A_{ij}]$ be the matrix of L in these ordered bases. Put $\Gamma_{\mathbb{R}} = (v_1, iv_1, \dots, v_n, iv_n)$ and $\Delta_{\mathbb{R}} = (w_1, iw_1, \dots, w_m, iw_m)$. Then the matrix of $L_{\mathbb{R}}$ in these ordered bases is obtained by replacing A_{ij} by the 2-by-2 block $\begin{pmatrix} \mathrm{Re}\, A_{ij} & -\mathrm{Im}\, A_{ij} \\ \mathrm{Im}\, A_{ij} & \mathrm{Re}\, A_{ij} \end{pmatrix}$.

12. Let $\Gamma_1 = (u_1, \dots, u_m)$ and $\Delta_1 = (v_1, \dots, v_n)$, and put

$$\Omega_1 = (u_1 \otimes v_1, u_1 \otimes v_2, \dots, u_1 \otimes v_n, u_2 \otimes v_1, \dots, u_2 \otimes v_n, \dots, u_m \otimes v_n).$$

Form Ω_2 from the ordered bases Γ_2 and Δ_2 similarly. Members of Ω_1 are indexed by pairs (i, j) with $1 \le i \le m$ and $1 \le j \le n$, and members of Ω_2 are indexed similarly by pairs (r, s). Then $C_{(r,s),(i,j)} = A_{ri} B_{sj}$.

13. Define F to be the vector space $\mathbb{K}U \oplus \mathbb{K}V$, and let l be the linear map $l : F \to T(E)$ given by $l(U) = Y$ and $l(V) = X^2 + XY + Y^2$. Let L be the extension of l to an algebra homomorphism $L : T(F) \to T(E)$ with $L(1) = 1$. The subalgebra in question is the image of L, and the affirmative answer to the question comes by showing that L is one-one. It is enough to show that the basis elements consisting of all iterated products $U^{i_1} \otimes V^{j_1} \otimes U^{i_2} \otimes \cdots \otimes V^{j_n}$ are carried by L to linearly independent elements. The image of this element is homogeneous of degree $\sum_{k=1}^{n}(i_k + 2j_k)$, and it is enough to consider only those images with the same homogeneity, i.e., with $\sum_{k=1}^{n}(i_k + 2j_k)$ constant. A failure of linear independence would mean that among these, the ones with the highest total power of X, namely with $\sum_{k=1}^{n} 2j_k$ maximal, must cancel together. These terms are monomials with $\sum i_k$ factors of Y and $\sum j_k$ factors of X^2, and all such monomials, being also monomials in X and Y, are linearly independent.

14. Let $\iota_E : E \to S(E)$ be the one-one linear map that embeds E as $S^1(E) \subseteq S(E)$, and define ι_F similarly. The composition $\iota_F \varphi$ is a linear map of E into the commutative associative algebra $S(F)$, and Proposition 6.23b yields a homomorphism $\Phi : S(E) \to S(F)$ of algebras with identity such that $\iota_F \varphi = \Phi\, \iota_E$. We take Φ as $S(\varphi)$, and this addresses (a). Part (c) is part of the construction of $S(\varphi)$. For (b), it is plain that $S(1_E) = 1_{S(E)}$. For compositions, suppose that $\psi : F \to G$ is linear and that $S(\psi)$ is formed similarly. Proposition 6.23b says that $S(\psi\varphi)$ is the unique homomorphism of $S(E)$ into $S(G)$ carrying 1 into 1 and satisfying $\iota_G \psi\varphi = S(\psi\varphi)\, \iota_E$. On the other hand, $S(\psi)S(\varphi)$ is another homomorphism of $S(E)$ into $S(G)$ carrying 1 into 1, and

it satisfies $\iota_G(\psi\varphi) = (\iota_G\psi)\varphi = (S(\psi)\,\iota_F)\varphi = S(\psi)(\iota_F\varphi) = S(\psi)(S(\varphi)\,\iota_E) = (S(\psi)S(\varphi))\iota_E$. Therefore $S(\psi\varphi) = S(\psi)S(\varphi)$ by uniqueness, and S is a functor.

15. The homomorphism $\widetilde{\Phi}$ carries each $T^n(E)$ into itself. Since $\widetilde{\Phi}$ carries commutators into commutators, $\widetilde{\Phi}(I) \subseteq I$. Thus $\widetilde{\Phi}(T^n(E) \cap I) \subseteq T^n(E) \cap I$. Also, $\widetilde{\Phi}$ commutes with the symmetrizer operator and hence carries $\widetilde{S}^n(E)$ into itself. We are given the equation $q\widetilde{\Phi}(x) = \Phi q(x)$ on all of $T^n(E)$. Since $\widetilde{\Phi}$ carries $\widetilde{S}^n(E)$ into itself, we can interpret this as saying that $\widetilde{\Phi}|_{\widetilde{S}^n(E)}$ is well defined, and then all the assertions in the problem have been addressed.

16. Fix an ordered basis and check the result directly for L's that correspond to elementary matrices. The determinant and the scalar effect on $\bigwedge^{\dim E}(E)$ both multiply under composition, and the result follows.

17. Part (a) is a consequence of uniqueness. The formula for (b) is $\Phi(g)P(v) = \Phi(g^{-1}v)$ for v in $\mathbb{K}^n$.

18. For (a), take $\mathcal{A}$ to be the category of commutative associative algebras over $\mathbb{K}$ with identity, $\mathcal{V}$ to be the category of vector spaces over $\mathbb{K}$, and $\mathcal{F} : \mathcal{A} \to \mathcal{V}$ to be the forgetful functor that takes an algebra and retains only the vector-space structure. If a vector space E is given, then (S, ι) is taken to be $(S(E), \iota_E)$, where $S(E)$ is the symmetric algebra of E and $\iota_E : E \to \mathcal{F}(S(E))$ is the identification of E with the first-order symmetric tensors.

For (b), take $\mathcal{V}$ again to be the category of vector spaces over $\mathbb{K}$. Define $\mathcal{A}$ to be the category whose objects are pairs (A, F) in which A is an associative algebra over $\mathbb{K}$ with identity and F is a vector subspace of A such that every element f of F has $f^2 = 0$ and whose morphisms $\varphi \in \mathrm{Morph}((A, F), (A_1, F_1))$ are algebra homomorphisms $\varphi : A \to A'$ such that $\varphi(F) \subseteq F_1$. The functor $\mathcal{F} : \mathcal{A} \to \mathcal{V}$ is to take the pair (A, F) to F and is to take the morphism φ to $\varphi\big|_F : F \to F_1$. If a vector space E is given, we take (S, ι) to be $((\bigwedge E, \bigwedge^1 E), \iota_E)$, where $\iota_E : E \to \bigwedge^1 E = \mathcal{F}(\bigwedge E, \bigwedge^1(E))$ is the identification of E with the first-order alternating tensors.

For (c), let the nonempty index set be J. Take $\mathcal{V} = \mathcal{C}^J$ and $\mathcal{A} = \mathcal{C}$. The functor $\mathcal{F} : \mathcal{C} \to \mathcal{C}^J$ is the "diagonal functor" taking an object A to the J-tuple whose j^{th} coordinate is A for every j; this functor takes any morphism $\varphi \in \mathrm{Morph}_{\mathcal{C}}(A, A')$ to the J-tuple whose j^{th} coordinate is φ for every j. The given E is to be a J-tuple of objects $\{X_j\}_{j\in J}$, S is to be the coproduct $\coprod_{j\in J} X_j$, and $\iota : \{X_j\}_{j\in J} \to \mathcal{F}(S)$ is to be the given J-tuple $\{i_j\}_{j\in J}$ of morphisms of X_j into X.

19. Let L be the unique member of $\mathrm{Morph}_{\mathcal{A}}(S, S')$ given as corresponding to ι' in $\mathrm{Morph}_{\mathcal{V}}(E, \mathcal{F}(S'))$, i.e., satisfying $\mathcal{F}(L)\iota = \iota'$. Similarly let L' be the unique member of $\mathrm{Morph}_{\mathcal{A}}(S, S')$ corresponding to ι in $\mathrm{Morph}_{\mathcal{V}}(E, \mathcal{F}(S))$, i.e., satisfying $\mathcal{F}(L')\iota' = \iota$. Then $L'L$ and 1_S are in $\mathrm{Morph}_{\mathcal{A}}(S, S)$ and have $\mathcal{F}(1_S)\iota = 1_{\mathcal{F}(S)}\iota = \iota$ and $\mathcal{F}(L'L)\iota = (\mathcal{F}(L')\mathcal{F}(L))\iota = \mathcal{F}(L')(\mathcal{F}(L)\iota) = \mathcal{F}(L')\iota' = \iota$. By uniqueness, $1_S = L'L$. Similarly $LL' = 1_{S'}$.

20. By definition, T_A satisfies $T_A(L) = \mathcal{F}(L)\iota$ for $L \in \mathrm{Morph}_{\mathcal{A}}(S, A)$. For φ in $\mathrm{Morph}_{\mathcal{A}}(A, A')$, we are to show that $G(\varphi)(T_A(L)) = T_{A'}(F(\varphi)(L))$. Sub-

stitution from the definitions gives $G(\varphi)(T_A(L)) = \mathcal{F}(\varphi)\mathcal{F}(L)\iota = \mathcal{F}(\varphi L)\iota$ and $T_{A'}(F(\varphi)(L)) = T_{A'}(\varphi L) = \mathcal{F}(\varphi L)\iota$. These are equal, and hence $\{T_A\}$ is a natural transformation. Since each T_A is one-one onto by hypothesis, the system $\{T_A\}$ is a natural isomorphism.

21. The previous problem shows that F is naturally isomorphic to G and that F' is naturally isomorphic to G. Hence F is naturally isomorphic to F'. The hypotheses of Proposition 6.16 are satisfied, and the conclusion is that the object S is isomorphic in $\mathcal{A}$ to the object S' by a specific isomorphism described in the proposition.

22. Let E and F be in $\mathrm{Obj}(\mathcal{V})$, and let φ be in $\mathrm{Morph}_{\mathcal{V}}(E, F)$. Then $\iota_F\varphi$ is in $\mathrm{Morph}_{\mathcal{V}}(E, \mathcal{F}(S(F)))$, and the universal mapping property of $(S(E), \iota_E)$ produces a unique Φ in $\mathrm{Morph}_{\mathcal{A}}(S(E), S(F))$ such that $\mathcal{F}(\Phi)\iota_E = \iota_F\varphi$. We define $S(\varphi) = \Phi$. There is no difficulty in checking that $S(1_E) = 1_{S(E)}$. Let us check that if we are given also ψ in $\mathrm{Morph}_{\mathcal{V}}(F, G)$, then $S(\psi)S(\varphi) = S(\psi\varphi)$. We know that $S(\psi\varphi)$ is the unique member of $\mathrm{Morph}_{\mathcal{A}}(S(E), S(G))$ satisfying $\iota_G\psi\varphi = \mathcal{F}(S(\psi\varphi))\,\iota_E$. On the other hand, $S(\psi)S(\varphi)$ is another member of $\mathrm{Morph}_{\mathcal{A}}(S(E), S(G))$, and it satisfies $\iota_G(\psi\varphi) = (\iota_G\psi)\varphi = (\mathcal{F}(S(\psi))\,\iota_F)\varphi = \mathcal{F}(S(\psi))(\iota_F\varphi) = \mathcal{F}(S(\psi))(\mathcal{F}(S(\varphi))\,\iota_E) = (\mathcal{F}(S(\psi))\mathcal{F}(S(\varphi)))\iota_E = \mathcal{F}(S(\psi)S(\varphi))\iota_E$. Therefore $S(\psi\varphi) = S(\psi)S(\varphi)$ by uniqueness, and S is a functor.

23. $\mathrm{Pfaff}(J) = 1$ because the only nonzero term comes from $\tau = 1$.

24. The terms in which σ contains a 1-cycle are each 0 because the diagonal entries of X are 0. The remaining terms in which σ contains some cycle of odd length will be grouped in disjoint pairs that add to 0. If such a σ is given, choose the smallest label $1, \dots, 2n$ that is moved by a cycle of odd length within σ, and let τ be that cycle. Let σ' be the product of τ^{-1} and the remaining cycles of σ. The resulting unordered pairs $\{\sigma, \sigma'\}$ are disjoint. For the indices i moved by τ, $x_{i,\sigma(i)} = x_{i,\tau(i)}$ while $x_{i,\sigma'(i)} = x_{i,\tau^{-1}(i)} = -x_{\tau^{-1}(i),i}$. Then $\prod_{\tau(i)\neq i} x_{i,\sigma(i)} = \prod_{\tau(i)\neq i} x_{i,\tau(i)}$ and we obtain $\prod_{\tau(i)\neq i} x_{i,\sigma'(i)} = \prod_{\tau(i)\neq i} x_{i,\tau^{-1}(i)} = (-1)^{\text{length }\tau} \prod_{\tau(i)\neq i} x_{\tau^{-1}(i),i} = (-1)^{\text{length }\tau} \prod_{\tau(i)\neq i} x_{i,\tau(i)} = (-1)^{\text{length }\tau} \prod_{\tau(i)\neq i} x_{i,\sigma(i)} = -\prod_{\tau(i)\neq i} x_{i,\sigma(i)}$. If $\tau(i) = i$, then $x_{i,\sigma(i)} = x_{i,\sigma'(i)}$. Thus $\prod_i x_{i,\sigma(i)} = -\prod_i x_{i,\sigma'(i)}$. Since $\operatorname{sgn}\sigma = \operatorname{sgn}\sigma'$, the terms for σ and σ' sum to 0.

25. If σ is good, let A_0 consist of the smallest index in each cycle of σ, let A be the union of all $\sigma^{2k}(A_0)$ for $k \geq 0$, and let B be the union of all $\sigma^{2k+1}(A_0)$ for all $k \geq 0$. Certainly $A \cup B = \{1, \dots, 2n\}$, $\sigma(A) = B$, and $\sigma(B) = A$. We have to prove that $A \cap B = \varnothing$. If the intersection is nonempty, we have $\sigma^{2k}(a_0) = \sigma^{2l+1}(a_0')$ for some a_0 and a_0' in A_0. Possibly by increasing l by an even multiple of the order of σ, we may assume that $l \geq k$. Then $\sigma^{2(l-k)+1}a_0' = a_0$. This says that a_0' and a_0 lie in the same cycle. Being least indices in cycles, they must be equal. Then some odd power of σ fixes a_0, and the cycle of σ whose least element is a_0 must have odd length, contradiction.

The definitions of A and B in terms of A_0 are forced by the conditions in the statement of the problem, and therefore A and B are unique.

26. Since $A \cup B = \{1, \dots, 2n\}$ and $A \cap B = \varnothing$, we have $y(\sigma)z(\sigma) = \prod_{l=1}^{2n} x_{i,\sigma(i)}$. The definitions of τ and τ' make $y(\sigma) = s(\tau)\prod_{k=1}^{n} x_{\tau(2k-1),\tau(2k)}$ and $z(\sigma) = s'(\tau')\prod_{k=1}^{n} x_{\tau'(2k-1),\tau'(2k)}$. The construction has made the integers $\tau(2k-1)$ increasing and has made the inequalities $\tau(2k-1) < \tau(2k)$ hold, and similarly for τ'. This proves the desired equality, apart from signs.

27. The previous problem shows that $(\operatorname{sgn}\sigma)\prod_{l=1}^{2n} x_{i,\sigma(i)}$ equals

$$(\operatorname{sgn}\sigma)s(\tau)s'(\tau')\prod_{k=1}^{n} x_{\tau(2k-1),\tau(2k)}\prod_{k=1}^{n} x_{\tau'(2k-1),\tau'(2k)}.$$

Thus we want to see that

$$(\operatorname{sgn}\sigma)s(\tau)s'(\tau') = (\operatorname{sgn}\tau)(\operatorname{sgn}\tau'). \tag{$*$}$$

In proving $(*)$, we retain the step in which factors x_{ij} of $y(\sigma)$ and $z(\sigma)$ are replaced by x_{ji} with a minus sign if $j < i$, but we may disregard the step in which the factors are then rearranged so that τ and τ' can be defined. In fact, this rearranging does not affect the signs of τ and τ'. The reason is that if ρ is in $\mathfrak{S}_n$ and if $\widetilde{\rho}$ in $\mathfrak{S}_{2n}$ is defined by $\widetilde{\rho}(2k-1) = 2\rho(k) - 1$ and $\widetilde{\rho}(2k) = 2\rho(k)$, then $\operatorname{sgn}\widetilde{\rho} = +1$; it is enough to check this fact when ρ is a consecutive transposition, and in this case $\widetilde{\rho}$ is the product of two transpositions and is even.

Turning to $(*)$, we first consider the case in which σ, when written as a disjoint product of cycles, takes the integers $1, \dots, 2n$ in order. In this case we compute directly that $\tau = 1$, that $s(\tau)$ involves no sign changes, and that τ' is the product of cycles of odd length, with an individual cycle of τ' permuting cyclically all but the last member of a cycle of σ. Thus τ' is even. In the adjustment of factors of $z(\sigma)$, one minus sign is introduced because of each cycle in σ and comes from the last and first indices in the cycle. Thus $s'(\tau')$ is $(-1)^p$, where p is the number of cycles in σ, and this is also the value of $\operatorname{sgn}\sigma$. Hence $(*)$ holds for this σ.

A general σ is conjugate in $\mathfrak{S}_{2n}$ to the one in the previous paragraph. Thus it is enough to show that if $(*)$ holds for σ, then it holds for $\sigma' = (a\ \ a+1)\sigma(a\ \ a+1)$. First suppose that $\sigma(a) \neq a+1$ and $\sigma(a+1) \neq a$. Then a factor of $y(\sigma)$ gets replaced with a minus sign for σ if and only if it gets replaced for σ', and similarly for $z(\sigma)$. Hence $s(\tau)$ and $s'(\tau')$ are unchanged in passing from σ to σ'. The effect on τ and τ', in view of the observation immediately after $(*)$, is to multiply each on the left by $(a\ \ a+1)$. Thus $\operatorname{sgn}\tau$ and $\operatorname{sgn}\tau'$ are each reversed. Since $\operatorname{sgn}\sigma = \operatorname{sgn}\sigma'$, $(*)$ remains valid for σ'.

Now suppose that $\sigma(a) = a+1$. We may assume that $\sigma(a+1) \neq a$ since otherwise $\sigma' = \sigma$. To fix the ideas, first suppose that a is in A. Then one factor in $y(\sigma)$ is $x_{a,a+1}$, and the corresponding factor of $y(\sigma')$ is $x_{a+1,a}$. As a result τ is unchanged under the passage from σ to σ', but the number of minus signs contributing to $s(\tau)$ is increased by 1 and $s(\tau)$ is therefore reversed. Meanwhile, τ' is left multiplied by $(a\ \ a+1)$, and $s'(\tau')$ is unchanged. Thus $(*)$ remains valid for σ'. If a instead is in

B, then the roles of τ and τ' are reversed in the above argument, but the conclusion about $(*)$ is not affected. Finally suppose that $\sigma(a+1) = a$ and $\sigma(a) \neq a+1$. Then the argument is the same except that the number of signs contributing to $s(\tau)$ or $s'(\tau')$ is decreased by 1. In any event, $(*)$ remains valid for σ'.

28. What is needed is an inverse construction that passes from the pair (τ, τ') to σ. Define $\omega \in \mathfrak{S}_{2n}$ to be the commuting product of the n transpositions $(2k-1\ \ 2k)$ for $1 \leq k \leq n$.

Assuming for the moment that we know that some index a is to be in A, we see from the definitions above that $b = \sigma(a)$ is to be given by $b = \tau(\omega(\tau^{-1}(a)))$ and b is to be in B. If, on the other hand, we know that some index b is to be in B, then $\sigma(b)$ is to be given by $\tau'(\omega(\tau'^{-1}(b)))$ and is to be in A. Thus the cycle within σ to which a belongs has to be given by applying alternately $\tau\omega\tau^{-1}$ and then $\tau'\omega\tau'^{-1}$.

The critical fact is that this cycle is necessarily even. In the contrary case we would have $\tau\omega\tau^{-1}(\tau'\omega\tau'^{-1}\tau\omega\tau^{-1})^k(a) = a$ for some k. If $k = 2l$, then this equality gives $(\tau\omega\tau^{-1}\tau'\omega\tau'^{-1})^l(\tau\omega\tau^{-1})(\tau'\omega\tau'^{-1}\tau\omega\tau^{-1})^l(a) = a$, which we can rewrite as $(\tau\omega\tau^{-1})(\tau'\omega\tau'^{-1}\tau\omega\tau^{-1})^l(a) = (\tau'\omega\tau'^{-1}\tau\omega\tau^{-1})^l(a)$; this equation is contradictory since $\tau\omega\tau^{-1}$ is a permutation that moves every index. If $k = 2l+1$, then this equality gives $(\tau\omega\tau^{-1})(\tau'\omega\tau'^{-1}\tau\omega\tau^{-1})^l(\tau'\omega\tau'^{-1})(\tau\omega\tau^{-1}\tau'\omega\tau'^{-1})^l(\tau\omega\tau^{-1})(a) = a$ and hence $(\tau'\omega\tau'^{-1})(\tau\omega\tau^{-1}\tau'\omega\tau'^{-1})^l(\tau\omega\tau^{-1})(a) = (\tau\omega\tau^{-1}\tau'\omega\tau'^{-1})^l(\tau\omega\tau^{-1})(a)$; this equation is contradictory since $\tau\omega\tau^{-1}$ is a permutation that moves every index.

What we know is that the smallest index in each cycle is to be in A. Thus we can use this process to construct σ from (τ, τ'), one cycle at a time. For the first cycle the index 1 is to be in A; for the next cycle the smallest remaining index is to be in A, and so on. We have seen that the constructed σ will be the product of even cycles, and we can define A as the union of the images of the even powers of σ on the least indices of each cycle, with B as the complement. In this way we have formed σ and its disjoint decomposition $\{1, \ldots, 2n\} = A \cup B$, and it is apparent that τ and τ' are indeed the permutations formed in the usual passage from σ to (τ, τ') via (A, B).

29. It is enough to prove that $\varphi|_{V_n} : V_n \to V_n^{\#}$ is an isomorphism for every n. We establish this property by induction on n, the trivial case for the induction being $n = -1$. Suppose that

$$\varphi|_{V_{n-1}} : V_{n-1} \to V_{n-1}^{\#} \quad \text{is an isomorphism.} \tag{$*$}$$

By assumption

$$\operatorname{gr}^n \varphi : (V_n / V_{n-1}) \to (V_n^{\#} / V_{n-1}^{\#}) \quad \text{is an isomorphism.} \tag{$**$}$$

If v is in $\ker(\varphi|_{V_n})$, then $(\operatorname{gr}^n \varphi)(v + V_{n-1}) = 0 + V_{n-1}^{\#}$, and $(**)$ shows that v is in V_{n-1}. By $(*)$, $v = 0$. Thus $\varphi|_{V_n}$ is one-one. Next suppose that $v^{\#}$ is in $V_n^{\#}$. By $(**)$ there exists v_n in V_n such that $(\operatorname{gr}^n \varphi)(v_n + V_{n-1}) = v^{\#} + V_{n-1}^{\#}$. Write $\varphi(v_n) = v^{\#} + v_{n-1}^{\#}$ with $v_{n-1}^{\#}$ in $V_{n-1}^{\#}$. By $(*)$ there exists v_{n-1} in V_{n-1} with $\varphi(v_{n-1}) = v_{n-1}^{\#}$. Then $\varphi(v_n - v_{n-1}) = v^{\#}$, and thus $\varphi|_{V_n}$ is onto. This completes the induction.

30. We define a product $(A_m/A_{m-1}) \times (A_n/A_{n-1}) \to A_{m+n}/A_{m+n-1}$ by

$$(a_m + A_{m-1})(a_n + A_{n-1}) = a_m a_n + A_{m+n-1}.$$

This is well defined since $a_m A_{n-1}$, $A_{m-1}a_n$, and $A_{m-1}A_{n-1}$ are all contained in A_{m+n-1}. It is clear that this multiplication is distributive and associative as far as it is defined. We extend the definition of multiplication to all of gr A by taking sums of products of homogeneous elements, and the result is an associative algebra. The identity is the element $1 + A_{-1}$ of A_0/A_{-1}.

31. $[x, x] = xx - xx = 0$, and also $[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = (xyz - xzy - yzx + zyx) + (yzx - yxz - zxy + xzy) + (zxy - zyx - xyz + yxz) = 0$.

32. In (a), let x and y be in $\mathfrak{g}$. Then we have

$$\begin{aligned}[x, y]^t A + A[x, y] &= (xy - yx)^t A + A(xy - yx) \\ &= y^t x^t A - x^t y^t A + Axy - Ayx \\ &= y^t(x^t A + Ax) - x^t(y^t A + Ay) + (x^t A + Ax)y - (y^t A + Ay)x = 0.\end{aligned}$$

Part (b) is the special case $A = I$.

33. Uniqueness follows from the fact that 1 and $\iota(\mathfrak{g})$ generate $U(\mathfrak{g})$. For existence let $\widetilde{L} : T(\mathfrak{g}) \to A$ be the extension given by the universal mapping property of $T(\mathfrak{g})$ in Proposition 6.22. To obtain L, we are to show that $\widetilde{L}$ annihilates the ideal I''. It is enough to consider $\widetilde{L}$ on a typical generator of I'', where we have

$$\begin{aligned}\widetilde{L}(\iota X \otimes \iota Y - \iota Y \otimes \iota X - \iota[X, Y]) &= \widetilde{L}(\iota X)\widetilde{L}(\iota Y) - \widetilde{L}(\iota Y)\widetilde{L}(\iota X) - \widetilde{L}(\iota[X, Y]) \\ &= l(X)l(Y) - l(Y)l(X) - l[X, Y] \\ &= 0.\end{aligned}$$

34. First one proves the following: if $Z_1, \dots, Z_p$ are in $\mathfrak{g}$ and σ is a permutation of $\{1, \dots, p\}$; then $(\iota Z_1) \cdots (\iota Z_p) - (\iota Z_{\sigma(1)}) \cdots (\iota Z_{\sigma(p)})$ is in $U_{p-1}(\mathfrak{g})$. In fact, it is enough to prove this statement when σ is the transposition of j with $j + 1$. In this case the statement follows from the identity $(\iota Z_j)(\iota Z_{j+1}) - (\iota Z_{j+1})(\iota Z_j) = \iota[Z_j, Z_{j+1}]$ by multiplying through on the left by $(\iota Z_1) \cdots (\iota Z_{j-1})$ and on the right by $(\iota Z_{j+2}) \cdots (\iota Z_p)$.

For the assertion in the problem, if we use *all* monomials with $\sum_m j_m \le p$, we certainly have a spanning set, since the obvious preimages in $T(\mathfrak{g})$ span $\oplus_{k \le p} T_k(\mathfrak{g})$. The result of the previous paragraph then implies inductively that the monomials with monotone increasing indices suffice.

35. We shall construct the map in the opposite direction without using the Poincaré–Birkhoff–Witt Theorem, appeal to the theorem to show that we have an isomorphism, and then compute what the map is in terms of a basis. Let $T_n(\mathfrak{g}) = \bigoplus_{k=0}^{n} T^k(\mathfrak{g})$ be the n^{th} member of the usual filtration of $T(\mathfrak{g})$. Define

$U_n(\mathfrak{g})$ to be the image in $U(\mathfrak{g})$ of $T_n(\mathfrak{g})$ under the passage $T(\mathfrak{g}) \to T(\mathfrak{g})/I''$. Form the composition

$$T_n(\mathfrak{g}) \to (T_n(\mathfrak{g}) + I'')/I'' = U_n(\mathfrak{g}) \to U_n(\mathfrak{g})/U_{n-1}(\mathfrak{g}).$$

This composition is onto and carries $T_{n-1}(\mathfrak{g})$ to 0. Since $T^n(\mathfrak{g})$ is a vector-space complement to $T_{n-1}(\mathfrak{g})$ in $T_n(\mathfrak{g})$, we obtain an onto linear map $T^n(\mathfrak{g}) \to U_n(\mathfrak{g})/U_{n-1}(\mathfrak{g})$. Taking the direct sum over n gives an onto linear map

$$\widetilde{\psi} : T(\mathfrak{g}) \to \text{gr } U(\mathfrak{g})$$

that respects the grading.

Let I be the two-sided ideal in $T(\mathfrak{g})$ such that $S(\mathfrak{g}) = T(\mathfrak{g})/I$. It is generated by all $X \otimes Y - Y \otimes X$ with X and Y in $T^1(\mathfrak{g})$. Let us show that the linear map $\widetilde{\psi} : T(\mathfrak{g}) \to \text{gr } U(\mathfrak{g})$ respects multiplication and annihilates the defining ideal I for $S(\mathfrak{g})$; then we can conclude that ψ descends to an algebra homomorphism

$$\psi : S(\mathfrak{g}) \to \text{gr } U(\mathfrak{g})$$

that respects the grading.

To do so, let x be in $T^r(\mathfrak{g})$ and let y be in $T^s(\mathfrak{g})$. Then $x + I''$ is in $U_r(\mathfrak{g})$, and we may regard $\widetilde{\psi}(x)$ as the coset $x + T_{r-1}(\mathfrak{g}) + I''$ in $U_r(\mathfrak{g})/U_{r-1}(\mathfrak{g})$, with 0 in all other coordinates of gr $U(\mathfrak{g})$ since x is homogeneous. Arguing in a similar fashion with y and xy, we obtain

$$\widetilde{\psi}(x) = x + T_{r-1}(\mathfrak{g}) + I'', \qquad \widetilde{\psi}(y) = y + T_{s-1}(\mathfrak{g}) + I'',$$
$$\text{and} \qquad \widetilde{\psi}(xy) = xy + T_{r+s-1}(\mathfrak{g}) + I''.$$

Since I'' is an ideal, $\widetilde{\psi}(x)\widetilde{\psi}(y) = \widetilde{\psi}(xy)$. General members x and y of $T(\mathfrak{g})$ are sums of homogeneous elements, and hence $\widetilde{\psi}$ respects multiplication.

Consequently $\ker \widetilde{\psi}$ is a two-sided ideal. To show that $\ker \widetilde{\psi} \supseteq I$, it is enough to show that $\ker \widetilde{\psi}$ contains all generators $X \otimes Y - Y \otimes X$. We have

$$\begin{aligned} \widetilde{\psi}(X \otimes Y - Y \otimes X) &= X \otimes Y - Y \otimes X + T_1(\mathfrak{g}) + I'' \\ &= [X, Y] + T_1(\mathfrak{g}) + I'' \\ &= T_1(\mathfrak{g}) + I'', \end{aligned}$$

and thus $\widetilde{\psi}$ maps the generator to 0. Hence $\widetilde{\psi}$ descends to a homomorphism ψ as asserted.

Finally we show that this homomorphism is an isomorphism. Let $\{X_i\}$ be an ordered basis of $\mathfrak{g}$. We know that the monomials $X_{i_1}^{j_1} \cdots X_{i_k}^{j_k}$ in $S(\mathfrak{g})$ with

$i_1 < \cdots < i_k$ and with $\sum_m j_m = n$ form a basis of $S^n(\mathfrak{g})$. Let us follow the effect of ψ on such a monomial. A preimage of this monomial in $T^n(\mathfrak{g})$ is the element

$$X_{i_1} \otimes \cdots \otimes X_{i_1} \otimes \cdots \otimes X_{i_k} \otimes \cdots \otimes X_{i_k},$$

in which there are j_m factors of X_{i_m} for $1 \le m \le k$. This element maps to the monomial in $U_n(\mathfrak{g})$ that we have denoted by $X_{i_1}^{j_1} \cdots X_{i_k}^{j_k}$, and then we pass to the quotient $U_n(\mathfrak{g})/U_{n-1}(\mathfrak{g})$. The Poincaré–Birkhoff–Witt Theorem shows that such monomials modulo $U_{n-1}(\mathfrak{g})$ form a basis of $U_n(\mathfrak{g})/U_{n-1}(\mathfrak{g})$. Consequently ψ is an isomorphism.

36. This is quite similar to Problem 33.

37. This is similar to Problem 34.

38. What is needed here is a description of a triple product of generators in terms of permuting indices and replacing repeated pairs of indices by a scalar; the description does not depend on the way that the parentheses are inserted in a triple product, and then associativity follows. The details are omitted.

39. Using the universal mapping property of Problem 36, construct an algebra homomorphism $L : \mathrm{Cliff}(E, \langle \cdot, \cdot \rangle) \to C$ carrying 1 into 1 and extending the mapping $e_i \mapsto e_i$. Since the e_i's and 1 generate C, L is onto C. Problem 37 shows that $\dim \mathrm{Cliff}(E, \langle \cdot, \cdot \rangle) \le 2^n$, and we know that $\dim C = 2^n$. Since L is onto, L must be one-one, as well as onto.

40. This is similar to Problem 35. The substitute for the Poincaré–Birkhoff–Witt Theorem is the fact established by Problem 39 that the spanning set of 2^n elements in Problem 37 is actually a basis.

41. The matrix that corresponds to X_0 has $r = -2$.

42. To see that $\widetilde{\tau}$ has the asserted properties, form the quotient map $T(H(V)) \to T(V)$ by factoring out the two-sided ideal generated by $X_0 - 1$. The composition $T(H(V)) \to W(V)$ is obtained by factoring out the two-sided ideal generated by $X_0 - 1$ and all $u \otimes v - v \otimes u - \langle u, v \rangle 1$, hence by all $u \otimes v - v \otimes u - \langle u, v \rangle X_0$ and by $X_0 - 1$. Thus $T(H(V)) \to W(V)$ factors into the standard quotient map $T(H(V)) \to U(H(V))$ followed by the quotient map of $U(H(V))$ by the ideal generated by $X_0 - 1$. By uniqueness in the universal mapping property for universal enveloping algebras, $\widetilde{\tau}$ is given by factoring out by $X_0 - 1$.

43. Let P be the extension of φ to an associative algebra homomorphism of $U(H(V))$ into A. Then $P(X_0) = 1$ since $\varphi(X_0) = 1$. The previous problem shows that P descends to $W(V)$, i.e., that there exists $\widetilde{\varphi}$ with $P = \widetilde{\varphi} \circ \widetilde{\tau}$. Restriction to V gives $\varphi = \widetilde{\varphi} \circ \iota$.

44. This is immediate from Problem 42 and the spanning in Problem 34.

46. The linear combination $L_j = \varphi(p_j) + 2\pi\varphi(q_j)$ of the two given linear mappings $\varphi(p_j) = \partial/\partial x_j$ and $\varphi(q_j) = m_j$ replaces $P(x)$ in $e^{-\pi|x|^2} P(x)$ by

$\partial P/\partial x_j$. Take a nonzero $e^{-\pi|x|^2}P(x)$ in an invariant subspace U, let $x_1^{k_1}\cdots x_n^{k_n}$ be a monomial of maximal total degree in $P(x)$, and apply $L_1^{k_1}\cdots L_n^{k_n}$ to $e^{-\pi|x|^2}P(x)$ to see that $e^{-\pi|x|^2}$ is in U. Then apply products of powers of the various m_j's to this to see that all of V is contained in U.

47. Let $r_i = p_i + 2\pi q_i$, so that $\varphi(r_i)\big(Pe^{-\pi|x|^2}\big) = (\partial P/\partial x_i)e^{-\pi|x|^2}$. It is enough to prove that no nontrivial linear combination of the members of the spanning set $q_1^{k_1}\cdots q_n^{k_n}r_1^{l_1}\cdots r_n^{l_n}$ maps to 0 under $\widetilde{\varphi}$. Let a linear combination of such terms map to 0 under $\widetilde{\varphi}$. Among all the terms that occur in the linear combination with nonzero coefficient, let $(L_1,\dots,L_n)$ be the largest tuple of exponents $(l_1,\dots,l_n)$ that occurs; here "largest" refers to the lexicographic ordering taking l_1 first, then l_2, and so on. Put $P(x_1,\dots,x_n) = x_1^{L_1}\cdots x_n^{L_n}$. If $(l_1,\dots,l_n) < (L_1,\dots,L_n)$ lexicographically, then $\widetilde{\varphi}(r_1^{l_1}\cdots r_n^{l_n})\big(Pe^{-\pi|x|^2}\big) = 0$. Thus $\widetilde{\varphi}(q_1^{k_1}\cdots q_n^{k_n}r_1^{l_1}\cdots r_n^{l_n})\big(Pe^{-\pi|x|^2}\big)$ is 0 if $(l_1,\dots,l_n) < (L_1,\dots,L_n)$ lexicographically and equals $x_1^{k_1}\cdots x_n^{k_n}L_1!\cdots L_n!e^{-\pi|x|^2}$ if $(l_1,\dots,l_n) = (L_1,\dots,L_n)$. The linear independence follows immediately.

48. This is similar to Problems 35 and 40. The key fact needed is the linear independence established in the previous problem.

52. In (a), for $[a,b,c]$ to be alternating means that $[a,a,c] = [a,b,a] = [b,a,a] = 0$. These say that $(aa)c - a(ac) = (ab)a - a(ba) = (ba)a - b(aa) = 0$. For (b), $[a,a,c] = [b,a,a] = 0$ and the 3-linearity together imply that $[a,b,a] = [a,b,a]+[b,b,a] = [a+b,b,a] = [a+b,b,a]+[a+b,a,a] = [a+b,a+b,a] = 0$.

53. For (a), $(1,0)(c,d) = (c,d)$ and $(a,b)(1,0) = (a,b)$ directly from the definition. Also, the definition $(a,b)^* = (a^*,-b)$ makes $(1,0)^* = (1,0)$, $(a,b)^{**} = (a^*,-b)^* = (a^{**},b) = (a,b)$, and $(c,d)^*(a,b)^* = (c^*,-d)(a^*,-b) = (c^*a^*-bd^*, -c^{**}b-a^*d) = ((c^*a^*-bd^*)^*, a^*d+cb)^* = (ac-db^*, a^*d+cb)^* = ((a,b)(c,d))^*$.

For (b), (c), and (d), we observe that

$$((a,b)(c,d))(e,f) = (ac\cdot e - db^*\cdot e - f\cdot d^*a + f\cdot b^*c^*,\ c^*a^*\cdot f - bd^*\cdot f + e\cdot a^*d + e\cdot cb)$$

and

$$(a,b)((c,d)(e,f)) = (a\cdot ce - a\cdot fd^* - c^*f\cdot b^* - ed\cdot b^*,\ a^*\cdot c^*f + a^*\cdot ed + ce\cdot b - fd^*\cdot b),$$

and the results are immediate.

In (e), (i) is the usual construction, and (ii) has $\mathbf{1} = (1,0)$, $\mathbf{i} = (i,0)$, $\mathbf{j} = (0,1)$, and $\mathbf{k} = (0,-i)$, with the identity of $\mathbb{H}$ written now as $\mathbf{1}$.

54. For (a), $(a,b)^*+(a,b) = (a^*,-b)+(a,b) = (a^*+a,0)$, which is a real multiple of $(1,0)$. Also, $(a,b)(a,b)^* = (a,b)(a^*,-b) = (aa^* + bb^*, a^*(-b) + a^*b) = (aa^*+bb^*,0)$, and this is a positive multiple of $(1,0)$ since aa^* and bb^* are ≥ 0 and at least one of them is positive. A similar argument applies to $(a,b)^*(a,b)$.

In (b), certainly (a, b) is bilinear over $\mathbb{R}$, the expression for (a, b) is manifestly symmetric, and we know that $(a, a) = aa^*$ is ≥ 0 with equality only for $a = 0$.

In (c), we are to prove that $(xx)y = x(xy)$ and $(yx)x = y(xx)$ in B. It is enough to prove the first identity since application of $*$ to it gives the second identity. We use $(c, d) = (a, b)$ and substitute into the displayed formulas above for Problem 53. We find that $((a, b)(a, b))(e, f)$ equals

$$(aa \cdot e - bb^* \cdot e - f \cdot b^*a - f \cdot b^*a^*, \;\; a^*a^* \cdot f - bb^* \cdot f + e \cdot a^*b + e \cdot ab)$$

and that $(a, b)((a, b)(e, f))$ equals

$$(a \cdot ae - a \cdot fb^* - a^*f \cdot b^* - eb \cdot b^*, \;\; a^* \cdot a^*f + a^* \cdot eb + ae \cdot b - fb^* \cdot b).$$

Taking into account the associativity of A, we see that it is enough to show that $(bb^*)e = e(bb^*)$, $fb^*(a + a^*) = (a + a^*)fb^*$, $(bb^*)f = f(bb^*)$, and $e(a + a^*) = (a + a^*)e$. These all follow from the fact that A is nicely normed.

55. Part (a) follows from (a) and (c) of the previous problem.

In (b), we have $(xx^*)y = (x(c1 - x))y = cxy - (xx)y = cxy - x(xy) = x(cy - xy) = x((c1 - x)y) = x(x^*y)$. The equality $x(yy^*) = (xy)y^*$ follows by applying $*$ and renaming the variables.

In (c), use of (b) and the definitions of the norm and $*$ gives $\|ab\|^2 a = ((ab)(ab)^*)a = (ab)((ab)^*a) = (ab)((b^*a^*)a) = (ab)(b^*(a^*a)) = \|a\|^2((ab)b^*) = \|a\|^2 a(bb^*) = \|a\|^2\|b\|^2 a$.

For (d), the norm equality of (c) implies that the $\mathbb{R}$ linear maps left-by-a and right-by-a are one-one, and the finite dimensionality of $\mathbb{O}$ allows us to conclude that they are onto. Hence they are invertible.

For (e), use of (b) gives $a(\|a\|^{-2}a^*b) = \|a\|^{-2}a(a^*b) = \|a\|^{-2}(aa^*)b = \|a\|^{-2}\|a\|^2 b = b$. This proves the result for left multiplication, and the argument for right multiplication is similar.

For (f), the table is as follows, with each entry representing the product of the element at the left (the row index) by the element at the top (the column index):

$(\mathbf{1}, 0)$	$(\mathbf{i}, 0)$	$(\mathbf{j}, 0)$	$(\mathbf{k}, 0)$	$(0, \mathbf{1})$	$(0, \mathbf{i})$	$(0, \mathbf{j})$	$(0, \mathbf{k})$
$(\mathbf{i}, 0)$	$-(\mathbf{1}, 0)$	$(\mathbf{k}, 0)$	$-(\mathbf{j}, 0)$	$-(0, \mathbf{i})$	$(0, \mathbf{1})$	$-(0, \mathbf{k})$	$(0, \mathbf{j})$
$(\mathbf{j}, 0)$	$-(\mathbf{k}, 0)$	$-(\mathbf{1}, 0)$	$(\mathbf{i}, 0)$	$-(0, \mathbf{j})$	$(0, \mathbf{k})$	$(0, \mathbf{1})$	$-(0, \mathbf{i})$
$(\mathbf{k}, 0)$	$(\mathbf{j}, 0)$	$-(\mathbf{i}, 0)$	$-(\mathbf{1}, 0)$	$-(0, \mathbf{k})$	$-(0, \mathbf{j})$	$(0, \mathbf{i})$	$(0, \mathbf{1})$
$(0, \mathbf{1})$	$(0, \mathbf{i})$	$(0, \mathbf{j})$	$(0, \mathbf{k})$	$-(0, \mathbf{1})$	$-(0, \mathbf{i})$	$-(0, \mathbf{j})$	$-(0, \mathbf{k})$
$(0, \mathbf{i})$	$-(\mathbf{1}, 0)$	$-(\mathbf{k}, 0)$	$(\mathbf{j}, 0)$	$(0, \mathbf{i})$	$-(0, \mathbf{1})$	$-(0, \mathbf{k})$	$(0, \mathbf{j})$
$(0, \mathbf{j})$	$(\mathbf{k}, 0)$	$-(\mathbf{1}, 0)$	$-(\mathbf{i}, 0)$	$(0, \mathbf{j})$	$(0, \mathbf{k})$	$-(0, \mathbf{1})$	$-(0, \mathbf{i})$
$(0, \mathbf{k})$	$-(\mathbf{j}, 0)$	$(\mathbf{i}, 0)$	$-(\mathbf{1}, 0)$	$(0, \mathbf{k})$	$-(0, \mathbf{j})$	$(0, \mathbf{i})$	$-(0, \mathbf{1})$

56. Although B is nicely normed, the steps of (b) in Problem 55 are not justified for it because we cannot conclude that B is alternative. Since the argument for (b) breaks down, so do the arguments for (c) and (d).

Chapter VII

1. The only integer < 60 that is not the product of powers of at most two primes is 30. Thus Burnside's Theorem assures us that the only possible order less than 60 for a nonabelian simple group is 30. The integer 30 is of the form $2pq$ with $p = 3$ and $q = 5$, and $q + 1 = 2p$. Part (b) of Problem 34 at the end of Chapter IV is applicable and shows that the group has a subgroup of index 2; subgroups of index 2 are always normal.

2. For (a) and (b), $(xyx^{-1}y^{-1})^{-1} = yxy^{-1}x^{-1}$ is a commutator, and so is $a(xyx^{-1}y^{-1})a^{-1} = (axa^{-1})(aya^{-1})(axa^{-1})^{-1}(aya^{-1})^{-1}$.

3. Let H be generated by a and b, and let K be generated by bab^2 and bab^3. Certainly $K \subseteq H$. Since bab^2 and bab^3 are in K, so is $(bab^2)^{-1}(bab^3) = b$ and then so is $(b^{-1})(bab^2)(b^{-2}) = a$. Hence $H \subseteq K$.

4. If H is characteristic, then in particular every inner automorphism $x \to gxg^{-1}$ carries H to itself, and H is normal. If $\varphi : G \to G$ is an automorphism and z is in Z_G, then the equality $\varphi(z)\varphi(g) = \varphi(zg) = \varphi(gz) = \varphi(g)\varphi(z)$ and the fact that φ is onto G show that $\varphi(z)$ is in Z_G. If $\psi : G \to G$ is an automorphism, then $\psi(xyx^{-1}y^{-1}) = \psi(x)\psi(y)(\psi(x))^{-1}(\psi(y))^{-1}$ shows that ψ carries commutators to commutators; hence ψ carries the generated subgroup G' to itself.

5. H_8, Z_{H_8}, and $\{1\}$ are characteristic. But the subgroups of order 4 are not, because, for example, there exists an automorphism of H_8 carrying $\mathbf{i}$ to $\mathbf{j}$.

6. Yes. The proof of Proposition 7.7, which takes $S = G$, gives a finite presentation.

7. In (a), $\begin{pmatrix} \sqrt{2} & 0 \\ 0 & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{2} & 0 \\ 0 & \frac{1}{\sqrt{2}} \end{pmatrix}^{-1} \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$.

In (b), we have also $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} e^s & 0 \\ 0 & e^{-s} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}^{-1} \begin{pmatrix} e^s & 0 \\ 0 & e^{-s} \end{pmatrix}^{-1} = \begin{pmatrix} e^{-2s} & 0 \\ 0 & e^{2s} \end{pmatrix}$ and $\begin{pmatrix} \frac{1}{\sqrt{2}} & 0 \\ 0 & \sqrt{2} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ r & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 \\ 0 & \sqrt{2} \end{pmatrix}^{-1} \begin{pmatrix} 1 & 0 \\ r & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 \\ r & 1 \end{pmatrix}$. Thus $\begin{pmatrix} 1 & 0 \\ r & 1 \end{pmatrix}$, $\begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$, and $\begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix}$ are in G' for $a > 0$. Since $\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ c/a & 1 \end{pmatrix} \begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix} \begin{pmatrix} 1 & b/a \\ 0 & 1 \end{pmatrix}$, the matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is in G' if $a > 0$. If $a < 0$, we have $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 & 0 \\ r & 1 \end{pmatrix} = \begin{pmatrix} a+br & b \\ c+dr & d \end{pmatrix}$; if $b \neq 0$, then $a + br > 0$ for suitable r and therefore the equality $\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a+br & b \\ c+dr & d \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -r & 1 \end{pmatrix}$ exhibits $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ as in G'. Similarly if $c \neq 0$, then $a + cr > 0$ for suitable r and hence $\begin{pmatrix} 1 & r \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a+cr & b+dr \\ c & d \end{pmatrix}$ exhibits $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ as in G'. Thus all members of G are in G' except possibly for $\begin{pmatrix} a & 0 \\ 0 & a^{-1} \end{pmatrix}$ with $a < 0$. So it is enough to prove that $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ is in G'. This follows since $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ has been shown to be in G' and has square equal to $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$.

In (c), suppose that $(xyx^{-1})y^{-1} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$. Then $xyx^{-1} = -y$. Taking the trace of both sides and using the fact that $\operatorname{Tr} xyx^{-1} = \operatorname{Tr} y$, we see that $\operatorname{Tr} y = -\operatorname{Tr} y$ and $\operatorname{Tr} y = 0$. Put $x = \begin{pmatrix} r & s \\ t & u \end{pmatrix}$ and $y = \begin{pmatrix} a & b \\ c & -a \end{pmatrix}$, and substitute into the equality $xy = -yx$. The entry-by-entry equations are $ra + sc = -ra - tb$, $rb = -ub$, $uc = -rc$, and $tb - ua = -sc + ua$. The first and fourth equations together say that $2ra = -tb - sc = -2ua$. Thus we have $(r + u)a = 0$, $(r + u)b = 0$, and $(r + u)c = 0$. Since at least one of a, b, c is nonzero, $r + u = 0$ and $x = \begin{pmatrix} r & s \\ t & -r \end{pmatrix}$. Writing out the equality $xy = -yx$, we obtain the necessary and sufficient condition

$$2ra = -sc - tb. \tag{$*$}$$

The determinant conditions are $-r^2 - st = 1$ and $-a^2 - bc = 1$. Multiplying $(*)$ by sc and substituting $st = -1 - r^2$ and $bc = -1 - a^2$, we obtain $2rsac = -s^2c^2 - (-1 - r^2)(-1 - a^2)$ and then $0 = -s^2c^2 - 2rsac - 1 - a^2 - r^2 - r^2a^2 = -(ra + sc)^2 - 1 - a^2 - r^2$, contradiction. Thus $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ is not a commutator.

8. By Proposition 7.8 the constructed group is a quotient of the group given by generators and relations. We actually have an isomorphism if each element of the group given by generators and relations is of the form $b^p a^q$ with $0 \le p \le 2$ and $0 \le q \le 8$ because the group given by generators and relations then has order ≤ 27. Right multiplication by a carries this set to itself. Right multiplication by b has $b^p a^q b = b^p b(b^{-1}a^q b) = b^{p+1}(b^{-1}ab)^q = b^p(a^4)^q = b^p a^{4q}$, and this equals a suitable element $b^{p'}a^{q'}$ with $0 \le p' \le 2$ and $0 \le q' \le 8$. Hence the group defined by generators and relations has at most 27 elements, and we have the desired isomorphism.

9. Let F_n be free on $x_1, y_1, \dots, x_n, y_n$, let $\varphi : F_n \to F_n/F_n'$ be the homomorphism of Corollary 7.5, and let $\Psi : F_n \to G_n$ be the given quotient homomorphism. Then $\ker \varphi \subseteq \ker \Psi$, and Proposition 4.11 shows that there exists a group homomorphism $\psi : G_n \to F_n/F_n'$ such that $\psi \circ \Psi = \varphi$. Since F_n/F_n' is abelian, ψ factors as $\bar{\psi} \circ q$, where $q : G_n \to G_n/G_n'$ is the quotient and $\bar{\psi} : G_n/G_n' \to F_n/F_n'$ is a homomorphism. Thus $\bar{\psi} \circ q \circ \Psi = \varphi$. Since φ is onto, $\bar{\psi}$ is onto; thus the image of $\bar{\psi}$ is isomorphic with F_n/F_n', which is free abelian of rank $2n$. The group G_n/G_n' is abelian and has a generating set of $2n$ generators, thus is a homomorphic image $\xi : A_n \to G_n/G_n'$, where A_n is free abelian with $2n$ generators. The composition $\bar{\psi} \circ \xi$ is a homomorphism from a free abelian group of rank $2n$ onto a free abelian group of rank $2n$. Taking into account the proof of Theorem 4.46, we see that $\bar{\psi} \circ \xi$ is one-one. Since ξ is onto G_n/G_n', $\bar{\psi}$ is one-one. Therefore G_n/G_n' is free abelian of rank $2n$.

10. Let F be a free group of rank n, let $q : F \to F/F'$ be the quotient homomorphism, let $x_1, \dots, x_k$ with $k < n$ be generators of F, let $\widetilde{F} = F(\{x_1, \dots, x_k\})$, and let $\Phi : \widetilde{F} \to F$ be the quotient homomorphism. The composition $q \circ \Phi$ is a homomorphism of $\widetilde{F}$ onto the abelian group F/F', and it factors through to a

homomorphism of $\widetilde{F}/\widetilde{F}'$ onto F/F'. Here the domain is abelian with k generators, and the image is free abelian with n generators, and there can be no such homomorphism.

11. For (a), we can use 1 and a. For (b), the proof of Theorem 7.10 says that we are to multiply each of these by a, b, c on the right and take the H part of the result. The H parts that are not 1 form a free basis. We have $1a = a$ and $1a\rho(a)^{-1} = 1$, $1b = ba^{-1}a$ and $1b\rho(b)^{-1} = ba^{-1}$, $1c = c = ca^{-1}a$ and $1c\rho(c)^{-1} = ca^{-1}$, $aa = a^2 1$ and $aa\rho(a^2)^{-1} = a^2$, $ab = ab1$ and $ab\rho(ab)^{-1} = ab$, and $ac = ac1$ and $ac\rho(ac)^{-1} = ac$. Thus a free basis of the generated subgroup is $\{ba^{-1}, ca^{-1}, a^2, ab, ac\}$.

12. The thing to prove, by induction on n, is that if $a_1 a_2 \cdots a_n$ is a reduced word in variables $u_0, u_1, u_2, \ldots$ and their inverses, and if we then substitute $x^k y x^{-k}$ for u_k and reduce in terms of x, y, then the reduced form involves a total of n factors of y or y^{-1}, the factor to the left of the first y or y^{-1} is x^p if $a_1 = u_p^{\pm 1}$, and the factor to the right of the last y or y^{-1} is x^{-q} if $a_n = u_q^{\pm 1}$.

13. The remarks with Proposition 7.15 show that the reduced words in $C_2 * C_2$ are all words whose terms are alternately x and y. Let H be a normal subgroup $\neq \{1\}$. Then H contains a conjugate of a nontrivial such word. Form the shortest such word $\neq 1$ in H. If the word begins and ends with x and has length > 1, we can conjugate by x and reduce the length by 2; similarly if it begins and ends with y and has length > 1, we can conjugate by y and reduce the length by 2. We conclude that the word has length 1. Then H contains x or y and is a quotient of either $\langle y;\ y^2\rangle$ or $\langle x;\ x^2\rangle$, which give C_2 and $\{1\}$.

Thus we may assume that a shortest nontrivial reduced word in H is a product $xy \cdots xy$ with $2n$ factors or a product $yx \cdots yx$ with $2n$ factors. Then G/H is a quotient of $\langle a, b;\ a^2, b^2, (ab)^n\rangle$, and we saw in an example in Section 2 that this group is D_n. We readily check that all quotients of D_n are of the form $\{1\}$, C_2, $C_2 \times C_2$, and D_m for certain values of $m \geq 3$.

14. Argument #1: When the irreducible representations are all 1-dimensional, Corollary 7.25 shows that the number of irreducible representations must be $|G|$, and Corollary 7.28 shows that the number of conjugacy classes must be $|G|$. Therefore each conjugacy class contains just one element, and G is abelian.

Argument #2: Theorem 7.24 shows that the irreducible representations separate points in G in the sense that for any pair x, y in the group, there is some irreducible R with $R(x) \neq R(y)$. When the irreducible representations are all 1-dimensional, the multiplicative characters separate points. Since every multiplicative character is trivial on the commutator subgroup, the commutator subgroup must be $\{1\}$. Then every pair x, y has $xyx^{-1}y^{-1} = 1$ and $xy = yx$.

15. This is immediate from Lemma 7.11.

16. For (a), every cochain f has the property that $mf = 0$. Hence the same thing is true of cocycles and of cohomology elements.

For (b), the cocycle condition for f says that

$$\begin{aligned}(-1)^n f(g_1,\dots,g_n) = {} & g_1(f(g_2,\dots,g_{n+1})) \\ & + \sum_{i=1}^{n-1} (-1)^i f(g_1,\dots,g_{i-1},g_ig_{i+1},g_{i+2},\dots,g_{n+1}) \\ & + (-1)^n f(g_1,\dots,g_{n-1},g_ng_{n+1}).\end{aligned}$$

Summing over g_{n+1} in G gives

$$\begin{aligned}(-1)^n|G|f(g_1,\dots,g_n) = {} & g_1(F(g_2,\dots,g_n)) \\ & + \sum_{i=1}^{n-1} (-1)^i F(g_1,\dots,g_{i-1},g_ig_{i+1},g_{i+2},\dots,g_n) \\ & + (-1)^n F(g_1,\dots,g_{n-1}).\end{aligned}$$

The right side we recognize as $(\delta_{n-1}F)(g_1,\dots,g_n)$, which is the value of a coboundary at $(g_1,\dots,g_n)$. Therefore $|G|f$ is a coboundary and becomes the 0 element in $H^2(G,N)$. Thus f, when regarded as an element of $H^2(G,N)$ has order dividing $|G|$.

17. The two parts of the previous problem show that every element of $H^2(G,N)$ is of finite order dividing both $|G|$ and $|G/N|$. Since $\mathrm{GCD}(|G|,|G/N|) = 1$, every element of $H^2(G,N)$ has order 1. Thus $H^2(G,N) = 0$, and the only extension is the semidirect product.

18. The only automorphism of C_2 is the trivial automorphism, and therefore τ is trivial. The two possibilities for G are $C_2 \times C_2$ and C_4. With $G = C_2 \times C_2$, the group E can be $C_2 \times C_2 \times C_2$ or H_8, and with $G = C_4$, E can be $C_2 \times C_4$ or C_8. For the cases $E = C_2 \times C_2 \times C_2$ and $E = C_2 \times C_4$, the extension is the direct product, and no further discussion is necessary. For the cases $E = H_8$ and $E = C_8$, the embedding of $N = C_2$ is unique, and we therefore get only one extension in each case. Thus there are exactly two inequivalent extensions for each choice of G.

19. If N embeds as a summand C_2, then the quotient E/N has one fewer summand C_2, is still the countable direct sum of copies of C_2 and C_4, and is therefore isomorphic to E. If N embeds as a 2-element subgroup of a summand C_4, then the quotient E/N has one fewer summand C_4 and one more summand C_2, is still the countable direct sum of copies of C_2 and C_4, and is therefore isomorphic to E.

The action τ has to be trivial because C_2 has only the trivial automorphism.

If an equivalence Φ of extensions were to exist, it would have to satisfy $\Phi\, i_1(x) = i_2(x)$ for the nontrivial element x of $N = C_2$. But $i_1(x)$ is an element of order 2 that is not the square of an element of order 4, while $i_2(x)$ is an element of order 2 that *is* the square of an element of order 4. Since Φ is an isomorphism, it has to carry nonsquares to nonsquares, and we cannot have $\Phi\, i_1(x) = i_2(x)$.

20. Let us write i_1 and i_2 for the inclusions of N into E_1 and E_2. For $(i_1(x), 1)$ to be in Q, $i_1(x)$ must be 1; hence x must be 1. Thus $x \mapsto (i_1(x), 1)Q$ is one-one. The image of φ is the same as the image of φ_1, which is G. Suppose that (e_1, e_2) is in $(E_1, E_2) \cap Q$. Then $\varphi_1(e_1) = \varphi_2(e_2)$ and $(e_1, e_2) = (i_1(x), i_2(x)^{-1})$ for some $x \in N$. Then $\varphi(e_1, e_2) = \varphi_1(i_1(x)) = 1$, and φ descends to the quotient.

If $(e_1, e_2)Q$ is in the kernel of the descended φ, then (e_1, e_2) is in the kernel of the original φ, and e_1 is in the kernel of φ_1. Therefore $e_1 = i_1(x)$ for some $x \in N$. Since $\varphi_2(e_2) = \varphi_1(e_1)$, e_2 is in the kernel of φ_2 and $e_2 = i_2(y)$ for some $y \in N$. The element $(i_1(y), i_2(y)^{-1})$ is in Q, and we therefore have $(i_1(x), i_2(y))Q = (i_1(x), i_2(y))(i_1(y), i_2(y)^{-1})Q = (i_1(xy), 1)Q$. Thus $(i_1(x), i_2(y))Q$ is exhibited as in the image of the embedded copy of N.

21. Since Q is normal, we have $(\bar{u}, \widetilde{u})(\bar{v}, \widetilde{v})Q = (a(u, v)\overline{uv}, b(u, v)\widetilde{uv})Q = (b(u, v), b(u, v)^{-1})(a(u, v)\overline{uv}, b(u, v)\widetilde{uv})Q = (b(u, v)a(u, v)\overline{uv}, 1\widetilde{uv})Q = (b(u, v)a(u, v), 1)(\overline{uv}, \widetilde{uv})Q)$. Thus the cocycle for $(E_1, E_2)Q$ is $\{b(u, v)a(u, v)\} = \{a(u, v)b(u, v)\}$.

22. Let $\Phi_1 : E_1 \to E_1'$ and $\Phi_2 : E_2 \to E_2'$ be isomorphisms exhibiting the equivalences of the extensions. Define $\Phi(e_1, e_2) = (\Phi(e_1), \Phi(e_2))Q'$, and check that this descends to the required isomorphism $\Phi : (E_1, E_2)/Q \to (E_1', E_2')/Q'$.

23. $\widehat{f}(\chi) = \sum_{t \in G} f(t)\overline{\chi(t)} = \sum_{\dot{t} \in G/H} \sum_{h \in H} f(t+h)\overline{\dot{\chi}(\dot{t})} = \sum_{\dot{t} \in G/H} F(\dot{t})\overline{\dot{\chi}(\dot{t})} = \widehat{F}(\dot{\chi})$.

24. Fourier inversion and Problem 23 give $F(x) = |G/H|^{-1} \sum_{\dot{\chi} \in \widehat{G/H}} \widehat{F}(\dot{\chi})\dot{\chi}(x) = |G/H|^{-1} \sum_{\dot{\chi} \in \widehat{G/H}} \widehat{f}(\chi)\dot{\chi}(x)$. Pulling back $\dot{\chi}$ to the member χ of $\widehat{G}$ with $\chi\big|_H = 1$ and substituting the definition of F, we obtain the desired result.

25. For (a), if $C = 0$, then all $a \in \mathbb{F}^n$ have $(a, 0) = 0$, and hence $C^\perp = \mathbb{F}^n$. For (b), the repetition code has $C = \{0, (1, \ldots, 1)\}$. The members a of $\mathbb{F}^n$ with $(a, (1, \ldots, 1)) = 0$ are the members of even weight, hence the members of the parity-check code. For (c), it is enough to check that $(a, c) = 0$ for each pair of members a, c of a basis of C, and this one can do by hand.

For (d), Proposition 6.3 shows that $n = \dim C + \dim C^\perp$. Since $C = C^\perp$, $\dim C = n/2$.

For (e), every member c of C is in $C^\perp$ and must in particular have $(c, c) = 0$. Therefore c has even weight.

For (f), let c and c' be in C, and write cc' for the entry-by-entry product (logical "and"). Then $\mathrm{wt}(c + c') = \mathrm{wt}(c) + \mathrm{wt}(c') - 2\mathrm{wt}(cc')$, and hence $\frac{1}{2}\mathrm{wt}(c + c') = \frac{1}{2}\mathrm{wt}(c) + \frac{1}{2}\mathrm{wt}(c') - \mathrm{wt}(cc')$. Considering this equality modulo 2 shows that it is enough to prove that $C \subseteq C^\perp$ implies that $\mathrm{wt}(cc')$ is even whenever c and c' are in C. Modulo 2, we have $\mathrm{wt}(cc') \equiv (c, c')$, and $(c, c') = 0$ since $C \subseteq C^\perp$.

26. In (a), every element of $\mathbb{F}^n$ has order at most 2, and thus χ takes only the values ± 1. Define $(a_\chi)_i$ to be 0 if $\chi(e_i) = +1$ and to be 1 if $\chi(e_i) = -1$. Then $\chi(e_i) = (-1)^{(a, e_i)}$ for each i. The two sides extend uniquely as homomorphisms of

$\mathbb{F}^n$ to $\{\pm 1\}$, and it follows that $\chi(c) = (-1)^{(a,c)}$ for all $c \in \mathbb{F}^n$. The remainder of (a) is routine.

In (b), let χ correspond to a. Then $\widehat{f}(a) = \widehat{f}(\chi) = \sum_{c\in\mathbb{F}^n} f(c)\overline{\chi(c)} = \sum_{c\in\mathbb{F}^n} f(c)(-1)^{(a,c)}$.

In (c), we have

$$\begin{aligned}\prod_i \widehat{f_i}(a_i) &= \prod_i \sum_{c_i\in\mathbb{F}} f_i(c_i)(-1)^{a_i c_i} \\ &= \sum_{c_1\in\mathbb{F}} f_1(c_1)(-1)^{a_1c_1}\cdots\sum_{c_n\in\mathbb{F}} f_n(c_n)(-1)^{a_nc_n} \\ &= \sum_{c\in\mathbb{F}^n} f_1(c_1)(-1)^{a_1c_1}\cdots f_n(c_n)(-1)^{a_nc_n} = \sum_{c\in\mathbb{F}^n} f(c)(-1)^{(a,c)} = \widehat{f}(a).\end{aligned}$$

27. In (a), $\widehat{f_0}(0) = \sum_{c_0\in\mathbb{F}} f_0(c_0)(-1)^{0c_0} = f_0(0)(+1) + f_0(1)(+1) = x + y$ and $\widehat{f_0}(1) = \sum_{c_0\in\mathbb{F}} f_0(c_0)(-1)^{1c_0} = f_0(0)(+1) + f_0(1)(-1) = x - y$.

In (b), Problem 26c gives

$$\begin{aligned}\widehat{f}(a) = \prod_{i=1}^n \widehat{f_0}(a_i) &= \Big(\prod_{i \text{ with } a_i=0}(x+y)\Big)\Big(\prod_{i \text{ with } a_i=1}(x-y)\Big) \\ &= (x+y)^{n-\mathrm{wt}(a)}(x-y)^{\mathrm{wt}(a)}.\end{aligned}$$

28. In (a), the members of $\widehat{G/H}$ lift exactly to the members ω of $\widehat{G}$ with $\omega\big|_H = 1$. Under the mapping of Problem 26a, any member χ of $\widehat{G}$ yields a unique member a_χ of $\mathbb{F}^n$ with $\chi(c) = (-1)^{(a_\chi,c)}$ for all $c \in \mathbb{F}^n$. If a_χ is in $C^\perp$, then this formula gives $\chi(c) = 1$, i.e., $\chi\big|_H = 1$. If a_χ is not in $C^\perp$, then $\chi(c_0) \neq 1$ for some $c_0 \in C$, i.e., $\chi\big|_H \neq 1$.

In (b), we apply the special case of Problem 24 mentioned in the educational note. Then the result is immediate, in view of (a).

In (c), we let $f(c) = x^{n-\mathrm{wt}(c)}y^{\mathrm{wt}(c)}$. Problem 27b says that $\widehat{f}(a) = (x+y)^{n-\mathrm{wt}(a)}(x-y)^{\mathrm{wt}(a)}$. Substituting into the formula of the previous part gives $\sum_{c\in C} x^{n-\mathrm{wt}(c)}y^{\mathrm{wt}(c)} = |C^\perp|^{-1}\sum_{a\in C^\perp}(x+y)^{n-\mathrm{wt}(a)}(x-y)^{\mathrm{wt}(a)}$, and this says that $W_C(x,y) = |C^\perp|^{-1}W_{C^\perp}(x+y, x-y)$.

In (e), parts (d) and (e) of Problem 25 show that the only monomials X^kY^l in $W_C(X,Y)$ with nonzero coefficients are those with k and l even. Therefore $W_C(X,Y)$ is invariant under the transformations $X \mapsto -X$ and $Y \mapsto -Y$. The MacWilliams identity shows that $W_C(X,Y)$, apart from a constant, is the same polynomial in $X+Y$ and $X-Y$. Therefore $W_C(X,Y)$ is invariant also under $(X+Y) \mapsto -(X+Y)$ and under $(X-Y) \mapsto -(X-Y)$. Thus $W_C(X,Y)$ is invariant under the group of symmetries of a regular octagon centered at 0 with one of its sides centered at $(1,0)$. This symmetry group is D_8.

29. The characters of G are the ones with $\chi_n(1) = \zeta_m^n$ for $0 \le n < m$. Such a character is trivial on H if and only if $\chi_n(q) = 1$, i.e., if and only if $\zeta_m^{nq} = 1$; this means that nq is a multiple of m, hence that n is a multiple of p.

The element 1 of H is the element q of G. Thus the question about the identification of the descended characters asks the value of $\chi_n(1)$ when n is a multiple jp of p. The value is $\chi_n(1) = \zeta_m^n = \zeta_{pq}^{jp} = \zeta_q^j$.

If we have computed F on G/H and want to compute $\widehat{F}$ from the definition of Fourier coefficients, we have to multiply each of the q values of F by the values of each of the q characters of G/H and then add. The number of multiplications is q^2. The actual computation of F from f involves p additions for each of the q values of $\dot{t}$, hence pq additions.

30. $\widehat{f}(\zeta_m^{jp+k}) = \sum_{i=0}^{m-1} f(i)\zeta_m^{-(jp+k)i} = \sum_{i=0}^{m-1}(f(i)\zeta_m^{-ki})\zeta_m^{-jp}$. The variant of f for the number k is then $i \mapsto f(i)\zeta_m^{-ki}$. Handling each value of k involves $m = pq$ steps to compute the variant of f and then the $q^2 + pq$ steps of Problem 29. Thus we have $q^2 + 2pq$ steps for each k, which we regard as of order $q^2 + pq$. This means $p(q^2 + pq)$ steps when all k's are counted, hence $pq(p + q)$ steps.

32. By inspection, $(\ell_{v_1}, \ell_{v_2})_{V'} = (v_1, v_2)_{\overline{V}}$ has the properties of an inner product. The definition is set up so that the linear mapping $\ell_v \mapsto v$ of V' into $\overline{V}$ preserves inner products.

33. The contragredient has $(R^c(x)\ell_v)(v') = \ell_v(R(x^{-1})v') = (R(x^{-1})v', v)_V = (v', R(x)v)_V = \ell_{R(x)v}(v')$. Hence $R^c(x)\ell_v = \ell_{R(x)v}$, and $(R^c(x)\ell_v, R^c(x)\ell'_v)_{V'} = (R(x)v, R(x)v')_{\overline{V}} = (R(x)v', R(x)v)_V = (v', v)_V = (v, v')_{\overline{V}} = (\ell_v, \ell_{v'})_{V'}$.

34. If $\{v_j\}$ is an orthonormal basis of V, then $\{\ell_{v_j}\}$ is an orthonormal basis of V' by Problem 32, and $(R^c(x)\ell_{v_j}, \ell_{v_j})_{V'} = (\ell_{R(x)v_j}, \ell_{v_j})_{V'} = (v_j, R(x)v_j)_V = \overline{(R(x)v_j, v_j)_V}$. Summing on j gives the desired equality of group characters.

35. In view of Problem 34 a necessary condition on a 1-dimensional representation for it to be equivalent to its contragredient is that it be real-valued. Hence the two nontrivial multiplicative characters of C_3 are not equivalent to their contragredients.

36. Following the notation in the discussion before Theorem 7.23, let $\rho_{ij}(x) = (R(x)u_j, u_i)$, let l be the left-regular representation, and let $\ell_v(u) = (u, v)_V$ be as above. Consider, for fixed j_0, the image of $R^c(g)\ell_{u_i}$ under the linear extension to V' of the map $E'(\ell_{u_k})(x) = (R(x)u_{j_0}, u_k)_V$. This is $E'(\ell_{\sum_k c_k u_k})(x) = E'\big(\sum_k \bar{c}_k \ell_{u_k}\big)(x)$ $= \sum_k \bar{c}_k E'(\ell_{u_k})(x) = \sum_k \bar{c}_k (R(x)u_{j_0}, u_k)_V = (R(x)u_{j_0}, \sum_k c_k u_k)_V$, and hence $E'(\ell_v)(x) = (R(x)u_{j_0}, v)_V$. Then the image of interest is

$$\begin{aligned} E'(R^c(g)\ell_{u_i})(x) &= E'(\ell_{R(g)u_i})(x) = (R(x)u_{j_0}, R(g)u_i)_V \\ &= (R(g^{-1}x)u_{j_0}, u_i)_V = (l(g)\rho_{ij_0})(x). \end{aligned}$$

Therefore l carries a column of matrix coefficients to itself and is equivalent on such a column to R^c.

37. Let $x = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ and $y = \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix}$, and let Γ be the subgroup generated by x and y. Observe that $-I = x^2$, $y^{-1} = \begin{pmatrix} -1 & -1 \\ 1 & 0 \end{pmatrix}$, and $yx = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}$ are in Γ.

Arguing by contradiction, suppose that $\Gamma \neq \mathrm{SL}(2,\mathbb{Z})$. Choose a matrix $z = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ in $\mathrm{SL}(2,\mathbb{Z})$ but not Γ such that $\max(|a|,|b|)$ is as small as possible. If $ab = 0$, then one of $|a|$ and $|b|$ is 1 and the other is 0 because the matrix has determinant 1. If $|a| = 0$, then zy^{-1} has top row $(\pm 1 \quad 0)$; so in either event we see that some member of $\mathrm{SL}(2,\mathbb{Z})$ outside Γ is of the form $\pm\left(\begin{smallmatrix} 1 & 0 \\ t & 1 \end{smallmatrix}\right)$. Since $x^2 = -I$ is in Γ and $yx = \left(\begin{smallmatrix} 1 & 0 \\ -1 & 1 \end{smallmatrix}\right)$ is in Γ, this is a contradiction.

Thus the matrix z cannot have $ab = 0$. Suppose that $ab > 0$. Then zy has top row $(-b \quad a-b)$, and zy^{-1} has top row $(-a+b \quad -a)$. The minimality of $\max(|a|,|b|)$ for z says that

$$\max(|a|,|b|) \le \max(|-b|,|a-b|) \quad \text{and} \quad \max(|a|,|b|) \le \max(|-a+b|,|-a|).$$

Now $|a-b| < \max(|a|,|b|)$ since $ab > 0$, and the only way that we can have the above inequalities is if $a = b$. In this case, zy is a member of $\mathrm{SL}(2,\mathbb{Z})$ outside Γ whose top-row entries have product 0, and we have seen that this is a contradiction.

Thus we must have $ab < 0$. Then zx has top row $(b \quad -a)$. The product of these entries is positive and the maximum of their absolute values is the same as that for z. So we are reduced to the situation in the previous paragraph, which we saw leads to a contradiction. We conclude that $\Gamma = \mathrm{SL}(2,\mathbb{Z})$.

38. In $\mathrm{PSL}(2,\mathbb{Z})$, we have $x^2 = y^3 = 1$, and Problem 37 shows that x and y generate $\mathrm{PSL}(2,\mathbb{Z})$. Proposition 7.8 therefore produces a homomorphism carrying $\langle X, Y; X^2, Y^3\rangle$ onto $\mathrm{PSL}(2,\mathbb{Z})$. Proposition 7.16 shows that $C_2 * C_3 \cong \langle X, Y; X^2, Y^3\rangle$, and the composition of these two maps yields the desired homomorphism Φ.

39. Let us drop the "mod $\pm I$" in order to simplify the notation. In (a), $yx = \left(\begin{smallmatrix} 0 & 1 \\ -1 & -1 \end{smallmatrix}\right)\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1 & 0 \\ -1 & 1 \end{smallmatrix}\right)$ and $y^{-1}x = \left(\begin{smallmatrix} -1 & -1 \\ 1 & 0 \end{smallmatrix}\right)\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right) = \left(\begin{smallmatrix} -1 & 1 \\ 0 & -1 \end{smallmatrix}\right)$. Then $zyx = \left(\begin{smallmatrix} a-b & b \\ c-d & d \end{smallmatrix}\right)$, and $\mu(zyx) = \max(|a-b|,|b|)$. If $ab \le 0$, then $|a-b| \ge |a|$ and hence $\mu(zyx) \ge \mu(z)$. Similarly $zy^{-1}x = \left(\begin{smallmatrix} -a & a-b \\ -c & c-d \end{smallmatrix}\right)$, and $\mu(zy^{-1}x) = \max(|a|,|a-b|)$. If $ab \le 0$, then $|a-b| \ge |b|$ and hence $\mu(zy^{-1}x) \ge \mu(z)$. The arguments with ν are similar.

In (b), we have $zx = \left(\begin{smallmatrix} b & -a \\ d & -c \end{smallmatrix}\right)$. Then $\mu(zx) = \max(|b|,|a|) = \mu(z)$ and $\nu(zx) = \max(|d|,|c|) = \nu(z)$.

In (c), the entries of z are limited to ± 1 and 0. We may take the first nonzero entry in the first column to be $+1$ by adjusting by $-I$ if necessary. Then the possibilities with determinant 1 are $\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} 1 & -1 \\ 0 & 1 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} 1 & -1 \\ 1 & 0 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} 1 & 0 \\ -1 & 1 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} 1 & 1 \\ -1 & 0 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 1 \end{smallmatrix}\right)$, and $\left(\begin{smallmatrix} 0 & -1 \\ 1 & -1 \end{smallmatrix}\right)$.

In (d), let us prove by induction on n that if $Z = a_1 \cdots a_n$ is reduced and ends in X, then $\Phi(Z) = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ has $ab \le 0$. The base cases of the induction are $n = 1$

and $n = 2$, where we have $Z = X$, $Z = YX$, and $Z = Y^{-1}X$; since $\Phi(Z)$ is $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} 1 & 0 \\ -1 & 1 \end{smallmatrix}\right)$, and $\left(\begin{smallmatrix} -1 & 1 \\ 0 & -1 \end{smallmatrix}\right)$ in the three cases, we have $ab \le 0$ for each. For the inductive step we pass from Z, which ends in X, to anything obtained by adjoining factors at the right in such a way that the new word is still reduced and has X at the right end. This means that Z is replaced by ZYX or by $ZY^{-1}X$. Suppose that $\Phi(Z) = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$. We are assuming that $ab \le 0$. According to the calculation in the solution of (a), the entries in the first row of $\Phi(ZYX)$ are $a - b$ and b, with product $(a-b)b = ab - b^2 \le ab \le 0$, and the entries in the first row of $\Phi(ZY^{-1}X)$ are $-a$ and $a - b$, with product $-a(a-b) = -a^2 + ab \le ab \le 0$. Thus the induction goes forward, and our assertion follows.

Now we can prove by induction that

$$\mu(\Phi(a_1 \cdots a_n)) \ge \mu(\Phi(a_1 \cdots a_{n-1})) \qquad (*)$$

if $Z = a_1 \cdots a_n = Z' a_n$ is reduced. The result is trivial for $n = 1$, and we let $n \ge 2$ be given and assume the inequality for words of length $< n$. Let a word of length $n \ge 2$ be given. If $a_n = X$, then $(*)$ is immediate from (b). If $a_n \ne X$, then $a_{n-1} = X$ and a_n is Y or Y^{-1}. Also, ZX is a reduced word. From the previous paragraph we know that the product of the entries in the first row of $\mu(\Phi(Z'))$ is ≤ 0. Applying (b) and then (a), we obtain $\mu(\Phi(Z)) = \mu(\Phi(ZX)) = \mu(\Phi(Z' a_n X)) \ge \mu(\Phi(Z'))$, and this proves $(*)$. Similar arguments apply to ν.

For (e), we are to prove that if W is a nonempty reduced word, then $\Phi(W)$ is not the identity of $\mathrm{PSL}(2, \mathbb{Z})$. Assuming the contrary, we may assume without loss of generality that W is as short as possible with this property. If $W = a_1 \cdots a_n$, and $\Phi(W)$ is the identity, then $\mu(\Phi(W)) = \mu(I) = 1$ and similarly $\nu(\Phi(W)) = 1$. By (d), we must have $\mu(\Phi(a_1 \cdots a_k)) = \nu(\Phi(a_1 \cdots a_k)) = 1$ for $1 \le k \le n$. Then, for each k with $1 \le k < n$, $\Phi(a_1 \cdots a_k)$ lies in the set of 10 matrices in (c) but is not the identity. The 10 matrices in (c) are obtained by applying Φ to the elements 1, XY, $Y^{-1}X$, XY^{-1}, XYX, YX, Y^{-1}, X, Y, and $XY^{-1}X$. The remaining words W of length 3 are YXY, YXY^{-1}, $Y^{-1}XY$, $Y^{-1}XY^{-1}$, and the ones of length 4 are $XYXY$, $XYXY^{-1}$, $XY^{-1}XY$, $XY^{-1}XY^{-1}$, $YXYX$, $YXY^{-1}X$, $Y^{-1}XYX$, $Y^{-1}XY^{-1}X$. We compute Φ directly on these 12 reduced words and obtain $\left(\begin{smallmatrix} 0 & 1 \\ -1 & -2 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} -1 & -1 \\ 2 & 1 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} -1 & -2 \\ 1 & 1 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} 2 & 1 \\ -1 & 0 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} 1 & 2 \\ 0 & 1 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} 2 & 1 \\ 1 & 1 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} 1 & 1 \\ 1 & 2 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} 1 & 0 \\ 2 & 1 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} 1 & 0 \\ -2 & 1 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} 1 & -1 \\ -1 & 2 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} 2 & -1 \\ -1 & 1 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} 1 & -2 \\ 0 & 1 \end{smallmatrix}\right)$. Consequently $\Phi(W)$ is not the identity for W of positive length ≤ 4. The inequality of (d) shows that $\mu(\Phi(W)) \ge 2$ if W has length > 4, and therefore $\Phi(W)$ is the identity only if W is the empty word.

40. The definition of $\widetilde{\sigma}_m$ is $\widetilde{\sigma}_m \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) = \left(\begin{smallmatrix} a+m\mathbb{Z} & b+m\mathbb{Z} \\ c+m\mathbb{Z} & d+m\mathbb{Z} \end{smallmatrix}\right)$. We readily check that $\widetilde{\sigma}_m$ respects multiplication and hence is a homomorphism into some group of matrices. Since $(a + m\mathbb{Z})(d + m\mathbb{Z}) - (b + m\mathbb{Z})(c + m\mathbb{Z}) = (ad - bc) + m\mathbb{Z} = 1 + m\mathbb{Z}$, the image group is contained in $\mathrm{SL}(2, \mathbb{Z}/m\mathbb{Z})$. The kernel is the set of matrices $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$

in $\mathrm{SL}(2,\mathbb{Z})$ with $a+m\mathbb{Z}=1+m\mathbb{Z}$, $b+m\mathbb{Z}=0+m\mathbb{Z}$, $c+m\mathbb{Z}=0+m\mathbb{Z}$, $d+m\mathbb{Z}=1+m\mathbb{Z}$, and these are exactly the matrices M in $\mathrm{SL}(2,\mathbb{Z})$ with every entry of $M-I$ divisible by m. Therefore $\ker\widetilde{\sigma}_m=\Gamma(m)$. This proves (a).

In (b), let $\gamma=\mathrm{GCD}(\alpha,m)$, so that $\alpha\gamma^{-1}$ and $m\gamma^{-1}$ are relatively prime. Applying Dirichlet's theorem on primes in arithmetic progressions, take $p>|\beta|$ to be a prime of the form $p=\alpha\gamma^{-1}+rm\gamma^{-1}$ for some r. Then $\alpha+rm=p\gamma$, and $\mathrm{GCD}(\alpha+rm,\beta)=\mathrm{GCD}(p\gamma,\beta)=\mathrm{GCD}(\gamma,\beta)=\mathrm{GCD}(\mathrm{GCD}(\alpha,m),\beta)=\mathrm{GCD}(\alpha,\beta,m)=1$.

For (c), corresponding to any member of $\mathrm{SL}(2,\mathbb{Z}/m\mathbb{Z})$ is a matrix $\left(\begin{smallmatrix}a&b\\c&d\end{smallmatrix}\right)$ with integer entries with $ad-bc\equiv 1 \bmod m$. If p is a prime dividing $a-b$ and $c-d$, then $ad-bc\equiv bd-bd\equiv 0 \mod p$, and hence p does not divide m. Therefore $\mathrm{GCD}(a-b,c-d,m)=1$. Applying (b), we obtain an integer r such that $\mathrm{GCD}(a+rm-b,c-d)=1$. Let us then work instead with $\left(\begin{smallmatrix}a+rm&b\\c&d\end{smallmatrix}\right)$. Adjusting notation to call this matrix $\left(\begin{smallmatrix}a&b\\c&d\end{smallmatrix}\right)$, we may assume that $\mathrm{GCD}(a-b,c-d)=1$. Since m divides $ad-bc-1$, there exist integers C and A with

$$(a-b)C+(d-c)A=\tfrac{1-(ad-bc)}{m}.$$

Then $\det\left(\begin{smallmatrix}a+mA&b+mA\\c+mC&d+mC\end{smallmatrix}\right)$ is equal to

$$(ad-bc)+(d-c)mA+(a-b)mC=(ad-bc)+m\left(\tfrac{1-(ad-bc)}{m}\right)=1,$$

and $\left(\begin{smallmatrix}a+mA&b+mA\\c+mC&d+mC\end{smallmatrix}\right)$ is a member of $\mathrm{SL}(2,\mathbb{Z})$ whose image under $\widetilde{\sigma}_m$ is the given matrix in $\mathrm{SL}(2,\mathbb{Z}/m\mathbb{Z})$.

41. For the remainder of the problems in this set, it will be convenient to regard the isomorphism $C_2*C_3\cong\langle X,Y;\ X^2,Y^3\rangle$ of Proposition 7.16 as an equality: $C_2*C_3=\langle X,Y;\ X^2,Y^3\rangle$.

In (a), Φ_m is well defined as a consequence of the second conclusion of Proposition 7.8.

In (b), it is immediate from Proposition 7.8 that the kernel of Φ_m is the smallest normal subgroup of C_2*C_3 containing the element $(XY)^m$. Under the isomorphism $\Phi: C_2*C_3\to\mathrm{PSL}(2,\mathbb{Z})$, we have $\Phi((XY)^m)=(xy)^m \bmod \pm I$. Since the smallest normal subgroup H_m of $\mathrm{PSL}(2,\mathbb{Z})$ containing $(xy)^m \bmod \pm I=\Phi((XY)^m)$ is Φ of the smallest normal subgroup of C_2*C_3 containing $(XY)^m$, we have $H_m=\Phi(\ker\Phi_m)$.

In (c), if passage to the quotient is denoted by q_m, Proposition 4.11 shows that the point needing verification is that the scalar matrices in $\mathrm{SL}(2,\mathbb{Z})$ lie in the kernel of $q_m\circ\widetilde{\sigma}_m$, and this follows since $\left(\begin{smallmatrix}-1&0\\0&-1\end{smallmatrix}\right)$ maps under $\widetilde{\sigma}_m$ to the matrix with entries taken modulo m and then maps to the identity under q_m.

In (d), K_m is a normal subgroup of $\mathrm{PSL}(2,\mathbb{Z})$, and it is thus enough to show that the element $(xy)^m \bmod \pm I$ of H_m is in K_m. Since $\left(\begin{smallmatrix}0&-1\\1&0\end{smallmatrix}\right)\left(\begin{smallmatrix}0&1\\-1&-1\end{smallmatrix}\right)=\left(\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right)$ and since the m^{th} power of this matrix is in $\Gamma(m)$, $(xy)^m \bmod \pm I$ is indeed in K_m.

For (e), part (d) shows that $\left(\begin{smallmatrix}1 & m\\ 0 & 1\end{smallmatrix}\right)$ mod $\pm I$ is in K_m, and its t^{th} power $\left(\begin{smallmatrix}1 & tm\\ 0 & 1\end{smallmatrix}\right)$ mod $\pm I$, for t an integer, has to be in K_m. Then $\left(\begin{smallmatrix}1 & 0\\ -tm & 1\end{smallmatrix}\right) = x\left(\begin{smallmatrix}1 & tm\\ 0 & 1\end{smallmatrix}\right)x^{-1}$ mod $\pm I$ is in K_m since K_m is normal, and so are $\left(\begin{smallmatrix}1+tm & tm\\ -tm & 1-tm\end{smallmatrix}\right) = y^{-1}\left(\begin{smallmatrix}1 & tm\\ 0 & 1\end{smallmatrix}\right)y$ and $\left(\begin{smallmatrix}1-tm & tm\\ -tm & 1+tm\end{smallmatrix}\right) = xy^{-1}\left(\begin{smallmatrix}1 & tm\\ 0 & 1\end{smallmatrix}\right)yx^{-1}$, for the same reason.

42. Let x and y be the listed images in the stated permutation groups of X and Y. The homomorphisms in this problem come from Proposition 7.8 since in each case $x^2 = 1$, $y^2 = 1$, and $(xy)^m$ can be verified to be 1. What needs to be verified in each case is that x and y generate the stated permutation group.

In (a), the image group has a subgroup of order 2 and a subgroup of order 3 and hence must be the whole 6-element $\mathfrak{S}_3$.

In (b), Lemma 4.41 shows that $(1\ 2\ 3)(1\ 2)(3\ 4)(1\ 2\ 3)^{-1} = (2\ 3)(1\ 4)$, and hence the image group has a subgroup of 4 even permutations and a subgroup of 3 even permutations, therefore must be all of $\mathfrak{A}_4$.

In (c), we have $(1\ 2)(2\ 3\ 4) = (1\ 2\ 3\ 4)$. Thus the image group contains $(1\ 2\ 3\ 4)^2 = (1\ 3)(2\ 4)$, $(2\ 3\ 4)(1\ 3)(2\ 4)(2\ 3\ 4)^{-1} = (1\ 4)(2\ 3)$, and $(2\ 3\ 4)(1\ 2)(2\ 3\ 4)^{-1} = (1\ 3)$, hence a subgroup of order 8 and a subgroup of order 3. Therefore it is all of $\mathfrak{S}_4$.

In (d), we have $(1\ 2)(3\ 4)(1\ 3\ 5) = (1\ 4\ 3\ 5\ 2)$. Thus the image group contains a subgroup of order 5, a subgroup of order 3, and a subgroup of order 2, all contained in $\mathfrak{A}_5$. The image group is not of order 30 because $\mathfrak{A}_5$ has no nontrivial normal subgroups, and hence it must be all of $\mathfrak{A}_5$.

43. As with Problem 39, let us drop the " mod $\pm I$" in order to simplify the notation. In (a), we can take $g_1 = \left(\begin{smallmatrix}1 & 0\\ 0 & 1\end{smallmatrix}\right)$, $g_2 = \left(\begin{smallmatrix}1 & -1\\ 0 & 1\end{smallmatrix}\right)$, $g_3 = \left(\begin{smallmatrix}0 & 1\\ -1 & -1\end{smallmatrix}\right)$, $g_4 = \left(\begin{smallmatrix}0 & 1\\ -1 & 0\end{smallmatrix}\right)$, $g_5 = \left(\begin{smallmatrix}-1 & -1\\ 1 & 0\end{smallmatrix}\right)$, $g_6 = \left(\begin{smallmatrix}1 & 0\\ -1 & 1\end{smallmatrix}\right)$.

For (b), first we compute the six values of g_ib_1 as $g_1b_1 = \left(\begin{smallmatrix}0 & 1\\ -1 & 0\end{smallmatrix}\right)$, $g_2b_1 = \left(\begin{smallmatrix}1 & 1\\ -1 & 0\end{smallmatrix}\right)$, $g_3b_1 = \left(\begin{smallmatrix}-1 & 0\\ 1 & -1\end{smallmatrix}\right)$, $g_4b_1 = \left(\begin{smallmatrix}-1 & 0\\ 0 & -1\end{smallmatrix}\right)$, $g_5b_1 = \left(\begin{smallmatrix}1 & -1\\ 0 & 1\end{smallmatrix}\right)$, $g_6b_1 = \left(\begin{smallmatrix}0 & 1\\ -1 & -1\end{smallmatrix}\right)$, and then we compute the six values of g_ib_2 as $g_1b_2 = \left(\begin{smallmatrix}0 & 1\\ -1 & -1\end{smallmatrix}\right)$, $g_2b_2 = \left(\begin{smallmatrix}1 & 2\\ -1 & -1\end{smallmatrix}\right)$, $g_3b_2 = \left(\begin{smallmatrix}-1 & -1\\ 1 & 0\end{smallmatrix}\right)$, $g_4b_2 = \left(\begin{smallmatrix}-1 & -1\\ 0 & -1\end{smallmatrix}\right)$, $g_5b_2 = \left(\begin{smallmatrix}1 & 0\\ 0 & 1\end{smallmatrix}\right)$, $g_6b_2 = \left(\begin{smallmatrix}0 & 1\\ -1 & -2\end{smallmatrix}\right)$. Next we locate each of these products in a coset, writing them with some g_i on the right. We find that, up to mod $\pm\, I$, the results are $g_1b_1 = g_4$, $g_2b_1 = g_5$, $g_3b_1 = g_6$, $g_4b_1 = g_1$, $g_5b_1 = g_2$, $g_6b_1 = g_3$, $g_1b_2 = g_3$, $g_2b_2 = \left(\begin{smallmatrix}3 & 2\\ -2 & -1\end{smallmatrix}\right)g_6$, $g_3b_2 = g_5$, $g_4b_2 = \left(\begin{smallmatrix}1 & 2\\ 0 & 1\end{smallmatrix}\right)g_2$, $g_5b_2 = g_1$, $g_6b_2 = \left(\begin{smallmatrix}1 & 0\\ -2 & 1\end{smallmatrix}\right)g_4$. The conclusion is that generators of K_2 are the three matrices $\left(\begin{smallmatrix}3 & 2\\ -2 & -1\end{smallmatrix}\right)$, $\left(\begin{smallmatrix}1 & 2\\ 0 & 1\end{smallmatrix}\right)$, $\left(\begin{smallmatrix}1 & 0\\ -2 & 1\end{smallmatrix}\right)$.

For (c), the second and third of the generators in (b) are in H_2 by Problem 41e. The equality $\left(\begin{smallmatrix}3 & 2\\ -2 & -1\end{smallmatrix}\right) = -\left(\begin{smallmatrix}1 & -2\\ 0 & 1\end{smallmatrix}\right)\left(\begin{smallmatrix}1 & 0\\ 2 & 1\end{smallmatrix}\right)$ exhibits the first of the generators as in

H_2. Hence all the generators are in H_2 and $K_2 \subseteq H_2$. Therefore $K_2 = H_2$.

For (d) with $m = 3$, we can take the 12 coset representatives to be $g_1 = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$, $g_2 = \left(\begin{smallmatrix} 1 & -1 \\ 0 & 1 \end{smallmatrix}\right)$, $g_3 = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$, $g_4 = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$, $g_5 = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 1 \end{smallmatrix}\right)$, $g_6 = \left(\begin{smallmatrix} 0 & 1 \\ -1 & -1 \end{smallmatrix}\right)$. $g_7 = \left(\begin{smallmatrix} 1 & -1 \\ 1 & 0 \end{smallmatrix}\right)$, $g_8 = \left(\begin{smallmatrix} 1 & 1 \\ 1 & 2 \end{smallmatrix}\right)$, $g_9 = \left(\begin{smallmatrix} 1 & 1 \\ -1 & 0 \end{smallmatrix}\right)$, $g_{10} = \left(\begin{smallmatrix} 1 & -1 \\ -1 & 2 \end{smallmatrix}\right)$, $g_{11} = \left(\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right)$, $g_{12} = \left(\begin{smallmatrix} 1 & 0 \\ -1 & 1 \end{smallmatrix}\right)$. Then we compute that $g_1 b_1 = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right) = g_4$, $g_2 b_1 = \left(\begin{smallmatrix} 1 & 1 \\ -1 & 0 \end{smallmatrix}\right) = g_9$, $g_3 b_1 = \left(\begin{smallmatrix} -1 & 1 \\ -1 & 0 \end{smallmatrix}\right) = g_7$, $g_4 b_1 = \left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right) = g_1$, $g_5 b_1 = \left(\begin{smallmatrix} -1 & 0 \\ -1 & -1 \end{smallmatrix}\right) = g_{11}$, $g_6 b_1 = \left(\begin{smallmatrix} -1 & 0 \\ 1 & -1 \end{smallmatrix}\right) = g_{12}$, $g_7 b_1 = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right) = g_3$, $g_8 b_1 = \left(\begin{smallmatrix} -1 & 1 \\ -2 & 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} -1 & 0 \\ -3 & -1 \end{smallmatrix}\right) g_{10}$, $g_9 b_1 = \left(\begin{smallmatrix} -1 & 1 \\ 0 & -1 \end{smallmatrix}\right) = g_2$, $g_{10} b_1 = \left(\begin{smallmatrix} 1 & 1 \\ -2 & -1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1 & 0 \\ -3 & 1 \end{smallmatrix}\right) g_8$, $g_{11} b_1 = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 1 \end{smallmatrix}\right) = g_5$, $g_{12} b_1 = \left(\begin{smallmatrix} 0 & 1 \\ -1 & -1 \end{smallmatrix}\right) = g_6$.

Also, $g_1 b_2 = \left(\begin{smallmatrix} 0 & 1 \\ -1 & -1 \end{smallmatrix}\right) = g_6$, $g_2 b_2 = \left(\begin{smallmatrix} 1 & 2 \\ -1 & -1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 4 & 3 \\ -3 & -2 \end{smallmatrix}\right) g_{10}$, $g_3 b_2 = \left(\begin{smallmatrix} -1 & 0 \\ -1 & -1 \end{smallmatrix}\right) = g_{11}$, $g_4 b_2 = \left(\begin{smallmatrix} -1 & -1 \\ 0 & -1 \end{smallmatrix}\right) = g_3$, $g_5 b_2 = \left(\begin{smallmatrix} -1 & -1 \\ -1 & -2 \end{smallmatrix}\right) = g_8$, $g_6 b_2 = \left(\begin{smallmatrix} -1 & -1 \\ 1 & 0 \end{smallmatrix}\right) = g_9$, $g_7 b_2 = \left(\begin{smallmatrix} 1 & 2 \\ 0 & 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1 & 3 \\ 0 & 1 \end{smallmatrix}\right) g_2$, $g_8 b_2 = \left(\begin{smallmatrix} -1 & 0 \\ -2 & -1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} -1 & 0 \\ -3 & -1 \end{smallmatrix}\right) g_{12}$, $g_9 b_2 = \left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right) = g_1$, $g_{10} b_2 = \left(\begin{smallmatrix} 1 & 2 \\ -2 & -3 \end{smallmatrix}\right) = \left(\begin{smallmatrix} -2 & 3 \\ 3 & -5 \end{smallmatrix}\right) g_7$, $g_{11} b_2 = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right) = g_4$, $g_{12} b_2 = \left(\begin{smallmatrix} 0 & 1 \\ -1 & -2 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1 & 0 \\ -3 & 1 \end{smallmatrix}\right) g_5$.

Thus generators of K_3 are $\left(\begin{smallmatrix} -1 & 0 \\ -3 & -1 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} 1 & 0 \\ -3 & 1 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} 4 & 3 \\ -3 & -2 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} 1 & 3 \\ 0 & 1 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} -1 & 0 \\ -3 & -1 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} -2 & 3 \\ 3 & -5 \end{smallmatrix}\right)$, $\left(\begin{smallmatrix} 1 & 0 \\ -3 & 1 \end{smallmatrix}\right)$. All but $\left(\begin{smallmatrix} 4 & 3 \\ -3 & -2 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} -2 & 3 \\ 3 & -5 \end{smallmatrix}\right)$ are certainly in H_3. The expressions $\left(\begin{smallmatrix} 4 & 3 \\ -3 & -2 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1+3 & 3 \\ -3 & 1-3 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} -2 & 3 \\ 3 & -5 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1-3 & -3 \\ 3 & 1+3 \end{smallmatrix}\right) \left(\begin{smallmatrix} 1 & -3 \\ 0 & 1 \end{smallmatrix}\right)$ show that these two generators are in H_3. Therefore $K_3 = H_3$.

44. Problem 41 produces a homomorphism σ_m of G_m onto $\mathrm{PSL}(2, \mathbb{Z}/m\mathbb{Z})$ with kernel isomorphic to K_m/H_m. The given fact $H_m = K_m$ for $2 \le m \le 5$ implies that σ_m is an isomorphism for these values of m. This proves the first isomorphism in each part. Problem 42 gives us homomorphisms of G_m for these m's onto the third group listed in each part. Composition with σ_m^{-1} then gives a homomorphism of $\mathrm{PSL}(2, \mathbb{Z}/m\mathbb{Z})$ onto the third group. In each case the statement of Problem 43 gives the number of elements in $\mathrm{PSL}(2, \mathbb{Z}/m\mathbb{Z})$, and this matches the number of elements in the third group. It follows that these homomorphisms are isomorphisms.

45. For (a), linearity gives $R_\theta T_{(a,b)} R_\theta^{-1}(x, y) = R_\theta(R_\theta^{-1}(x, y) + (a, b)) = R_\theta R_\theta^{-1}(x, y) + R_\theta(a, b) = (x, y) + R_\theta(a, b) = T_{R_\theta(a,b)}(x, y)$.

For (b), the result of (a) says that we get a semidirect product. Let us show that the two sets—the elements of the semidirect product and the union of the translations and rotations—coincide. In one direction a rotation about (x_0, y_0) is of the form $(x, y) \mapsto R_\theta(x - x_0, y - y_0) + (x_0, y_0) = R_\theta(x, y) + (a, b) = T_{(a,b)} R_\theta(x, y)$, where $(a, b) = -R_\theta(x_0, y_0) + (x_0, y_0)$. Hence it is in the semidirect product. In the reverse direction suppose that $T_{(a,b)} R_\theta$ is in the semidirect product and is not a translation. Then θ is not a multiple of 2π, and we can put $(x_0, y_0) = (1 - R_\theta)^{-1}(a, b)$. Then we

have $T_{(a,b)}R_\theta(x, y) = R_\theta(x, y)+(a, b) = R_\theta(x-x_0, y-y_0)+R_\theta(x_0, y_0)+(a, b) = R_\theta(x - x_0, y - y_0) + R_\theta(1 - R_\theta)^{-1}(a, b) + (a, b) = R_\theta(x - x_0, y - y_0)-(1-R_\theta)(1-R_\theta)^{-1}(a, b)+(a, b)+(1-R_\theta)^{-1}(a, b) = R_\theta(x-x_0, y-y_0)+(x_0, y_0)$. Hence $T_{(a,b)}R_\theta$ is a rotation about (x_0, y_0).

46. In (a), we need to show only that $r_c = r_a r_b$. In (b), we need to show that $r_b r_a r_b r_a r_b$ is a translation but not the identity. Then it follows from (b) that the group G generated by r_a and r_b is infinite. Since (a) and Proposition 7.8 yield a homomorphism of $G_6 = \langle X, Y; X^2, Y^3, (XY)^6\rangle$ onto the infinite group G, it follows that G_6 is infinite. Since $\mathrm{PSL}(\mathbb{Z}/6\mathbb{Z})$ is finite, (c) follows.

To establish the two facts that need checking, we may, without loss of generality, take T to be the triangle with vertices $a = (0, 0)$, $b = (0, -1)$, and $c = (\sqrt{3}, 0)$. The formulas for r_a, r_b, and r_c are $r_a(x, y) = (-x, -y)$,

$$\begin{aligned} r_b(x, y) &= (x\cos\tfrac{2\pi}{3} - (y + 1)\sin\tfrac{2\pi}{3},\ x\sin\tfrac{2\pi}{3} + (y + 1)\cos\tfrac{2\pi}{3} - 1) \\ &= \big(-x/2 - y\sqrt{3}/2 - \sqrt{3}/2,\ x\sqrt{3}/2 - y/2 - 1/2 - 1\big), \end{aligned}$$

and

$$\begin{aligned} r_c(x, y) &= ((x-\sqrt{3})\cos\tfrac{\pi}{3} + y\sin\tfrac{\pi}{3} + \sqrt{3}, -(x-\sqrt{3})\sin\tfrac{\pi}{3} + y\cos\tfrac{\pi}{3}) \\ &= \big((x-\sqrt{3})/2 + y\sqrt{3}/2 + \sqrt{3},\ -(x-\sqrt{3})\sqrt{3}/2 + y/2\big). \end{aligned}$$

Then $r_a r_b(x, y) = -r_b(x, y) = r_c(x, y)$ by inspection.

To verify that $r_b r_a r_b r_a r_b$ is a translation, we write $r_b r_a r_b r_a r_b(x, y) = r_b r_c^2(x, y)$. The formula above for r_c gives

$$\begin{aligned} r_c^2(x, y) &= ((x-\sqrt{3})\cos\tfrac{2\pi}{3} + y\sin\tfrac{2\pi}{3} + \sqrt{3}, -(x-\sqrt{3})\sin\tfrac{2\pi}{3} + y\cos\tfrac{2\pi}{3}) \\ &= \big(-(x-\sqrt{3})/2 + y\sqrt{3}/2 + \sqrt{3},\ -(x-\sqrt{3})\sqrt{3}/2 - y/2\big). \end{aligned}$$

Then the first coordinate of $r_b r_c^2(x, y)$ is $-\frac{1}{2}(-(x - \sqrt{3})/2 + y\sqrt{3}/2 + \sqrt{3})+((x - \sqrt{3})\sqrt{3}/2 + y/2)\sqrt{3}/2 - \sqrt{3}/2 = x - 2\sqrt{3}$, while the second coordinate is $(-(x-\sqrt{3})/2 + y\sqrt{3}/2 + \sqrt{3})\sqrt{3}/2 + ((x-\sqrt{3})\sqrt{3}/2 + y/2)/2 - 3/2 = y$. So $r_b r_c^2(x, y) = (x - 2\sqrt{3}, y)$ is a translation.

47. We may suppose that the representations are unitary. Let $\{v_{1,i}\}$ and $\{v_{2,j}\}$ be orthonormal bases of V_1 and V_2. Then

$$\begin{aligned} (\chi_{R_1} * \chi_{R_2})(x) &= \sum_y \chi_{R_1}(xy^{-1})\chi_{R_2}(y) \\ &= \sum_{y,i,j} (R_1(xy^{-1})v_{1,i}, v_{1,i})(R_2(y)v_{2,j}, v_{2,j}) \\ &= \sum_{y,i,j,k} (R_1(x)(R_1(y^{-1})v_{1,i}, v_{1,k})v_{1,k}, v_{1,i})(R_2(y)v_{2,j}, v_{2,j}) \\ &= \sum_{i,j,k} (R_1(x)v_{1,k}, v_{1,i}) \sum_y \overline{(R_1(y)v_{1,k}, v_{1,i})}(R_2(y)v_{2,j}, v_{2,j}) \end{aligned}$$

For (a), the inside sum is 0, and the argument is complete. For (b), let $R_1 = R_2$ and $v_{2,j} = v_{1,j}$. Then the right side of the display continues as

$$\begin{aligned} &= \sum_{i,j,k} (R_1(x)v_{1,k}, v_{1,i})|G|d_{R_1}^{-1}(v_{1,j}, v_{1,k})\overline{(v_{1,j}, v_{1,i})} \\ &= |G|d_{R_1}^{-1} \sum_{i,j,k} (R_1(x)v_{1,k}, v_{1,i})\delta_{jk}\delta_{ji} \\ &= |G|d_{R_1}^{-1} \sum_{i} (R_1(x)v_{1,i}, v_{1,i}) = |G|d_{R_1}^{-1}\chi_{R_1}(x). \end{aligned}$$

48. We have $E_\alpha E_\beta = |G|^{-2}d_\alpha d_\beta R(\overline{\chi_\alpha})R(\overline{\chi_\beta}) = |G|^{-2}d_\alpha d_\beta R(\overline{\chi_\alpha * \chi_\beta})$. Problem 47a shows that this is 0 if R_α and R_β are inequivalent; this proves (b). Problem 47b shows that the computation with $R_\alpha = R_\beta$ continues as $= |G|^{-1}d_\alpha R(\overline{\chi_\alpha}) = E_\alpha$; this proves (a).

49. Let S be the set of all finite-dimensional irreducible invariant subspaces V_s of V. Call a subset T of S "independent" if the sum $\sum_{t\in T} V_t$ is direct. This condition means that for every finite subset $\{t_1, \dots, t_n\}$ of T and every set of elements $v_i \in V_{t_i}$, the equation

$$v_1 + \cdots + v_n = 0$$

implies that each v_i is 0. From this formulation it follows that the union of any increasing chain of independent subsets of S is itself independent. By Zorn's Lemma there is a maximal independent subset T_0 of S. By definition the sum $V_0 = \sum_{t\in T_0} V_t$ is direct. Consequently the problem is to show that V_0 is all of V. Since every member of V lies in a finite direct sum of finite-dimensional irreducible invariant subspaces of V, it suffices to show that each V_s is contained in V_0. If s is in T_0, this conclusion is obvious. Thus suppose s is not in T_0. By the maximality of T_0, $T_0 \cup \{s\}$ is not independent. Consequently the sum $V_0 + V_s$ is not direct, and it follows that $V_0 \cap V_s \neq 0$. But this intersection is an invariant subspace of V_s. Since V_s is irreducible, a nonzero invariant subspace must be all of V_s. Thus V_s is contained in V_0, as we wished to show.

50. Let us impose an inner product on V_0 that makes $R\big|_{V_0}$ unitary. Let $\{v_1, \dots, v_n\}$ be an orthonormal basis of V_0. If we write $R(x)v_j = \sum_{i=1}^n R_{ij}(x)v_i$, then $R_{ij}(x) = (R(x)v_j, v_i)$. Consequently the character χ_α of $R\big|_{V_0}$ is given by $\chi_\alpha(x) = \sum_i R_{ii}(x)$. Then we have

$$E_\alpha v_j = |G|^{-1}d_\alpha \sum_{x\in G} \overline{\chi_\alpha(x)}R(x)v_j = |G|^{-1}d_\alpha \sum_{x\in G}\sum_{i,k} \overline{R_{kk}(x)}R_{ij}(x)v_i = v_j,$$

and E_α is the identity on V_0.

51. Problem 49 allows us to write V as the direct sum of possibly infinitely many finite-dimensional irreducible invariant subspaces $V = \bigoplus_\gamma V_\gamma$. If any v in V is given, we can write $v = \sum_\gamma v_\gamma$ with only finitely many terms nonzero. Applying

E_α and using Problem 50, we see that $E_\alpha v$ is the sum of those v_γ such that $R\big|_{V_\gamma}$ is equivalent to R_α. Thus each nonzero v_γ has the property that $E_\alpha v_\gamma = v_\gamma$ for some α.

On the other hand, this equality cannot hold for two distinct α's. In fact, if R_α and R_β are inequivalent and we have $E_\alpha v_\gamma = v_\gamma$ and $E_\beta v_\gamma = v_\gamma$, then application of E_α to the second equality gives $E_\alpha E_\beta v_\gamma = E_\alpha v_\gamma = v_\gamma$. But $E_\alpha E_\beta = 0$ by Problem 48b, and hence $v_\gamma = 0$.

The conclusion is that for each nonzero v_γ, there is one and only one E_α such that $E_\alpha v_\gamma \neq 0$, and that α has $E_\alpha v_\gamma = v_\gamma$. Applying $\sum_\alpha E_\alpha$ to $v = \sum_\gamma v_\gamma$, we obtain $\sum_\alpha E_\alpha v = \sum_{\gamma,\alpha} E_\alpha v_\gamma = \sum_\gamma v_\gamma = v$. Thus $\sum_\alpha E_\alpha = I$. Problem 50 shows that E_α is the identity on any finite sum of vectors lying in finite-dimensional irreducible invariant subspaces equivalent to R_α. The direct-sum decomposition just proved shows that E_α is 0 on any vector in the direct sum of the images of the other E_β's. Thus the image of E_α is as asserted.

52. For α as given and for any v in V, we have $E_\alpha v = |G|^{-1} \sum_{x\in G} \overline{\omega(x)} R(x)v$. The members of the image of E_α are exactly the vectors v for which $E_\alpha v = v$, hence exactly the vectors v for which $|G|^{-1} \sum_{x\in G} \overline{\omega(x)} R(x)v = v$. Applying $R(y)$ to both sides gives $R(y)v = |G|^{-1} \sum_{x\in G} \overline{\omega(x)} R(yx)v = |G|^{-1} \sum_{x\in G} \overline{\omega(y^{-1}x)} R(x)v = \overline{\omega(y^{-1})} |G|^{-1} \sum_{x\in G} \overline{\omega(x)} R(x)v = \overline{\omega(y^{-1})} v = \omega(y)v$.

Chapter VIII

1. In (a), φ fixes 1 and must therefore fix the subfield generated by 1; this is $\mathbb{Q}$. For (b), $\varphi(a^2) = \varphi(a)^2$. For (c), if $a \leq b$, then $b - a = c^2$ for some c. Hence $\varphi(b) - \varphi(a) = \varphi(c)^2$, and $\varphi(a) \leq \varphi(b)$. For (d), let r be any real, let $\epsilon > 0$ be given, and choose rationals q_1 and q_2 with $q_1 \leq r \leq q_2$ and $q_2 - q_1 < \epsilon$. Then $q_1 = \varphi(q_1) \leq \varphi(r) \leq \varphi(q_2) = q_2$ by (a) and (c). Hence $|\varphi(r) - r| < \epsilon$. Since ϵ is arbitrary, $\varphi(r) = r$.

2. $(1 + r)^{-1} = 1 - r + r^2 - r^3 + \cdots \pm r^{n-1}$ if $r^n = 0$.

3. This follows from the universal mapping property of the field of fractions.

4. Suppose that X divides $A(X)B(X)$, i.e., $A(X)B(X) = XC(X)$. If a_0 and b_0 are the constant terms of $A(X)$ and $B(X)$, we then have $a_0 b_0 = 0$. If $a_0 = 0$, then X divides $A(X)$; if $b_0 = 0$, then X divides $B(X)$. Hence X is prime.

5. In (a), take (X) as the ideal. It is prime by Problem 4. Suppose that a is a member of R with no inverse in R. then (X) is not maximal since (a, X) strictly contains it and does not contain 1. For (b), we can use (a, X).

6. In (a), I_{x_0} is certainly an ideal. Suppose J is an ideal with $I_{x_0} \subsetneqq J$. Choose f in J that is not in I_{x_0}. The function $x - x_0$ is in I_{x_0}. Therefore $g = f^2 + (x - x_0)^2$ is in J. This function is everywhere > 0, and consequently $1/g$ is in R. Hence $1 = (1/g)g$ is in J, and J cannot be proper. So I_{x_0} is maximal.

Part (b) uses the Heine–Borel Theorem. For each point p in $[0, 1]$, choose a function f_p in I with $f_p(p) \neq 0$. By continuity, f_p is nonvanishing on some open set N_p containing p. As p varies, these open sets N_p cover $[0, 1]$. The Heine–Borel Theorem produces finitely many $N_{p_1}, \ldots, N_{p_k}$ that cover $[0, 1]$. Then f_{p_j} is nonvanishing on N_{p_j}. If x is a member of $[0, 1]$, then x is in some N_{p_j}, and f_{p_j} does not vanish at x. Thus the functions $f_{p_1}, \ldots, f_{p_k}$ have no common zero.

For (c), suppose that the maximal ideal I is not some I_{x_0}. Using (b), we form the function $g = f_{p_1}^2 + \cdots + f_{p_k}^2$. This is in I and is everywhere positive. The function $1/g$ is therefore in R, and $1 = (1/g)g$ is in I. Hence $I = R$, in contradiction to the fact that I is proper.

7. In (a), I_∞ is an ideal, and it is properly contained in the proper ideal of all members of R vanishing at $-\infty$. Part (b) follows from Proposition 8.8. The reason for (c) is that for each x_0 in $\mathbb{R}$, there is a member of R that is nonzero at x_0 and vanishes at infinity; this function has to be in I, and thus I cannot equal I_{x_0}.

8. For (a), let $a + bi$ be a nonzero member of I. Then $(a + b\sqrt{-5})(a - b\sqrt{-5}) = a^2 + 5b^2$ is a positive integer in I.

For (b), I is an additive subgroup of $\mathbb{Z} + \mathbb{Z}\sqrt{-5}$, which is free abelian of rank 2. Therefore I is free abelian of rank 1 or 2. We can rule out rank 1 because I contains a nonzero integer and also the product of that integer and $\sqrt{-5}$.

For (c), a $\mathbb{Z}$ basis of I consists of $x_1 = a_1 + b_1\sqrt{-5}$ and $x_2 = a_2 + b_2\sqrt{-5}$. Put $y_1 = rx_1 + sx_2 = (ra_1 + sa_2) + (rb_1 + sb_2)\sqrt{-5}$ and $y_2 = tx_1 + ux_2$, and aim to have y_1, y_2 form a $\mathbb{Z}$ basis with y_1 not involving $\sqrt{-5}$. We thus want $rb_1 + sb_2 = 0$, and the most economical way of achieving this equality is to put $d = \mathrm{GCD}(b_1, b_2)$ and to take $r = b_2 d^{-1}$ and $s = -b_1 d^{-1}$. Then $\mathrm{GCD}(r, s) = 1$, and we can choose t and u with $ru - st = 1$. With these choices we have $\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} r & s \\ t & u \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$. Since $\det \begin{pmatrix} r & s \\ t & u \end{pmatrix} = 1$, this change is invertible. In other words, y_1 and y_2 form a $\mathbb{Z}$ basis in which y_1 is some nonzero integer n. We may assume that $n > 0$. Let m be the smallest positive integer in I. Then n must be a multiple of m by an application of the division algorithm. Since y_1 and y_2 form a $\mathbb{Z}$ basis of I, we see that n equals m.

9. It is straightforward to see that P is an ideal and that $xy \in P$ implies $x \in P$ or $y \in P$. The ideal P is proper since the presence of 1 in $\varphi^{-1}(P')$ would mean that $\varphi(1) = 1$ is in P'. But P' is proper, and thus 1 is not in P'.

10. (a) $\{(r, 0) \mid r \in \mathbb{R}\}$ and $\{(0, r) \mid r \in \mathbb{R}\}$.
(b) (X).
(c) $(X - 1)$ and $(X - 2)$.
(d) (0).

11. For (a), $\mathbb{Q}[X]/I$ is a field and hence is a unique factorization domain. For (b), one can give a counterexample. The ring $\mathbb{Z}[\sqrt{-5}]$ is an integral domain and is the quotient of $\mathbb{Z}[X]$ by the ideal $(X^2 + 5)$; therefore $I = (X^2 + 5)$ is prime. On the other hand, $\mathbb{Z}[\sqrt{-5}]$ is not a unique factorization domain.

12. For (a), choose x and y with $xd + yc = 1$. Dividing by n gives $xc^{-1} + yd^{-1} = n^{-1}$. Then (a) follows by multiplying through by m. Part (b) uses an induction. Group n as $(p_1^{k_1} \cdots p_{r-1}^{k_{r-1}})p_r^{k_r}$ and apply (a) to write $mn^{-1} = a(p_1^{k_1} \cdots p_{r-1}^{k_{r-1}})^{-1} + bp_r^{-k_r}$. Repeat the process with $a(p_1^{k_1} \cdots p_{r-1}^{k_{r-1}})^{-1}$, and continue.

13. For (a), proceed as in the argument in Section 4 until near the end, obtaining x and y just as in that construction. Then $\delta(x + y\sqrt{-2}) = x^2 + 2y^2 \le \frac{1}{4} + 2 \cdot \frac{1}{4} = \frac{3}{4}$. Then we have $\delta(r + s\sqrt{-2}) < \delta(c + d\sqrt{-2})$, and the argument goes through.

For (b), we would get $\delta(x + y\sqrt{-3}) = x^2 + 3y^2 \le \frac{1}{4} + 3 \cdot \frac{1}{4} = 1$, and then the step $\delta(r + s\sqrt{-3}) < \delta(c + d\sqrt{-3})$ fails.

14. The map extends to an R module homomorphism by the universal mapping property of RG, and it is one-one onto by inspection. To check that it respects multiplication, it is enough to show that the product g_1g_2 in RG maps to $f_{g_1} * f_{g_2}$, i.e., that $f_{g_1} * f_{g_2} = f_{g_1g_2}$. The computation is $(f_{g_1} * f_{g_2})(x) = \sum_{y \in G} f_{g_1}(xy^{-1}) f_{g_2}(y) = f_{g_1}(xg_2^{-1})$, and this is 1 if and only if $xg_2^{-1} = g_1$, i.e., if and only if $x = g_1g_2$. For other values of x, it is 0. Therefore $(f_{g_1} * f_{g_2})(x) = f_{g_1g_2}(x)$ for all x.

15. Let the monic polynomial in question be $P(X)$. We prove by induction on m that any polynomial $A(X)$ in I of degree m is a multiple of $P(X)$. The base case of the induction is all polynomials of degree $< n$ in I; only 0 fits this description. Assume the result for all degrees $< m$, and let $A(X)$ be any polynomial in I, say with leading term a_mX^m, $a_m \neq 0$. Then $a_mX^{m-n}P(X)$ is in I, and so is $B(X) = A(X) - a_mX^{m-n}P(X)$. The coefficient of X^m in $B(X)$ is 0, and hence $B(X) = 0$ or else $\deg B(X) < m$. If $B(X) = 0$, then $A(X) = a_mX^{m-n}P(X)$, and $A(X)$ is a multiple of $P(X)$. If $\deg B(X) < m$, then induction gives $B(X) = C(X)P(X)$, and therefore $A(X) = (a_mX^{m-n} + C(X))P(X)$. So again $A(X)$ is a multiple of $P(X)$.

16. Let $p_1, \ldots, p_n$ be n distinct positive primes in $\mathbb{Z}$, put $q_k = p_1 \cdots p_k$ for $0 \le k \le n$, and take $I_{n+1} = (q_n, q_{n-1}X, q_{n-2}X^2, \ldots, q_0X^n)$. This can be written with $n + 1$ generators but not with fewer than that.

17. In (a), certainly $\ker\varphi \supseteq (y^2 - x^3)$. In the reverse direction, suppose that $\sum_{n=0}^{N} P_n(x)y^n$ is in $\ker\varphi$. Since $y^2 \equiv x^3 \bmod (y^2 - x^3)$, we can reduce this element of $\ker\varphi$ to the form $Q_0(x) + Q_1(x)y$. Substituting with t gives $Q_0(t^2) + Q_1(t^2)t^3 = 0$. The first term involves only even powers of t, and the second term involves only odd powers. Thus each is 0 separately. We are thus to determine what members $Q_0(x)$ and $Q_1(x)y$ of $\mathbb{K}[x, y]$ are in $\ker\varphi$. For $Q_0(t^2)$ to be 0, every coefficient of Q_0 must be 0. For $Q_1(t^2)t^3$ to be 0, every coefficient of Q_1 must be 0. Therefore only 0 is of the stated form, and every member of $\ker\varphi$ lies in $(y^2 - x^3)$.

For (b), image φ contains t^2, t^3, and every power t^n such that $n = 2a + 3b$ with a and b nonnegative integers. It follows that image φ consists of all linear combinations of powers t^n for $n \ge 2$.

19. Write $A(X) = B(X)Q(X)$ in $F[X]$, and let $A(X) = c(A)(c(A)^{-1}A(X))$, $B(X) = c(B)(c(B)^{-1}B(X))$, and $Q(X) = c(Q)(c(Q)^{-1}Q(X))$ be the decomposi-

tions of Proposition 8.19. Then we have

$$c(A)(c(A)^{-1}A(X)) = c(B)c(Q)\big((c(B)^{-1}B(X))(c(Q)^{-1}Q(X))\big).$$

By Gauss's Lemma and the uniqueness in Proposition 8.19, we obtain $c(A)^{-1}A(X) = (c(B)^{-1}B(X))(c(Q)^{-1}Q(X))$, apart from unit factors. Therefore the member $B_0(X) = c(B)^{-1}B(X)$ of $R[X]$ is exhibited as dividing $A_0(X) = c(A)^{-1}A(X)$ with a quotient $c(Q)^{-1}Q(X)$ in $R[X]$.

20. Let R be a finite integral domain, and let $a \neq 0$. Multiplication by a is one-one since R is an integral domain, and it must be onto R by the finiteness. Therefore there is some b with $ab = 1$, and we have produced an inverse for a.

21. Let $R' = R/(p)$. Suppose that $A(X) = B(X)C(X)$ nontrivially in $R[X]$ with $B(X) = b_kX^k + \cdots + b_0$, $C(X) = b_lX^l + \cdots + c_0$, and $k + l = N$. Since p divides a_0 but p^2 does not, p divides exactly one of b_0 and c_0, say the former. In $R'[X]$, we have $A(X) \equiv a_NX^N$, $C(X) \equiv c_lX^l + \cdots + c_0$, and $A(X) = B(X)C(X)$. Now X is prime in $R'[X]$ by Problem 4, and X^N divides $B(X)C(X)$ in $R'[X]$. Using the defining property of a prime, one power at a time, we find that X^N divides $B(X)$. Since $\deg B < N$, we must have $B(X) \equiv 0$ in $R'[X]$. Thus p divides b_k in R, and p divides a_N, contradiction.

22. In (a), we regard $WZ - XY$ as a first-degree polynomial in W, with Z being a prime in the ring of coefficients. A nontrivial factorization of $WZ - XY$ must be of the form $A(X, Y, Z)\big(B(X, Y, Z)W + C(X, Y, Z)\big)$ with $Z = A(X, Y, Z)B(X, Y, Z)$. Since Z is prime, one of these factors must be a unit, hence a scalar. If $A(X, Y, Z)$ is a scalar, then the factorization of $WZ - XY$ is trivial. Otherwise we may assume that the factorization is $WZ - XY = Z\big(W + C(X, Y, Z)\big)$. Then Z divides XY, and we arrive at a contradiction since Z does not appear in XY.

In (b), we expand in cofactors about the top row. Using induction, we see that we can regard the determinant $\det[X_{ij}]$ as a first-degree polynomial in X_{11} with an irreducible coefficient $P(X_{22}, X_{23}, \ldots, X_{nn})$. A nontrivial factorization must be of the form $\det[X_{ij}] = PX_{11} + Q = A(BX_{11} + C)$, where Q, A, B, C are polynomials in the remaining indeterminates. Then $AB = P$ and P irreducible implies that A or B is a unit, hence a scalar. If A is a scalar, our factorization of $\det[X_{ij}]$ is trivial. Otherwise we may assume that the factorization is $\det[X_{ij}] = PX_{11} + Q = P(X_{11} + C)$. Then P must divide Q. Taking the degrees of homogeneity into account, we see that Q must be the product of P and a homogeneous polynomial of degree 1. Every term of P is of the form $\prod_{i=2}^{n} X_{2,\sigma(2)}$ for some permutation σ of $\{2, \ldots, n\}$, and thus such a factor must appear in every term of Q. However, the only terms of $\det[X_{ij}]$ that contain a factor $\prod_{i=2}^{n} X_{2,\sigma(2)}$ also contain the factor X_{11}, and this factor is absent in Q. Thus the assumed reducibility has led to a contradiction.

23. The ideal of $\mathbb{Z}[X]$ generated by $A(X)$ and $B(X)$ consists of all polynomials $A(X)C(X) + B(X)D(X)$ with $C(X)$ and $D(X)$ in $\mathbb{Z}[X]$. If such an expression equals some integer n, then a GCD within $\mathbb{Q}[X]$ of $A(X)$ and $B(X)$ divides $A(X)$ and $B(X)$

and hence must divide n. It is therefore of degree 0 and is a unit in $\mathbb{Q}[X]$. Thus $A(X)$ and $B(X)$ are relatively prime in $\mathbb{Q}[X]$.

Conversely if $A(X)$ and $B(X)$ are members of $\mathbb{Z}[X]$ that are relatively prime in $\mathbb{Q}[X]$, we can find $P(X)$ and $Q(X)$ in $\mathbb{Q}[X]$ with $A(X)P(X) + B(X)Q(X) = 1$. Multiplying by a common denominator of the coefficients of $P(X)$ and $Q(X)$, we obtain a relation $A(X)C(X) + B(X)D(X) = n$ with all polynomials in $\mathbb{Z}[X]$. Thus n is in the ideal of $\mathbb{Z}[X]$ generated by $A(X)$ and $B(X)$.

24. We are given $\begin{pmatrix} 1+i & 2-i \\ 3 & 5i \end{pmatrix}\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ with coefficient matrix $C = \begin{pmatrix} 1+i & 2-i \\ 3 & 5i \end{pmatrix}$. Left multiplication on C by a matrix with determinant a unit does not change the total set of conditions on $\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$, and right multiplication by such a matrix changes the generators but not the module they generate. In the first column of C, we observe that $\mathrm{GCD}(1+i, 3) = 1$ because $1+i$ divides 2 and $\mathrm{GCD}(2, 3) = 1$. Then we have $-(1-i)(1+i)+1\cdot 3 = 1$, and we are led to the matrix $A = \begin{pmatrix} -(1-i) & 1 \\ -3 & 1+i \end{pmatrix}$, which has determinant 1. We can thus replace C by $AC = \begin{pmatrix} 1 & -1+8i \\ 0 & -11+8i \end{pmatrix}$. An invertible column operation replaces the upper right entry by 0. Thus we are led to the diagonal matrix $\begin{pmatrix} 1 & 0 \\ 0 & -11+8i \end{pmatrix}$. In other words, we may assume that the $\mathbb{Z}[i]$ module was given to us with generators t_1, t_2 satisfying $t_1 = 0$ and $(-11+8i)t_2 = 0$. Therefore the given $\mathbb{Z}[i]$ module is cyclic and is $\mathbb{Z}[i]$ isomorphic to $\mathbb{Z}[i]/(-11+8i)$.

25. In (a), $\delta(z) = z\bar{z}$. Then $\delta(zw) = zw\bar{z}\bar{w} = z\bar{z}w\bar{w} = \delta(z)\delta(w)$.

In (b), we start with two nonzero members α and β of R. We are to find γ and ρ in R with $\alpha = \beta\gamma + \rho$ and $\delta(\rho) < \delta(\beta)$. It is the same to find γ and ρ with $\alpha/\beta = \gamma + \rho/\beta$ and $\delta(\rho/\beta) < 1$. Apply the hypothesis with $z = \alpha/\beta$, and let γ be the element r such that $\delta(z-r) < 1$. Then ρ may be defined as $\beta(z-r)$, and all the conditions are satisfied.

26. Given $z = x + y\sqrt{-m}$, define $r = a + \frac{1}{2}b(1+\sqrt{-m})$ in $\mathbb{Z}[\frac{1}{2}(1+\sqrt{-m})]$ by choosing b to be an integer with $|2y - b| \le \frac{1}{2}$ and then choosing a to be an integer with $|x - a - \frac{1}{2}b| \le \frac{1}{2}$. Since $|y - \frac{1}{2}b| \le \frac{1}{4}$, we then have

$$\delta(z-r) = (x - a - \tfrac{1}{2}b)^2 + m(y - \tfrac{1}{2}b)^2 \le \tfrac{1}{4} + m\,\tfrac{1}{16} \le \tfrac{1}{4} + \tfrac{11}{16} < 1.$$

27. In (a), complex conjugation is an automorphism of $\mathbb{Z}[i]$ and must therefore carry primes to primes.

In (b), we know that $(a+bi)(a-bi)$ is the integer $N(a+bi)$. Suppose that $N(a+bi) = mn$ nontrivially with $\mathrm{GCD}(m, n) = 1$. Since $a+bi$ is prime, it divides one of m and n. Say that $m = (a+bi)(c+di)$. Then $m^2 = N(m) = N(a+bi)N(c+di) = mnN(c+di)$. Any prime number dividing n must divide the left side m^2, and hence there can be no such prime. We conclude that $N(a+bi)$ does not have nontrivial relatively prime divisors. Hence it is a power of some prime number p.

In (c), let $N(a+bi) = p^k$. The left side is the product of two primes of $\mathbb{Z}[i]$. If p is the product of l primes of $\mathbb{Z}[i]$, then p^k is the product of kl primes. Then we must have $kl = 2$, and k must divide 2.

In (d), suppose $N(a+bi) = p^2$, so that $k = 2$ in (c). Then $l = 1$, and p is prime in $\mathbb{Z}[i]$.

28. The equation $N(a+bi) = p$ says that $a^2 + b^2 = p$. The right side is $\equiv 3 \bmod 4$, but 3 is not the sum of two squares modulo 4. Hence $N(a+bi) = p$ is impossible when $p \equiv 3 \bmod 4$. Problem 27c then forces $N(a+bi) = p^2$, and Problem 27d says that p is prime in $\mathbb{Z}[i]$.

29. If $N(a+bi) = 2$, then $|a| = |b| = 1$, and we obtain $1+i$ and its associates. If $N(a+bi) = 4$, then $a = \pm 2$ with $b = 0$ or else $a = 0$ with $b = \pm 2$; in these cases $a + bi$ is an associate of 2, which is $(1+i)(1-i)$ and is not prime in $\mathbb{Z}[i]$.

30. The multiplicative group of $\mathbb{F}_p$ is cyclic of order $p-1$. If p is of the form $4n+1$, then $\mathbb{F}_p^\times$ has order $4n$. The n^{th} power of a generator then has to be an integer whose square is $\equiv -1 \bmod p$.

31. For (a), we obtain φ_1 by mapping $\mathbb{Z}[X]$ to $\mathbb{F}_p[X]$ with a substitution homomorphism and following this with a passage to the quotient. Similarly φ_2 is obtained from the substitution homomorphism $\mathbb{Z}[X] \to \mathbb{Z}[i]$ followed by the passage to the quotient.

For (b), the kernel of φ_1 consists of all polynomials that are multiples of X^2+1 when their coefficients are taken modulo p. This is $p\mathbb{Z}[X] + (X^2+1)\mathbb{Z}[X] = (p, X^2+1)$. The kernel of φ_2 consists of all polynomials with the property that when taken modulo X^2+1, they are multiples of p. This too is the ideal (p, X^2+1).

For (c), Problem 30 shows that the polynomial X^2+1 factors nontrivially in $\mathbb{F}_p[X]$. Therefore X^2+1 is not prime, the ideal (X^2+1) is not prime, and $F_p[X]/(X^2+1)$ is not an integral domain. By (b), $\mathbb{Z}[i]/(p)$ is not an integral domain, and the ideal (p) is not prime. Hence p is not prime in $\mathbb{Z}[i]$. By (c) and (d) in Problem 27, p is of the form $N(a+bi)$ for some prime $a+bi$ in $\mathbb{Z}[i]$.

For (d), if we have $p = N(a+bi) = N(a'+b'i)$, we obtain two prime factorizations of p in $\mathbb{Z}[i]$ as $p = (a+bi)(a-bi) = (a'+b'i)(a'-b'i)$, and unique factorization in $\mathbb{Z}[i]$ implies that $a'+b'i$ is an associate of $a+bi$ or $a-bi$.

32. For (a), multiply C on the left by the matrix A that is the identity except in the first column, where the i^{th} entry is C_{ii}.

For (b) and (c), the step of row reduction leads to a first column that is 0 in all entries but the first, where it is $\text{GCD}(C_{11}, \dots, C_{nn})$. In other words, the new entry in position (1, 1) divides all entries in the new C. Therefore one step of column reduction leaves the entry unchanged in position (1, 1), leaves the remainder of the first column equal to 0, and makes the remainder of the first row equal to 0. What is left in the rows and columns other than the first is a matrix whose entries are all divisible by $\text{GCD}(C_{11}, \dots, C_{nn})$. Hence we can induct on the size.

33. In (a), changing notation slightly from Lemma 8.26, write $AE = DB$ with $\det A$ and $\det B$ in $R^\times$. Over the field of fractions of R, the m-by-n matrices E and D

must have the same rank since A and B are invertible, and consequently D and E have the same number of nonzero diagonal entries. Thus for some l with $0 \le l \le k$, we are given that D_{jj} divides $D_{j+1,j+1}$ and E_{jj} divides $E_{j+1,j+1}$ whenever $1 \le j < l$. Fix i with $1 \le i \le l$, and consider all possible i-by-i determinants that can be formed using the first i rows of B and one of the $\binom{n}{i}$ sets of i columns. Since $\det B$ is in $R^\times$, it follows from the expansion-by-cofactors formula that these determinants have GCD equal to 1. Each corresponding determinant for DB equals $D_{11} \cdots D_{ii}$ times such a determinant, and hence the GCD for DB is $D_{11} \cdots D_{ii}$.

Meanwhile, the GCD of the determinants for A is also 1, and, because of the divisibility property of the diagonal entries of E, $E_{11} \cdots E_{ii}$ divides each of the determinants for AE. Hence $E_{11} \cdots E_{ii}$ divides the GCD of the determinants for AE, which equals the GCD of the determinants for DB, which equals $D_{11} \cdots D_{ii}$. Thus $E_{11} \cdots E_{ii}$ divides $D_{11} \cdots D_{ii}$.

Arguing similarly with the determinants formed from the first i columns of A, AE, B, and BD, we see that $D_{11} \cdots D_{ii}$ divides $E_{11} \cdots E_{ii}$. Therefore $D_{11} \cdots D_{ii}$ and $E_{11} \cdots E_{ii}$ are associates for $1 \le i \le l$. Since none of the factors in question is 0, we see that each of the first l diagonal entries of D is an associate of the corresponding diagonal entry of E. This proves the desired uniqueness.

34. For (a), we have seen in this setting that the decomposition of V as a direct sum of cyclic $\mathbb{K}[X]$ modules means a decomposition of V as a direct sum of vector subspaces, each of which is invariant under L. Also, if V_0 is one of these vector subspaces, the cyclic nature of the module means that there is some vector v_0 in V_0 such that $\mathbb{K}[X]v_0 = V_0$, and the diagonal entry of the matrix D in the proof of Theorem 8.25 is a polynomial $M[X]$ such that $V_0 \cong K[X]/(M(X))$ as a $K[X]$ module. Referring to Problems 26–31 of Chapter V, we see that v_0 is a cyclic vector for the cyclic subspace V_0, and $M[X]$ is the minimal polynomial of L on this subspace.

The divisibility property of the minimal polynomials and also the uniqueness assertion now follow from what has been proved in Problems 32–33. We know from Problem 28 in Chapter V that the data of a cyclic subspace and the minimal polynomial yield a particular matrix for the linear mapping and hence determine the linear mapping on that subspace up to similarity. Consequently the uniqueness statement that has just been observed says that L is determined up to similarity by the integer r and the sequence of minimal polynomials.

35. Let A and B be members of $M_n(\mathbb{K})$. Form the data for each from the rational canonical form in Problem 34. Now consider everything as involving vector spaces over the larger field $\mathbb{L}$. We are given that the two matrices are similar over $\mathbb{L}$, i.e., are conjugate via $\mathrm{GL}(n, \mathbb{L})$. Problem 34 shows that the respective decompositions have the same data. The two matrices still have the same data when we again consider the field to be $\mathbb{K}$. Hence they are similar over $\mathbb{K}$, i.e., are conjugate via $\mathrm{GL}(n, \mathbb{K})$.

36. The fact that the homomorphisms are isomorphisms follows from the composition rule.

38. In (b), we can write any member of $F[X_1, \ldots, X_n, X]$ as

$$A_n(X_1, \ldots, X_n)X^n + \cdots + A_1(X_1, \ldots, X_n)X + A_0(X_1, \ldots, X_n),$$

and σ^{**} acts by having σ^* act on each coefficient. Invariance under all σ^{**}'s therefore means that each coefficient is invariant under all σ^*'s and hence is a symmetric polynomial.

39. In (a), if, for example $i < j$ and $k_i < k_j$, then the monomial $aX_1^{k_1} \cdots X_n^{k_n}$ is increased in the ordering by replacing the factors $X_i^{k_i} X_j^{k_j}$ by $X_i^{k_j} X_j^{k_i}$.

For (b), we need only take the largest monomial in each E_i, raise it to the c_i power, and multiply the results.

For (c), let the largest monomial in A be $aX_1^{k_1} \cdots X_n^{k_n}$. To define M, choose $r = a$ and define $c_j = k_j - k_{j+1}$ for $1 \le j < n$ and $c_n = k_n$.

For (d), the construction in (c) yields 0 coefficient for $X_1^{k_1} \cdots X_n^{k_n}$, and $A - rM$ has no larger monomials. So if $A - rM = 0$, the largest monomial is below that monomial $X_1^{k_1} \cdots X_n^{k_n}$.

For (e), iteration of the construction in (c) and (d) shows that any homogeneous symmetric polynomial equals a homogeneous polynomial in the elementary symmetric polynomials. Problem 37 shows that any symmetric polynomial is a linear combination of homogeneous symmetric polynomials, and hence every symmetric polynomial is a polynomial in the elementary symmetric polynomials.

40. Suppose that z_0 and w_0 in $\mathbb{C}^m$ have $P(z_0) \neq 0$ and $P(w_0) \neq 0$. As a function of $t \in \mathbb{C}$, $P(z_0 + t(w_0 - z_0))$ is a polynomial function nonvanishing at $t = 0$ and $t = 1$. The subset of $t \in \mathbb{C}$ where it vanishes is finite, and its complement in $\mathbb{C}$ is necessarily pathwise connected and therefore connected. Thus z_0 and w_0 lie in a connected subset of $\mathbb{C}^m$ where P is nonvanishing. Taking the union of these connected sets with z_0 fixed and w_0 varying, we see that the set of $w_0 \in \mathbb{C}^m$ where $P(w_0) \neq 0$ is connected.

41. For (a), two applications of the formula relating Pfaffians and determinants gives us $\text{Pfaff}(A^t X A)^2 = \det(A^t X A) = (\det A)^2 \det X = (\det A)^2 \text{Pfaff}(X)^2$. Taking the square root gives the desired result.

For (b), we fix X with $\text{Pfaff}(X) \neq 0$ and allow A to vary. On the set where $\det A \neq 0$, the function $A \mapsto \text{Pfaff}(A^t X A)/\det A$ is a continuous function with image in the two-point set $\{\pm\text{Pfaff}(X)\}$, by (a). The domain of the function is connected by Problem 40, and therefore the image has to be connected. Hence the function has to be constant. Checking the value of the function at $A = I$, we see that the function has to be constantly equal to $\text{Pfaff}(X)$.

42. Form the ring $S = \mathbb{Z}[\{A_{ij}\}, \{X_{ij}\}]$. We can then regard $\text{Pfaff}(A^t X A)$ and $(\det A)\text{Pfaff}(X)$ as two polynomials with entries in S. If we fix arbitrary elements $a_{ij} \in \mathbb{Z}$ for all i and j and also $x_{ij} \in \mathbb{Z}$ for $i < j$, then Proposition 4.30 gives us a unique substitution homomorphism $\Psi \to \mathbb{Z}$ such that $\Psi(1) = 1$, $\Psi(A_{ij}) = a_{ij}$,

and $\Psi(X_{ij}) = x_{ij}$. Assemble the a_{ij} and x_{ij} into matrices $a = [a_{ij}]$ and $x = [x_{ij}]$ with x alternating. Problem 41b shows that the identity in question holds when the entries are in $\mathbb{C}$, and in particular it holds when the entries are in $\mathbb{Z}$. Therefore $\text{Pfaff}(a^t x a) = (\det a)\text{Pfaff}(x)$. Since $\mathbb{Z}$ is an integral domain and since a and x are arbitrary with x alternating, Corollary 4.32 allows us to conclude that $\text{Pfaff}(A^t X A) = (\det A)\text{Pfaff}(X)$ as an equality in S.

To pass from S to $\mathbb{K}$, let $1_{\mathbb{K}}$ be the identity of $\mathbb{K}$, and let $\varphi_1 : \mathbb{Z} \to \mathbb{K}$ be the unique homomorphism of rings such that $\varphi_1(1) = 1_{\mathbb{K}}$. If we fix arbitrary elements a_{ij} of $\mathbb{K}$ for all i and j, as well as arbitrary elements x_{ij} of $\mathbb{K}$ for $i < j$, then Proposition 4.30 gives us a unique substitution homomorphism $\Phi : S \to \mathbb{K}$ such that $\Phi(1) = \varphi_1(1) = 1_{\mathbb{K}}$, $\Phi(A_{ij}) = a_{ij}$ for all i and j, and $\Phi(X_{ij}) = x_{ij}$ whenever $i < j$. Applying Φ to our identity in S, we obtain $\text{Pfaff}(a^t x a) = (\det a)\text{Pfaff}(x)$ as an equality in $\mathbb{K}$.

43. From Problem 42 and the hypothesis on g, we have $1 = \text{Pfaff}(J) = \text{Pfaff}(g^t J g) = (\det g)\text{Pfaff}(J) = \det g$. Hence $\det g = 1$.

45. For (a), if $\varphi : R \to R/P^k$ is the quotient homomorphism, then φ^{-1} of any ideal of R/P^k is an ideal I of R containing P^k. If Q is a prime ideal dividing I, then Q divides P^k, and it follows that $Q = P$. Thus the only possibilities for I are the powers P^i of P, necessarily stopping with $i = k$.

For (b), we know that π^i lies in P^i but not P^{i+1}. For $1 \le i \le k-1$, it follows that the principal ideal $(\pi^i + P^k)/P^k$ is contained in the ideal P^i/P^k but not in P^{i+1}/P^k. Since the ideals P^j/P^k for $j \le k$ are nested and there are no other ideals in R/P^k, we must have $(\pi^i + P^k)/P^k = P^i/P^k$. Thus P^i/P^k is principal.

46. Corollary 8.63 and Problem 44 together show that every ideal of R/I is principal if it can be shown that every ideal of R/P^k is principal when P is a nonzero prime ideal. The two parts of Problem 45 together show that every ideal of R/P^k is principal.

47. We may assume that $(a) \subsetneqq I$ since otherwise the result follows with $b = 0$. Since $a \neq 0$, the ideal $I/(a)$ in $R/(a)$ is a principal ideal by Problem 46c. If b_0 is a generator of this ideal, then $\big(R/(a)\big)b_0 = I/(a)$. Since b_0 is in $I/(a)$, we can write it as $b_0 = b + (a)$ for some b in I. Every member of $I/(a)$ is then of the form $\big(r + (a)\big)\big(b + (a)\big) = rb + (a)$, and we conclude that every member of I is of the form $rb + sa$ with r and s in R.

48. Any R submodule of R is an ideal.

49. Write $M = Rx_1 + \cdots + Rx_n$ with $x_1, \ldots, x_n$ in F. Each x_i is of the form $r_i s_i^{-1}$ with r_i and s_i in R and with $s_i \neq 0$. Then aM lies in R for $a = \prod_{i=1}^n s_i$. So aM is an ideal in R, by Problem 48. If N is a second fractional ideal, choose $b \neq 0$ such that bN is an ideal in R. Then $(aM)(bN)$ is an ideal in R, and the formula $MN = (ab)^{-1}(aM)(bN)$ shows that MN is a fractional ideal.

50. Since I is a finitely generated R module, we can write $I = Ra_1 + \cdots + Ra_n$ with all a_i in R. The condition for $x \in F$ to be in I^{-1} is that $xI \subseteq R$, and it is

necessary and sufficient that xa_i be in R for all i. Thus it is necessary that x be in $(a_1 \cdots a_n)^{-1}R$. Consequently I^{-1} is an R submodule of the singly generated R module $(a_1 \cdots a_n)^{-1}R$. Since R is Noetherian, I^{-1} is finitely generated.

51. If I is maximal among the nonzero ideals of R for which there is no fractional ideal M of F with $IM = R$, then Lemma 8.58 shows that I is not prime. Choose a nonzero prime ideal P with $I \subsetneqq P$. Then Lemma 8.58 and the definitions give $I \subseteq IP^{-1} \subseteq II^{-1} \subseteq R$. We cannot have $I = IP^{-1}$ since otherwise $IP = (IP^{-1})P = I(P^{-1}P) = I$ and Proposition 8.52 gives $I = 0$. By maximality of I, we can find some fractional ideal N with $(IP^{-1})N = R$. Then $I(P^{-1}N) = R$, and we can take $M = P^{-1}N$, by Problem 49.

52. Every member x of M has $xI \subseteq R$, and thus $M \subseteq I^{-1}$. On the other hand, if x is in I^{-1}, then $xI \subseteq R$, $x = xIM \subseteq RM = M$, and x is in M.

53. If M is a fractional ideal, then Problem 49 produces $c \neq 0$ in F with $cM \subseteq R$, and Problem 48 shows that cM is an ideal of R. Using Problem 52, we can write $M = (c)^{-1}(c)M = (c)^{-1}(cM)$. This proves that $M = IJ^{-1}$ for ideals I and J. Then (a) follows from Theorem 8.55 and Problem 52, and (b) follows from Problem 52.

Chapter IX

1. The equation for r gives $r^3 = 3r - 4$ and $r^4 = 3r^2 - 4r$. Therefore the inverse has $1 = (r^2+r+1)(ar^2+br+c) = ar^4+(a+b)r^3+(a+b+c)r^2+(b+c)r+c = r^2(4a+b+c)+r(-a+4b+c)+1(-4a-4b+c)$, and we are led to the system of linear equations

$$\begin{aligned} 4a + b + c &= 0, \\ -a + 4b + c &= 0, \\ -4a - 4b + c &= 1. \end{aligned}$$

Then $(a, b, c) = (-\frac{3}{49}, -\frac{5}{49}, \frac{17}{49})$, and $(r^2 + r + 1)^{-1} = -\frac{3}{49}r^2 - \frac{5}{49}r + \frac{17}{49}$.

2. Multiplication by a nonzero r is a one-one F linear mapping from the F vector space R onto itself. Since $\dim_F R < \infty$, this linear mapping must be onto. The element s such that $rs = 1$ is a multiplicative inverse of r.

3. Let z_0 be a nonreal element of $\mathbb{K}$. Then the closure of the $\mathbb{Q}$ vector space $\mathbb{Q} + \mathbb{Q}z_0$ contains $\mathbb{R} + \mathbb{R}z_0 = \mathbb{C}$.

4. If $y = F(x)/G(x)$, then $G(x)y = F(x)$. Arranging the terms as powers of x with coefficients of the form $ay + b$ with a and b in $\Bbbk$, we see that x is a root of a polynomial in one indeterminate over $\Bbbk(y)$. Therefore x is algebraic over $\Bbbk(y)$.

5. The condition is that N be the square of an integer. For any other N, $X^2 - N$ is irreducible over $\mathbb{Q}$, and $[\mathbb{Q}(\sqrt{N}) : \mathbb{Q}] = 2$. Since 2 does not divide 3, $\mathbb{Q}(\sqrt{N})$ cannot be a subfield of $\mathbb{Q}(\sqrt[3]{2})$.

6. $X \mapsto Y + 1$.

7. No, since 8 is not a power of 4. See Corollary 9.19.

8. Let g be a generator of the cyclic group $\mathbb{K}^\times$, and let q be the order of $\mathbb{K}$. Then

$$g \cdot g^2 \cdot g^3 \cdots g^{q-1} = g^{1+2+3+\cdots+(q-1)} = g^{\frac{1}{2}q(q-1)}.$$

If q is even, then this is $(g^{q-1})^{q/2} = 1^{q/2} = 1 = -1$. If q is odd, it is $(g^{\frac{1}{2}(q-1)})^q = (-1)^q = -1$.

9. Proof 1: Let $F(X) = X^n + c_{n-1}X^{n-1} + \cdots + c_0$ be the minimal polynomial of r. We are given that n is odd. Write the equation $F(r) = 0$ as

$$r(r^{n-1} + c_{n-2}r^{n-3} + \cdots + c_1) = -c_{n-1}r^{n-1} - c_{n-3}r^{n-3} - \cdots - c_0.$$

Then r is expressed as an element of $\Bbbk(r^2)$ unless $r^{n-1} + c_{n-2}r^{n-3} + \cdots + c_1 = 0$. But this expression cannot be 0 because this polynomial has degree $n-1$ and the minimal polynomial for r has degree n.

Proof 2: The element r of $\mathbb{K}$ is a root of the polynomial $X^2 - r^2$ in $\Bbbk(r^2)[X]$, and hence $[\Bbbk(r) : \Bbbk(r^2)] \le 2$. Since $[\Bbbk(r) : \Bbbk] = [\Bbbk(r) : \Bbbk(r^2)]\,[\Bbbk(r^2) : \Bbbk]$ with the left side odd by assumption, $[\Bbbk(r) : \Bbbk(r^2)]$ has to be odd. Thus it is 1.

10. Let $d_r = [\Bbbk(r) : \Bbbk]$ and $d_s = [\Bbbk(s) : \Bbbk]$. Since $\mathbb{K}$ contains $\Bbbk(r)$ and $\Bbbk(s)$, we see that d_r and d_s divide $[\mathbb{K} : \Bbbk]$. Since $\mathrm{GCD}(d_r, d_s) = 1$, $d_r d_s$ divides $[\mathbb{K} : \Bbbk]$. The minimal polynomial $M(X)$ of r over $\Bbbk$ is a polynomial over $\Bbbk(s)$ such that $M(r) = 0$. Thus the minimal polynomial $N(X)$ of r over $\Bbbk(s)$ divides $M(X)$. If c is the degree of $N(X)$, we then have $c \le d_r$. Since $d_r d_s$ divides $[\mathbb{K} : \Bbbk]$, we obtain

$$d_r d_s \le [\mathbb{K} : \Bbbk] = [\Bbbk(r, s) : \Bbbk] = [\Bbbk(s)(r) : \Bbbk] = c[\Bbbk(s) : \Bbbk] = cd_s \le d_r d_s.$$

Equality must hold throughout. Equality at the right end says that $c = d_r$, and this proves (a). Equality at the left end says that $d_r d_s = [\mathbb{K} : \Bbbk]$, and this proves (b).

11. In (a), we have $\gamma = \beta + c\alpha = \beta(1 + c\omega)$. Here $r = 1 + c\omega$ lies in $\mathbb{Q}(\sqrt{-3}\,)$, and so does r^3. Therefore r^3 is a root of a quadratic polynomial $Y^2 + pY + q$. Then $\gamma^6 + a\gamma^3 + b = r^6\beta^6 + ar^3\beta^3 + b = 4r^6 + 2ar^3 + b = 4(r^6 + \frac{1}{2}ar^3 + \frac{1}{4}b)$, and the right side is 0 if a and b are chosen such that $p = \frac{1}{2}a$ and $q = \frac{1}{4}b$.

In (b), $\gamma = \beta + \alpha = \beta(1 + \omega)$, and $\gamma^3 = \beta^3(\frac{1}{2}(1 + \sqrt{-3}\,))^3 = 2(-1) = -2$. Then γ satisfies $\gamma^3 + 2 = 0$, and this is irreducible since -2 is not a cube in $\mathbb{Q}$.

In (c), the field $\mathbb{Q}(\gamma)$ contains $\gamma^3 = \beta^3(\frac{1}{2}(3 - \sqrt{-3}\,))^3 = \frac{1}{4}(3 - \sqrt{-3}\,)^3 = \frac{1}{4}(27 - 9\sqrt{-3} + 3 \cdot 3(-3) - (-3)\sqrt{-3}\,) = -\frac{3}{2}\sqrt{-3}$. Thus $\mathbb{Q}(\sqrt{-3}\,)$ is a subfield of $\mathbb{Q}(\gamma)$, and 2 divides $[\mathbb{Q}(\gamma) : \mathbb{Q}]$. Since $\mathbb{Q}(\sqrt{-3}\,)$ is a subfield, $\beta = \gamma(1-\omega)^{-1}$ lies in $\mathbb{Q}(\gamma)$. Thus $\mathbb{Q}(\sqrt[3]{2}\,)$ is a subfield of $\mathbb{Q}(\gamma)$, and 3 divides $[\mathbb{Q}(\gamma) : \mathbb{Q}]$. Consequently 6 divides $[\mathbb{Q}(\gamma) : \mathbb{Q}]$, and the minimal polynomial of γ has degree ≥ 6. By (a), it has degree exactly 6.

12. Let the characteristic be p. If $F(X)$ has $F'(X) = 0$, then all the exponents of X appearing in $F(X)$ are multiples of p. Let $F(X) = a_n X^{np} + a_{n-1} X^{(n-1)p} + \cdots + a_1 X^p + a_0$. Since the Frobenius map is onto in the case of a finite field, we can choose members $c_n, \ldots, c_0$ of $\Bbbk$ such that $c_n^p = a_n$, $c_{n-1}^p = a_{n-1}, \ldots, c_0^p = a_0$. Put $G(X) = c_n X^n + c_{n-1} X^{n-1} + \cdots + c_0$. Then $F(X) = G(X)^p$, and $F(X)$ is reducible.

13. In (a), if $F(X) = G(X)H(X)$ is reducible and r_1 is a root of $G(X)$, then $\sigma(r_1)$ is a root of $G(X)$ for any $\sigma \in \mathrm{Gal}(\mathbb{K}/\Bbbk)$. Consequently the orbit of r_1 under $\mathrm{Gal}(\mathbb{K}/\Bbbk)$ is a proper subset of the set of roots of $F(X)$. Conversely if $F(X)$ is irreducible and r_j is given, then the uniqueness of simple extensions gives us a $\Bbbk$ isomorphism of $\Bbbk(r_1)$ onto $\Bbbk(r_j)$. Theorem 9.13′ shows that this isomorphism extends to a $\Bbbk$ automorphism of $\mathbb{K}$, and hence $\mathrm{Gal}(\mathbb{K}/\Bbbk)$ is transitive on the set of roots of $F(X)$.

In (b), the transitivity follows from (a) and the irreducibility of $\Phi_8(X)$ over $\mathbb{Q}$. Let $\zeta = e^{2\pi i/8}$. The roots of $\Phi_8(X) = X^4 + 1$ are $\zeta, \zeta^3, \zeta^5, \zeta^7$. So if σ is in $\mathrm{Gal}(\mathbb{K}/\mathbb{Q})$, then $\sigma(\zeta) = \zeta^k$ with k odd. Then $\sigma^2(\zeta) = \sigma(\zeta^k) = \sigma(\zeta)^k = (\zeta^k)^k = \zeta^{k^2}$. Since the square of any odd integer is congruent to 1 modulo 8, $\sigma^2(\zeta) = \zeta$. Thus each σ has $\sigma^2 = 1$, and $\mathrm{Gal}(\mathbb{K}/\mathbb{Q})$ cannot contain a 4-cycle.

In (c), the irreducibility of $F(X)$ implies that $F(X)$ is the minimal polynomial of r_1. Hence $[\Bbbk(r_1) : \Bbbk] = n$. Since $\Bbbk(r_1) \subseteq \mathbb{K}$, $[\Bbbk(r_1) : \Bbbk]$ must divide $[\mathbb{K} : \Bbbk]$, and n divides $[\mathbb{K} : \Bbbk]$. Therefore n divides the equal integer $\mathrm{Gal}(\mathbb{K}/\Bbbk)$. If n is prime, then the fact that n divides the order of $\mathrm{Gal}(\mathbb{K}/\Bbbk)$ implies that $\mathrm{Gal}(\mathbb{K}/\Bbbk)$ contains an element of order n, by Sylow's Theorems. The only elements of order n in $\mathfrak{S}_n$ are the n-cycles, and hence $\mathrm{Gal}(\mathbb{K}/\Bbbk)$ contains at least one n-cycle.

14. In (a), we have $\mathbb{L}_{k+1} = \mathbb{L}_k(\sqrt{a_{k+1}})$, and hence $[\mathbb{L}_{k+1} : \mathbb{L}_k]$ equals 1 or 2. By induction, $[\mathbb{L}_k : \mathbb{Q}]$ is a power of 2, and the power is at most the number of steps in the induction, namely k.

In (b), associate to each subset S of $\{1, \ldots, k\}$ the element $v_S = \prod_{j \in S} \sqrt{a_j}$ in $\mathbb{L}_k$. The product of any two such elements is an integer multiple of a third such element, and hence the elements v_S span $\mathbb{L}_k$ linearly over $\mathbb{Q}$. Since there are 2^k such elements, they form a vector-space basis. The extension $\mathbb{L}_k/\mathbb{Q}$ is separable, being in characteristic 0, and it is normal as the splitting field of $\prod_{j=1}^k (X^2 - a_j)$. So it is a finite Galois extension. Any member σ of $\mathrm{Gal}(\mathbb{L}_k/\mathbb{Q})$ must permute the roots of each $X^2 - a_j$ and hence must send $\sqrt{a_j}$ to $\pm\sqrt{a_j}$. On the other hand, σ is determined by its effect on each $\sqrt{a_j}$. Since $\mathrm{Gal}(\mathbb{L}_k/\mathbb{Q})$ has order 2^k, there exists for each subset S of $\{1, \ldots, k\}$ one and only one σ such that $\sigma(\sqrt{a_j}) = -\sqrt{a_j}$ for $j \in S$ and $\sigma(\sqrt{a_j}) = +\sqrt{a_j}$ for $j \notin S$. The group $\mathrm{Gal}(\mathbb{L}_k/\mathbb{Q})$ consists exactly of these elements.

In (c), let σ_j be the member of $\mathrm{Gal}(\mathbb{L}_k/\mathbb{Q})$ with $\sigma_j(\sqrt{a_i}) = -\sqrt{a_i}$ for $i = j$ and $\sigma_j(\sqrt{a_i}) = +\sqrt{a_i}$ for $i \neq j$. Then $\sigma_j(v_S) = -v_S$ if j is in S, and $\sigma_j(v_S) = +v_S$ if j is not in S.

Arguing by contradiction, let $\sqrt{a_{k+1}} = \sum c_S v_S$ with each c_S in $\mathbb{Q}$. If

$\sigma_j(\sqrt{a_{k+1}}) = \sqrt{a_{k+1}}$, then we have

$$\sum_{\text{all } S} c_S v_S = \sqrt{a_{k+1}} = \sigma_j(\sqrt{a_{k+1}}) = \sum_{\text{all } S} c_S \sigma_j(v_S) = -\sum_{S \text{ with } j \in S} c_S v_S + \sum_{S \text{ with } j \notin S} c_S v_S,$$

and it follows that $c_S = 0$ whenever j is in S. On the other hand, if $\sigma_j(\sqrt{a_{k+1}}) = -\sqrt{a_{k+1}}$, then we have

$$\sum_{\text{all } S} c_S v_S = \sqrt{a_{k+1}} = -\sigma_j(\sqrt{a_{k+1}}) = -\sum_{\text{all } S} c_S \sigma_j(v_S) = \sum_{S \text{ with } j \in S} c_S v_S - \sum_{S \text{ with } j \notin S} c_S v_S,$$

and it follows that $c_S = 0$ whenever j is not in S.

Define $S_0 = \{j \mid \sigma_j(\sqrt{a_{k+1}}) = -\sqrt{a_{k+1}}\}$. From the above it follows that $c_S = 0$ whenever some member of S_0 is not in S, and that $c_S = 0$ whenever some member of the complement of S_0 is in S. In other words, $c_S = 0$ except for c_{S_0}. We conclude that $\sqrt{a_{k+1}} = c_{S_0} v_{S_0} = c_{S_0} \sqrt{\prod_{j \in S_0} a_j}$ and hence that $a_{k+1} = c_{s_0}^2 \prod_{j \in S_0} a_j$. This contradicts the hypothesis that $\{a_1, \ldots, a_n\}$ are relatively prime and square free. Hence $\sqrt{a_{k+1}}$ does not lie in $\mathbb{L}_k$. This proves (c), and we obtain $[\mathbb{L}_{k+1} : \mathbb{L}_k] = 2$. By induction we see that $[\mathbb{L} : \mathbb{Q}] = 2^n$. This proves (d).

15. For (a) and (b), Lemma 9.45 shows that $X^p - a$ is irreducible over $\mathbb{Q}$. Hence $[\mathbb{Q}(r) : \mathbb{Q}] = p$. Let ζ be a primitive p^{th} root of 1. Then $[\mathbb{Q}(\zeta) : \mathbb{Q}] = p - 1$ is relatively prime to $[\mathbb{Q}(r) : \mathbb{Q}] = p$. Problem 10a shows that $\Phi_p(X)$ is irreducible in $\mathbb{Q}(r)$. Since ζ and r generate $\mathbb{K}$, Problem 10b shows that $[\mathbb{K} : \mathbb{Q}] = [\mathbb{Q}(r) : \mathbb{Q}]\,[\mathbb{Q}(\zeta) : \mathbb{Q}] = p(p-1)$.

In (c), the Galois correspondence between intermediate fields and subgroups of $G = \mathrm{Gal}(\mathbb{K}/\mathbb{Q})$ associates $\mathbb{Q}(\zeta)$ to the subgroup $N = \mathrm{Gal}(\mathbb{K}/\mathbb{Q}(\zeta))$, and it associates $\mathbb{Q}(r)$ to the subgroup $H = \mathrm{Gal}(\mathbb{K}/\mathbb{Q}(r))$. Since $\mathbb{Q}(\zeta)/\mathbb{Q}$ is a normal extension, N is a normal subgroup of G. Any member of $H \cap N$ fixes r and ζ, hence fixes all of $\mathbb{K}$; thus $H \cap N = \{1\}$. The order of N is $[\mathbb{K} : \mathbb{Q}(\zeta)] = p$, and the order of H is $[\mathbb{K} : \mathbb{Q}(r)] = p - 1$. Therefore $|G| = |H||N|$, and G is a semidirect product with N normal.

Proposition 4.44 says that the action of an internal semidirect product is given by $\tau_h(n) = hnh^{-1}$. Let us identify τ_h. Let $h \in H = \mathrm{Gal}(\mathbb{K}/\mathbb{Q}(r))$ have $h(r) = r$ and $h(\zeta) = \zeta^k$, and let n in $N = \mathrm{Gal}(\mathbb{K}/\mathbb{Q}(\zeta))$ have $n(\zeta) = \zeta$ and $n(r) = r\zeta^l$. Then $hnh^{-1}(r) = hn(r) = h(r\zeta^l) = r\zeta^{kl}$, and $hnh^{-1}(\zeta) = hn(\zeta^{k'}) = h(\zeta^{k'}) = \zeta$. So if n sends r to $r\zeta^l$ and $h(\zeta) = \zeta^k$, then hnh^{-1} is the member of N sending r to ζ^{kl}.

This n is the member of N corresponding to $l \in \mathbb{F}_p$, and this h is the member of H corresponding to $k \in \mathbb{F}_p^\times$. We have just shown that hnh^{-1} is the member of N corresponding to $kl \in \mathbb{F}_p$. Hence the action corresponds to multiplication of $\mathbb{F}_p^\times$ on additive $\mathbb{F}_p$.

16. $[\mathbb{K} : \Bbbk] = \mathrm{Gal}(\mathbb{K}/\Bbbk)$, and $\mathrm{Gal}(\mathbb{K}/\Bbbk)$ is a subgroup of $\mathfrak{S}_n$. Being a subgroup, its order divides the order of $\mathfrak{S}_n$, which is $n!$.

17. In (a), the most general element of $\mathbb{K}$ is of the form $x + yr$ with x and y in $\Bbbk$, and its square is $(x^2 + y^2r^2) + 2xyr$. This is in $\Bbbk$ if and only if $xy = 0$, i.e., if and only if $x + yr$ is in $\Bbbk$ or in $r\Bbbk$. In other words, the only squares in $\mathbb{K}$ that lie in $\Bbbk$ are the obvious ones.

In (b), the same remarks apply unless the characteristic is 2. If the characteristic is 2, then $(x + yr)^2 = x^2 + y^2r^2$ and this is in $\Bbbk$ for all x and y. Hence every element of $\mathbb{K}$ is a square.

18. The finite group G may be regarded as a subgroup of the symmetric group $\mathfrak{S}_n$ for $n = |G|$. It was shown in Example 3 of Section 17 that there exists a finite Galois extension $\mathbb{K}$ of $\mathbb{Q}$ with Galois group $\mathfrak{S}_n$. Let $\Bbbk$ be the fixed field of G within $\mathfrak{S}_n$. Then $\mathrm{Gal}(\mathbb{K}/\Bbbk) = G$.

19. The polynomial in question in fixed by every element of the Galois group. Hence its coefficients are in the subfield of $\mathbb{K}$ fixed by all elements of $\mathrm{Gal}(\mathbb{K}/\Bbbk)$. This is $\Bbbk$.

20. For (a), define $F(X) = \prod_{j=1}^{n} (X - x_j)$. For φ in H, we have $F^{\varphi}(X) = \prod_{j=1}^{n} (X - \varphi(x_j)) = \prod_{j=1}^{n} (X - x_j) = F(X)$. Thus $F(X)$ is in $\mathbb{K}^H[X]$. Let $M(X)$ be the minimal polynomial of x_1 over $\mathbb{K}^H$. Since $F(x_1) = 0$, $M(X)$ divides $F(X)$. On the other hand, the equalities $M^{\varphi}(X) = M(X)$ and $M(x_1) = 0$ imply that $M(x_j) = 0$ for each j. Thus $M(X)$ has degree at least n, and we conclude that $F(X) = M(X)$.

In (b), n is the number of elements in an orbit of H and hence divides $|H|$.

In (c), when the isotropy subgroup of H at x_1 is trivial, $n = |H|$. Therefore $[\mathbb{K}^H(x_1) : \mathbb{K}^H] = n = |H| = [\mathbb{K} : \mathbb{K}^H]$, the last equality following from Corollary 9.37. Since $\mathbb{K}^H(x_1) \subseteq \mathbb{K}$, it follows that $\mathbb{K}^H(x_1) = \mathbb{K}$.

21. For (a), let $\varphi(z) = \frac{az+b}{cz+d}$ with $ad - bc \neq 0$. Then we have a substitution homomorphism of $\mathbb{C}[X]$ into $\mathbb{C}(z)$ fixing $\mathbb{C}$ and sending X into z. Since the range is a field, this factors through the field of fractions of $\mathbb{C}[X]$ to give a field mapping $\mathbb{C}(X) \to \mathbb{C}(z)$. We can regard the result as a map of $\mathbb{C}(z)$ into itself, and we write the map of $\mathbb{C}(z)$ into itself as $\Phi_{\varphi^{-1}}$. The formula is $\Phi_{\varphi^{-1}}(r) = r \circ \varphi$ for $r = r(z)$ in $\mathbb{C}(z)$. Then $\Phi_{\psi\varphi}(r) = r \circ (\psi\varphi)^{-1} = (r \circ \varphi^{-1}) \circ \psi^{-1} = \Phi_{\psi}(r \circ \varphi^{-1}) = \Phi_{\varphi}(\Phi_{\psi}(r))$, and hence $\Phi_{\psi\varphi} = \Phi_{\psi} \circ \Phi_{\varphi}$. From this it follows that $\Phi_{\varphi^{-1}}$ is a two-sided inverse of Φ_{φ}. Hence Φ_{φ} is an automorphism.

For (b), $\Phi_{\sigma}(w(z)) = w(\sigma^{-1}(z)) = (-z)^2 + (-z)^{-2} = z^2 + z^{-2} = w(z)$, and $\Phi_{\tau}(w(z)) = w(\tau^{-1}(z)) = (1/z)^2 + (1/z)^{-2} = z^2 + z^{-2} = w(z)$. Since $\Phi_{\varphi\psi} = \Phi_{\varphi}\Phi_{\psi}$ by (a), it follows that every element of H fixes w. Since each Φ_{φ} is a field automorphism, $\mathbb{C}(w)$ lies in $\mathbb{K}^H$.

For (c), we know from (b) that $\mathbb{C}(w) \subseteq \mathbb{K}^H$. The orbit of z under H has 4 elements, and Problem 20a shows that the minimal polynomial of z over $\mathbb{K}^H$ has degree 4 and is equal to

$$F(X) = (X-z)(X+z)(X-z^{-1})(X+z^{-1}) = (X^2-z^2)(X^2-z^{-2}) = X^4 - w(z)X^2 + 1.$$

The polynomial $F(X)$ is irreducible over $\mathbb{K}^H$, and its formula shows that its coefficients are in the smaller field $\mathbb{C}(w)$. Hence it is irreducible over $\mathbb{C}(w)$ and is the minimal polynomial of z over $\mathbb{C}(w)$.

For (d), (c) shows that $[\mathbb{K}^H(z) : \mathbb{C}(w)] = 4$. Problem 20c shows that $\mathbb{K} = \mathbb{K}^H(z)$, and hence $[\mathbb{K} : \mathbb{C}(w)] = 4$. Since $[\mathbb{K} : \mathbb{C}(w)] = [\mathbb{K} : \mathbb{K}^H]\,[\mathbb{K}^H : \mathbb{C}(w)]$ and since $[\mathbb{K} : \mathbb{K}^H] = 4$ by Corollary 9.37, $\mathbb{K}^H = \mathbb{C}(w)$.

22. For (a), let $\mathbb{L} = \mathbb{K}(\sqrt{u}\,)$ and $\mathbb{K} = \Bbbk(\sqrt{v}\,)$. The minimal polynomial of $\sqrt{u}$ over $\mathbb{K}$ is $X^2 - u$, and this must divide the minimal polynomial of $\sqrt{u}$ over $\Bbbk$. The degree of the latter polynomial equals $[\Bbbk(\sqrt{u}\,) : \Bbbk]$, which must divide 4. Hence it must be 2 or 4. If it is 2, then $X^2 - u$ lies in $\Bbbk[X]$, and u is in $\Bbbk$. We return to this case in a moment. Suppose that the minimal polynomial of $\sqrt{u}$ over $\Bbbk$ has degree 4. Let us write $u = r + s\sqrt{v}$ for some r and s in $\Bbbk$. Then $\sqrt{u}$ is a root of $(X^2 - r - s\sqrt{v}\,)(X^2 - r + s\sqrt{v}) = (X^2 - r)^2 - s^2 v = X^4 - 2rX^2 + (r^2 - s^2 v)$, which is a quartic polynomial in $\Bbbk[X]$. Since the minimal polynomial over $\Bbbk$ has degree 4, this is the minimal polynomial and is irreducible. Thus (a) holds with $r = \sqrt{u}$.

The remaining case is that u is in $\Bbbk$ but $\sqrt{u}$ is not in $\Bbbk$. Consider $\pm\sqrt{u} \pm \sqrt{v}$. None of these is in $\Bbbk$. The computation

$$\begin{aligned}
&(X + \sqrt{u} + \sqrt{v}))(X + \sqrt{u} - \sqrt{v})(X - \sqrt{u} + \sqrt{v})(X - \sqrt{u} - \sqrt{v}) \\
&\qquad = ((X + \sqrt{u})^2 - v)((X - \sqrt{u})^2 - v) \\
&\qquad = (X^2 + u - v + 2X\sqrt{u})(X^2 + u - v - 2X\sqrt{u}) \\
&\qquad = (X^2 + u - v)^2 - 4uX^2 = X^4 + 2uX^2 - 2vX^2 + (u - v)^2 - 4uX^2 \\
&\qquad = X^4 - 2(u + v)X^2 + (u - v)^2 = X^4 + bX^2 + c
\end{aligned}$$

shows that these are all roots of a quartic polynomial in $\Bbbk[X]$ of the correct kind, and the question concerns its irreducibility over $\Bbbk$. As in the previous paragraph, reducibility implies that it is the product of two irreducible quadratic members of $\Bbbk[X]$. Then the product of two of the first-order factors is in $\Bbbk[X]$, and the sum of those two roots must be in $\Bbbk$. The six possible sums of pairs of roots are $\pm\sqrt{u}$, $\pm\sqrt{v}$, and 0 twice. Since $\sqrt{u}$ and $\sqrt{v}$ are not in $\Bbbk$, the irreducible quadratic must be $X^2 - (\sqrt{u} + \sqrt{v})^2$ or $X^2 - (\sqrt{u} - \sqrt{v})^2$. However, the fact that $\sqrt{u}$ is not in $\mathbb{K} = \Bbbk(\sqrt{v}\,)$ implies that neither of $\sqrt{u} \pm \sqrt{v}$ is in $\Bbbk$. Thus the quartic polynomial is indeed irreducible. This completes (a).

In (b), we have $4 = [\mathbb{L} : \Bbbk] = [\mathbb{L} : \Bbbk(r)]\,[\Bbbk(r) : \Bbbk] = 4[\mathbb{L} : \Bbbk(r)]$. Thus $[\mathbb{L} : \Bbbk(r)] = 1$, and $\mathbb{L} = \mathbb{K}(r)$.

In (c), suppose that $c = t^2$ for the given $F(X)$. Find members u and v of $\Bbbk$ with $-2(u + v) = b$ and $u - v = t$. Then the displayed computation in (a) shows that $\pm\sqrt{u} \pm \sqrt{v}$ are the roots of $X^4 - 2(u + v)X^2 + (u - v)^2 = X^4 + bX^2 + c$. The given root r must be one of these. Say that $r = \sqrt{u} + \sqrt{v}$ without loss of generality. Since $[\Bbbk(r) : \Bbbk] = 4$ and $[\mathbb{L} : \Bbbk] = 4$ and $\Bbbk(r) \subseteq \mathbb{L}$, we have $\mathbb{L} = \Bbbk(r)$. On the other hand, $\Bbbk(r) \subseteq \Bbbk(\sqrt{u}, \sqrt{v}\,)$, and $[\Bbbk(\sqrt{u}, \sqrt{v}\,) : \Bbbk] = [\Bbbk(\sqrt{u}, \sqrt{v}) : \Bbbk(\sqrt{u})]\,[\Bbbk(\sqrt{u}) : \Bbbk] \leq$

$2 \cdot 2 = 4$. Hence $\Bbbk(\sqrt{u}, \sqrt{v}\,) = \Bbbk(r) = \mathbb{L}$. Then all four roots $\pm\sqrt{u} \pm \sqrt{v}$ of $F(X)$ lie in $\mathbb{L}$, $\mathbb{L}$ is the splitting field of $F(X)$ over $\Bbbk$, and $\mathbb{L}/\Bbbk$ is normal. The Galois group is generated by one element that sends $\sqrt{u}$ to $-\sqrt{u}$ and fixes $\sqrt{v}$, and by a second element that fixes $\sqrt{u}$ and sends $\sqrt{v}$ to $-\sqrt{v}$. Hence it is $C_2 \times C_2$.

Conversely suppose that $\mathbb{L}/\Bbbk$ is normal with Galois group $G = \mathrm{Gal}(\mathbb{L}/\Bbbk) = C_2 \times C_2$. Let an irreducible polynomial $X^4 + bX^2 + c$ in $\Bbbk[X]$ with a root r in $\mathbb{L}$ be given. Since $\mathbb{L}/\Bbbk$ is normal, $X^4 + bX^2 + c$ splits in $\mathbb{L}$. Let the four roots be $\pm r$ and $\pm s$. The square u of any of these roots satisfies $u^2 + bu + c = 0$ and therefore lies in a quadratic extension within $\mathbb{K}$, the same quadratic extension for each root. Let us define $\mathbb{K}$ to be this extension. Then $\mathbb{K} = \Bbbk(\sqrt{b^2 - 4c}\,)$. Because of the structure of G, there exists exactly one element σ in G whose fixed field is $\mathbb{K}$. The minimal polynomial of $\pm r$ over $\mathbb{K}$ is $X^2 + \frac{1}{2}b \pm \frac{1}{2}\sqrt{b - 4c}$ for one of the two choices of sign, and the minimal polynomial of $\pm s$ over $\mathbb{K}$ is the one for the other choice of sign. The element σ must then permute the roots of each of these polynomials, and it follows that $\sigma(r) = \pm r$ and $\sigma(s) = \pm s$. Since neither r nor s is in $\mathbb{K}$, we must in fact have $\sigma(r) = -r$ and $\sigma(s) = -s$. Therefore $\sigma(rs) = rs$. One of the other two nontrivial members τ of G has $\tau(r) = s$. Since $\tau^2 = \tau$, we have $\tau(s) = r$. Thus $\tau(rs) = rs$, and we see that every member of G fixes rs. Consequently rs is in $\Bbbk$. Since rs is equal for some choice of signs to

$$\pm\sqrt{-\tfrac{1}{2}b + \tfrac{1}{2}\sqrt{b^2 - 4c}}\sqrt{-\tfrac{1}{2}b - \tfrac{1}{2}\sqrt{b^2 - 4c}} = \pm\sqrt{\tfrac{1}{4}b^2 - \tfrac{1}{4}(b^2 - 4c)} = \pm\sqrt{c},$$

$\sqrt{c}$ is in $\Bbbk$. In other words, c is the square of a member of $\Bbbk$, as asserted.

In (d), suppose that $c^{-1}(b^2 - 4c)$ for the given $F(X)$ is a square in $\Bbbk$. Arguing with r^2 as in (c), we see that $\mathbb{K} = \Bbbk(\sqrt{b^2 - 4c}\,)$. Making the same computation as in the display just above, we see that $rs = \sqrt{c}$. Since $c^{-1}(b^2 - 4c)$ is a square in $\Bbbk$, $\sqrt{c}$ lies in $\mathbb{K}$. One of the roots, say r, lies in $\mathbb{L}$, and the product $rs = \sqrt{c}$ lies in $\mathbb{K}$, hence in $\mathbb{L}$. We conclude that $\pm r$ and $\pm s$ all lie in $\mathbb{L}$. In other words, $\mathbb{L}$ is the splitting field of $F(X)$ over $\Bbbk$ and is normal. Thus $\mathbb{L}/\Bbbk$ is normal. The Galois group must be either $C_2 \times C_2$ or C_4. If it is $C_2 \times C_2$, then (c) shows that $\sqrt{c}$ lies in $\Bbbk$. Under our assumption that $c^{-1}(b^2 - 4c)$ is a square, $\sqrt{b^2 - 4c}$ lies in $\Bbbk$. Consequently $F(X)$ is reducible, contradiction. We conclude that the Galois group is C_4.

Conversely suppose that $\mathbb{L}/\Bbbk$ is normal with Galois group $G = \mathrm{Gal}(\mathbb{L}/\Bbbk) = C_4$. Let an irreducible polynomial $X^4 + bX^2 + c$ in $\Bbbk[X]$ with a root r in $\mathbb{L}$ be given. Arguing with r^2, we see that r^2 lies in $\mathbb{K} = \Bbbk(\sqrt{b^2 - 4c}\,)$. Since $\mathbb{L}$ is generated by $\Bbbk$ and r, a generator of G cannot send r into $\pm r$. On the other hand, some element of G has to send r into $-r$ since $-r$ is a root of the given polynomial. Therefore $\sigma^2(r) = -r$. Then we have $\sigma(r\sigma(r)) = \sigma(r)\sigma^2(r) = -r\sigma(r)$, and we see that $\sigma^2(r\sigma(r)) = r\sigma(r)$. Consequently $r\sigma(r)$ lies in $\mathbb{K}$. Computing as in (c), we find that $\sqrt{c}$ lies in $\mathbb{K}$. This member of $\mathbb{K}$ has its square in $\Bbbk$, and Problem 17 shows that $\sqrt{c}$ lies in $\Bbbk$ or in the set of products $\Bbbk\sqrt{b^2 - 4c}$. By (c), $\sqrt{c}$ cannot lie in $\Bbbk$, and therefore $\sqrt{c} = d\sqrt{b^2 - 4c}$ for some d in $\Bbbk$. Hence $c^{-1}(b^2 - 4c) = d^{-2}$ for an element d of $\Bbbk$.

For (e), one can take $\mathbb{L} = \mathbb{K}(\sqrt[4]{2})$ and $\mathbb{K} = \mathbb{Q}(\sqrt{2})$. We can easily see directly that $\mathbb{L}$ is not normal. But let us use (c) and (d). The minimal polynomial $F(X)$ in question is $X^4 + 2$, with $b = 0$ and $c = 2$. The conditions in (c) and (d) say that $\mathbb{L}/\Bbbk$ is normal if and only if either 2 is a square in $\mathbb{Q}$ or -1 is a square in $\mathbb{Q}$. Neither condition is satisfied, and hence $\mathbb{L}/\Bbbk$ is not normal.

23. A cubic will be irreducible if it is divisible by no degree-one factor over $\mathbb{Q}$, hence if it has no root in $\mathbb{Q}$. Since these cubics are monic and are in $\mathbb{Z}[X]$, they will be irreducible if they have no integer root. An integer root must divide the constant term, and we check that neither of ± 1 is a root in either case. Hence both cubics are irreducible. By Problem 13 the Galois group in each case is a transitive subgroup of $\mathfrak{S}_3$, hence is $\mathfrak{S}_3$ or $\mathfrak{A}_3$. The discriminant $-4p^3 - 27q^2$ is 81 in the first case and -31 in the second case; this is a square in the first case but not in the second case. Thus $X^3 - 3X + 1$ has Galois group $\mathfrak{A}_3$, and $X^3 + X + 1$ has Galois group $\mathfrak{S}_3$.

24. The extension field is either $\mathbb{K}$ itself, in which case the Galois group remains $\mathfrak{S}_3$, or it is $\mathbb{L} = \mathbb{K}[\sqrt{-3}\,]$. Since $\mathbb{K}/\mathbb{Q}$ is normal, $\mathrm{Gal}(\mathbb{L}/\mathbb{K})$ is a normal subgroup of $\mathrm{Gal}(\mathbb{L}/\mathbb{Q})$ of order 2 with quotient isomorphic to $\mathrm{Gal}(\mathbb{K}/\mathbb{Q}) = \mathfrak{S}_3$. The groups of order 12 are classified in Problems 45–48 at the end of Chapter IV. Two such groups are abelian, one is $\mathfrak{A}_4$, and one is $D_6 \cong C_2 \times \mathfrak{S}_3$.

Write a general element of $\mathbb{L}$ as $a + b\sqrt{-3}$. Define $\tau(a + b\sqrt{-3}\,) = a - b\sqrt{-3}$. This is the nontrivial member of the 2-element group $\mathrm{Gal}(\mathbb{L}/\mathbb{K})$. If σ is in $\mathrm{Gal}(\mathbb{K}/\mathbb{Q})$, then σ extends to a member $\overline{\sigma}$ of $\mathrm{Gal}(\mathbb{L}/\mathbb{Q})$ by the definition $\overline{\sigma}(a + b\sqrt{-3}) = \sigma(a) + \sigma(b)\sqrt{-3}$. In fact, $\overline{\sigma}$ respects addition. To see that it respects multiplication, we compute

$$\begin{aligned}
\overline{\sigma}(a + b\sqrt{-3}\,)\overline{\sigma}(c + d\sqrt{-3}\,) &= \big(\sigma(a) + \sigma(b)\sqrt{-3}\big)\big(\sigma(c) + \sigma(d)\sqrt{-3}\big) \\
&= \big(\sigma(a)\sigma(c) - 3\sigma(b)\sigma(d)\big) + \big(\sigma(b)\sigma(c) + \sigma(a)\sigma(d)\big)\sqrt{-3} \\
&= \sigma(ac - 3bd) + \sigma(bc + ad)\sqrt{-3} \\
&= \overline{\sigma}\big((ac - 3bd) + (bc + ad)\sqrt{-3}\big) \\
&= \overline{\sigma}\big((a + b\sqrt{-3}\,)(c + d\sqrt{-3}\,)\big).
\end{aligned}$$

It follows that $\mathrm{Gal}(\mathbb{L}/\mathbb{Q})$ is the direct product $C_2 \times \mathfrak{S}_3$, the subgroup C_2 being $\mathrm{Gal}(\mathbb{L}/\mathbb{K})$.

25. Yes. Let $\mathbb{L}$ be the intermediate field corresponding to the subgroup $\{(1), (1\ 2)\}$. Since the subgroup is not normal, $\mathbb{L}/\Bbbk$ is not normal. Let r be any element of $\mathbb{L}$ not in $\Bbbk$. Then the minimal polynomial of r over $\Bbbk$ has degree 3, and it does not split in $\mathbb{L}$ since $\mathbb{L}/\Bbbk$ is not normal. Its splitting field has to be something between $\mathbb{L}$ and $\mathbb{K}$, and the only choice is $\mathbb{K}$.

26. Yes, substitute and check it.

28. In (a), direct expansion of the right side gives $(X - r)(X^2 + rX + (r^2 + p)) = X^3 + pX - r^3 - pr$. Since $-r^3 - pr = q$, the assertion follows.

For (b), let us check that $r^2(-4p^3 - 27q^2) = (-3r^2 - 4p)(3q + 2pr)^2$, from which the assertion follows. In fact, the right side equals

$$\begin{aligned}
&-(3r^2 + 4p)(9q^2 + 12pqr + 4p^2r^2) \\
&\quad = -(36pq^2 + 48p^2qr + 27q^2r^2 + 16p^3r^2 + 36pqr^3 + 12p^2r^4) \\
&\quad = -r^2(4p^3 + 27q^2) - 12p^3r^2 - 36pq^2 - 48p^2qr - 36pqr^3 - 12p^2r^4 \\
&\quad = -r^2(4p^3 + 27q^2) - 12p^3r^2 - 36pq^2 - 48p^2qr \\
&\qquad - 36pq(-pr - q) - 12p^2r(-pr - q) \\
&\quad = -r^2(4p^3 + 27q^2).
\end{aligned}$$

29. No. For example, $F(X)$ could have three real roots, and then $\mathbb{K}$ would be a subfield of $\mathbb{R}$. A concrete example is $X^3 - 12X + 1$, which is < 0 at -4, is > 0 at 0, is < 0 at 1, and is > 0 at 4; the Intermediate Value Theorem shows that $F(X)$ has three real roots.

30. The group in question is a subgroup of $\mathfrak{S}_5$. It is transitive because of the irreducibility, and it is a subgroup of $\mathfrak{A}_5$ since the discriminant is a square. Problem 13c shows that it contains a 5-cycle. The other cycle structures in $\mathfrak{A}_5$ are the 3-cycles and the pairs of 2-cycles. If a 3-cycle is present, then the group is all of $\mathfrak{A}_5$ because 15 divides its order, all groups of order 15 are cyclic, and $\mathfrak{A}_5$ contains no subgroup of order 30, being simple.

Suppose there are no 3-cycles. A Sylow 2-subgroup may be taken to be a subgroup of $H = \{(1), (1\ 2)(3\ 4), (1\ 3)(2\ 4), (1\ 4)(2\ 3)\}$, and it acts on the group of powers of a 5-cycle. The only nontrivial action of a 2-element group on a 5-element group carries elements to their inverses. Since no nontrivial element of H commutes with a 5-cycle (because $\mathfrak{S}_5$ has no elements of order 10), the Sylow 2-subgroup contains at most two elements. If it is trivial, then the group in question is of order 5, consisting of the powers of a 5-cycle. If the Sylow 2-subgroup has 2 elements, we obtain a semidirect product of a 2-element group with the powers of the 5-cycle, and the result has to be isomorphic to the dihedral group D_5.

Thus the only possibilities are C_5, D_5, and $\mathfrak{A}_5$.

31. Computation shows that the discriminant is $2^{12}7^219^2$, which is a square. By Proposition 9.63 the Galois group is a subgroup of $\mathfrak{A}_5$. Modulo 3, the given polynomial is $2 + 2x + x^5$ and is irreducible. By Theorem 9.64 the Galois group contains a 5-cycle. The given polynomial factors as $(7 + x)(7 + 10x + 7x^2 + x^3)$ modulo 11, and Theorem 9.64 shows that the Galois group contains a 3-cycle. The 5-cycle and 3-cycle generate all of $\mathfrak{A}_5$, and thus the Galois group is $\mathfrak{A}_5$.

32. Write e and f for e_1 and f_1. The proof of Theorem 9.64 showed that $f' = f$. Then $e'f' = |G_P| = |G|/g = efg/g = ef = ef'$, and $e' = e$.

33. If $\mathfrak{p}T = \prod_i P_i^{e(P_i|\mathfrak{p})}$ and $P_iU = \prod_j Q_{ij}^{e(Q_{ij}|P_i)}$, then $\mathfrak{p}U = \prod_i (P_i^{e(P_i|\mathfrak{p})}U) = \prod_i (P_iU)^{e(P_i|\mathfrak{p})} = \prod_i (\prod_j Q_{ij}^{e(Q_{ij}|P_i)})^{e(P_i|\mathfrak{p})}$. Hence $e(P_i|\mathfrak{p}) = e(Q_{ij}|P_i))e(P_i|\mathfrak{p})$. The formula for the f's follows from Corollary 9.7.

34. Corollary 9.58 shows that the norm and the trace are the product and sum of $a + b\sqrt{m}$ and $a - b\sqrt{m}$. Hence they are $a^2 - b^2 m$ and $2a$. This proves (a).

In (b), the minimal polynomial of $r = a + b\sqrt{m}$ has degree 2 if $b \neq 0$, and this is the same as the degree of the field polynomial. Hence the two polynomials are equal, and the minimal polynomial is $X^2 - (\operatorname{Tr} r)X + N(r)$. An algebraic integer is an algebraic element whose minimal polynomial over $\mathbb{Q}$ has integer coefficients, and (b) follows.

In (c), if $r = a + b\sqrt{m}$ is a unit with inverse s, then $N(r)N(s) = N(rs) = N(1) = 1$ shows that $N(r)$ is a unit with inverse $N(s)$. Conversely if r is in T with $N(r) = \pm 1$, then $r(a - b\sqrt{m}) = \pm 1$, and $\pm(a - b\sqrt{m})$ is an inverse element in T.

For (d), $\sqrt{2} - 1$ is a unit in the algebraic integers of $\mathbb{Q}[\sqrt{2}]$. Its inverse is $\sqrt{2} + 1$.

35. With respect to the ordered basis $\left(1, \sqrt[3]{2}, (\sqrt[3]{2})^2\right)$, the matrix of multiplication by $a + b\sqrt[3]{2} + c(\sqrt[3]{2})^2$ is

$$\begin{pmatrix} a & 2c & 2b \\ b & a & 2c \\ c & b & a \end{pmatrix}.$$

The trace and norm are the trace and determinant of this matrix, namely $3a$ and $a^3 + 2b^3 + 4c^3 - 6abc$.

36. In (a), if ξ is any number algebraic over $\mathbb{Q}$ of degree r, then the norm relative to $\mathbb{Q}(r)/\mathbb{Q}$ of ξ is $(-1)^r M(0)$, where $M(X)$ is the minimal polynomial of ξ over $\mathbb{Q}$. Since $M(1 - (1 - \xi)) = 0$, the minimal polynomial of $1 - \xi$ is the polynomial $M(1 - X)$ adjusted so as to be monic. That is, it is $P(X) = (-1)^r M(1 - X)$. Hence the norm of $1 - \xi$ is $(-1)^r P(0) = (-1)^{2r} M(1) = M(1)$. In the case of the given ζ, the minimal polynomial of ζ is $\Phi_n(X)$, and therefore the norm of $1 - \zeta$ is $\Phi_n(1)$.

For (b), division of both sides of the identity $\prod_{d|n} \Phi_d(X) = X^n - 1$ by $X - 1$ gives $\prod_{d|n,\ d>1} \Phi_d(X) = X^{n-1} + X^{n-2} + \cdots + 1$. Therefore $\prod_{d|n,\ d>1} \Phi_d(1) = n$.

If n is a prime power, say with $n = p^k$, let us see by induction on k that $\Phi_n(1) = p$. The base case of the induction is $k = 1$, and the result of the previous paragraph applies. Assuming that $\Phi_n(1) = p$ for $n = p^k$, we have $p^{k+1} = \prod_{l=1}^{k+1} \Phi_{p^l}(1) = \Phi_{p^{k+1}}(1) \prod_{l=1}^{k} p$. Therefore $\Phi_{p^{k+1}}(1) = p$, and the induction is complete.

Inducting on n, let us now show that $\Phi_n(1) = 1$ if n is divisible by more than one positive prime. The base case of the induction is $n = 2$. Assume that $n = p_1^{k_1} \cdots p_r^{k_r}$ and that the result is known for integers less than n. We may assume that n is divisible by at least two positive primes. Then

$$n = \prod_{d|n,\ d>1} \Phi_d(1) = \prod_{s=1}^{r} \Big(\prod_{l=1}^{k_s} \Phi_{p_s^l}(1) \Big) \prod_{\text{other } d} \Phi_d(1),$$

where the "other d" are the divisors of n that are divisible by at least two primes. These include n itself. So one of the corresponding factors is $\Phi_n(1)$, and the others are 1 by the inductive hypothesis. The factor in parentheses is $p_l^{k_l}$ by the result of the previous paragraph, and the product of the factors in parentheses is n. Therefore $\Phi_n(1) = 1$, and the induction is complete.

37. For (a), the imaginary part of $p^{-1}x \pm p^{-1}\sqrt{-1}$ is not an integer, and therefore $p^{-1}(x \pm \sqrt{-1}\,)$ is not a Gaussian integer. Consequently p does not divide either of $x \pm \sqrt{-1}$. Since p divides $x^2 + 1$ in $\mathbb{Z}$ and hence in $\mathbb{Z}[\sqrt{-1}\,]$, p is not prime in $\mathbb{Z}[\sqrt{-1}\,]$.

For (b), it follows since p is not prime that $p = \alpha\beta$ nontrivially in $\mathbb{Z}[\sqrt{-1}\,]$. Then $p^2 = N(p) = N(\alpha)N(\beta)$. Problem 34c shows that nontrivial factorization implies that $N(\alpha)$ and $N(\beta)$ are not units. Thus they are both p. If $\alpha = a + b\sqrt{-1}$, then the equation $p = N(\alpha)$ says that $p = a^2 + b^2$.

38. Let N be the norm function in $\mathbb{Q}(\sqrt{-2})$. Since p divides $x^2 + 2 = (x + \sqrt{-2}\,)(x - \sqrt{-2}\,)$ and since neither of $p^{-1}(x \pm \sqrt{-2}\,)$ is of the form $a + b\sqrt{-2}$ with a and b in $\mathbb{Z}$, p is not prime in $\mathbb{Z}[\sqrt{-2}]$. Write $p = \alpha\beta$ nontrivially. Then $p^2 = N(p) = N(\alpha)N(\beta)$ and $N(\alpha) = N(\beta) = p$. If $\alpha = a + b\sqrt{-2}$, then $p = N(\alpha)$ says that $p = a^2 + 2b^2$.

39. This is similar to Problem 38 except that the members of the ring are of the form $a + b\sqrt{-3}$ with a, b in $\mathbb{Z}$ or a, b in $\mathbb{Z} + \frac{1}{2}$. Thus $p = N(\alpha)$ says that $p = a^2 + 3b^2$ either with a, b in $\mathbb{Z}$ or with a, b in $\mathbb{Z} + \frac{1}{2}$. In the latter case, let $\omega^{\pm 1} = \frac{1}{2}(-1 - \sqrt{-3}\,)$. These have $N(\omega^{\pm 1}) = 1$. Therefore

$$\begin{aligned} p = N(\alpha) &= N(\alpha\omega^{\pm 1}) = N\big((a + b\sqrt{-3}\,)(-\tfrac{1}{2} \pm \tfrac{1}{2}\sqrt{-3}\,)\big) \\ &= N\big(\tfrac{1}{2}(-a \mp 3b) + \tfrac{1}{2}(\pm a - b)\sqrt{-3}\big) \\ &= \big(\tfrac{1}{2}(-a \mp 3b)\big)^2 + 3\big(\tfrac{1}{2}(\pm a - b)\sqrt{-3}\big)^2. \end{aligned}$$

Since a, b are in $\mathbb{Z} + \frac{1}{2}$, one of $a + b$ and $a - b$ is even, and the other is odd, the sum $2a$ being odd. If $a + b$ is even, then $a - 3b$ is even since their difference $4b$ is even, and vice versa. Hence one of the two choices of sign exhibits p as $c^2 + 3d^2$ with c, d in $\mathbb{Z}$.

40. Write $\mathbb{L}' = \Bbbk(x)$ by the Theorem of the Primitive Element, and let $\mathbb{K}$ be a splitting field of the minimal polynomial of x over $\Bbbk$. Then $\mathbb{K}$ is a finite Galois extension of $\Bbbk$ by Corollary 9.30, and we have $\Bbbk \subseteq \mathbb{L} \subseteq \mathbb{L}' \subseteq \mathbb{K}$. For a in $\mathbb{L}'$ and b in $\mathbb{L}$, Corollary 9.58 says that $N_{\mathbb{L}'/\Bbbk}(a) = \prod_{\sigma \in G/H'} \sigma(a)$, $N_{\mathbb{L}/\Bbbk}(b) = \prod_{\sigma \in G/H'} \sigma(b)$, and $N_{\mathbb{L}'/\mathbb{L}}(a) = \prod_{\tau \in H/H'} \tau(a)$. Hence $N_{\mathbb{L}/\Bbbk}(N_{\mathbb{L}'/\mathbb{L}}(a)) = \prod_{\sigma \in G/H} \sigma(N_{\mathbb{L}'/\mathbb{L}}(a)) = \prod_{\sigma \in G/H} \sigma\big(\prod_{\tau \in H/H'} \tau(a)\big) = \prod_{\sigma \in G/H} \prod_{\tau \in H/H'} \sigma\tau(a) = \prod_{\sigma \in G/H'} \sigma(a) = N_{\mathbb{L}'/\Bbbk}(a)$. The formula for traces follows similarly by replacing the products by sums in the above computation.

41. Since P is symmetric, $P(X_{\sigma(1)}, \dots, X_{\sigma(n)}) = P(X_1, \dots, X_n)$ for every permutation σ. Therefore $P(r_{\sigma(1)}, \dots, r_{\sigma(n)}) = P(r_1, \dots, r_n)$ for every σ. Problem 39e at the end of Chapter VIII implies that $P(r_1, \dots, r_n) = Q(s_1, \dots, s_n)$ for a polynomial $Q(X_1, \dots, X_n)$ in $\Bbbk[X_1, \dots, X_n]$, where $s_1, \dots, s_n$ are the elementary symmetric polynomials in $r_1, \dots, r_n$. The elements $s_1, \dots, s_n$ are the coefficients of $F(X)$, up to sign, and hence are in $\Bbbk$. Therefore $P(r_1, \dots, r_n) = Q(s_1, \dots, s_n)$ is in $\Bbbk$.

42. Inspection of the formula gives $H_1(X) = \prod_{i=1}^m G(X - r_i)$. For each i, we can expand $G(X - r_i)$ in powers of X as

$$G(X - r_i) = X^n + b_{n-1}(r_i)X^{n-1} + \cdots + b_1(r_i)X + b_0(r_i),$$

and each of $b_{n-1}, \ldots, b_0$ is a member of $\Bbbk[X]$. When we multiply these for $1 \le i \le m$, each power of X in the product has a coefficient that is unchanged if we permute $r_1, \ldots, r_n$. Problem 41 says that the coefficient of each power of X is therefore in $\Bbbk$. Thus $H_1(X)$ is in $\Bbbk[X]$. A similar argument shows that $H_2(X)$ is in $\Bbbk[X]$.

43. For (a), we use $F(X) = X^2 - 2$ and $G(X) = X^2 - 3$ in the previous problem. Then $\sqrt{2} + \sqrt{3}$ is a root of

$$(X - (\sqrt{2} + \sqrt{3}))(X - (\sqrt{2} - \sqrt{3}))(X - (-\sqrt{2} + \sqrt{3}))(X - (-\sqrt{2} - \sqrt{3})),$$

which must have coefficients in $\mathbb{Q}$.

44. Proposition 4.40 extends the action by an element σ in $\mathfrak{S}_n$ uniquely from the set $\{r_1, \ldots, r_n\}$ to $\Bbbk[r_1, \ldots, r_n]$ fixing $\Bbbk$. The extended σ is a one-one homomorphism of $\Bbbk[r_1, \ldots, r_n]$ into itself, hence into $\Bbbk(r_1, \ldots, r_n)$. It extends uniquely to a field mapping of $\Bbbk(r_1, \ldots, r_n)$ into itself by Proposition 8.6. The homomorphism corresponding to a composition is the composition of the homomorphisms, and consequently the homomorphism corresponding to σ^{-1} is a two-sided inverse of the homomorphism corresponding to σ. Thus the extension of σ is an automorphism, as required.

Conclusion (a) is immediate from Problem 20a. For (b), since $\mathbb{K}$ is generated by $\Bbbk$ and $r_1, \ldots, r_n$, $\mathbb{K}$ is certainly generated by $\mathbb{K}^{\mathfrak{S}_n}$ and $r_1, \ldots, r_n$. We have arranged that $F(X)$ splits over $\mathbb{K}$, and hence $\mathbb{K}$ is the splitting field. Conclusion (d) follows from Corollary 9.37 once (c) is proved. Thus we are to prove (c).

The argument for (c) is similar to that in Problem 21. Since $F(X)$ is in $\mathbb{K}^{\mathfrak{S}_n}$, its coefficients are in $\mathbb{K}^{\mathfrak{S}_n}$. Thus $\Bbbk(u_1, \ldots, u_n) \subseteq \mathbb{K}^{\mathfrak{S}_n}$. Consequently Corollary 9.37 gives $n! = [\mathbb{K} : \mathbb{K}^{\mathfrak{S}_n}] \le [\mathbb{K} : \Bbbk(u_1, \ldots, u_n)]$. Problem 16 shows that the right side divides $n!$. Therefore equality holds throughout, and we see that $[\mathbb{K} : \mathbb{K}^{\mathfrak{S}_n}] = [\mathbb{K} : \Bbbk(u_1, \ldots, u_n)]$. Since $\Bbbk(u_1, \ldots, u_n) \subseteq \mathbb{K}^{\mathfrak{S}_n}$, we must have $\Bbbk(u_1, \ldots, u_n) = \mathbb{K}^{\mathfrak{S}_n}$.

45. For (a), we have

$$\begin{aligned}
c_1 &= \sum_i \theta_i = 2\sum_{i<j} s_i s_j = 2p,\\
c_2 &= \sum_{i<j} \theta_i\theta_j = \sum_{i<j} s_i^2 s_j^2 + 3\sum_{i<j<k} s_i s_j s_k (s_i + s_j + s_k) + 6 s_1 s_2 s_3 s_4,\\
c_3 &= \theta_1\theta_2\theta_3 = \sum_{\substack{i,j,k\\ \text{unequal}}} s_i^3 s_j^2 s_k + 2 s_1 s_2 s_3 s_4 \Big(\sum_i s_i^2\Big)\\
&\qquad + 2\sum_{i<j<k} s_i^2 s_j^2 s_k^2 + 4 s_1 s_2 s_3 s_4 \Big(\sum_{i<j} s_i s_j\Big).
\end{aligned}$$

Part (b) is a calculation with symmetric polynomials and is omitted. For (c), we have

$$\begin{aligned}\theta_1 - \theta_2 &= -(s_1 - s_4)(s_2 - s_3),\\ \theta_1 - \theta_3 &= -(s_1 - s_3)(s_2 - s_4),\\ \theta_2 - \theta_3 &= -(s_1 - s_2)(s_3 - s_4).\end{aligned}$$

The square of the product of the left sides is the discriminant of the cubic resolvent, and the square of the product of the right sides is the discriminant of the given quartic.

46. In (a), the subgroups in question are

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

and $\mathfrak{A}_4$. In (b), one considers the possibilities for a Sylow 2-subgroup and is led to conclude that the only possibilities for the subgroups in question are the powers of a 4-cycle, the dihedral group (generated by H and $(1\ \ 2\ \ 3\ \ 4)$), and $\mathfrak{S}_4$. (The group H and any 2-cycle generate $\mathfrak{S}_4$, and thus the dihedral group cannot be generated by H and a 2-cycle.)

47. In (a), the discriminant reduces when $q = 0$ to $16p^4r - 128p^2r^2 + 256r^3 = 16r(p^4 - 8p^2r + 16r^2) = 16r(p^2 - 4r)^2$. This is 0 if $r = 0$ or $r = p^2/4$. If it is nonzero, it is a square if and only if r is a square. Hence in all cases it is a square if and only if r is a square.

In (b), let $Y = X^2$. The equation is $Y^2 + pY + r = 0$, which can be solved with a square root. For each of the two solutions, we can then solve for X with a square root. Hence all the roots lie in an extension obtained by adjoining at most three square roots. Thus $[\mathbb{K} : \mathbb{Q}]$ divides 8, and $|G|$ divides 8. Consequently G cannot have any element of order 3.

In (c), the irreducibility shows that the possibilities for G are as in Problem 46. Since r is a square, the discriminant is a square, by (a). Proposition 9.63 shows that the possibilities are as in Problem 46a. Part (b) rules out $\mathfrak{A}_4$, and then (c) follows.

In (d), r nonsquare and $F(X)$ irreducible implies that G is a transitive subgroup of $\mathfrak{S}_4$ but not a subgroup of $\mathfrak{A}_4$, by (a). Problem 46b shows that G is $\mathfrak{S}_4$, or the powers of a 4-cycle, or the dihedral group D_4. By (b), there is no element of order 3, and $\mathfrak{S}_4$ is therefore ruled out.

48. The polynomial remains irreducible when reduced modulo 2, and a prime factorization modulo 3 is $(X+2)(X^3+X^2+X+2)$. Thus G is a transitive subgroup of $\mathfrak{S}_4$ containing a 3-cycle. The discriminant is 257, not square. By Problem 46b, $G = \mathfrak{S}_4$.

49. Part (a) is just a computation; the answer is $2^{12}3^4$. The factorization in (b) is routine to check, and the only issue is the irreducibility of the cubic factor. For a cubic polynomial, irreducibility follows if the polynomial has no root in the field. Thus we need only verify that none of 0, 1, 2, 3, 4 is a root modulo 5.

For (c), the conclusion of (b) shows that the only possible reducibility over $\mathbb{Q}$ is into a degree-one factor and a cubic factor. For $X^4 + 8X + 12$ to have a degree-one factor, it must have a rational root, and this root must be an integer dividing 12. Let r be an integer dividing 12. If r is even, then $r^4 + 8r$ is divisible by 16, but 12 is not; so an even r cannot be a root. We are left with ± 1 and ± 3 as the possibilities, and we check that none of these is a root.

In (d), $F(X)$ is irreducible, and G is transitive. It is a subgroup of $\mathfrak{A}_4$ since the discriminant is a square. By (b) and Theorem 9.64, G contains a 3-cycle. Problem 46a shows therefore that $G = \mathfrak{A}_4$.

50. We saw in Problem 49 that $G = \mathfrak{A}_4$ for $X^4 + 8X + 12$, in Problem 47c that $G = \{(1), (1\ \ 2)(3\ \ 4), (1\ \ 3)(2\ \ 4), (1\ \ 4)(2\ \ 3)\}$ for $X^4 + 1$, in Problem 48 that $G = \mathfrak{S}_4$ for $X^4 + X + 1$, and via Eisenstein's criterion that $G = C_4$ for $\Phi_5(X)$. Since $X^4 - 2$ does not split in $\mathbb{Q}(\sqrt[4]{2})$, the Galois group in this case cannot be of order 4, and Problem 47d shows that G must be D_4 in this case.

51. For (a), let C correspond to a set of polynomials I of degree at most $n-1$. If C is cyclic, then I is at least a vector space over $\mathbb{F}$. If $F(X) = c_0 + c_1X + \cdots + c_{n-1}X^{n-1}$ is in I, then $XF(X) = c_0X + c_1X^2 + \cdots + c_{n-1}X^n$ is congruent modulo $(X^n - 1)$ to $c_{n-1} + c_0X + \cdots + c_{n-2}X^{n-1}$, which is in I since C is cyclic. Hence I is closed under multiplication by X mod $(X^n - 1)$ and hence under arbitrary multiplications modulo $(X^n - 1)$. Therefore I is an ideal in $\mathbb{F}[X]/(X^n - 1)$.

Conversely if I is an ideal in $\mathbb{F}[X]/(X^n - 1)$, then it is a vector space and is closed under multiplication by X mod $(X^n - 1)$ in $\mathbb{F}[X]/(X^n - 1)$. If $F(X) = c_0 + c_1X + \cdots + c_{n-1}X^{n-1}$ is in I, then $XF(X) = c_0X + c_1X^2 + \cdots + c_{n-1}X^n$ mod I has to be in I, and the corresponding member of C is $(c_{n-1}, c_0, c_1, \ldots, c_{n-2})$. Hence C is cyclic.

For the remaining parts, we identify the cyclic code C with the corresponding ideal I in $\mathbb{F}[X]/(X^n - 1)$. In (b), let the lowest degree of a member of I be $n-k$, and let $G(X)$ be a member of I of this degree. If there is a second member of this same degree, then their difference has lower degree since both polynomials are monic, and the difference must be in I, contradiction. Thus $G(X)$ is uniquely defined. Regard $G(X)$ as a member of $\mathbb{F}[X]$ of degree $n - k$, and let $M(X) = \text{GCD}(G(X), X^n - 1)$. Then we can choose $A(X)$ and $B(X)$ in $\mathbb{F}[X]$ with $A(X)(X^n - 1) + B(X)G(X) = M(X)$. Passing to $\mathbb{F}[X]/(X^n - 1)$, we have $B(X)G(X) \equiv M(X)$ mod $(X^n - 1)$. Therefore $M(X)$ is in the ideal I. Since the degree of $M(X)$ is at most $\deg G(X)$ and since $G(X)$ has the minimum degree among the nonzero members of I, either $M(X) = 0$ or $M(X) = G(X)$. The conclusion $M(X) = 0$ is ruled out since $M(X)$ is a greatest common divisor of nonzero polynomials, and thus $M(X) = G(X)$. Therefore $G(X)$ divides $X^n - 1$.

Let $\widetilde{I}$ be the inverse image of I in $\mathbb{F}[X]$. This is an ideal, it contains $G(X)$, and it contains no nonzero element of degree $< \deg G(X)$. Since $\widetilde{I}$ has to be principal, $\widetilde{I} = (G(X))$. In other words, $\widetilde{I}$ consists of all products of $G(X)$ by a member of $\mathbb{F}[X]$. If $F(X)G(X)$ is such a product, then the division algorithm gives $F(X)G(X) =$

$B(X)(X^n - 1) + R(X)$ with $R(X) = 0$ or $\deg R < n$. Since $G(X)$ divides $X^n - 1$, $G(X)$ divides $R(X)$. Therefore every member of $\widetilde{I}$ is congruent modulo $X^n - 1$ to a product $G(X)S(X)$ that is 0 or has degree $< n$. Then (c) is clear.

For (d), (b) showed that $G(X)$ divides $X^n - 1$ in $\mathbb{F}[X]$. Write $X^n - 1 = G(X)H(X)$. If $B(X)$ in $\mathbb{F}[X]/(X^n - 1)$ corresponds to a member of C, then (b) shows that $B(X) = F(X)G(X)$ for some $F(X)$ in $\mathbb{F}[X]$. Multiplying by $H(X)$ gives $B(X)H(X) = F(X)G(X)H(X) = F(X)(X^n - 1)$. Hence $B(X)H(X) \equiv 0 \bmod (X^n - 1)$. Conversely if $B(X)H(X) = A(X)(X^n - 1)$, then $B(X)H(X) = A(X)G(X)H(X)$, and $B(X) = A(X)G(X)$.

52. In (a), if r_1, r_2, r_3 denote the rows and if $v_1 = r_1 + r_3$, $v_2 = r_2$, and $v_3 = r_3$, then v_1, v_2, v_3 form a basis for the row space, and they cycle into one another when the columns are shifted in cyclic fashion. Consequently the code is cyclic. Part (b) involves looking at the 7 nonzero members of the space, and one can just do that directly.

In (c), one such matrix is

$$\mathcal{H} = \begin{pmatrix} 0\,0\,0\,1\,1\,0\,1 \\ 0\,0\,1\,1\,0\,1\,0 \\ 0\,1\,1\,0\,1\,0\,0 \\ 1\,1\,0\,1\,0\,0\,0 \end{pmatrix}.$$

A little check shows that the matrix product $\mathcal{H}\mathcal{G}^t$ is the 4-by-3 zero matrix, and hence $\mathcal{H}v = 0$ for each v in C. Thus C is contained in the null space of $\mathcal{H}$. The rank of $\mathcal{H}$ is 4 since the rows are certainly linearly independent. Since the sum of the rank and the dimension of the null space is the number of columns, namely 7, the dimension of the null space is 3. Therefore the null space is C and is no larger.

For (d), the general matrix $\mathcal{H}$ is to have n columns and $n - k$ rows. The entries of the top row are the coefficients of $H(X)$ with the constant term at the right, the coefficient of X in the next-to-last position, and so on. In each successive row these coefficients are shifted one position to the left.

Let $G(X) = g_0 + g_1X + \cdots + g_{n-k}X^{n-k}$ and $H(X) = h_0 + h_1X + \cdots + X^k$. We know that $\{0, XG(X), X^2G(X), \ldots, X^{k-1}G(X)\}$ is a basis of C. In terms of members of $\mathbb{F}^n$ the l^{th} such vector has the entries $g_0, g_1, \ldots, g_{n-k}$ beginning in the l^{th} position. The $(1, j)^{\text{th}}$ entry of $\mathcal{H}$ is h_{n-j} with 0's elsewhere in the row, and the $(i, j)^{\text{th}}$ entry is $h_{n-j-i+1}$ with 0's elsewhere in the row. The product of the i^{th} row of $\mathcal{H}$ and the l^{th} basis vector of C is $\sum_{j=n-k-i+1}^{n-i+1} h_{n-j-i+1}g_{j-l}$, which is the coefficient of $X^{n-i+1-l}$ in $G(X)H(X)$. Here $1 \le i \le n-k$ and $1 \le l \le k$, so that $2 \le i+l \le n$. Thus the power of X in question varies from 1 to $n-1$. Since $G(X)H(X) = X^n - 1$, the coefficient is 0. Thus C lies in the null space of $\mathcal{H}$. The same argument with rank as in the previous paragraph shows that C is exactly the null space.

53. Since $X^n - 1$ has derivative nX^{n-1}, we have $\text{GCD}(X^n - 1, nX^{n-1}) = 1$ when n is odd. Lemma 9.26 then shows that X^{n-1} is separable. If n is even, write $n = 2k$. Then $X^n - 1 = (X^k - 1)^2$ in characteristic 2 by Lemma 9.18, and hence every root has multiplicity at least 2.

54. In (a), we have $0 = P(\alpha^j) = c_0 + c_1\alpha^j + c_2\alpha^{2j} + \cdots + c_{n-1}\alpha^{(n-1)j}$ for $r \le j \le r+s$, and therefore the column vector $(c_0, c_1, \ldots, c_{n-1})$ satisfies

$$\begin{pmatrix} \alpha^r & \alpha^{2r} & \cdots & \alpha^{(n-1)r} \\ \alpha^{r+1} & \alpha^{2(r+1)} & \cdots & \alpha^{(n-1)(r+1)} \\ & & \vdots & \\ \alpha^{r+s} & \alpha^{2(r+s)} & \cdots & \alpha^{(n-1)(r+s)} \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \\ c_2 \\ \vdots \\ c_{n-1} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

In (b), since $s+1 \le n$, the number $s+1$ of rows is $\le$ the number n of columns. Any square submatrix of size $s+1$ is a Vandermonde matrix after factoring a power of α from each column and transposing, and the determinant of the square submatrix is therefore the product of a power of α and the differences $\alpha^{r+j} - \alpha^{r+i}$ with $j > i$. Since α is nonzero and since two powers of α can be equal only when the exponents differ by a multiple of n, the determinant of the square submatrix is nonzero.

In (c), suppose that $s+1$ or fewer of the coefficients $c_0, c_1, \ldots, c_{n-1}$ are nonzero. Choose $s+1$ of them, say c_{i_j} for $1 \le j \le s+1$, such that the remaining ones are 0. If we discard the others from the matrix equation in (a) and discard the corresponding columns of the coefficient matrix, then the matrix equation is still valid since we have discarded only 0's from the given equations. The resulting system is square with an invertible coefficient matrix, and hence the unique solution has $c_{i_j} = 0$ for all j. But then $P(X) = 0$, in contradiction to the assumption that $F(X) \ne 0$.

In (d), if some nonzero member $P(X)$ of C has weight less than $s+2$, then (c) leads to a contradiction. Hence every nonzero weight is $\ge s+2$, and $\delta(C) \ge s+2$.

55. Since α is a root of $X^n - 1$, so is every α^j. Since F_j is the minimal polynomial of α^j, F_j divides $X^n - 1$. Also, $1 + X = X - 1$ divides X^{n-1}, and no F_j equals $X - 1$, since $\alpha^j \ne 1$ for $1 \le j \le 2e$ when $2e < n$. Therefore $G(X)$ divides $X^n - 1$. Applying Problem 54 with $r = 0$ and $s = 2e$, we see that the code C generated by $G(X)$ has $\delta(C) \ge s + 2 = 2e + 2$.

56. In (a), if an irreducible polynomial $F(X)$ of degree d has a root β in $\mathbb{K}$, then $\mathbb{K} \supseteq \mathbb{F}(\beta) \supseteq \mathbb{F}$, and $[\mathbb{F}(\beta) : \mathbb{F}] = d$ must divide $[\mathbb{K} : \mathbb{F}] = m$. In the previous problem it follows that each $F_j(X)$ has degree dividing m, hence degree $\le m$. The worst case for the degree of $G(X)$ is that the LCM equals the product, and then the degree of $G(X)$ is the sum of 1 (from $1 + X$) and the sum of the degrees of the $F_j(X)$'s. Hence $\deg G \le 2em + 1$ in all cases.

In (b), let $n_r = 2^r - 1$, and let $\mathbb{K}$ be a field with 2^r elements. Theorem 9.14 shows that $\mathbb{K}$ is a splitting field for $X^{2^r} - X$ over $\mathbb{F}$. Hence it is a splitting field for $X^{n_r} - 1$ over $\mathbb{F}$. Let $e = r$, so that $e < n_r/2$ as soon as $r \ge 3$. Using this e in the previous problem, we obtain a cyclic code C_r in $\mathbb{F}^{n_r}$ with $\delta(C_r) \ge 2r + 2$. According to (a), the generating polynomial $G_r(X)$ has degree at most $2er + 1 = 2r^2 + 1$. Therefore $k_r = \dim C_r = n_r - \deg G_r \ge n_r - 2r^2 - 1 = 2^r - 2r^2 - 2$. Then k_r/n_r tends to 1, and $\delta(C_r)$ tends to infinity, as required.

57. In (a), the polynomial $F_1(X)$ splits over $\mathbb{K}$ because every finite extension of a finite field is Galois. The Galois group $\mathrm{Gal}(\mathbb{K}/\mathbb{F})$ consists of the powers of the

Frobenius isomorphism $x \mapsto x^2$, by Proposition 9.40, and is transitive on the roots of $F_1(X)$, by Problem 13a. Hence all the roots are of the form α^{2^k}, and all these elements are roots. Taking $k = 0, 1, 2, 3$, we get distinct roots, which is necessary since $\mathbb{K}/\mathbb{F}$ is separable.

For (b), we start from $1 + \alpha + \alpha^4 = 0$ and compute the powers of α in terms of $1, \alpha, \alpha^2, \alpha^3$. The interest is in only the powers $\alpha^0, \alpha^3, \alpha^6, \alpha^9, \alpha^{12}$, but some of the intermediate powers help in the computation. We have

$$\begin{aligned}
\alpha^3 &= \alpha^3,\\
\alpha^4 &= 1 + \alpha,\\
\alpha^5 &= \alpha + \alpha^2,\\
\alpha^6 &= \alpha^2 + \alpha^3,\\
\alpha^9 &= \alpha^3\alpha^6 = \alpha^3(\alpha^2 + \alpha^3) = \alpha^5 + \alpha^6 = \alpha + \alpha^3,\\
\alpha^{12} &= (\alpha^2 + \alpha^3)^2 = \alpha^4 + \alpha^6 = 1 + \alpha + \alpha^2 + \alpha^3.
\end{aligned}$$

Then we form the equation $a + b\alpha^3 + c\alpha^6 + d\alpha^9 + \alpha^{12} = 0$, substitute from above, and equate coefficients. The result is a homogeneous system of four linear equations with five unknowns in $\mathbb{F}$. Solving, we find that the space of solutions is 1-dimensional with $a = b = c = d = e$. Therefore the minimal polynomial of α^3 has degree 4 and is $1 + \alpha + \alpha^2 + \alpha^3 + \alpha^4$.

In (c), we apply Problem 55 with $n = 15$ and $e = 2$. Part (a) shows that $F_1 = F_2 = F_4$, and part (b) computed F_3 as something else of degree 4. Therefore $G(X) = (1+X)\mathrm{LCM}(F_1, F_2, F_3, F_4) = (1+X)\mathrm{LCM}(F_1F_3) = (1+X)F_1(X)F_3(X)$, which has degree 9. Then $\dim C = 15 - 9 = 6$, and Problem 55 gives $\delta(C) \geq 2e + 2 = 6$.

59. In (a), Problems 12–13 are applicable when the scalars are extended to $\mathbb{K}$ because the minimal polynomial becomes a product of first-degree factors. The existence in the conclusion is immediate by applying (a) through (d) in Problem 12 to $L \otimes 1$, and the uniqueness is immediate from Problem 13.

In (b), fix a basis $\{v_i\}$ of V over $\Bbbk$. Any member of $V^{\mathbb{K}}$ has a unique expansion as $\sum_i (v_i \otimes c_i)$ with each c_i in $\mathbb{K}$. Since $\varphi(1) = 1$, application of the given identity to $v \otimes 1$ gives

$$\mathcal{T}(v \otimes 1) = \mathcal{T}(1 \otimes \varphi)(v \otimes 1) = (1 \otimes \varphi)\mathcal{T}(v \otimes 1).$$

If we expand $\mathcal{T}(v \otimes 1)$ as $\sum_i (v_i \otimes c_i)$, the displayed equation says that

$$\textstyle\sum_i (v_i \otimes c_i) = (1 \otimes \varphi) \sum_i (v_i \otimes c_i) = \sum_i (v_i \otimes \varphi(c_i)).$$

Hence $\varphi(c_i) = c_i$ for all i. Since φ is arbitrary, Theorem 9.38 implies that c_i is in $\Bbbk$ for all i. Thus $\sum_i (v_i \otimes c_i)$ is in V. If we write Tv for this element of V, then T is a $\Bbbk$ linear map of V to itself such that $\mathcal{T} = T \otimes 1$.

In (c), we multiply the identity $L \otimes 1 = \mathcal{S} + \mathcal{N}$ on the left by $1 \otimes \varphi^{-1}$ and on the right by $1 \otimes \varphi$ to obtain

$$L \otimes 1 = (1 \otimes \varphi^{-1})(L \otimes 1)(1 \otimes \varphi) = (1 \otimes \varphi^{-1})\mathcal{S}(1 \otimes \varphi) + (1 \otimes \varphi^{-1})\mathcal{N}(1 \otimes \varphi).$$

The equation $\big((1\otimes\varphi^{-1})\mathcal{N}(1\otimes\varphi)\big)^n = (1\otimes\varphi^{-1})\mathcal{N}^n(1\otimes\varphi)$ shows that $\mathcal{N}$ is nilpotent. Since $\big((1 \otimes \varphi^{-1})\mathcal{S}(1 \otimes \varphi)\big)\big((1 \otimes \varphi^{-1})\mathcal{N}(1 \otimes \varphi)\big) = (1 \otimes \varphi^{-1})\mathcal{S}\mathcal{N}(1 \otimes \varphi) = (1 \otimes \varphi^{-1})\mathcal{N}\mathcal{S}(1 \otimes \varphi) = \big((1 \otimes \varphi^{-1})\mathcal{N}(1 \otimes \varphi)\big)\big((1 \otimes \varphi^{-1})\mathcal{S}(1 \otimes \varphi)\big)$, the maps $(1 \otimes \varphi^{-1})\mathcal{S}(1 \otimes \varphi)$ and $(1 \otimes \varphi^{-1})\mathcal{N}(1 \otimes \varphi)$ commute. Finally if $\sum_i (v_i \otimes c_i)$ is an eigenvector of $\mathcal{S}$ with eigenvalue λ, we have $\mathcal{S}\big(\sum_i v_i \otimes c_i)\big) = \sum_i v_i \otimes \lambda c_i$. Therefore $(1 \otimes \varphi^{-1})\mathcal{S}(1 \otimes \varphi)\big(\sum_i (v_i \otimes \varphi^{-1}(c_i))\big) = (1 \otimes \varphi^{-1})\sum_i (v_i \otimes \lambda c_i) = \sum_i (v_i \otimes \varphi^{-1}(\lambda)\varphi^{-1}(c_i))$, and $\sum_i (v_i \otimes \varphi^{-1}(c_i))$ is an eigenvector of $(1\otimes\varphi^{-1})\mathcal{S}(1\otimes\varphi)$ with eigenvalue $\varphi^{-1}(\lambda)$. Then it follows that $(1 \otimes \varphi^{-1})\mathcal{S}(1 \otimes \varphi)$ has a basis of eigenvectors. By uniqueness of the decomposition $L \otimes 1 = \mathcal{S} + \mathcal{N}$, we must have $(1 \otimes \varphi^{-1})\mathcal{S}(1 \otimes \varphi) = \mathcal{S}$ and $(1 \otimes \varphi^{-1})\mathcal{N}(1 \otimes \varphi) = \mathcal{N}$. Since φ is arbitrary in $\mathrm{Gal}(\mathbb{K}/\Bbbk)$, (b) shows that $\mathcal{S} = S \otimes 1$ and $\mathcal{N} = N \otimes 1$.

In (d), $(N^n \otimes 1) = (N \otimes 1)^n = \mathcal{N}^n$, and $\mathcal{N}$ nilpotent implies N nilpotent. Similarly $\mathcal{S}\mathcal{N} = \mathcal{N}\mathcal{S}$ implies $SN = NS$. Then the fact that $S^{\mathbb{K}} = S \otimes 1 = \mathcal{S}$ has a basis of eigenvectors implies that S is semisimple.

In (e), $S \otimes 1$ and $N \otimes 1$ can be expressed uniquely as polynomials in $L \otimes 1$ that are 0 or have degree less than the degree of the minimal polynomial of $L \otimes 1$; the coefficients of these polynomials are in $\mathbb{K}$. Application of a member φ to a polynomial expression $S \otimes 1 = P(L \otimes 1)$ just affects the coefficients and gives another polynomial expression for S unless φ fixes each coefficient. By uniqueness and Theorem 9.38, we see that the coefficients are in $\Bbbk$. A similar argument applies to $N \otimes 1$.

60. This is proved by the same argument as for Problem 13 in Chapter V.

61. The splitting field for the minimal polynomial is $\mathbb{C}$. According to the procedure in the solution of Problem 59, we are first supposed to find a decomposition over $\mathbb{C}$. In a suitable basis we know that A is the sum of a diagonal matrix and a strictly upper triangular matrix, and this is the Jordan–Chevalley decomposition. Section V.6 shows how to find the Jordan form and the basis over $\mathbb{C}$ in which it is realized. We transform the D and N back separately to find the semisimple and nilpotent components of A relative to the standard basis. The result is that

$$S = \begin{pmatrix} 0 & -1 & 0 & 1/2 \\ 1 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \qquad \text{and} \qquad N = \begin{pmatrix} 0 & 0 & 0 & -1/2 \\ 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

62. In (a), if there were a basis of eigenvectors over $\mathbb{K}$, then the fact that the eigenvalues are equal would mean that A is similar to a scalar matrix. This is manifestly not so. Thus A is not semisimple.

In (b), a matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ that commutes with A is necessarily of the form $\left(\begin{smallmatrix} a & cx \\ c & a \end{smallmatrix}\right)$ and has characteristic polynomial $X^2 + a^2 + c^2x$ since the characteristic is 2. If the

characteristic polynomial reduces to 0, and then $a = c = 0$. In this case, A is the 0 matrix.

In (c), suppose that $A = S + N$ is a Jordan–Chevalley decomposition. Then (a) says that A is not semisimple and hence cannot be S. On the other hand, (b) says that N has to be 0 and therefore that $A = S$ is the only possibility. The result is a contradiction, and thus there is no Jordan–Chevalley decomposition.

63. This comes down to what is happening in Problem 12 in Chapter V. In terms of matrices, the problem reduces to the case that a square matrix A is upper triangular with a certain nonzero scalar c in every diagonal entry. Then $D = cI$, and U is taken to be $D^{-1}A$.

64. In (e), for characteristic $p > 0$, -1 is the sum of $p - 1$ copies of 1. Hence $\Bbbk$ cannot be formally real.

65. In (e), we have $(b^{-1} - a^{-1})ab = a - b$. The right side is in P, and so are a and b. Thus the remaining factor, $b^{-1} - a^{-1}$, has to be in P. In (f), the sum of $a(c-d) > 0$ and $(a-b)d > 0$ is $ac-bd > 0$. In (g), expansion of $(a-b)(c-d) > 0$ gives $ac + bd > ad + bc$.

66. The definition is that $\frac{a_m x^m + \cdots + a_0}{b_n x^n + \cdots + b_0}$ is positive if $a_m b_n^{-1}$ is in P. It is routine to check that the set P' of positive elements of $\Bbbk(x)$ is closed under addition and multiplication, and certainly every nonzero element is in exactly one of P' and $-P'$.

67. In (a), one ordering has $a + b\sqrt{2}$ in P if $a + b\sqrt{2} > 0$ in the ordinary sense, and the other has $a + \sqrt{2}$ in P if $a - b\sqrt{2} > 0$ in the ordinary sense.

In (b), for any element $a + b\sqrt{c}$ with $a^2 > b^2c$, define $a + b\sqrt{c}$ to be in P' if and only if a is in P. For any element $a + b\sqrt{c}$ with $b^2c > a^2$, define $a + b\sqrt{c}$ to be in P' if and only if b is in P. The only element left undecided by this process is 0, which is not to be in P'. The elements $a + b\sqrt{c}$ in P' with $a^2 > b^2c$ will be said to be of type I, while those with $a^2 < b^2c$ will be said to be of type II. It is clear that each nonzero element x of $\mathbb{K}$ is in exactly one of P' and $-P'$, and we have to verify that P' is closed under addition and multiplication.

The verification is a little complicated. It uses parts (f) and (g) of Problem 65 repeatedly. Consider addition. There are cases. Case 1 is that $a + b\sqrt{c}$ and $a' + b'\sqrt{c}$ are in P' with both of type I. If the sum is of type II, then addition of $a^2 > b^2c$, $a'^2 > b'^2c$, and $(b+b')^2c > (a+a')^2$ gives $bb'c > aa'$ upon cancellation. Squaring and taking into account that $aa' > 0$, we obtain $(b^2c)(b'^2c) > a^2a'^2$. On the other hand, $a^2 > b^2c$ and $a'^2 > b'^2c$ together imply $a^2a'^2 > (b^2c)(b'^2c)$, contradiction. Thus the sum is of type I. Since a and a' are in P, so is $a + a'$. Thus the sum is in P'.

Case 2 is that $a + b\sqrt{c}$ and $a' + b'\sqrt{c}$ are in P' with both of type II. If the sum is of type I, then addition of $a^2 < b^2c$, $a'^2 < b'^2c$, and $(b + b')^2c < (a + a')^2$ gives $bb'c < aa'$ upon cancellation. Squaring and taking into account that $bb' > 0$, we obtain $(b^2c)(b'^2c) < a^2a'^2$. On the other hand, $a^2 < b^2c$ and $a'^2 < b'^2c$ together imply $a^2a'^2 < (b^2c)(b'^2c)$, contradiction. Thus the sum is of type II. Since b and b' are in P, so is $b + b'$. Thus the sum is in P'.

Case 3 is that $a + b\sqrt{c}$ is of type I and $a' + b'\sqrt{c}$ is of type II (or vice versa). The argument now depends on the type of the sum. Case 3A is that the sum is of type I. Adding $a^2 > b^2c$, $b'^2c > a'^2$, and $(a+a')^2 > (b+b')^2c$ and canceling gives $a(a+a') > b(b+b')c$. We want to see that $a + a' > 0$. If $a + a' < 0$, then the left side is negative, and hence both sides are negative. Thus the squares of the two sides are related in the opposite order: $a^2(a+a')^2 < b^2(b+b')^2c^2$. Here the right side is $< a^2(b+b')^2c$, and we get $(a+a')^2 < (b+b')^2c$, in contradiction to the fact that the sum is of type I. So $a + a'$ is > 0, and the sum is in P'. Case 3B is that the sum is of type II. Adding $a^2 > b^2c$, $b'^2c > a'^2$, and $(b+b')^2c > (a+a')^2$ and canceling gives $b'(b+b')c > a'(a+a')$. We want to see that $b + b' > 0$. If $b + b' < 0$, then both sides are negative. Thus the squares of the two sides are related in the opposite order: $b'^2(b+b')^2c^2 < a'^2(a+a')^2$. Here the right side is $< b'^2c(a+a')^2$, and thus $(b+b')^2c < (a+a')^2$, in contradiction to the fact that the sum is of type II. So $b + b'$ is > 0, and the sum is in P'.

This completes the verification that P' is closed under addition. We now consider multiplication, again dividing matters into cases. Case 1 is that $a+b\sqrt{c}$ and $a'+b'\sqrt{c}$ are in P' with both of type I. Applying Problem 65g to the inequalities $a^2 > b^2c$ and $a'^2 > b'^2c$, we obtain $a^2a'^2+b^2b'^2c^2 > a^2b'^2c+a'^2b^2c$, which says that the product is of type I. We are to show that $aa' + bb'c$ is > 0. From $a^2 > b^2c$ and $a'^2 > b'^2c$, we obtain $0 < a^2a'^2 - b^2b'^2c^2 = (aa' + bb'c)(aa' - bb'c)$. Thus $aa' + bb'c$ and $aa' - bb'c$ are both > 0 or both < 0, and they are the same as their sum, which is $2aa'$. Since a and a' are in P, we have $aa' > 0$, we conclude that the product is in P'.

Case 2 is that $a + b\sqrt{c}$ and $a' + b'\sqrt{c}$ are in P' with both of type II. Applying Problem 65g to the inequalities $b^2c > a^2$ and $b'^2c > a'^2$, we obtain $a^2a'^2+b^2b'^2c^2 > a^2b'^2c+a'^2b^2c$, which says that the product is of type I. We are to show that $aa'+bb'c$ is > 0. From $a^2 < b^2c$ and $a'^2 < b'^2c$, we see that $0 < b^2b'^2c^2 - a^2a'^2 = (bb'c + aa')(bb'c - aa')$. Thus $bb'c + aa'$ and $bb'c - aa'$ are both > 0 or both < 0, and they are the same as their sum, which is $2bb'$. Since b and b' are in P, we have $bb' > 0$, we conclude that the product is in P'.

Case 3 is that $a+b\sqrt{c}$ is of type I and $a'+b'\sqrt{c}$ is of type II (or vice versa). From $(a^2-b^2c)(b'^2c-a'^2) > 0$, we obtain $c(a'^2b^2+a^2b'^2) > a^2a'^2+b^2b'^2c^2$. Addition of $2aa'bb'c$ to both sides yields $(ab' + a'b)^2c > (aa' + bb'c)^2$, an inequality that shows the product to be of type II. To show that the product is in P', we are to show that $ab' + a'b > 0$. The product of $a^2 > b^2c$ and $b'^2c > a'^2$ gives $a^2b'^2 > b^2a'^2$ upon cancellation of c, c being positive. Then $(ab' + ba')(ab' - ba') > 0$, and the two factors have the same sign. Now $a > 0$ and $b' > 0$ since the given elements are in P'. Thus $ab' > 0$. Arguing by contradiction, suppose that $ab' < ba'$. Then $ab' > 0$ implies $(ab')^2 < (ba')^2$, in contradiction to $a^2b'^2 > b^2a'^2$. We conclude that $ab' > ba'$, hence that $ab' - ba' > 0$. Thus $ab' + ba' > 0$, as required.

These steps complete all the verifications that P', as we have defined it, is a positive system. It remains to define a second version of P' and to carry out the verifications for it. For the definition, there is no change if $a^2 > b^2c$, but if $b^2c > a$, then $a+b\sqrt{c}$

is to be in P' if and only if $-b$ is in P. The verifications are essentially unchanged except that the roles of b and $-b$ are interchanged throughout.

68. In (a), the integer n, if it exists, cannot be 0 because $\Bbbk$ is formally real by Problem 64d. So $n \geq 1$. We write $\xi_j = a_j + b_j\sqrt{c_n}$ with each a_j and b_j in $\Bbbk(\sqrt{c_1}, \ldots, \sqrt{c_{n-1}})$ and expand out the squares.

In (b), let $\Bbbk' = \Bbbk(\sqrt{c_1}, \ldots, \sqrt{c_{n-1}})$. If the coefficient of $\sqrt{c_n}$ is 0, then $(*)$ becomes an equality in $\Bbbk'$ that exhibits $\Bbbk'$ as not formally real, in contradiction to the definition of n. If the coefficient of $\sqrt{c_n}$ is not 0, then $(*)$ exhibits $\sqrt{c_n}$ as a member of $\Bbbk'$, again in contradiction to the definition of n. The conclusion is that $\mathbb{K}$ is formally real.

69. Order the formally real subfields of $\overline{\Bbbk}$ by inclusion upward. The set of such subfields is nonempty since $\Bbbk$ is one. The union of a chain of such subfields is again such a subfield because any expression of a sum equal to -1 has to be valid in a finite such union. By Zorn's Lemma, there is a maximal element $\mathbb{K}$. By maximality, $\mathbb{K}$ is a real closed field.

70. In (a), if c is not a square in $\Bbbk$, then $\Bbbk(\sqrt{c})$ is a proper algebraic extension of $\Bbbk$. Since $\Bbbk$ is maximal among formally real subfields of $\overline{\Bbbk}$, $\Bbbk(\sqrt{c})$ is not formally real. Therefore -1 is a sum of squares in $\Bbbk(\sqrt{c})$, as indicated.

In (b), expansion gives $-1 = \sum a_j^2 + c\sum b_j^2 + 2\sqrt{c}\sum a_j b_j$. Equating coefficients of 1 and $\sqrt{c}$ shows that $-1 = \sum a_j^2 + c\sum b_j^2$. We cannot have $\sum b_j^2 = 0$ because otherwise we would have $-1 = \sum a_j^2$ and $\Bbbk$ would not be formally real. Thus $-c = (1 + \sum a_j^2)/\sum b_j^2$, and $-c$ is exhibited as a sum of squares, hence a member of P. Thus if c is not a square in $\Bbbk$, then c cannot be a sum of squares in $\Bbbk$. The contrapositive is: every sum of squares in $\Bbbk$ is a square in $\Bbbk$.

In (c), the equality $-c = (1 + \sum a_j^2)/\sum b_j^2$, in view of (b), exhibits $-c$ as the quotient of two squares, hence as a square.

In (d), let P be the set of nonzero squares. We see from (a) through (c) that every nonzero element is in P or in $-P$. By (b), every sum of squares is a square; thus P is closed under addition. It is clear that P is closed under multiplication. Thus F becomes an ordered field. Problem 64b shows that every nonzero square has to be in P, and thus P is the only possibility for the set of positive elements.

71. In (a), let n be the least odd positive integer such that some polynomial over $\Bbbk$ of degree n has no root in $\Bbbk$. If this polynomial were reducible, some factor of it would have smaller odd degree and would have a root. So the polynomial in question has to be irreducible.

In (b), if -1 is a sum of squares in $\Bbbk(\alpha)$, then we have $-1 = \sum_{j=1}^{k} R_j(\alpha)^2$ for suitable polynomials $R_j(X)$ in $\Bbbk[X]$, necessarily of degree $\leq n-1$. In other words, $\sum_{j=1}^{k} R_j(X)^2 + 1$ is a member of $\Bbbk[X]$ that vanishes at α. Since $Q(X)$ is the minimal polynomial of α, $Q(X)$ divides $\sum_{j=1}^{k} R_j(X)^2 + 1$, and we obtain $-1 = \sum_{j=1}^{k} R_j(X)^2 + Q(X)A(X)$ for a suitable polynomal $A(X)$ in $\Bbbk[X]$ of degree $\leq n-2$.

In (c), the equality of the coefficients of X^{2n-2} in the polynomial identity of (b) shows that the -1 equals the sum of squares of the leading coefficients of the R_j's plus the coefficient of X^{n-2} in $A(X)$. The coefficient of X^{n-2} in $A(X)$ cannot be 0, or else -1 would be exhibited as a sum of squares in $\Bbbk$. Thus $A(X)$ has degree exactly $n-2$, which is odd. The inductive hypothesis applies to $A(X)$ and says that $A(X)$ has a root r. We evaluate the polynomial identity from (b) at r, take into account that $A(r) = 0$, and obtain $-1 = \sum_{j=1}^{k} R_j(r)^2$. Again we have a contradiction to the fact that $\Bbbk$ is formally real, and thus the minimal integer n in (a) cannot exist.

72. The indicated proof goes through without essential change.

73. Problem 68 shows that $\Bbbk$ is contained in a certain formally real subfield $\mathbb{L}$ of $\overline{\Bbbk}$, Problem 69 shows that $\mathbb{L}$ is contained in a real closed subfield $\mathbb{K}$ of $\overline{\Bbbk}$, and Problem 70 shows that $\mathbb{K}$ becomes an ordered field. The set of positive elements for $\mathbb{K}$ includes all squares by Problem 64b, and all the members of $\Bbbk$ in P have become squares in $\mathbb{L}$ by definition of $\mathbb{L}$. Therefore all members of P are squares in $\mathbb{K}$. The fact that $\overline{\Bbbk} = \mathbb{K}(\sqrt{-1}\,)$ follows from Problem 72.

Chapter X

1. If R is a field, then the only ideals are 0 and R, and they certainly satisfy the descending chain condition. Conversely if the ideals satisfy the descending chain condition, then there is a minimal nonzero ideal I. Fix $m \neq 0$ in I. For any nonzero element $a \in I$, $Ra = I$ since I is a simple module. If $x \neq 0$ is in R, we apply this observation to xm, which is nonzero since R is an integral domain. Since $Rxm = I$, there exists y in R with $yxm = m$. Then $(1 - yx)m = 0$. Since R is an integral domain and $m \neq 0$, we obtain $1 - yx = 0$. Therefore $y = x^{-1}$.

2. In (a), let $C_2 = \{\pm 1\}$. Define $r(1) = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ and $r(-1) = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$. Then r is a representation since $1 + 1 = 0$ in $\mathbb{F}$. The subspace $U = \mathbb{F}\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$ is invariant. If there were a complementary invariant subspace, there would be an eigenvector of $r(-1)$ not in U. However, the roots of the characteristic polynomial are both 1, and a second eigenvector would mean that $r(-1)$ is the identity, which it is not. For (b), the representation in (a) makes $\mathbb{F}^2$ into a unital left R module, the R submodules being the invariant subspaces. There is no complementary R submodule to U, and hence $\mathbb{F}^2$ is not semisimple as an R module.

3. If $\{a_s\}$ is a set of generators of M as a right R module and $\{b_t\}$ is a set of generators of N as a left R module, then $\{a_s \otimes b_t\}$ is a set of generators of $M \otimes_R N$ as an abelian group. Then (a) follows from this fact and the fact that 1 generates both $\mathbb{Z}/k\mathbb{Z}$ and $\mathbb{Z}/l\mathbb{Z}$.

In (b), if $l = dk$ for some d and if b has $b = qk + r$ with $0 \leq r < |k|$, then $a1 \otimes b1 = aqk(1 \otimes 1) + (a1 \otimes r1) = aq(k1 \otimes 1) + (a1 \otimes r1) = a1 \otimes r1$, and it

follows that the map $a1 \otimes b1 \mapsto a1 \otimes (b \bmod k)1$ is a well-defined group isomorphism of $(\mathbb{Z}/k\mathbb{Z}) \otimes_{\mathbb{Z}} (\mathbb{Z}/l\mathbb{Z})$ onto $(\mathbb{Z}/k\mathbb{Z}) \otimes_{\mathbb{Z}} (\mathbb{Z}/k\mathbb{Z})$.

In (c), let $b(x1, y1) = xy \bmod k$ for $x, y \in \mathbb{Z}/k\mathbb{Z}$. This is $\mathbb{Z}$ bilinear from $\mathbb{Z}/k\mathbb{Z} \times \mathbb{Z}/k\mathbb{Z}$ into $\mathbb{Z}/k\mathbb{Z}$ and extends to a group homomorphism $L : \mathbb{Z}/k\mathbb{Z} \otimes_{\mathbb{Z}} \mathbb{Z}/k\mathbb{Z} \to \mathbb{Z}/k\mathbb{Z}$ with $L(x1 \otimes y1) = xy \bmod k$. In particular, $L(1 \otimes 1) = 1 \bmod k$. Therefore k divides the order of $1 \otimes 1$, and $\mathbb{Z}/k\mathbb{Z} \times \mathbb{Z}/k\mathbb{Z}$ has at least $|k|$ elements.

In (d), we have $0 = k1 \otimes 1 = k(1 \otimes 1)$ and $0 = 1 \otimes l1 = l(1 \otimes 1)$. If $xk + yl = d$, then $d(1 \otimes 1) = x(k(1 \otimes 1)) + y(l(1 \otimes 1)) = 0$. Hence $1 \otimes 1$ has order dividing d. By (c), $1 \otimes 1$ has order at least $|d|$. The result follows.

4. In (a), each $\ker \varphi^n$ is an R submodule of M, and these R submodules form an ascending chain. Hence they are the same from some point on. Similarly each $\operatorname{image} \varphi^n$ is an R submodule of M, and these form a descending chain. Hence they are the same from some point on.

In (b), if x is in $\mathcal{K} \cap \mathcal{I}$, then $\varphi^N x = 0$ and $x = \varphi^N y$ for some y. Then $0 = \varphi^N x = \varphi^{2N} y$. Since y is in $\ker \varphi^{2N} = \ker \varphi^N$, we obtain $0 = \varphi^N y = x$, and $x = 0$.

In (c), if x is in M, then $\varphi^N x$ is in $\operatorname{image} \varphi^N = \operatorname{image} \varphi^{2N}$. Hence $\varphi^N x = \varphi^{2N} z = \varphi^N(\varphi^N z)$ for some $z \in M$, and $\varphi^N x = \varphi^N y$ with $y = \varphi^N z$.

For (d), if x is in M, let y be as in (c), and write $x = (x - y) + y$. Then $\varphi^N(x - y) = \varphi^N x - \varphi^N y = 0$ and $y = \varphi^N z$ show that $x - y$ is in $\mathcal{K}$ and y is in $\mathcal{I}$. Thus $M = \mathcal{K} + \mathcal{I}$. Since $\mathcal{K} \cap \mathcal{I} = 0$ by (b), $M = \mathcal{K} \oplus \mathcal{I}$.

In (e), we know that $\varphi(\operatorname{image} \varphi^n) = \operatorname{image} \varphi^{n+1}$ for all n. Taking $n > N$, we see that $\varphi(\mathcal{I}) = \mathcal{I}$. From (b), $\ker(\varphi\big|_{\mathcal{I}}) \subseteq \mathcal{K} \cap \mathcal{I} = 0$. Therefore φ is one-one from $\mathcal{I}$ onto itself. In addition, $\varphi(\ker \varphi^n) \subseteq \ker \varphi^{n-1}$ for all n. Taking $n > N$ shows that $\varphi(\mathcal{K}) \subseteq \mathcal{K}$. For x in $\mathcal{K}$, we have $\varphi^N x = 0$. Therefore $(\varphi\big|_{\mathcal{K}})^N = 0$.

5. If (i) holds, then $\psi\big|_{N'}$ is one-one from N' onto P. Let σ be its inverse. Then $\sigma : P \to N'$ is one-one with $\psi\sigma = 1_P$. So (ii) holds.

If (ii) holds, then any n in N has the property that $n - \sigma\psi(n)$ has $\psi(n - \sigma\psi(n)) = \psi(n) - 1_P\psi(n) = 0$ and is therefore in $\operatorname{image} \varphi$. Write $n - \sigma\psi(n) = \varphi(m)$ for some m depending on n; m is unique since φ is one-one. If $\tau : N \to M$ is defined by $\tau(n) = m$, then τ is an R homomorphism by the uniqueness of m. Consider $\tau(\varphi(m))$ for m in M. The element $n = \varphi(m)$ has $n - \sigma\psi(n) = \varphi(m) - \sigma\psi\varphi(m) = \varphi(m) - \sigma(0) = \varphi(m)$, and the definition of τ says that $\tau(\varphi(m)) = m$. Hence $\tau\varphi = 1_M$, and (iii) holds.

If (iii) holds, then $N' = \ker \tau$ is an R submodule of N. If n is in $N' \cap \operatorname{image} \varphi$, then $n = \varphi(m)$ for some $m \in M$ and also $0 = \tau(n) = \tau\varphi(m) = 1_M(m) = m$. So $n = 0$, and $N' \cap \operatorname{image} \varphi = 0$. If $n \in N$ is given, write $n = (n - \varphi\tau(n)) + \varphi\tau(n)$. Then $\varphi\tau(n)$ is certainly in $\operatorname{image} \varphi$, and $\tau(n - \varphi\tau(n)) = \tau(n) - 1_M\tau(n) = 0$ shows that $n - \varphi\tau(n)$ is in N'. Therefore $N = N' \oplus \operatorname{image} \varphi$. Since $\operatorname{image} \varphi = \ker \psi$, we see that $N = N' \oplus \ker \psi$ and that (i) holds.

6. For (a), the conjugation mapping C on R, carrying 1 to itself and carrying $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$ to their negatives, respects addition and satisfies $C(xy) = C(y)C(x)$. Hence it exhibits R and R^o as isomorphic. Then the result follows from Proposition 10.14.

For (b), again by Proposition 10.14, we need a noncommutative ring R with identity such that R is not isomorphic to R^o. Let $\mathbb{F}$ be a field with two elements, and let R be the 8-element ring consisting of all matrices $\begin{pmatrix} a & b \\ 0 & c \end{pmatrix}$ with a, b, c in $\mathbb{F}$. Define x to be the matrix with $a = 1$ and $b = c = 0$, and define y to be the matrix with $b = 1$ and $a = c = 0$. Computation shows that $x^2 = x$, $y^2 = 0$, $xy = y$, and $yx = 0$. A ring isomorphism of R with R^o is the same as an additive isomorphism that reverses the order of multiplication, and we call this an "antiautomorphism" of R. Suppose that an antiautomorphism φ of R exists. We must have $\varphi(1) = 1$. Suppose that $\varphi(x) = u$ and $\varphi(y) = v$. Then $u = \varphi(x) = \varphi(x^2) = \varphi(x)^2 = u^2$ and $0 = \varphi(y^2) = \varphi(y)^2 = v^2$. Expanding u and v in terms of the basis $\{1, x, y\}$ and computing, we find that $u = k1 + lx$ and $v = my$ with k, l, m in $\mathbb{F}$. Since φ reverses the order of multiplication, we have $uv = \varphi(x)\varphi(y) = \varphi(yx) = \varphi(0) = 0$. Thus $0 = (k1 + lx)(my) = km1 + lmxy = (km)1 + (lm)y$, and $km = lm = 0$. Therefore either $m = 0$ or $k = l = 0$. In the first case, $\varphi(y) = v = my = 0$; in the second case $\varphi(x) = u = k1 + lx = 0$. In either case, φ fails to be one-one. We conclude that no antiautomorphism φ of R exists.

7. Take the sum of all simple R submodules of M.

8. Example 4 in Section 5 shows that $A \otimes_{\mathbb{F}} \mathbb{K}$ is a vector space over $\mathbb{K}$ in such a way that $k_0(a \otimes k) = a \otimes k_0 k$. It is therefore enough to show that the multiplication is $\mathbb{K}$ linear in each variable of the product. Additivity is known, and it is enough to check that $k_0\big((a_1 \otimes k_1)(a_2 \otimes k_2)\big) = \big(k_0(a_1 \otimes k_1)\big)(a_2 \otimes k_2) = (a_1 \otimes k_1)\big(k_0(a_2 \otimes k_2)\big)$. Since scalar multiplication by k_0 equals left multiplication by $1 \otimes k_0$, the left equality is immediate from associativity of multiplication, and the right equality follows from associativity and from the formula $(a_1 \otimes k_1)(1 \otimes k_0) = a_1 \otimes k_1 k_0 = a_1 \otimes k_0 k_1 = (1 \otimes k_0)(a_1 \otimes k_1)$.

9. Define $\mu(x)(y) = [x, y]$ for x and y in $\mathfrak{g}$, and let $\nu(c)(d) = cd$ for c and d in $\mathbb{L}$. Then $\mu(x) : \mathfrak{g} \to \mathfrak{g}$ and $\nu(c) : \mathbb{L} \to \mathbb{L}$ are $\mathbb{K}$ linear. Therefore $b(x, c) = \mu(x) \otimes \nu(c)$ is $\mathbb{K}$ bilinear from $\mathfrak{g} \times \mathbb{L}$ into the $\mathbb{K}$ vector space $\mathrm{End}_{\mathbb{K}}(\mathfrak{g} \otimes_{\mathbb{K}} \mathbb{L})$, and it extends to a $\mathbb{K}$ linear mapping $L : \mathfrak{g} \otimes_{\mathbb{K}} \mathbb{L} \to \mathrm{End}_{\mathbb{K}}(\mathfrak{g} \otimes_{\mathbb{K}} \mathbb{L})$. Define $[X, Y] = L(X)(Y)$.

With the Lie algebra multiplication now well defined in $\mathfrak{g} \otimes_{\mathbb{K}} \mathbb{L}$, one readily checks the two required properties. Therefore $\mathfrak{g} \otimes_{\mathbb{K}} \mathbb{L}$ is a Lie algebra over $\mathbb{K}$ satisfying the two required identities.

Meanwhile, we know that $\mathfrak{g} \otimes_{\mathbb{K}} \mathbb{L}$ is a vector space over $\mathbb{L}$ because of a change of rings. To complete the proof, we need to show that the multiplication is $\mathbb{L}$ linear, not just $\mathbb{K}$ linear. It is enough to check $\mathbb{L}$ linearity in the second variable because of the alternating property. Let s be in $\mathbb{L}$, and let $x \otimes c$ and $y \otimes d$ be elements of $\mathfrak{g} \otimes_{\mathbb{K}} \mathbb{L}$. Then we have $[x \otimes c, s(y \otimes d)] = [x \otimes c, y \otimes sd] = [x, y] \otimes csd = s([x, y] \otimes cd) = s[x \otimes c,\ y \otimes d]$. Forming $\mathbb{K}$ linear combinations, we obtain the desired $\mathbb{L}$ linearity in the second variable of the Lie algebra product.

10. This problem will follow from the uniqueness of the tensor product as given in Theorem 10.18 if it is shown that $((A \otimes_{\mathbb{Z}} B)/H,\ qb_2)$ is a tensor product of A and B over R. Thus let $\beta : A \times B \to G$ be an R bilinear function from $A \times B$ into an abelian

group G. Since β is automatically $\mathbb{Z}$ bilinear, there exists a group homomorphism $\varphi : A \otimes_{\mathbb{Z}} B \to G$ such that $\varphi(a \otimes b) = \beta(a, b)$ for all $a \in A$ and $b \in B$. Then $\varphi(ar \otimes b - a \otimes rb) = \varphi(ar \otimes b) - \varphi(a \otimes rb) = \beta(ar, b) - \beta(a, rb)$. The right side is in H, and hence φ descends to a group homomorphism $\overline{\varphi} : (A \otimes_{\mathbb{Z}} B)/H \to G$ such that $\overline{\varphi}q = \varphi$. Then $\beta(a, b) = \varphi(a \otimes b) = \overline{\varphi}qb_2(a, b)$ shows that $\overline{\varphi}(qb_2) = \beta$. Thus $\overline{\varphi}$ is the required additive extension of β. For uniqueness, suppose $\overline{\varphi}'$ is a second additive extension of β. Then $\overline{\varphi}'qb_2(a, b) = \overline{\varphi}qb_2(a, b)$ for all $a \in A$ and $b \in B$, and hence $\overline{\varphi}'q(a \otimes b) = \overline{\varphi}q(a \otimes b)$. The elements $a \otimes b$ generate $A \otimes_{\mathbb{Z}} B$, and hence $\overline{\varphi}'q = \overline{\varphi}q$ on $A \otimes_{\mathbb{Z}} B$. Since q maps onto $(A \otimes_{\mathbb{Z}} B)/H$, $\overline{\varphi}' = \overline{\varphi}$ on $(A \otimes_{\mathbb{Z}} B)/H$.

11. We are to show that if C is a commutative associative R algebra with identity and if $\varphi_1 : A_1 \to C$ and $\varphi_2 : A_2 \to C$ are homomorphisms of commutative associative R algebras with identity, then there exists a unique homomorphism $\varphi : A_1 \otimes_R A_2 \to C$ of R algebras with identity such that $\varphi i_1 = \varphi_1$ and $\varphi i_2 = \varphi_2$. Define $b(a_1, a_2) = \varphi_1(a_1)\varphi_2(a_2)$. This is R bilinear into C because $b(a_1 r, a_2) = \varphi_1(a_1 r)\varphi_2(a_2) = \varphi_1(a_1)r\varphi_2(a_2) = \varphi_1(a_1)\varphi_2(ra_2) = b(a_2, ra_2)$, and hence there exists a unique homomorphism $\varphi : A_1 \otimes_R A_2 \to C$ of abelian groups such that $\varphi(a_1 \otimes a_2) = b(a_1, a_2) = \varphi_1(a_1)\varphi_2(a_2)$. Then $\varphi i_1(a_1) = \varphi(a_1 \otimes 1) = \varphi_1(a_1)\varphi_2(1) = \varphi_1(a_1)1 = \varphi_1$, and $\varphi i_1 = \varphi_1$. Similarly $\varphi i_2 = \varphi_2$. To complete the proof, it is enough to show that the homomorphism φ of abelian groups is a homomorphism of R algebras. The fact that φ is a homomorphism of R modules is immediate from Corollary 10.19. Also, $\varphi(1 \otimes 1) = \varphi_1(1)\varphi_2(1) = 1$ shows that φ carries identity to identity. Finally the computation $\varphi\big((a_1 \otimes a_2)(a_1' \otimes a_2')\big) = \varphi(a_1 a_1' \otimes a_2 a_2') = \varphi_1(a_1 a_1')\varphi_2(a_2 a_2') = \varphi_1(a_1)\varphi_1(a_1')\varphi_2(a_2)\varphi_2(a_2') = \varphi_1(a_1)\varphi_2(a_2)\varphi_1(a_1')\varphi_2(a_2') = \varphi(a_1 \otimes a_2)\varphi(a_1' \otimes a_2')$ shows that φ respects multiplication on a set of additive generators of $A_1 \otimes_R A_2$.

12. Part (a) is immediate from Proposition 10.1. If ψ is a nonzero map in M^E, then $\psi(E)$ is a submodule of M isomorphic to E. Hence $\psi(E) \subseteq M_E$ by construction, and (b) follows. Part (c) is immediate from (b).

13. With $d \in D_E = \mathrm{Hom}_R(E, E)$, we can form $\psi d = \psi \circ d$ if ψ is in $\mathrm{Hom}_R(E, M_E)$, and we can form $de = d(e)$ if e is in E. These definitions give the required unital D_E module structures for (a) and (b). The members of $D_E = \mathrm{Hom}_R(E, E)$ commute with the left R action on E by definition, and this is (c).

14. In view of (c) in the previous problem, the left action of R on E can be regarded as a right R^o action on E in such a way that it commutes with the left D_E action on E. In other words, E is a unital (D_E, R^o) bimodule. Corollary 10.19b shows that $M^E \otimes_{D_E} E$ becomes a unital right R^o module, hence a unital left R module.

15. Define a map $b : M^E \times E \to M$, additive in each variable, by $b(\psi, e) = \psi(e)$. For d in D_E, this has $b(\psi \circ d, e) = (\psi \circ d)(e) = \psi(d(e)) = b(\psi, d(e))$. Hence b is D_E bilinear and has an additive extension $\Phi : M^E \otimes_{D_E} E \to M$ with $\Phi(\psi \otimes e) = \psi(e)$.

The map Φ is R linear since $\Phi(r(\psi \otimes e)) = \Phi(\psi \otimes re) = \psi(re) = r(\psi(e)) = r(\Phi(\psi \otimes e))$. Since ψ is in M^E, $\psi(e)$ is in M_E; thus Φ has image in M_E.

To see that Φ is onto M_E, write $M_E = \bigoplus_{s\in T} M_s$ with each M_s simple, and fix an isomorphism $\alpha_s \in \mathrm{Hom}_R(E, M_s)$ for each $s \in T$. For any element $m \in M_E$, we can find a finite subset T' of T such that $m = \sum_{s\in T'} m_s$ with $m_s \in M_s$. If we let $e_s = \alpha_s^{-1}(m_s)$, then $\Phi\big(\sum_{s\in T'} \alpha_s \otimes e_s\big) = m$. Thus Φ maps onto M_E.

To see that Φ is one-one, we observe from Problem 12 and Lemma 10.3 that

$$M^E = \mathrm{Hom}_R(E, M) = \mathrm{Hom}_R(E, M_E) = \mathrm{Hom}_R\Big(E, \bigoplus_{s\in T} M_s\Big) = \bigoplus_{s\in T} \mathrm{Hom}_R(E, M_s).$$

Each summand on the right side is isomorphic to D_E. That is, the collection of isomorphisms $\{\alpha_s\}_{s\in T}$ from the previous paragraph is a basis of M^E as a right D_E vector space. Consequently every element of $M^E \otimes_{D_E} E$ may be written as a finite sum $\sum \alpha_s \otimes e_s$ with $e_s \in E$. The image of the element $\sum \alpha_s \otimes e_s$ is $\sum \alpha_s(e_s)$. If this is 0, then each $\alpha_s(e_s)$ is 0 because of the independence of the M_s's. Since α_s is an isomorphism, it follows that $e_s = 0$ for each s. Therefore $\sum \alpha_s \otimes e_s = 0$. Thus Φ is one-one.

16. The composition in one order is

$$N \mapsto \mathrm{Hom}_R(E, N) \mapsto \mathrm{Hom}_R(E, N) \otimes_{D_E} E. \qquad (*)$$

For $N = M_E$, the map Φ, when applied to the composition, recovers M_E, since Problem 15 says that Φ is onto. For general N, we can write $M_E = N \oplus N'$. When we apply Φ to $(*)$ for N and N' separately, we recover R submodules of N and N', respectively. To have a match for all of M_E, we must recover all of N and N'.

The composition in the other order is

$$W \mapsto W \otimes_{D_E} E \mapsto \mathrm{Hom}_R(E, W \otimes_{D_E} E). \qquad (**)$$

For $W = M^E$, the image corresponds under the map $\mathrm{Hom}(1, \Phi)$ to $\mathrm{Hom}_R(E, M_E) = M^E$. For general W, we can write $M^E = W \oplus W'$. When we apply $\mathrm{Hom}(1, \Phi)$ to $(**)$ for W, we get an R submodule of M^E that contains W. In fact, for any $w \in W$, $\mathrm{Hom}_R(E, E \otimes_{D_E} E)$ contains the map $e \mapsto w \otimes e$. Composing with Φ gives $e \mapsto w(e)$. Thus the members of W are in the image. Similarly the members of W' are in the image for W'. The direct sum of the images must be M^E, and thus the images must be exactly W and W'.

17. The computation

$$\varphi(\Phi_M(\psi \otimes e)) = \varphi(\psi(e)) = (\varphi \circ \psi)(e)) = \Phi_N((\varphi \circ \psi) \otimes e) = \Phi_N(\varphi^E(\psi) \otimes e)$$

proves the formula in the last line of the statement of the problem. For the inverse, suppose we are given a map $\tau \in \mathrm{Hom}_{D_E}(M^E, N^E)$. Then τ induces an R linear map

$$\tau'_E : M^E \otimes_{D_E} E \to N^E \otimes_{D_E} E$$

defined by

$$\tau'_E(\psi \otimes e) = \tau(\psi) \otimes e.$$

Composition with the isomorphism of Problem 15 gives an R homomorphism

$$\tau_E = \Phi_N \circ \tau'_E \circ \Phi_M^{-1} : M_E \to N_E.$$

We show that $\varphi \mapsto \varphi^E$ and $\tau \mapsto \tau_E$ are inverses. If a map φ in $\mathrm{Hom}_R(M_E, N_E)$ is given, we are to calculate $(\varphi^E)_E \in \mathrm{Hom}_R(M_E, N_E)$. It is enough to find the effect of $(\varphi^E)_E$ on elements $\Phi_M(\psi \otimes e)$ with $\psi \in M^E$ and $e \in E$. For such an element,

$$\begin{aligned}(\varphi^E)_E(\Phi_M(\psi \otimes e)) &= \Phi_N((\varphi^E)'(\psi \otimes e)) = \Phi_N(\varphi^E(\psi) \otimes e)\\ &= \varphi^E(\psi)(e) = \varphi(\psi(e)) = \varphi(\Phi_M(\psi \otimes e)).\end{aligned}$$

Thus $(\varphi^E)_E = \varphi$. Similarly for $\tau \in \mathrm{Hom}_{D_E}(M^E, N^E)$, we find that $(\tau_E)^E = \tau$. Thus $\varphi \mapsto \varphi^E$ and $\tau \mapsto \tau_E$ are inverses.

18. Let us write $M = \bigoplus_{s\in S} M_s$ with each M_s semisimple. Each M_s is contained in some M_E, and hence $M = \sum_{E\in\mathcal{E}} M_E$. Let us see that the sum is direct. If M_E has nonzero intersection with $M_{E_1} + \cdots + M_{E_n}$, where $E_1, \ldots, E_n$ are simple R modules with no two isomorphic, then there is a nonzero R linear map from E into $M_{E_1} + \cdots + M_{E_n}$. We can write each M_{E_j} as a sum of simple R submodules isomorphic to E_j, and Proposition 10.1 shows that

$$M_{E_1} + \cdots + M_{E_n} = \bigoplus_{s\in T} M'_s$$

with each M'_s isomorphic to one of $E_1, \ldots, E_n$. If all of $E_1, \ldots, E_n$ are nonisomorphic with E, then Lemma 10.3 and Proposition 10.4a show that

$$\mathrm{Hom}_R(E, M_{E_1} + \cdots + M_{E_n}) = 0,$$

contradiction. We conclude that the sum $M = \sum_{E\in\mathcal{E}} M_E$ is direct. This proves the equality at the left in the displayed formula of the problem, and the isomorphism on the right in that display follows from Problem 15.

19. If N is a left R submodule of M, then $N_E \subseteq M_E$ for every E. Conversely the previous problem shows that a system of N_E's defines an R submodule N. Thus this problem is a restatement of Problem 16.

20. We have

$$\mathrm{Hom}_R(M, N) \cong \prod_{E\in\mathcal{E}} \mathrm{Hom}_R(M_E, N) = \prod_{E\in\mathcal{E}} \mathrm{Hom}_R(M_E, N_E),$$

and the rest follows from Problem 17.