

## Solutions (page 7)

**Solution 1**  $\begin{pmatrix} \rho_1 \\ \rho_2 \end{pmatrix}^{-1} = \begin{pmatrix} -5 & 4 \\ 4 & -3 \end{pmatrix}$  as  $\begin{vmatrix} 3 & 4 \\ 4 & 5 \end{vmatrix} = -1$  So  $\rho_1, \rho_2$  is a  $\mathbb{Z}$ -basis of  $\mathbb{Z}^2$  and

$$(m_1, m_2) = (10, 7) \begin{pmatrix} -5 & 4 \\ 4 & -3 \end{pmatrix} = (-22, 19). \text{ No, as } \begin{vmatrix} 3 & 5 \\ 5 & 6 \end{vmatrix} = -7 \neq \pm 1.$$

**Solution 2** The 6 elements of  $\mathbb{Z}^2/K$  are:  $1g_0 = K + e_1 + e_2$ ,  $2g_0 = K + 2e_2$ ,  $3g_0 = K + e_1$ ,  $4g_0 = K + e_2$ ,  $5g_0 = K + e_1 + 2e_2$ ,  $6g_0 = K$ . So  $\mathbb{Z}^2/K = \langle g_0 \rangle$  is cyclic with generator  $g_0$  and invariant factor 6.

**Solution 3** Each of  $(\bar{1}, \bar{0}), (\bar{1}, \bar{2}), (\bar{0}, \bar{2})$  has order 2 and generates a  $C_2$  type subgroup. The elements  $(\bar{0}, \bar{1}), (\bar{0}, \bar{3}), (\bar{1}, \bar{1}), (\bar{1}, \bar{3})$  each have order 4.  $H = \langle (\bar{1}, \bar{0}) \rangle \oplus \langle (\bar{0}, \bar{2}) \rangle$  has isomorphism type  $C_2 \oplus C_2$ .  $\langle (\bar{0}, \bar{1}) \rangle = \langle (\bar{0}, \bar{3}) \rangle$  and  $\langle (\bar{1}, \bar{1}) \rangle = \langle (\bar{1}, \bar{3}) \rangle$  are  $C_4$  type subgroups.  $\langle (\bar{0}, \bar{0}) \rangle$  and  $G'$  are subgroups of type  $C_1$  and  $C_2 \oplus C_4$  respectively.  $H_1$  is either  $\langle (\bar{1}, \bar{2}) \rangle$  or  $\langle (\bar{1}, \bar{0}) \rangle$ ,  $H_2$  is either  $\langle (\bar{0}, \bar{1}) \rangle$  or  $\langle (\bar{1}, \bar{1}) \rangle$ .

## Solutions 1.1 (page 16)

### Solution 1

$\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ -3 & 1 \end{pmatrix}, \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ -3 & 1 \end{pmatrix}$ . (a)  $r_2 - 3r_1, c_1 - 3c_2$  are paired. (b)  $r_1 + 2r_2, c_2 - 2c_1$  are conjugate. The eros  $-r_1, r_1 + lr_i$  where  $l \in \mathbb{Z}$  ( $i > 1$ ), leave rows 2, 3, ... unchanged.

### Solution 2

$$(i) D = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix} \text{ and, for example, } P = \begin{pmatrix} -1 & 1 \\ 3 & -2 \end{pmatrix}, Q = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

$$(ii) D = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix} \text{ and, for example, } P = \begin{pmatrix} 1 & 0 \\ 3 & -1 \end{pmatrix}, Q = \begin{pmatrix} 7 & 12 \\ 3 & 5 \end{pmatrix}.$$

$$(iii) D = \begin{pmatrix} 1 & 0 \\ 0 & 150 \end{pmatrix} \text{ and, for example, } P = \begin{pmatrix} 1 & 1 \\ 15 & 14 \end{pmatrix}, Q = \begin{pmatrix} 21 & 10 \\ 2 & 1 \end{pmatrix}.$$

### Solution 3

$$(i) D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 18 \end{pmatrix}, P = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 0 & -1 \\ 7 & 1 & -5 \end{pmatrix}, Q = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 5 \\ 0 & 0 & 1 \end{pmatrix}, x = (0, 0, 0).$$

$$(ii) D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, P = \begin{pmatrix} 1 & 0 & 0 \\ 2 & -1 & 0 \\ -1 & 2 & -1 \end{pmatrix}, Q = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}, x = (-l, 2l, -l), l \in \mathbb{Z}.$$

$$(iii) D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 12 \end{pmatrix}, P = \begin{pmatrix} 1 & 1 & 0 \\ 2 & 3 & 1 \\ 4 & 6 & 3 \end{pmatrix}, Q = \begin{pmatrix} 3 & 2 & 0 \\ 3 & 3 & 2 \\ 1 & 1 & 1 \end{pmatrix}, x = (0, 0, 0).$$

Note that the *eros*  $r_1 + r_2$  and  $r_2 + r_1$  are different!

#### Solution 4

$$(i) D = (6, 0), Q = \begin{pmatrix} 12 & 7 \\ 5 & 3 \end{pmatrix}, 72/6 = 12, 42/6 = 7, \det Q = 1, a = 3, b = -5.$$

$$(ii) D = (1, 0), Q = \begin{pmatrix} 34 & 55 \\ 13 & 21 \end{pmatrix}, \det Q = -1, a = -21, b = 13.$$

$$(iii) D = (119, 0), Q = \begin{pmatrix} 63 & 46 \\ 25 & 19 \end{pmatrix}, 7497/119 = 63, 5474/119 = 46, \det Q = 1, a = 19, b = -26.$$

#### Solution 5

(a) Apply  $c_1 + c_2, c_2 - c_1$ .

(b) Applying the sequence to  $A = (a, b)$  gives  $(b, -a)$ . Applying it to  $(b, -a)$  gives  $(-a, -b)$  and applying it to  $(-a, -b)$  gives  $(-b, a)$ . One of these has non-negative entries.

(c) Use (b) to get  $(a, b)$  with  $a \geq 0, b \geq 0$ . Assuming  $b > 0$  as we can, the Euclidean algorithm applies  $c_1 - qc_2$  to  $(a, b)$  and  $c_2 - qc_1$  to  $(b, a)$ , where  $a = bq + r, a \geq b > 0$ , ultimately ending with  $(0, d)$  or  $(d, 0) = D$ . Use part (a) if needed to finish. Then  $Q$  is the product of the corresponding elementary matrices all of which have determinant 1, and so  $\det Q = 1$ .

(d)  $P = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$  where  $ad - bc = 1$ . So  $\gcd\{a, b\} = 1$  and  $(a, b)$  reduces to  $(1, 0)$  using only *ecos* of

type (iii). So  $P \equiv \begin{pmatrix} 1 & 0 \\ e & 1 \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$  on applying these *ecos* and finally  $c_1 - ec_2$ , as determinants are unchanged by such *ecos*.

(e) Apply  $c_2 - c_1, c_1 - 3c_2, c_1 + c_2, c_2 - c_1, c_1 - c_2$

#### Solution 6

(a)  $P^T = P$  for *eros* of types (i) and (ii). Otherwise  $P, P^T$  correspond to  $r_i + lr_j, r_j + lr_i$ .

(b) As  $D$  is symmetric, transposing  $PA = DQ$  gives  $A^T P^T = Q^T D$ . Hence

$(Q^T PA)^T = A^T P^T Q = Q^T DQ = Q^T PA$  showing  $Q^T PA$  symmetric. As  $Q^T P$  is a product of elementary matrices, every square matrix  $A$  over  $\mathbb{Z}$  can be made symmetric by applying *eros*.

(c) The sequence  $c_2 - 3c_1, c_1 - 2c_2, c_1 \leftrightarrow c_2, r_2 + 19r_1$ , reduces  $A$  to  $D = \text{diag}(1, 46)$ .

Postmultiplication by  $Q^{-1}$  carries out the above *ecos*, and so premultiplication by  $Q^T P$  carries out

$$r_2 + 19r_1, r_1 \leftrightarrow r_2, r_1 + 2r_2, r_2 + 3r_1 \text{ and changes } A \text{ into } Q^T PA = \begin{pmatrix} 50 & 152 \\ 152 & 463 \end{pmatrix}.$$

### Solutions 1.2 (page 30)

#### Solution 1

$$(a) TA = AT \text{ gives } A = \begin{pmatrix} a & b \\ b & a \end{pmatrix}$$

(b)  $e_j P_1 = e_j + l e_k$ ,  $e_i P_1 = e_i$  ( $i \neq j$ ) Postmultiply by  $A$ :  $e_j P_1 A = e_j A + l e_k A$ , i.e. row  $j$  of  $P_1 A$  is row  $j$  of  $A + l(\text{row } k \text{ of } A)$ . Also  $e_i P_1 A = e_i A$ , that is row  $i$  of  $P_1 A$  is row  $i$  of  $A$  for  $i \neq j$ . So  $P_1 A$  is the result of applying  $r_j + l r_k$  to  $A$ .

(c)  $Q_1 e_j^T = e_k^T$ ,  $Q_1 e_k^T = e_j^T$ ,  $Q_1 e_i^T = e_i^T$  ( $i \neq j, k$ ). Premultiply by  $A$ :  $A Q_1 e_j^T = A e_k^T$ ,  $A Q_1 e_k^T = A e_j^T$ ,  $A Q_1 e_i^T = A e_i^T$ , i.e. columns  $j$  and  $k$  of  $A Q_1$  are columns  $k$  and  $j$  respectively of  $A$ , column  $i$  of  $A Q_1$  is column  $i$  of  $A$  ( $i \neq j, k$ ). So  $A Q_1$  is the result of applying  $c_j \leftrightarrow c_k$  to  $A$ .

(d) Let  $P$  and  $Q$  denote respectively the  $s \times s$  and  $t \times t$  identity matrices over  $\mathbb{Z}$ . Then  $PAQ^{-1} = A$  showing (i)  $A \equiv A$  for all  $s \times t$  matrices  $A$  over  $\mathbb{Z}$ . Suppose  $A \equiv B$ . There are invertible  $P$  and  $Q$  over  $\mathbb{Z}$  with  $PAQ^{-1} = B$ . Then  $P^{-1}B(Q^{-1})^{-1} = A$  showing (ii)  $A \equiv B \Rightarrow B \equiv A$  since  $P^{-1}$  and  $Q^{-1}$  are invertible over  $\mathbb{Z}$ . Suppose  $A \equiv B$  and  $B \equiv C$ . There are invertible  $P_1, P_2, Q_1, Q_2$  over  $\mathbb{Z}$  with  $P_1 A Q_1^{-1} = B$  and  $P_2 B Q_2^{-1} = C$ . Then  $(P_2 P_1) A (Q_2 Q_1)^{-1} = C$  showing (iii)  $A \equiv B$  and  $B \equiv C \Rightarrow A \equiv C$  as  $P_2 P_1$  and  $Q_2 Q_1$  are invertible over  $\mathbb{Z}$ . So  $\equiv$  is an equivalence relation.

**Solution 2**  $D = \text{diag}(a, b, c)$  where  $(a, b, c)$  is

$(1, 1, 1), (1, 1, 2), (1, 1, 3), (1, 1, 4), (1, 1, 6), (1, 1, 12), (1, 2, 2), (1, 2, 6)$  as  $\det D = abc$ .

(a) Just one,  $D = \text{diag}(1, 1, \dots, 1, 105)$  as  $105 = 3 \times 5 \times 7$  is a product of different primes.

(b)  $\text{diag}(1, \dots, 1, 1, 100), \text{diag}(1, \dots, 1, 2, 50), \text{diag}(1, \dots, 1, 5, 20), \text{diag}(1, \dots, 1, 10, 10)$

i.e. four if  $s > 1$ . Just one for  $s = 1$ .

**Solution 3**  $c_j - (a_{1j}/a_{11})c_1$  for  $2 \leq j \leq t$ .

**Solution 4**

(a)  $c_1 \leftrightarrow c_2$ ,  $r_1 \leftrightarrow r_2$ ,  $r_2 + r_4$ ,  $c_4 - 2c_2$ ,  $c_2 - c_4$ ,  $c_4 - c_2$ ,  $r_4 + 5r_2$  giving  $D = \text{diag}(1, 5, 5, 50)$ .

(b)  $A \equiv \text{diag}(1, 60, 50, 200) \equiv \text{diag}(1, 10, 300, 200) \equiv \text{diag}(1, 10, 100, 600) = D$ , each equivalence as in (1.10) using 5 elementary operations.

(c) Let the operation  $\sigma_i$  create a new diagonal matrix from a given one with non-negative entries by replacing the  $i$ th and  $(i+1)$ st diagonal entries by their gcd and lcm as in (1.10). Each  $\sigma_i$  is the product of at most 5 elementary operations and applying  $\sigma_1, \sigma_2, \dots, \sigma_s$  to  $D_1$  produces  $S(D_1)$ .

(d) The proof of (1.10) goes through unchanged in the case  $l$  and  $m$  both negative. Suppose  $l$  and  $m$  have opposite signs. Let  $\gcd\{l, m\} = d = al + bm$  where  $a, b \in \mathbb{Z}$ . The following sequence of elementary operations changes  $\text{diag}(l, m)$  into Smith normal form  $\text{diag}(d, lm/d)$ :

$$c_2 + ac_1, r_1 + br_2, c_1 \leftrightarrow c_2, c_2 - (l/d)c_1, r_2 - (m/d)r_1.$$

As  $\gcd\{18, -24\} = 6 = (-1)18 + (-1)(-24)$  we see  $a = b = -1$ . So

$$\begin{aligned} \begin{pmatrix} 18 & 0 \\ 0 & -24 \end{pmatrix} &\equiv \begin{pmatrix} 18 & -18 \\ 0 & -24 \end{pmatrix} \equiv \begin{pmatrix} 18 & 6 \\ 0 & -24 \end{pmatrix} \equiv \begin{pmatrix} 6 & 18 \\ -24 & 0 \end{pmatrix} \\ &\equiv \begin{pmatrix} 6 & 0 \\ -24 & 72 \end{pmatrix} \equiv \begin{pmatrix} 6 & 0 \\ 0 & 72 \end{pmatrix}. \end{aligned}$$

Let the operation  $\sigma_{ij}$  for  $i < j$  replace the  $(i, i)$ -entry and the  $(j, j)$ -entry in a diagonal matrix by their gcd and lcm respectively. Each  $\sigma_{ij}$  is expressible as the product of at most 5 elementary operations. The sequence  $\sigma_{12}, \sigma_{13}, \dots, \sigma_{1s}, \sigma_{23}, \dots, \sigma_{2s}, \dots, \sigma_{s-1 s}$  of  $s(s-1)/2$  operations  $\sigma_{ij}$  creates  $D$  in Smith normal form and is the composition of at most  $5 \times s(s-1)/2 = (5/2)s(s-1)$  elementary operations.

### Solution 5

(a) Using the method of (1.9),  $s-1$  applications of (1.7) produce a matrix  $B_1 = (b_{ij})$  with  $e_1 B_1 = e_1$ . The  $s-1$  eros  $r_i - b_{i1} r_1$  for  $1 < i \leq s$  produce  $B$  with  $e_1 = e_1 B$ ,  $B e_1^T = e_1^T$ . Now use induction on  $s$ . If  $|P| = -1$  then the ero  $-r_s$  finally gives  $I$ .

(b) The sequence  $c_1 - 3c_2, c_1 \leftrightarrow c_2, r_2 - 6r_1, -r_2$  reduces  $P$  to  $I$ .

The sequence  $c_2 - 2c_1, c_3 - 3c_1, r_2 - 2r_1, r_3 - 3r_1, c_2 - 3c_3, c_2 \leftrightarrow c_3, r_3 - 6r_2, -r_3$  reduces  $Q$  to  $I$ .

(c)  $P = \begin{pmatrix} 1 & 0 \\ 6 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 3 & 1 \end{pmatrix}.$

(d) The eros  $r_2 - 2r_1, r_3 - 3r_1, r_3 - 6r_2, -r_3, r_2 \leftrightarrow r_3, r_3 - 3r_2, r_1 - 3r_3, r_1 - 2r_2$  reduce  $Q$  to  $I$ .

(e) The ecos  $c_2 - 2c_1, c_3 - 3c_1, c_2 - 3c_3, c_2 \leftrightarrow c_3, -c_3, c_2 - 6c_3, c_1 - 3c_3, c_1 - 2c_2$  reduce  $Q$  to  $I$ .

### Solution 6

(a)  $B_1 = \begin{pmatrix} 390 & 0 \\ 330 & 7 \end{pmatrix}, B_2 = \begin{pmatrix} 30 & 42 \\ 0 & -91 \end{pmatrix}, B_3 = \begin{pmatrix} 6 & 0 \\ 182 & -455 \end{pmatrix}, B_4 = \begin{pmatrix} 2 & -455 \\ 0 & 1365 \end{pmatrix},$   
 $B_5 = \begin{pmatrix} 1 & 0 \\ -1365 & 2730 \end{pmatrix}, B_6 = \begin{pmatrix} 1 & 0 \\ 0 & -2730 \end{pmatrix}, D = \begin{pmatrix} 1 & 0 \\ 0 & 2730 \end{pmatrix}.$

(b)  $a_2 = 10, a_3 = 90, a_4 = 3690$ . On dividing  $a_{n+1}$  by  $2a_n$  the remainder is  $a_n$ , i.e.

$a_{n+1} = 2a_n(2a_{n-1}) + a_n$ . So  $\gcd\{a_{n+1}, 2a_n\} = \gcd\{2a_n, a_n\} = a_n$ . As

$a_{n+1} = a_n(4a_{n-1} + 1), a_n = a_{n-1}(4a_{n-2} + 1), \dots, a_2 = 2(4a_0 + 1)$  on substituting for the factors

$a_n, a_{n-1}, \dots, a_2$  we obtain the factorisation  $a_{n+1} = 2(4a_{n-1} + 1)(4a_{n-2} + 1) \cdots (4a_0 + 1)$ .

Hence  $a_n/a_{n-r+1} = (4a_{n-2} + 1) \cdots (4a_{n-r+1} + 1)(4a_{n-r} + 1)$ ; as each factor is congruent to 1 modulo  $4a_{n-r}$  so also is their product  $a_n/a_{n-r+1}$ . Therefore  $a_n/a_{n-r+1} = q_{n-r}a_{n-r} + 1$  where  $q_{n-r}$  is a positive multiple of 4 for  $r > 1$ .

$$\begin{aligned} \begin{pmatrix} 41 \times 90 & 2 \times 90 \\ 0 & -1 \end{pmatrix} &\equiv \begin{pmatrix} 90 & 2 \times 90 \\ 20 & -1 \end{pmatrix} \equiv \begin{pmatrix} 90 & 0 \\ 20 & -41 \end{pmatrix} = B_1 \equiv \begin{pmatrix} 10 & 4 \times 41 \\ 20 & -41 \end{pmatrix} \\ &\equiv \begin{pmatrix} 10 & 4 \times 41 \\ 0 & -9 \times 41 \end{pmatrix} = B_2 \equiv \begin{pmatrix} 10 & 4 \\ 0 & -9 \times 41 \end{pmatrix} \equiv \begin{pmatrix} 2 & 4 \\ 18 \times 41 & -9 \times 41 \end{pmatrix} \\ &\equiv \begin{pmatrix} 2 & 0 \\ 18 \times 41 & -45 \times 41 \end{pmatrix} = B_3 \equiv \begin{pmatrix} 2 & 0 \\ 0 & -45 \times 41 \end{pmatrix} = B_4. \end{aligned}$$

Applying the *ecos*  $c_1 - 2a_{n-2}c_2$ ,  $c_2 - 2c_1$  to  $A$  produces  $B_1 = \begin{pmatrix} a_{n-1} & 0 \\ 2a_{n-2} & -a_n/a_{n-1} \end{pmatrix}$ . Applying the *eros*  $r_1 - 2a_{n-3}r_2$ ,  $r_2 - 2r_1$  to  $B_1$  produces  $B_2 = \begin{pmatrix} a_{n-2} & 2a_{n-3} \\ 0 & -a_n/a_{n-2} \end{pmatrix}$ . Suppose  $n > r > 2$  and that  $B_{r-1}$  is as stated. Take  $r$  even and so

$$B_{r-1} = \begin{pmatrix} a_{n-r+1} & 0 \\ 2a_{n-r}(a_n/a_{n-r+2}) & -(a_n/a_{n-r+1}) \end{pmatrix}.$$

The method of (1.9) tells us to apply the Euclidean algorithm to the integers in column 1 of  $B_{r-1}$ : apply the *ero*  $r_2 - 2a_{n-r}q_{n-r+1}r_1$  (giving  $(2,1)$ -entry  $2a_{n-r}$  and leaving the other entries unchanged) for  $r > 2$  followed by the *eros*  $r_1 - 2a_{n-r-1}r_2$ ,  $r_2 - 2r_1$  obtaining  $B_r$  for  $r \geq 4$ . The case  $r$  odd,  $r \geq 3$  is simply the matrix transpose of this and so the induction is completed. Therefore  $B_{n-1}$  or  $B_{n-1}^T$  equals

$$\begin{pmatrix} a_1 & 2a_0(a_n/a_2) \\ 0 & -a_n/a_1 \end{pmatrix} = \begin{pmatrix} 2 & a_n/5 \\ 0 & -a_n/2 \end{pmatrix}$$

according as  $n$  is odd or even. Either  $c_2 - (a_n/10)c_1$  or  $r_2 - (a_n/10)r_1$  gives  $B_n = \text{diag}(2, -a_n/2)$ . Using the method of (1.9), for  $n \geq 3$  two *ecos* change  $A_n$  into  $B_1$ , two *eros* change  $B_1$  into  $B_2$ , three *ecos* change  $B_2$  into  $B_3$ , three *eros* change  $B_3$  into  $B_4, \dots$ , three elementary operations change  $B_{n-2}$  into  $B_{n-1}$  and finally just one elementary operation changes  $B_{n-1}$  into  $B_n$ . A further 5 elementary operations change  $B_n$  into  $S(A_n) = \text{diag}(1, a_n)$ , namely  $r_1 - r_2$ ,  $c_2 - (q_0/2)c_1$  where  $a_n/2 = a_n/a_1 = q_0 + 1$ ,  $c_1 - 2c_2$ ,  $c_1 \leftrightarrow c_2$ ,  $r_2 + (a_n/2)r_1$ . So a total of  $10 + 3(n-3) = 3n+1$  elementary operations are needed to carry out the algorithm of (1.11).

On the other hand the four elementary operations  $r_1 + 2a_{n-1}r_2$ ,  $-r_2$ ,  $r_1 \leftrightarrow r_2$ ,  $c_1 \leftrightarrow c_2$  change  $A_n$  into  $S(A_n) = \text{diag}(1, a_n)$ . So  $A_n$  is an ‘awkward’ matrix as far as the algorithm (1.11) is concerned!

### Solution 7

(a) Suppose  $(P, Q), (P', Q') \in Z(D)$ . Then  $PP'D = PDQ' = DQQ'$  shows

$(P, Q)(P', Q') = (PP', QQ') \in Z(D)$ . Also

$PD = DQ \Rightarrow DQ^{-1} = P^{-1}D \Rightarrow (P, Q)^{-1} = (P^{-1}, Q^{-1}) \in Z(D)$ . As  $ID = DI$  we conclude that  $Z(D)$  is a subgroup of  $G$ .

(b) Suppose  $(P, Q) \in Z(D)$ . Comparing entries in  $PD = DQ$  gives  $p_{ij}d_j = d_iq_{ij}$  for all  $i, j$ . So  $i \leq j$  gives  $(d_j/d_i)p_{ij} = q_{ij}$ . As  $p_{ji}d_i = d_jq_{ji}$  for  $i \leq j$  we obtain  $p_{ji} = (d_j/d_i)q_{ji}$ .

(c)  $P'A = DQ'$  and  $P''A = DQ''$  give  $A = (P')^{-1}DQ' = (P'')^{-1}DQ''$  and so  $DQ'(Q'')^{-1} = P'(P'')^{-1}D$ , that is,  $(P', Q')(P'', Q'')^{-1} \in Z(D)$ . (Looking ahead this means that  $(P', Q')$  and  $(P'', Q'')$  belong to the same left coset of  $Z(D)$  in  $G$ .)

$$(d) \left( \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}, \begin{pmatrix} 2 & 3 \\ 1 & 1 \end{pmatrix} \right) \left( \begin{pmatrix} 2 & -1 \\ -5 & 3 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right)^{-1} = \left( \begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix}, \begin{pmatrix} 3 & 2 \\ 1 & 1 \end{pmatrix} \right).$$

(e) Suppose  $D = \text{diag}(d_1, \dots, d_r, 0, \dots, 0)$  where  $d_r > 0$ ,  $t > r$ . Partition  $P = \left( \begin{array}{c|c} P_1 & P_2 \\ \hline P_3 & P_4 \end{array} \right)$  and

$Q = \left( \begin{array}{c|c} Q_1 & Q_2 \\ \hline Q_3 & Q_4 \end{array} \right)$  where  $P_1$  and  $Q_1$  are  $r \times r$  matrices. Then

$(P, Q) \in Z(D) \Leftrightarrow (P_1, Q_1) \in Z(\text{diag}(d_1, \dots, d_r))$ ,  $P_3 = 0$ ,  $Q_2 = 0$ ,  $P_2$  and  $Q_3$  arbitrary,  $P_4$  and  $Q_4$  are invertible  $(t-r) \times (t-r)$  matrices over  $\mathbb{Z}$ .

### Solutions 1.3 (page 43)

#### Solution 1

(a) (i)  $d = \gcd\{21, \gcd\{75, 175\}\} = \gcd\{21, 25\} = 1 = 6 \times 21 + 10 \times 75 - 5 \times 175$ .

(ii)  $d = \gcd\{\gcd\{42, 66\}, \gcd\{154, 231\}\} = \gcd\{6, 77\} = 1 = -39 \times 42 + 26 \times 66 + 1 \times 154 - 1 \times 231$ .

(b) Let  $d_1 = \gcd X_1$ ,  $d_2 = \gcd X_2$  and  $d = \gcd\{d_1, d_2\}$ . For  $l \in X_1 \cup X_2$  either  $d_1 \mid l$  or  $d_2 \mid l$  and so certainly  $d \mid l$ . Let  $d' \mid l$  for all  $l \in X_1 \cup X_2$ . Then  $d' \mid l$  for all  $l \in X_1$  and so  $d' \mid d_1$ . Also  $d' \mid l$  for all  $l \in X_2$  and so  $d' \mid d_2$ . Hence  $d' \mid d$ . So  $d$  is the non-negative gcd of  $X_1 \cup X_2$  by (1.16), i.e.

$$\gcd\{\gcd X_1, \gcd X_2\} = \gcd\{d_1, d_2\} = d = \gcd X_1 \cup X_2.$$

(c) Write  $d = \gcd\{l_1, l_2, \dots, l_k\}$  and  $\pi = n_1 n_2 \cdots n_k$ . Then  $m_0 = \pi/d = n_i l_i/d = n_i (l_i/d)$ , i.e.  $n_i \mid m_0$  as  $d \mid l_i$  for  $1 \leq i \leq k$ , showing (i). Let  $m \in \mathbb{Z}$  satisfy  $n_i \mid m$  for  $1 \leq i \leq k$ . So there are  $m_i \in \mathbb{Z}$  with  $n_i m_i = m$  for  $1 \leq i \leq k$ . By (1.16) there are  $a_1, a_2, \dots, a_k \in \mathbb{Z}$  with  $d = a_1 l_1 + a_2 l_2 + \dots + a_k l_k$ . So  $m = m \times 1 = a_1 l_1 m/d + a_2 l_2 m/d + \dots + a_k l_k m/d = a_1 l_1 n_1 m_1 + a_2 l_2 n_2 m_2 + \dots + a_k l_k n_k m_k = a_1 \pi m_1/d + a_2 \pi m_2/d + \dots + a_k \pi m_k/d = (a_1 m_1 + a_2 m_2 + \dots + a_k m_k) \pi/d$  as  $l_i n_i = \pi$  for  $1 \leq i \leq k$ . So  $m_0 \mid m$  as  $a_1 m_1 + a_2 m_2 + \dots + a_k m_k \in \mathbb{Z}$  showing (ii). Therefore  $m_0 = \text{lcm}\{n_1, n_2, \dots, n_k\}$ .

#### Solution 2

(a)  $K = \langle 3 \rangle$  as  $\gcd\{63, 231, 429\} = \gcd\{\gcd\{63, 231\}, 429\} = \gcd\{21, 429\} = 3$ .

$$\{k \in K : -10 < k < 10\} = \{-9, -6, -3, 0, 3, 6, 9\}.$$

(b)  $\langle d \rangle \subseteq \langle d' \rangle \Rightarrow d \in \langle d' \rangle \Rightarrow d' \mid d \Rightarrow d = qd' (q \in \mathbb{Z}) \Rightarrow bd = bqd' (\text{for all } b \in \mathbb{Z}) \Rightarrow \langle d \rangle \subseteq \langle d' \rangle$ .

The ideals of  $\mathbb{Z}$  which contain  $\langle 30 \rangle$  are generated by non-negative divisors of 30 and so are  $\langle 1 \rangle, \langle 2 \rangle, \langle 3 \rangle, \langle 5 \rangle, \langle 6 \rangle, \langle 10 \rangle, \langle 15 \rangle, \langle 30 \rangle$ .  $\mathbb{Z} = \langle 1 \rangle$  and  $\langle p \rangle$  are the only ideals of  $\mathbb{Z}$  containing  $\langle p \rangle$ , as 1 and  $p$  are the only non-negative divisors of  $p$ . The ideals of  $\mathbb{Z}$  containing  $\langle 64 \rangle$  are  $\langle 64 \rangle \subset \langle 32 \rangle \subset \langle 16 \rangle \subset \langle 8 \rangle \subset \langle 4 \rangle \subset \langle 2 \rangle \subset \langle 1 \rangle$  in ascending order.

(c) Let  $k \in K$ . Then  $1k = k \in K$ . Suppose  $nk \in K$  for some positive integer  $n$ . Then  $(n+1)k = nk + k \in K$  as  $K$  is closed under addition. So  $nk \in K$  for all positive integers  $n$ .  $0k = 0 \in K$ . Also  $(-n)k = -nk \in K$  as  $K$  is closed under negation. So  $mk \in K$  for all  $m \in \mathbb{Z}$  and hence  $K$  is an ideal of  $\mathbb{Z}$ .

(d) We show that  $K_1 \cap K_2$  satisfies the ideal conditions (i)–(iv) using the fact that  $K_1$  and  $K_2$  individually satisfy these conditions. Consider  $k, k' \in K_1 \cap K_2$ . Then  $k, k' \in K_1$  and  $k, k' \in K_2$ . As  $K_1$  and  $K_2$  are closed under addition we see  $k + k' \in K_1$  and  $k + k' \in K_2$ . Hence  $k + k' \in K_1 \cap K_2$ , verifying (i) for  $K_1 \cap K_2$ . As  $0 \in K_1$  and  $0 \in K_2$  we see  $0 \in K_1 \cap K_2$ , verifying (ii) for  $K_1 \cap K_2$ . As  $-k \in K_1$  and  $-k \in K_2$  we see  $-k \in K_1 \cap K_2$ , verifying (iii) for  $K_1 \cap K_2$ . As  $bk \in K_1$  and  $bk \in K_2$  we see  $bk \in K_1 \cap K_2$  for all  $b \in \mathbb{Z}$ , verifying (iv) for  $K_1 \cap K_2$ . Therefore  $K_1 \cap K_2$  is an ideal of  $\mathbb{Z}$ .

We show that  $K_1 + K_2$  satisfies the ideal conditions (i)–(iv) using the fact that  $K_1$  and  $K_2$  individually satisfy these conditions. Consider  $k, k' \in K_1 + K_2$ . There are  $k_1, k'_1 \in K_1$  and  $k_2, k'_2 \in K_2$  with  $k = k_1 + k_2$ ,  $k' = k'_1 + k'_2$ . Then  $k_1 + k'_1 \in K_1$  and  $k_2 + k'_2 \in K_2$  and so  $k + k' = (k_1 + k_2) + (k'_1 + k'_2) = (k_1 + k'_1) + (k_2 + k'_2) \in K_1 + K_2$ , showing that  $K_1 + K_2$  is closed under addition, i.e. (i) holds. As  $0 \in K_1$  and  $0 \in K_2$  we see  $0 = 0 + 0 \in K_1 + K_2$ , verifying (ii). As  $-k_1 \in K_1$  and  $-k_2 \in K_2$  we obtain  $-k = (-k_1) + (-k_2) \in K_1 + K_2$ , verifying (iii). For  $b \in \mathbb{Z}$  we know  $bk_1 \in K_1$  and  $bk_2 \in K_2$ . Hence  $bk = b(k_1 + k_2) = bk_1 + bk_2 \in K_1 + K_2$ , verifying (iv). So  $K_1 + K_2$  is an ideal of  $\mathbb{Z}$ .

By (1.15) there are non-negative integers  $d_1$  and  $d_2$  with  $K_1 = \langle d_1 \rangle$  and  $K_2 = \langle d_2 \rangle$ . Also by (1.15) there are non-negative integers  $d$  and  $l$  with  $K_1 + K_2 = \langle d \rangle$  and  $K_1 \cap K_2 = \langle l \rangle$ . As  $0 \in K_2$  we see  $k_1 = k_1 + 0 \in K_1 + K_2$  for all  $k_1 \in K_1$ , i.e.  $K_1 \subseteq K_1 + K_2$ , i.e.  $\langle d_1 \rangle \subseteq \langle d \rangle$ . By part (b) above  $d \mid d_1$ . In the same way  $d \mid d_2$  and so  $d$  is a common divisor of  $d_1$  and  $d_2$ . Let  $d'$  be a common divisor of  $d_1$  and  $d_2$ . Then  $\langle d_1 \rangle \subseteq \langle d' \rangle$  and  $\langle d_2 \rangle \subseteq \langle d' \rangle$  by (b) above and hence  $\langle d \rangle = K_1 + K_2 = \langle d_1 \rangle + \langle d_2 \rangle \subseteq \langle d' \rangle$  as the ideal  $\langle d' \rangle$  is closed under addition. By (b) above  $d' \mid d$ . Therefore  $d = \gcd\{d_1, d_2\}$  by (1.16). In a similar way  $\langle l \rangle = K_1 \cap K_2 \subseteq K_1 = \langle d_1 \rangle$  showing  $d_1 \mid l$  by (b) above. Also  $d_2 \mid l$  and so  $l$  is a common multiple of  $d_1$  and  $d_2$ . Let  $l'$  be a common multiple of  $d_1$  and  $d_2$ . Then  $\langle l' \rangle \subseteq \langle d_1 \rangle$  and  $\langle l' \rangle \subseteq \langle d_2 \rangle$  by (b) above. Hence  $\langle l' \rangle \subseteq \langle d_1 \rangle \cap \langle d_2 \rangle = K_1 \cap K_2 = \langle l \rangle$ . By (b) above  $l \mid l'$ . Therefore  $l = \text{lcm}\{d_1, d_2\}$  by Question 1 (c) above with  $k = 2$ .

### Solution 3

(a) The number of  $l$ -minors is  $\binom{5}{l} \times \binom{6}{l}$ , i.e. 30, 150, 200, 75, 6 for  $l = 1, 2, 3, 4, 5$  respectively.

(b)(i)  $g_1(A) = 1$ ,  $g_2(A) = 2$ ,  $S(A) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \end{pmatrix}$ .

(ii)  $g_1(A) = 3$ ,  $g_2(A) = 54$ ,  $S(A) = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 18 & 0 \end{pmatrix}$ .

(iii)  $g_1(A) = 1$ ,  $g_2(A) = 30$ ,  $g_3(A) = 900$ ,  $S(A) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 30 & 0 \\ 0 & 0 & 30 \end{pmatrix}$ .

(c)

$$\begin{pmatrix} 1 & 1 & 1 \\ 2 & 5 & 2 \\ 4 & 4 & 7 \end{pmatrix} \begin{pmatrix} 27 & -3 & -3 \\ -6 & 3 & 0 \\ -12 & 0 & 3 \end{pmatrix} = \begin{pmatrix} 9 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 9 \end{pmatrix} = \begin{pmatrix} 27 & -3 & -3 \\ -6 & 3 & 0 \\ -12 & 0 & 3 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 2 & 5 & 2 \\ 4 & 4 & 7 \end{pmatrix}.$$

The method of (1.11) gives  $P = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -4 & 0 & 1 \end{pmatrix}$  and  $Q = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$  satisfying

$PAQ^{-1} = \text{diag}(1, 3, 3) = S(A)$ .  $Q(\text{adj } A)P^{-1} = \text{diag}(9, 3, 3) = \text{adj } S(A)$ . Interchanging rows 1 and 3, and also columns 1 and 3, gives  $S(\text{adj } A) = \text{diag}(3, 3, 9)$ .

(d)  $(PAQ^{-1})(Q(\text{adj } A)P^{-1}) = P(A\text{adj } A)P^{-1} = P(\det A)IP^{-1} = (\det A)I$ . Pre-multiplying by  $(PAQ^{-1})^{-1} = (\text{diag}(d_1, \dots, d_t))^{-1}$  gives  $Q(\text{adj } A)P^{-1} = \text{diag}(\det A/d_1, \dots, \det A/d_t)$ . Taking determinants of  $PAQ^{-1} = S(A)$  gives  $\det A = \pm d_1 \cdots d_t$  and as  $\det A > 0$  we see  $\det A = d_1 \cdots d_t$ . So the  $(i, i)$ -entry in  $Q(\text{adj } A)P^{-1}$  is  $d_1 \cdots d_t / d_i$  which is the  $(i, i)$ -entry in the diagonal matrix  $\text{adj } S(A)$ . Hence  $Q(\text{adj } A)P^{-1} = \text{adj } S(A)$  which differs from  $S(\text{adj } A)$  only in that the diagonal entries appear in the opposite order, i.e.  $\text{adj } A$  has  $i$ th invariant factor  $d_1 \cdots d_t / d_{t+1-i}$ .

$Q(\text{adj } A)P^{-1} = -\text{adj } S(A)$  for  $\det A < 0$ ; otherwise no change.

$S(A) = \text{diag}(d_1, \dots, d_{t-1}, 0)$ ,  $d_{t-1} > 0$  in the case  $\text{rank } A = t-1$ . As  $S(A)$  and  $Q(\text{adj } A)P^{-1}$  have zero product, i.e.  $S(A)Q(\text{adj } A)P^{-1} = 0I = Q(\text{adj } A)P^{-1}S(A)$ . Comparing entries gives

$Q(\text{adj } A)P^{-1} = \text{diag}(0, \dots, 0, x)$ . By (1.20)

$x = \pm g_1(\text{adj } A) = \pm g_{t-1}(A) = \pm g_{t-1}(S(A)) = \pm d_1 d_2 \cdots d_{t-1}$ . Hence

$S(\text{adj } A) = \text{diag}(d_1 d_2 \cdots d_{t-1}, 0, 0, \dots, 0)$ .

In the case  $\text{rank } A \leq t-2$  we have  $S(A) = \text{diag}(d_1, \dots, d_{t-2}, 0, 0)$  and so  $g_{t-1}(A) = g_{t-1}(S(A)) = 0$ . So  $\text{adj } A = 0 = S(\text{adj } A)$ .

(e) Take  $t = 2$  and let  $S(A) = \text{diag}(d_1, d_2)$ . Then by (d) above  $S(\text{adj } A) = \text{diag}(d_1, d_2)$  also. So  $A \equiv \text{adj } A$  for all  $2 \times 2$  matrices  $A$  over  $\mathbb{Z}$ .

Take  $t = 3$  and write  $S(A) = \text{diag}(d_1, d_2, d_3)$ . Then  $S(\text{adj } A) = \text{diag}(d_1 d_2, d_1 d_3, d_2 d_3)$ . For  $d_1 = 0$  we see  $A \equiv \text{adj } A$  as  $A = 0 = \text{adj } A$  since  $d_2 = d_3 = 0$ . For  $d_1 > 0, d_2 = 0$  we have

$\text{rank } A = 1, \text{rank } (\text{adj } A) = 0$  as  $d_3 = 0$  and so  $A \not\equiv \text{adj } A$ . For  $d_2 > 0, d_3 = 0$  we see

$\text{rank } A = 2, \text{rank } (\text{adj } A) = 1$  as  $d_1 > 0$  and so  $A \not\equiv \text{adj } A$ . For  $d_3 > 0$  then  $A$  and  $\text{adj } A$  both have

$\text{rank } 3$ . Suppose  $A \equiv \text{adj } A$ . Then  $S(A) = S(\text{adj } A)$ , i.e.  $d_1 = d_1 d_2, d_2 = d_1 d_3$  giving  $d_1 = d_2 = d_3 = 1$ .

So  $A$  is invertible over  $\mathbb{Z}$ . Conversely each matrix  $A$  which is invertible over  $\mathbb{Z}$  satisfies  $A \equiv \text{adj } A$  as  $\text{adj } A$  is also invertible over  $\mathbb{Z}$ , i.e.  $S(A) = I = S(\text{adj } A)$ .

Suppose  $t \geq 3$ . The argument used in the above paragraph generalises to show

$A \equiv \text{adj } A \Leftrightarrow$  either  $A = 0$  or  $A$  is invertible over  $\mathbb{Z}$ .

#### Solution 4

(a)



$$\det BB^T = \begin{vmatrix} 21 & 39 \\ 39 & 83 \end{vmatrix} = 222. \quad \sum_Y \det B_Y \det_Y B^T = (\det B_{\{1,2\}})^2 + (\det B_{\{1,3\}})^2 + (\det B_{\{2,3\}})^2 =$$

$$\begin{vmatrix} 2 & 1 \\ 3 & 5 \end{vmatrix}^2 + \begin{vmatrix} 2 & 4 \\ 3 & 7 \end{vmatrix}^2 + \begin{vmatrix} 1 & 4 \\ 5 & 7 \end{vmatrix}^2 = 7^2 + 2^2 + (-13)^2 = 222.$$

(b)

$$\begin{vmatrix} 1 & 2 \\ 2 & x \end{vmatrix}^2 + \begin{vmatrix} 1 & 3 \\ 2 & y \end{vmatrix}^2 + \begin{vmatrix} 2 & 3 \\ x & y \end{vmatrix}^2 = \det BB^T = 0. \quad \text{Each determinant is zero. So } x=4, y=6.$$

(c)  $\det BB^T = \sum_M M^2$  where  $M$  runs through the  $s$ -minors of  $B$ . Hence

$$\det BB^T = 0 \Leftrightarrow \text{each } M = 0 \Leftrightarrow g_s(B) = 0.$$

(d)  $BC = B'C'$  and so  $\det BC = \det B'C' = (\det B')(\det C') = 0 \times 0 = 0$  by (1.18) as column  $l$  of  $B'$  is zero and row  $l$  of  $C'$  is zero.**Solution 5**(a)  $(PA)_Y = PA_Y \equiv A_Y$ . So  $g_l((PA)_Y) = g_l(A_Y)$  by (1.20).Suppose  $A$  can be changed into  $S(A) = D = \text{diag}(d_1, d_2, \dots, d_t)$  using *eros* only. By (1.4) there is an  $s \times s$  invertible matrix  $P$  over  $\mathbb{Z}$  with  $PA = D$ . Then  $g_l(A_Y) = g_l(D_Y) = \prod_{j \in Y} d_j$ , the product of theinvariant factors  $d_j$  for  $j \in Y$  where  $l = |Y|$ . In particular  $g_1(A_{\{j\}}) = d_j$  for  $1 \leq j \leq t$  and

$$g(A_{\{1,2,\dots,l\}}) = d_1 d_2 \cdots d_l \text{ for } 2 \leq l \leq t.$$

For the converse we use induction on  $t$  and (1.13). Suppose there is  $P$  invertible over  $\mathbb{Z}$  with $PA_Y = \text{diag}(d_1, d_2, \dots, d_{t-1})$  where  $Y = \{1, 2, \dots, t-1\}$ . If  $d_t = 0$  then  $g_1(A_{\{t\}}) = d_t = 0$  and so  $A_{\{t\}} = 0$  giving  $PA = \text{diag}(d_1, d_2, \dots, d_{t-1}, 0) = S(A)$ . So we may assume  $d_t \neq 0$ . Write  $b_{it}$  for the $(i, t)$ -entry in  $PA$  and  $d'_t = \gcd\{b_{1t}, b_{2t}, \dots, b_{st}\}$ . By (1.7) transposed there are *eros* leaving row  $i$  of  $PA$  unchanged for  $i < t$  and creating a matrix  $P'PA$  with only one non-zero entry, namely the $(t, t)$ -entry  $d'_t$ , in row  $i$  for  $t \leq i \leq s$  where  $P'$  is an invertible  $s \times s$  matrix over  $\mathbb{Z}$ . As  $P'PA$  has only one non-zero  $t$ -minor, namely  $d_1 d_2 \cdots d_{t-1} d'_t$  we see

$$d_1 d_2 \cdots d_{t-1} d'_t = g_t(P'PA) = g_t(A) = d_1 d_2 \cdots d_{t-1} d_t. \text{ So } d'_t = d_t. \text{ As } g_1(A_{\{t\}}) = d_t, \text{ the } \textit{eros}$$

 $r_i - (b_{it}/d_t)r_t$  for  $1 \leq i < t$  reduce all other entries in column  $t$  to zero without changing  $P'PA_Y$ . Sothere is an invertible  $s \times s$  matrix  $P''$  over  $\mathbb{Z}$  with  $P''P'PA = \text{diag}(d_1, d_2, \dots, d_t) = S(A)$  which completes the induction. So  $A$  is reducible to  $S(A)$  using *eros* only on applying (1.14) to  $P''P'P$ .(b) (i) No, as the column gcds are 1, 2, 4 but  $g_3(A) \neq \pm 1 \times 2 \times 4$  as  $g_3(A) = |\det A| = 48$ .(ii) Yes, as  $r_2 - 2r_1, r_3 - r_1, r_1 - 2r_2, r_1 - 2r_3$  produce  $S(A) = \text{diag}(1, 2, 4)$ .(c) By (1.20)  $d_1 d_2 \cdots d_s = g_s(A) = 1$ . So each  $d_i = 1$  and  $S(A) = (I_s \mid 0)$  where  $I_s$  is the  $s \times s$  identity matrix. There is an invertible  $s \times s$  matrix  $P_1$  over  $\mathbb{Z}$  and an invertible  $t \times t$  matrix  $Q_1$  over  $\mathbb{Z}$ with  $P_1 A Q_1^{-1} = S(A) = (I_s \mid 0)$ . So  $A Q_1^{-1} = P_1^{-1} (I_s \mid 0) = (P_1^{-1} \mid 0) = (I_s \mid 0) \begin{pmatrix} P_1^{-1} & 0 \\ 0 & I_{t-s} \end{pmatrix}$  where  $I_{t-s}$

is the  $(t-s) \times (t-s)$  identity matrix. So  $Q = \left( \begin{array}{c|c} P_1^{-1} & 0 \\ \hline 0 & I_{t-s} \end{array} \right) Q_1$  is invertible over  $\mathbb{Z}$  and satisfies

$A = (I_s \mid 0)Q$ . So  $A$  can be reduced to  $S(A) = (I_s \mid 0)$  using *ecos* only by (1.14). Also  $A$  is the submatrix of  $Q$  consisting of its first  $s$  rows.

(i) Reducing  $(6, 10, 15)$  to its Smith normal form  $(1, 0, 0)$  produces  $Q = \begin{pmatrix} 6 & 10 & 15 \\ 1 & 2 & 0 \\ 3 & 5 & 7 \end{pmatrix}$ .

(ii) Reducing  $\begin{pmatrix} 10 & 9 & 10 \\ 15 & 15 & 16 \end{pmatrix}$  to  $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$  produces  $Q = \begin{pmatrix} 10 & 9 & 10 \\ 15 & 15 & 16 \\ 1 & 1 & 1 \end{pmatrix}$ .

### Solution 6

(a) As  $\det A$  and  $\det B$  are non-zero we see  $\det AB$  is also non-zero. As  $\det A = \pm \prod_{k=1}^t d_k(A)$  all the invariant factors of  $A$  (and for the same reason those of  $B$  and  $AB$ ) are positive since none are zero. By (1.21) with  $r=s=t$  we see  $d_k(A)$  and  $d_k(B)$  are coprime divisors of  $d_k(AB)$ , since  $d_k(A)$  and  $d_k(B)$  are divisors of  $\det A$  and  $\det B$  respectively. Hence  $d_k(A)d_k(B)$  is a divisor of  $d_k(AB)$  for  $1 \leq k \leq t$ . Write  $d_k(AB) = \mu_k d_k(A)d_k(B)$ . From  $\det(AB) = \det A \det B$  we deduce, on taking

moduli,  $\prod_{k=1}^t d_k(AB) = \prod_{k=1}^t d_k(A)d_k(B)$ . So  $\prod_{k=1}^t \mu_k = 1$  and hence each  $\mu_k = 1$  as  $\mu_k$  is a positive integer. So  $d_k(AB) = d_k(A)d_k(B)$  for  $1 \leq k \leq t$  which gives  $S(AB) = S(A)S(B)$ .

(b)  $\det A = \pm 8$ ,  $\det B = \pm 3$ . By (a) above,  $S(AB) = S(A)S(B) = \text{diag}(2, 12)$ .

Secondly write  $S(AB) = \text{diag}(d_1, d_2)$ . Then  $2 \mid d_1$ ,  $6 \mid d_2$ ,  $4 \mid d_2$  by (1.21) and  $d_1 d_2 = 2 \times 6 \times 1 \times 4$ . So either  $S(AB) = \text{diag}(2, 24)$  or  $S(AB) = \text{diag}(4, 12)$ .

(c) Let  $d_k(A) = p_1^{n_{k1}} \cdots p_l^{n_{kl}}$  where the exponents are non-negative. Let

$$A_1 = P^{-1} \text{diag}(p_1^{n_{11}}, \dots, p_l^{n_{l1}}), \quad A_j = \text{diag}(p_j^{n_{1j}}, \dots, p_j^{n_{lj}}) \text{ for } 1 < j < l, \quad A_l = \text{diag}(p_l^{n_{1l}}, \dots, p_l^{n_{ll}})Q.$$

Then  $A = A_1 A_2 \cdots A_l$ ,  $d_k(A_j) = p_j^{n_{kj}}$  and  $S(A_j) = \text{diag}(p_j^{n_{1j}}, \dots, p_j^{n_{lj}})$  for  $1 \leq j \leq l$ . By part (a) above  $A_j$  must have this last property as  $S(A) = S(A_1)S(A_2) \cdots S(A_l)$  and so all choices for  $A_j$  are equivalent (i.e.  $A_j$  is unique up to equivalence).

(d)  $A = P^{-1} \text{diag}(2, 28)Q$  where  $P^{-1} = \begin{pmatrix} 4 & -1 \\ 5 & -1 \end{pmatrix}$  and  $Q = \begin{pmatrix} 4 & 1 \\ 1 & 0 \end{pmatrix}$ . So take

$$A_1 = P^{-1} \text{diag}(2, 4) = \begin{pmatrix} 8 & -4 \\ 10 & -4 \end{pmatrix} \text{ and } A_2 = \text{diag}(1, 7)Q = \begin{pmatrix} 4 & 1 \\ 7 & 0 \end{pmatrix}. \text{ Then } A = A_1 A_2. \text{ As}$$

$S(A_1 A_2) = S(A_1)S(A_2) = S(A_2)S(A_1) = S(A_2 A_1)$  we see that  $A_1 A_2$  and  $A_2 A_1$  are necessarily equivalent.

## Solutions 2.1 (page 58)

### Solution 1

(a) The addition table of the additive group  $\mathbb{Z}_5$  is:

+	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$
$\bar{0}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$
$\bar{1}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$	$\bar{0}$
$\bar{2}$	$\bar{2}$	$\bar{3}$	$\bar{4}$	$\bar{0}$	$\bar{1}$
$\bar{3}$	$\bar{3}$	$\bar{4}$	$\bar{0}$	$\bar{1}$	$\bar{2}$
$\bar{4}$	$\bar{4}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$

As  $0(\bar{4}) = \bar{0}$ ,  $1(\bar{4}) = \bar{4}$ ,  $2(\bar{4}) = \bar{3}$ ,  $3(\bar{4}) = \bar{2}$ ,  $4(\bar{4}) = \bar{1}$  every element of  $\mathbb{Z}_5$  is an integer multiple of  $\bar{4}$ , and so  $\bar{4}$  generates  $\mathbb{Z}_5$ . In the same way each of  $\bar{1}, \bar{2}, \bar{3}, \bar{4}$  generates  $\mathbb{Z}_5$ . The two subgroups of  $\mathbb{Z}_5$  are  $\{\bar{0}\}$  and  $\mathbb{Z}_5$  itself.

(b) The addition table of the  $\mathbb{Z}$ -module  $\mathbb{Z}_6$  is:

+	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$	$\bar{5}$
$\bar{0}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$	$\bar{5}$
$\bar{1}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$	$\bar{5}$	$\bar{0}$
$\bar{2}$	$\bar{2}$	$\bar{3}$	$\bar{4}$	$\bar{5}$	$\bar{0}$	$\bar{1}$
$\bar{3}$	$\bar{3}$	$\bar{4}$	$\bar{5}$	$\bar{0}$	$\bar{1}$	$\bar{2}$
$\bar{4}$	$\bar{4}$	$\bar{5}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$
$\bar{5}$	$\bar{5}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$

$27(\bar{2}) = \bar{0}$ ,  $-17(\bar{4}) = 17(-\bar{4}) = 17(\bar{2}) = \bar{4}$ ,  $15(\bar{3}) + 13(\bar{4}) = \bar{1}$ . By (2.2) the submodules  $H$  of  $\mathbb{Z}_6$  are  $\{\bar{0}\}$ ,  $\{\bar{0}, \bar{3}\}$ ,  $\{\bar{0}, \bar{2}, \bar{4}\}$ ,  $\mathbb{Z}_6$ . The corresponding submodules  $K$  of  $\mathbb{Z}$  are  $\langle 6 \rangle$ ,  $\langle 3 \rangle$ ,  $\langle 2 \rangle$ ,  $\langle 1 \rangle$ . Generators of the submodules  $H$  are  $\bar{0}, \bar{3}, \bar{2}, \bar{1}$  respectively.  $\bar{1}$  generates  $\mathbb{Z}_6$  and  $\bar{5}$  generates  $\mathbb{Z}_6$ .

(c)(i)  $\bar{14}, \bar{7}, \bar{0}$  are the integer multiples of  $\bar{14}$  in  $\mathbb{Z}_{21}$ . (ii)  $\bar{15}, \bar{9}, \bar{3}, \bar{18}, \bar{12}, \bar{6}, \bar{0}$  are the integer multiples of  $\bar{15}$  in  $\mathbb{Z}_{21}$ . The orders of  $\bar{14}, \bar{15}$  are 3, 7 respectively. Each of the 12 remaining elements generates  $\mathbb{Z}_{21}$  as the only submodules of  $\mathbb{Z}_{21}$  are  $\langle \bar{0} \rangle = \langle \bar{21} \rangle$ ,  $\langle \bar{14} \rangle = \langle \bar{7} \rangle$ ,  $\langle \bar{15} \rangle = \langle \bar{3} \rangle$  and  $\mathbb{Z}_{21} = \langle \bar{1} \rangle$  by (2.2).

### Solution 2

(a)  $\gcd\{91, 289\} = 1$ . As  $\bar{1}$  has order 289 in the  $\mathbb{Z}$ -module  $\mathbb{Z}_{289}$ , by (2.7) the order of  $\overline{91} = 91(\bar{1})$  is  $289/\gcd\{91, 289\} = 289$ . So  $\overline{91}$  generates  $\mathbb{Z}_{289}$ . As  $\gcd\{51, 289\} = 17$ , the order of  $\bar{51}$  is  $289/17 = 17$  and so  $\bar{51}$  does not generate the  $\mathbb{Z}$ -module  $\mathbb{Z}_{289}$ .

(b) As  $\bar{1}$  has order  $n$  in the  $\mathbb{Z}$ -module  $\mathbb{Z}_n$ , by (2.7) the order of  $\bar{m} = m(\bar{1})$  in  $\mathbb{Z}_n$  is  $n/\gcd\{m, n\}$ . Hence  $\bar{m}$  generates  $\mathbb{Z}_n \Leftrightarrow \bar{m}$  has order  $n$  in  $\mathbb{Z}_n \Leftrightarrow \gcd\{m, n\} = 1$ .

(c)  $\bar{5}, \bar{10}, \bar{15}, \bar{20}, \bar{25}$  are the non-generators of  $\mathbb{Z}_{25}$ . Yes, they are precisely the elements of the submodule  $\langle \bar{5} \rangle$ .  $\mathbb{Z}_{125}$  contains 25 elements  $\bar{r}$  with  $\gcd\{r, 125\} \neq 1$ , i.e.  $5 \mid r$ . By (b) above each of the remaining  $125 - 25 = 100$  elements are generators of  $\mathbb{Z}_{125}$ .

(d) (i)  $p-1$ , (ii)  $p^2-p$ , (iii)  $p^3-p^2$ , (iv)  $p^l-p^{l-1}$ .

### Solution 3

(a)  $(\bar{2})^4 = \bar{2}^4 = \bar{16} = \bar{3}$  as  $16 \equiv 3 \pmod{13}$ .  $(\bar{2})^6 = (\bar{2})^2 \times (\bar{2})^4 = \bar{4} \times \bar{3} = \bar{12} = \bar{-1}$  as  $12 \equiv -1 \pmod{13}$ . So  $(\bar{2})^{12} = (\bar{2})^6 \times (\bar{2})^6 = (\bar{-1})^2 = \bar{1}$  showing that the order  $n$  of  $\bar{2}$  satisfies  $n \mid 12$ . But  $n$  is not a divisor of either  $4 = 12/3$  or  $6 = 12/2$  since  $(\bar{2})^4 \neq \bar{1}$ ,  $(\bar{2})^6 \neq \bar{1}$ . Hence  $n = 12$  as 2 and 3 are the only prime divisors of 12. The integer powers of  $\bar{2}$  are

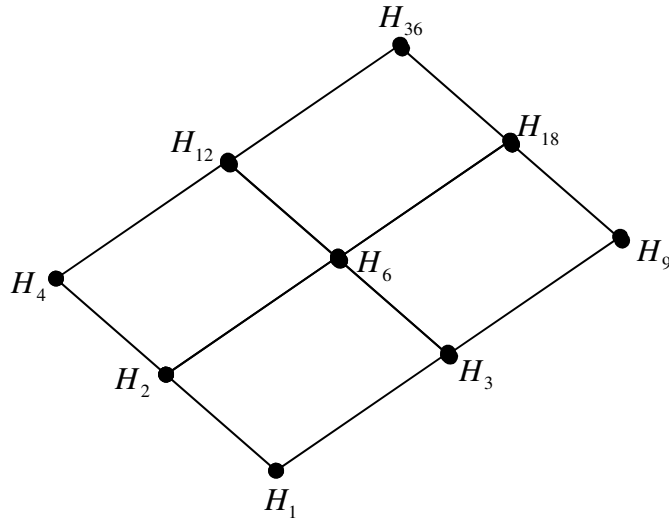
$$(\bar{2})^1 = \bar{2}, (\bar{2})^2 = \bar{4}, (\bar{2})^3 = \bar{8}, (\bar{2})^4 = \bar{3}, (\bar{2})^5 = \bar{6}, (\bar{2})^6 = \bar{12},$$

$$(\bar{2})^7 = \bar{11}, (\bar{2})^8 = \bar{9}, (\bar{2})^9 = \bar{5}, (\bar{2})^{10} = \bar{10}, (\bar{2})^{11} = \bar{7}, (\bar{2})^{12} = \bar{1}.$$

These account for all the elements of  $\mathbb{Z}_{13}^*$  and so  $\bar{2}$  generates  $\mathbb{Z}_{13}^*$ . Using (2.7) in multiplicative notation,  $(\bar{2})^l$  has order  $12/\gcd\{l, 12\}$ . So  $(\bar{2})^l$  generates  $\mathbb{Z}_{13}^* \Leftrightarrow (\bar{2})^l$  has order 12  $\Leftrightarrow 12/\gcd\{l, 12\} = 12 \Leftrightarrow \gcd\{l, 12\} = 1$ . There are four integers  $l$  with this property namely 1, 5, 7, 11. Hence the elements  $\bar{r}$  which generate  $\mathbb{Z}_{13}^*$  are  $\bar{2}, (\bar{2})^5 = \bar{6}, (\bar{2})^7 = \bar{11}$  and  $(\bar{2})^{11} = \bar{7}$ .

(b)  $\bar{2}$  does not generate  $\mathbb{Z}_{17}^*$  as  $(\bar{2})^4 = \bar{1}$ . However  $\bar{3}$  satisfies  $(\bar{3})^4 = \bar{-4}$  as  $81 \equiv -4 \pmod{17}$  and so  $(\bar{3})^8 = (\bar{3})^4 \times (\bar{3})^4 = (\bar{-4})^2 = \bar{16} = \bar{-1}$ . Hence  $\bar{3}$  has order 16 as  $(\bar{3})^{16} = (\bar{3})^8 \times (\bar{3})^8 = (\bar{-1})^2 = \bar{1}$  but  $(\bar{3})^8 \neq \bar{1}$ . So  $\bar{3} = g$  generates  $\mathbb{Z}_{17}^*$ . The 5 subgroups of  $\mathbb{Z}_{17}^*$  are  $\{\bar{1}\}, \{\bar{-1}, \bar{1}\}, \{\bar{-4}, \bar{-1}, \bar{1}, \bar{4}\}, \{\bar{-8}, \bar{-4}, \bar{-2}, \bar{-1}, \bar{1}, \bar{2}, \bar{4}, \bar{8}\}$  and  $\mathbb{Z}_{17}^*$  itself, being generated by  $(\bar{3})^{16} = \bar{1}, (\bar{3})^8 = \bar{-1}, (\bar{3})^4 = \bar{-4}, (\bar{3})^2 = \bar{-8}$  and  $\bar{3}$  respectively.  $\mathbb{Z}_{17}^*$  has 8 generators namely  $\bar{-7}, \bar{-6}, \bar{-5}, \bar{-3}, \bar{3}, \bar{5}, \bar{6}, \bar{7}$ , the 8 elements of  $\mathbb{Z}_{17}^*$  not in the subgroup of order 8.

(c)  $2^8 \equiv -3 \pmod{37}$  as  $2^8 + 3 = 259 = 37 \times 7$ . So  $2^{12} = 2^8 \times 2^4 \equiv (-3) \times 16 \equiv -11 \pmod{37}$  and so  $(\bar{2})^{12} = (\bar{2})^{36/3} \neq \bar{1}$  in  $\mathbb{Z}_{37}^*$ . Also  $2^{18} = 2^8 \times 2^8 \times 2^2 \equiv (-3)^2 \times 4 \equiv -1 \pmod{37}$  and so  $(\bar{2})^{18} = (\bar{2})^{36/2} = \bar{-1} \neq \bar{1}$ . But  $(\bar{2})^{36} = (\bar{2})^{18} \times (\bar{2})^{18} = (\bar{-1})^2 = \bar{1}$ . Hence  $\bar{2}$  in  $\mathbb{Z}_{37}^*$  has multiplicative order 36 and so  $\bar{2}$  generates  $\mathbb{Z}_{37}^*$ .



Let  $H_d$  denote the subgroup of  $\mathbb{Z}_{37}^*$  generated by  $(\bar{2})^{36/d}$  for each of the 9 positive divisors  $d$  of 36. Then  $|H_d| = d$  and the lattice of subgroups of  $\mathbb{Z}_{37}^*$  has diagram as shown. The generators of  $\mathbb{Z}_{37}^*$  are the  $36 - (18 + 12 - 6) = 12$  elements in  $\mathbb{Z}_{37}^*$  but not in either  $H_{18}$  or  $H_{12}$ .

(d)  $2^5 \equiv -9 \pmod{41}$  and so  $2^{10} \equiv (-9)^2 = 81 \equiv -1 \pmod{41}$ . Hence  $2^{20} \equiv 1 \pmod{41}$  and so  $\bar{2}$  has order 20. The congruences  $3^4 \equiv -1 \pmod{41}$  and  $3^8 = 3^4 \times 3^4 \equiv 1 \pmod{41}$  show that  $\bar{3}$  has order 8. By (2.7) the element  $\bar{4} = (\bar{2})^2$  has order  $20/\gcd\{2, 20\} = 10$ . Also  $5^2 \equiv -2^4 \pmod{41}$  and so  $5^4 \equiv 2^8 \not\equiv 1 \pmod{41}$ ,  $5^{10} \equiv -2^{20} \equiv -1 \pmod{41}$ . Hence  $\bar{5}$  has order 20. As  $\bar{6} = \bar{2} \times \bar{3}$  we obtain  $(\bar{6})^{40} = (\bar{2})^{40} \times (\bar{3})^{40} = \bar{1} \times \bar{1} = \bar{1}$ . Also  $(\bar{6})^{40/2} = (\bar{2})^{20} \times (\bar{3})^{20} = \bar{1} \times (\bar{3})^4 = -\bar{1} \neq \bar{1}$  and  $(\bar{6})^{40/5} = (\bar{2})^8 \times (\bar{3})^8 = \bar{10} \times \bar{1} = \bar{10} \neq \bar{1}$ . So  $\bar{6}$  has order 40 and so  $\bar{6}$  generates  $\mathbb{Z}_{41}^*$  as  $|\mathbb{Z}_{41}^*| = 40$ .

#### Solution 4

(a) Let  $g_1, g_2 \in G$ . Then  $(g_1 + g_2)\theta = c(g_1 + g_2) = cg_1 + cg_2 = (g_1)\theta + (g_2)\theta$ . For  $g \in G$ ,  $m \in \mathbb{Z}$ ,  $(mg)\theta = c(mg) = (cm)g = (mc)g = m(cg) = m((g)\theta)$ . So  $\theta$  is  $\mathbb{Z}$ -linear. As  $(g_0)\theta \in G = \langle g_0 \rangle$  there is an integer  $c$  with  $(g_0)\theta = cg_0$ . For  $g \in G$  there is  $m \in \mathbb{Z}$  with  $g = mg_0$ . Hence  $(g)\theta = (mg_0)\theta = m((g_0)\theta) = m(cg_0) = c(mg_0) = cg$ . So there is an integer  $c$  as stated. Let  $c' \in \mathbb{Z}$  satisfy  $(g)\theta = c'g$  for all  $g \in G$ . Then

$$(c - c')g_0 = cg_0 - c'g_0 = (g_0)\theta - (g_0)\theta = 0$$

showing that  $c - c' \in \langle n \rangle$ . Hence  $n \mid (c - c')$ , i.e.  $c \equiv c' \pmod{n}$ , i.e.  $c$  is unique modulo  $n$ . In particular  $c$  is unique for  $n = 0$  and  $c$  is arbitrary for  $n = 1$ . Suppose that  $\theta$  is an automorphism of  $G$ . As  $\theta$  is surjective there is  $a \in \mathbb{Z}$  with  $(ag_0)\theta = g_0$ , i.e.  $cag_0 = g_0$ , i.e.  $(ca - 1)g_0 = 0$ , i.e.  $ca - 1 \in \langle n \rangle$ , i.e.  $ca - 1 = bn$  for some  $b \in \mathbb{Z}$ . Hence  $ca - bn = 1$  showing  $\gcd\{c, n\} = 1$ .

Conversely suppose  $\gcd\{c, n\} = 1$ . There are integers  $a, b$  with  $ca - bn = 1$ . Reversing the above steps gives  $(ag_0)\theta = g_0$  and hence  $(mag_0)\theta = mg_0$  for all  $m \in \mathbb{Z}$ , showing  $\theta$  to be surjective. Suppose  $(mg_0)\theta = (m'g_0)\theta$  for some  $m, m' \in \mathbb{Z}$ . Then  $cmg_0 = cm'g_0$  and so  $cm - cm' \in \langle n \rangle$ , i.e.  $n \mid c(m - m')$ . Hence  $n \mid m - m'$  as  $\gcd\{c, n\} = 1$ . So  $m - m' \in \langle n \rangle$ . As  $\langle n \rangle$  is the order ideal of  $g_0$  we conclude  $(m - m')g_0 = 0$ , i.e.  $mg_0 = m'g_0$  showing that  $\theta$  is injective. So  $\theta$  is an automorphism being bijective.

The additive group  $\mathbb{Z}$  is generated by the integer 1 with order ideal  $\langle 0 \rangle$ ; so  $n = 0$  and  $\gcd\{c, 0\} = 1 \Leftrightarrow c = \pm 1$ . So  $\mathbb{Z}$  has exactly two automorphisms namely  $m \rightarrow m$  and  $m \rightarrow -m$  for all  $m \in \mathbb{Z}$ .

For  $n > 0$  the  $\mathbb{Z}$ -module  $\mathbb{Z}_n$  is cyclic being generated by  $\bar{1}$  with order ideal  $\langle n \rangle$ . By the first part every  $\mathbb{Z}$ -linear mapping  $\theta: \mathbb{Z}_n \rightarrow \mathbb{Z}_n$  is of the form  $(\bar{m})\theta = c\bar{m}$  for some integer  $c$  and all  $\bar{m} \in \mathbb{Z}_n$ . As  $c$  is unique modulo  $n$  we may write  $c\bar{m} = \bar{c}\bar{m}$  unambiguously. It follows directly from the first part with  $G = \mathbb{Z}_n$ ,  $g_0 = \bar{1}$ , that  $\theta$  is an automorphism of  $\mathbb{Z}_n \Leftrightarrow \gcd\{c, n\} = 1$ . So the additive group  $\mathbb{Z}_9$  has 6 automorphisms corresponding to the 6 invertible elements  $\bar{c}$  of  $\mathbb{Z}_9$  namely  $\bar{1}, \bar{2}, \bar{4}, \bar{5}, \bar{7}, \bar{8}$ , i.e. the elements  $\bar{c}$  with  $\gcd\{c, 9\} = 1$ . Yes, all these automorphisms are powers of  $\theta_2$  since  $(\bar{2})^1 = \bar{2}$ ,  $(\bar{2})^2 = \bar{4}$ ,  $(\bar{2})^3 = \bar{8}$ ,  $(\bar{2})^4 = \bar{16} = \bar{7}$ ,  $(\bar{2})^5 = \bar{32} = \bar{5}$ ,  $(\bar{2})^6 = \bar{64} = \bar{1}$ , i.e.  $\bar{2}$  generates the multiplicative group of invertible elements of  $\mathbb{Z}_9$ . So  $(\bar{m})\theta_2^3 = \bar{8}\bar{m}$ ,  $(\bar{m})\theta_2^4 = \bar{7}\bar{m}$  etc.

(b) As  $ng_0 = 0$ , applying  $\varphi$  gives  $n((g_0)\varphi) = (ng_0)\varphi = (0)\varphi = 0'$  showing  $n \in \langle n' \rangle$ , i.e.  $n' \mid n$ . Suppose first that  $\theta: G \rightarrow G'$  is  $\mathbb{Z}$ -linear and  $(g_0)\theta = g'_0$ . Then  $(mg_0)\theta = m((g_0)\theta) = mg'_0$  for all  $m \in \mathbb{Z}$  and so there is at most one such  $\theta$ . Consider  $\theta: G \rightarrow G'$  given by  $(mg_0)\theta = mg'_0$  for all

$m \in \mathbb{Z}$ . Let  $m_1 g_0 = m_2 g_0$ . Then  $n \mid (m_1 - m_2)$  as  $m_1 - m_2 \in \langle n \rangle$  since  $(m_1 - m_2)g_0 = 0$ . As  $d \mid n$  we deduce  $d \mid (m_1 - m_2)$ . So  $(m_1 - m_2)g'_0 = 0$  as  $\langle d \rangle$  is the order ideal of  $g'_0$ . So  $m_1 g'_0 = m_2 g'_0$  showing that  $\theta$  is unambiguously defined. Also  $\theta$  is additive as

$$(mg_0 + m'g_0)\theta = ((m+m')g_0)\theta = (m+m')g'_0 = mg'_0 + m'g'_0 = (mg_0)\theta + (m'g_0)\theta \text{ for } m, m' \in \mathbb{Z}.$$

As  $(m(m'g_0))\theta = ((mm')g_0)\theta = (mm')g'_0 = m(m'g'_0) = m((m'g_0)\theta)$  we see  $\theta$  is  $\mathbb{Z}$ -linear.

(c) With  $g_1 = g_2 = 0$  in (2.3) we obtain  $(0+0)\theta = (0)\theta + (0)\theta$ , i.e.  $(0)\theta = (0)\theta + (0)\theta$  as

$0+0=0$ . Add  $-(0)\theta$ , the negative in  $G'$  of  $(0)\theta$ , to both sides obtaining

$$0' = -(0)\theta + (0)\theta = -(0)\theta + (0)\theta + (0)\theta = 0' + (0)\theta = (0)\theta. \text{ Apply } \theta \text{ to } -g + g = 0 \text{ and use (2.3)}$$

to obtain  $(-g)\theta + (g)\theta = (-g + g)\theta = (0)\theta = 0'$  which means  $-(g)\theta = (-g)\theta$  for all  $g \in G$ . The

integer  $m$  is in the order ideal of  $\bar{r} \Leftrightarrow m\bar{r} = \bar{0}$  in  $\mathbb{Z}_n \Leftrightarrow mr = qn$  for some

$$q \in \mathbb{Z} \Leftrightarrow m(r/\gcd\{r, n\}) = q(n/\gcd\{r, n\}) \Leftrightarrow (n/\gcd\{r, n\}) \mid m. \text{ Therefore } \langle n/\gcd\{r, n\} \rangle \text{ is the}$$

order ideal of  $\bar{r}$  in  $\mathbb{Z}_n$ . For  $\bar{r} \in \mathbb{Z}_n$  there is a unique  $\mathbb{Z}$ -linear mapping  $\theta: \mathbb{Z}_m \rightarrow \mathbb{Z}_n$  with

$$(\bar{1})\theta = \bar{r} \Leftrightarrow (n/\gcd\{r, n\}) \mid m. \text{ Hence } (n/\gcd\{m, n\}) \mid \gcd\{r, n\} \text{ and so } (n/\gcd\{m, n\}) \mid r \text{ as}$$

$\gcd\{r, n\} \mid r$ . Conversely  $(n/\gcd\{m, n\}) \mid r \Rightarrow (n/\gcd\{r, n\}) \mid m$  in the same way. So there are

$\gcd\{m, n\}$  choices for  $\bar{r} \in \mathbb{Z}_n$  namely  $r = l(n/\gcd\{m, n\})$  for  $1 \leq l \leq \gcd\{m, n\}$ .

(d)  $(g_1 + g_2)\theta\theta' = ((g_1)\theta + (g_2)\theta)\theta' = (g_1)\theta\theta' + (g_2)\theta\theta'$  for all  $g_1, g_2 \in G$  and

$$(mg)\theta\theta' = ((mg)\theta)\theta' = (m((g)\theta))\theta' = m(((g)\theta)\theta') = m((g)\theta\theta') \text{ for all } m \in \mathbb{Z}, g \in G. \text{ So } \theta\theta' \text{ is}$$

$\mathbb{Z}$ -linear. Suppose  $\theta$  bijective. Then

$$(g'_1 + g'_2)\theta^{-1}\theta = g'_1 + g'_2 = (g'_1)\theta^{-1}\theta + (g'_2)\theta^{-1}\theta = ((g'_1)\theta^{-1} + (g'_2)\theta^{-1})\theta \text{ as } \theta \text{ is } \mathbb{Z}\text{-linear. As } \theta \text{ is}$$

injective  $(g'_1 + g'_2)\theta^{-1} = (g'_1)\theta^{-1} + (g'_2)\theta^{-1}$  for all  $g'_1, g'_2 \in G'$ . Also

$$((mg')\theta^{-1})\theta = (mg')\theta^{-1}\theta = mg' = m((g')\theta^{-1}\theta) = (m((g')\theta^{-1}))\theta \text{ and as } \theta \text{ is injective}$$

$$(mg')\theta^{-1} = m((g')\theta^{-1}) \text{ for all } m \in \mathbb{Z}, g' \in G'. \text{ So } \theta^{-1} \text{ is } \mathbb{Z}\text{-linear. Let } \theta, \varphi, \psi \text{ be}$$

automorphisms of  $G$ . Then  $\theta\varphi \in \text{Aut } G$  by the above theory with  $G' = G'' = G$  and  $\theta' = \varphi$ . Also

$$(\theta\varphi)\psi = \theta(\varphi\psi) \text{ as composition of mappings is associative. The identity } \iota: G \rightarrow G \text{ is in } \text{Aut } G$$

and  $\iota\theta = \theta = \theta\iota$  for all  $\theta$  in  $\text{Aut } G$ . For each  $\theta \in \text{Aut } G$  we see  $\theta^{-1} \in \text{Aut } G$  and  $\theta^{-1}\theta = \iota = \theta\theta^{-1}$ .

Hence  $\text{Aut } G$  is a group.

From (a) above  $\text{Aut } \mathbb{Z}_9$  is cyclic of order 6 with generator  $\theta_2$ . However  $\text{Aut } \mathbb{Z}_8 = \{\theta_1, \theta_3, \theta_5, \theta_7\}$

is not cyclic as  $\theta_3^2 = \theta_5^2 = \theta_7^2 = \theta_1$ , the identity element of  $\text{Aut } \mathbb{Z}_8$ .

### Solution 5

(a)  $H$  is closed under addition since  $2g + 2g' = 2(g + g')$  as  $G$  is closed under addition

$(g, g' \in G)$ .  $H$  contains the zero  $0$  of  $G$  as  $2 \times 0 = 0 + 0 = 0$ .  $H$  is closed under negation since

$$-2g = 2(-g) \text{ as } G \text{ is closed under negation. Let } k, k' \in K. \text{ Then } 2(k + k') = 2k + 2k' = 0 + 0 = 0.$$

So  $k + k' \in K$ .  $2 \times 0 = 0$  and so  $0 \in K$ .  $2(-k) = -2k = -0 = 0$ . So  $-k \in K$ . Therefore  $H$  and  $K$

are subgroups of  $G$  and hence are submodules of  $G$ , i.e.  $mh \in H$  and  $mk \in K$  for all

$$m \in \mathbb{Z}, h \in H, k \in K.$$

(i) Take  $G = \mathbb{Z}_2$ . Then  $H = \{0\}$ ,  $K = \mathbb{Z}_2$  and so  $H \subset K$ . (ii) Take  $G = \mathbb{Z}_4$ . Then

$$H = \{\bar{0}, \bar{2}\} = K. \text{ (iii) Take } G = \mathbb{Z}_8. \text{ Then } H = \{\bar{0}, \bar{2}, \bar{4}, \bar{6}\}, K = \{\bar{0}, \bar{4}\} \text{ and so } K \subset H.$$

(iv) Take  $G = \mathbb{Z}_6$ . Then  $H = \{\bar{0}, \bar{2}, \bar{4}\}$ ,  $K = \{\bar{0}, \bar{3}\}$  and so  $H \not\subset K$  and  $K \not\subset H$ .

(b)  $K$  is a  $\mathbb{Z}$ -module and scalar multiplication by elements  $\bar{m}$  of  $\mathbb{Z}_2$  is unambiguously defined by  $\bar{m}k = mk$  as  $2k = 0$  for all  $k \in K$ . Hence  $K$  is a vector space over  $\mathbb{Z}_2$ . For  $n$  odd,  $K = \{\bar{0}\}$  and  $\dim K = 0$ . For  $n$  even,  $K = \{\bar{0}, \overline{n/2}\}$  and so  $\dim K = 1$ .

### Solution 6

(a)  $(m_1q_1 + m_2q_2) + (m'_1q_1 + m'_2q_2) = (m_1 + m'_1)q_1 + (m_2 + m'_2)q_2$ . So  $\langle q_1, q_2 \rangle$  is closed under addition.  $0 = 0q_1 + 0q_2 \in \langle q_1, q_2 \rangle$ .  $-(m_1q_1 + m_2q_2) = (-m_1)q_1 + (-m_2)q_2$ . So  $\langle q_1, q_2 \rangle$  is closed under negation.  $m(m_1q_1 + m_2q_2) = (mm_1)q_1 + (mm_2)q_2$ . So  $\langle q_1, q_2 \rangle$  is closed under integer multiplication. Therefore  $\langle q_1, q_2 \rangle$  is a submodule of the  $\mathbb{Z}$ -module  $\mathbb{Q}$ .

(b)  $1/6 = 3/2 - 2(2/3) \in \langle 3/2, 2/3 \rangle$ . Hence  $\langle 1/6 \rangle \subseteq \langle 3/2, 2/3 \rangle$ . As  $3/2 = 9(1/6)$ ,  $2/3 = 4(1/6)$  we see  $\langle 3/2, 2/3 \rangle \subseteq \langle 1/6 \rangle$  since  $m_1(3/2) + m_2(2/3) = (9m_1 + 4m_2)(1/6)$ . So  $\langle 1/6 \rangle = \langle 3/2, 2/3 \rangle$  showing that  $\langle 3/2, 2/3 \rangle$  is cyclic with generator  $1/6$ .

(c) As  $\gcd\{a'_1, a'_2\} = 1$  and  $\gcd\{a'_1, b'_1\} = 1$  we deduce that  $\gcd\{a'_1, a'_2b'_1\} = 1$ . Similarly  $\gcd\{b'_2, a'_2b'_1\} = 1$  and so  $\gcd\{a'_1b'_2, a'_2b'_1\} = 1$ . Therefore  $sa'_1b'_2 + ta'_2b'_1 = 1$  for  $s, t \in \mathbb{Z}$ . Hence  $q_0 = \gcd\{a_1, a_2\}(sa'_1b'_2 + ta'_2b'_1)\gcd\{b_1, b_2\}/b_1b_2 = (sa_1b_2 + ta_2b_1)/b_1b_2 = sq_1 + tq_2 \in \langle q_1, q_2 \rangle$  and so  $\langle q_0 \rangle \subseteq \langle q_1, q_2 \rangle$ . As  $q_i/q_0 = (a_i/\gcd\{a_1, a_2\})(\text{lcm}\{b_1, b_2\}/b_i) \in \mathbb{Z}$  we deduce that  $q_1, q_2 \in \langle q_0 \rangle$  and so  $\langle q_1, q_2 \rangle \subseteq \langle q_0 \rangle$ . Hence  $\langle q_1, q_2 \rangle = \langle q_0 \rangle$  is cyclic with generator  $q_0$  and  $\langle q_1, q_2, q_3 \rangle = \langle q_0, q_3 \rangle$  is also cyclic.

(d) As  $3 = \gcd\{6, 75\}$  and  $7 = \gcd\{35, 56\}$  we obtain  $\text{lcm}\{35, 56\} = 35 \times 56 / 7 = 280$ . So  $3/280$  generates  $\langle 6/35, 75/56 \rangle$  by (c) above. As  $1 = \gcd\{3, 8\}$  and  $5 = \gcd\{280, 15\}$  we see that  $\text{lcm}\{280, 15\} = 840$  and so  $1/840$  generates  $\langle 3/280, 8/15 \rangle = \langle 6/35, 75/56, 8/15 \rangle$ .  $\mathbb{Z} \cap \langle 3/280 \rangle = \langle 3 \rangle$  and so  $\mathbb{Z} \not\subseteq \langle 6/35, 75/56 \rangle$ .  $\mathbb{Z} \cap \langle 1/840 \rangle = \langle 1 \rangle = \mathbb{Z}$  and so  $\mathbb{Z} \subseteq \langle 6/35, 75/56, 8/15 \rangle$ .

### Solution 7

(a) (i) Let  $h, h' \in H_1 \cap H_2$ . Then  $h, h' \in H_i$  ( $i = 1, 2$ ) and so  $h + h' \in H_i$  as  $H_i$  is closed under addition. So  $h + h' \in H_1 \cap H_2$  showing that  $H_1 \cap H_2$  is closed under addition.  $0 \in H_i$  ( $i = 1, 2$ ) and so  $0 \in H_1 \cap H_2$ .  $-h \in H_i$  ( $i = 1, 2$ ) as  $H_i$  is closed under negation and so  $-h \in H_1 \cap H_2$ .

Therefore  $H_1 \cap H_2$  is a subgroup of  $G$ .

(ii)  $(h_1 + h_2) + (h'_1 + h'_2) = (h_1 + h'_1) + (h_2 + h'_2) \in H_1 + H_2$  for all  $h_i, h'_i \in H_i$  ( $i = 1, 2$ ).

$0 = 0 + 0 \in H_1 + H_2$ .  $-(h_1 + h_2) = (-h_1) + (-h_2) \in H_1 + H_2$ . So  $H_1 + H_2$  is a subgroup of  $G$ .

(iii)  $\Rightarrow$  Suppose not. Pick  $h_1 \in H_1, h_1 \notin H_2, h_2 \in H_2, h_2 \notin H_1$ . Then  $h_1 + h_2 \in H_1 \cup H_2$  as  $H_1 \cup H_2$  is closed under addition. But  $h_1 + h_2 \in H_1$  implies  $h_2 = -h_1 + (h_1 + h_2) \in H_1$  (a contradiction) and  $h_1 + h_2 \in H_2$  implies  $h_1 = (h_1 + h_2) - h_2 \in H_2$  (a contradiction).

$\Leftarrow H_1 \cup H_2$  is either  $H_2$  or  $H_1$ , both of which are subgroups of  $G$ .

(b) Subgroups of the additive group  $\mathbb{Z}$  are cyclic, being principal ideals of the ring  $\mathbb{Z}$  by (1.15). So  $H_1 \cap H_2 = \langle 300 \rangle$ ,  $H_1 + H_2 = \langle 10 \rangle$ . More generally  $\langle m_1 \rangle \cap \langle m_2 \rangle$  and  $\langle m_1 \rangle + \langle m_2 \rangle$  are cyclic being generated by  $\text{lcm}\{m_1, m_2\} = m_1m_2/\gcd\{m_1, m_2\}$  and  $\gcd\{m_1, m_2\}$  respectively.

### Solution 8

(a) For  $n = 3$  we have  $s_3 = (g_1 + g_2) + g_3 = g_1 + (g_2 + g_3)$  by the associative law. Take  $n > 3$  and suppose inductively the result to be true for all ordered sets of less than  $n$  elements of  $G$ . Each summation of  $g_1, g_2, \dots, g_n$  in order decomposes  $h_i + h'_{n-i}$  for some  $i$  with  $1 \leq i < n$  where  $h_i$  is a summation of  $g_1, g_2, \dots, g_i$  in order and  $h'_{n-i}$  is a summation of  $g_{i+1}, g_{i+2}, \dots, g_n$  in order. By induction  $h_i = s_i$  and  $h'_{n-i} = s'_{n-i}$  where  $s'_{n-i} = (\dots((g_{i+1} + g_{i+2}) + g_{i+3}) \dots) + g_n = s'_{n-i-1} + g_n$  say. Hence  $h_i + h'_{n-i} = s_i + (s'_{n-i-1} + g_n) = (s_i + s'_{n-i-1}) + g_n$ . As  $s_i + s'_{n-i-1}$  is a summation of  $g_1, g_2, \dots, g_{n-1}$  we deduce  $s_i + s'_{n-i-1} = s_{n-1}$  by induction. Therefore  $h_i + h'_{n-i} = s_{n-1} + g_n = s_n$  which completes the induction. Each summation of  $g_1, g_2, \dots, g_n$  in order is equal to  $s_n$ . So the generalised associative law of addition holds.

(b) By the commutative law  $g_1 + g_2 = g_2 + g_1$ . Take  $n > 2$  and suppose the result is true for all sets of less than  $n$  elements of  $G$ . Each summation of  $g_1, g_2, \dots, g_n$  decomposes  $h_i + h'_{n-i}$  for some  $i$  with  $1 \leq i < n$  where  $h_i$  is a summation of  $g_j, j \in X, |X| = i$  and  $h'_{n-i}$  is a summation of  $g_j, j \in Y, |Y| = n - i, X \cap Y = \emptyset$ . Interchanging  $h_i$  and  $h'_{n-i}$  if necessary, we may assume  $n \in Y$ .

By induction  $h'_{n-i} = h'_{n-i-1} + g_n$  where  $h'_{n-i-1}$  is a summation of  $g_j$  for  $j \in Y \setminus \{n\}$  and so

$h_i + h'_{n-i-1} = s_{n-1}$  by induction. The induction is completed by

$$h_i + h'_{n-i} = h_i + (h'_{n-i-1} + g_n) = (h_i + h'_{n-i-1}) + g_n = s_{n-1} + g_n = s_n.$$

(c) For  $m \geq 0$  by (b) above  $m(g_1 + g_2) = mg_1 + mg_2$  on adding up the  $2m$  elements  $g_i, g_i, \dots, g_i$  ( $i = 1, 2$ ) in two ways. For  $m < 0$  write  $m = -n$ . Then

$$m(g_1 + g_2) = -n(g_1 + g_2) = -ng_1 + (-ng_2) = mg_1 + mg_2. \text{ If } m_1 m_2 = 0 \text{ then}$$

$(m_1 + m_2)g = m_1 g + m_2 g$ . By symmetry we may assume  $m_1 \geq m_2$ . For  $m_1 > 0, m_2 > 0$  using (a) above with  $g_i = g$ ,  $(m_1 + m_2)g = s_{m_1+m_2} = s_{m_1} + s_{m_2} = m_1 g + m_2 g$ . For

$$m_1 = -n_1 < 0, m_2 = -n_2 < 0 \text{ we have } (m_1 + m_2)g = -(n_1 + n_2)g = -n_1 g + (-n_2 g) = m_1 g + m_2 g.$$

For  $m_1 > 0, m_2 = -n_2 < 0, m_1 + m_2 > 0$ ,

$$(m_1 + m_2)g = s_{m_1+m_2} = s_{m_1} - s_{n_2} = m_1 g - n_2 g = m_1 g + m_2 g.$$

For  $m_1 > 0, m_2 = -n_2 < 0, m_1 + m_2 = -n < 0$ ,

$$(m_1 + m_2)g = -ng = -s_n = -s_{n_2-m_1} = -(s_{n_2} - s_{m_1}) = s_{m_1} - s_{n_2} = m_1 g - n_2 g = m_1 g + m_2 g. \text{ Now}$$

$$(m_1 m_2)g = 0 = m_1(m_2 g) \text{ for } m_1 m_2 = 0. \text{ For } m_1 > 0, m_2 > 0, \text{ by (a) above,}$$

$$(m_1 m_2)g = s_{m_1 m_2} = m_1(m_2 g). \text{ Hence for } m_1 = -n_1 < 0, m_2 = -n_2 < 0,$$

$$(m_1 m_2)g = ((-n_1)(-n_2))g = (n_1 n_2)g = n_1(n_2 g) = (-n_1)(-n_2 g) = m_1(m_2 g).$$

For  $m_1 > 0, m_2 = -n_2 < 0$ ,

$$(m_1 m_2)g = (-m_1 n_2)g = -((m_1 n_2)g) = -(m_1(n_2 g)) = m_1(-n_2 g) = m_1(m_2 g)$$

For  $m_1 = -n_1 < 0, m_2 > 0$ ,

$$(m_1 m_2)g = (-n_1 m_2)g = -((n_1 m_2)g) = -(n_1(m_2 g)) = (-n_1)(m_2 g) = m_1(m_2 g).$$



## Solutions 2.2 (page 72)

### Solution 1

- (a)  $K = K + \bar{0} = \{\bar{0}, \bar{4}\}$ ,  $K + \bar{1} = \{\bar{1}, \bar{5}\}$ ,  $K + \bar{2} = \{\bar{2}, \bar{6}\}$ ,  $K + \bar{3} = \{\bar{3}, \bar{7}\}$ .  $K + \bar{1}$  generates  $G/K$  as  $r(K + \bar{1}) = K + \bar{r}$  for  $0 \leq r < 4$ .  $G/K$  has isomorphism type  $C_4$ .
- (b)  $K = K + \bar{0} = \{\bar{0}, \bar{3}, \bar{6}, \bar{9}\}$ ,  $K + \bar{1} = \{\bar{1}, \bar{4}, \bar{7}, \bar{10}\}$ ,  $K + \bar{2} = \{\bar{2}, \bar{5}, \bar{8}, \bar{11}\}$ .  $K + \bar{1}$  generates  $G/K$  as  $r(K + \bar{1}) = K + \bar{r}$ ,  $0 \leq r < 3$ .  $G/K$  has isomorphism type  $C_3$ .
- (c)  $K = \{\bar{18}, \bar{12}, \bar{6}, \bar{0}\}$ . So  $|K| = 4$ .  $|G/K| = |G|/|K| = 24/4 = 6$ .  $K + \bar{1}$  generates  $G/K$  which has isomorphism type  $C_6$ .
- (d)  $|K| = n/d$  where  $d = \gcd\{m, n\}$ .  $|G/K| = |G|/|K| = n/(n/d) = d$ .  $G/K$  is cyclic being generated by  $K + \bar{1}$  and of isomorphism type  $C_d$ .

### Solution 2

- (a) The cyclic subgroup  $\langle \bar{d} \rangle$  has order  $n/d$  by (2.7), and so  $\langle \bar{d} \rangle$  has index  $n/(n/d) = d$  in  $\mathbb{Z}_n$ . Conversely let  $K$  be a subgroup of index  $d$  in  $\mathbb{Z}_n$ . So  $|K| = n/d$ . By (2.2)  $K = \langle \bar{d} \rangle$  and so the additive abelian group  $\mathbb{Z}_n$  has a unique subgroup of index  $d$ .
- (b)  $\langle d \rangle$  has index  $d$  in  $\mathbb{Z}$  as the cosets of  $\langle d \rangle$  in  $\mathbb{Z}$  are  $\langle d \rangle + r$  for  $0 \leq r < d$ . Every non-zero subgroup  $K$  of the additive group  $\mathbb{Z}$  is an ideal of the ring  $\mathbb{Z}$ , and so is of the type  $\langle d \rangle$ , where  $d$  is a positive integer, by (1.15). So  $\langle d \rangle$  is the only subgroup of  $\mathbb{Z}$  having index  $d$ . No, but  $\langle 0 \rangle$  is the only subgroup of infinite index in  $\mathbb{Z}$ .
- (c) Let  $G = \langle g_0 \rangle$ . Every coset of  $K$  in  $G$  is of the form  $K + mg_0 = m(K + g_0)$  for some  $m \in \mathbb{Z}$ . So  $G/K = \langle K + g_0 \rangle$  is cyclic.

### Solution 3

- (a)  $\mathbb{Z} + 1/3$  has order 3 as  $\mathbb{Z} + 1/3 \neq \mathbb{Z}$ ,  $2(\mathbb{Z} + 1/3) = \mathbb{Z} + 2/3 \neq \mathbb{Z}$ , but  $3(\mathbb{Z} + 1/3) = \mathbb{Z} + 1 = \mathbb{Z}$  the zero element of  $\mathbb{Q}/\mathbb{Z}$ . Similarly  $\mathbb{Z} + 5/8$  has order 8. Also  $\mathbb{Z} + m/n$ , where  $\gcd\{m, n\} = 1$ ,  $n \geq 1$ , has order  $n$ . So every element of  $\mathbb{Q}/\mathbb{Z}$  has finite order.
- (b)  $a(\mathbb{Z} + m/n) \in K$ . But  $a(\mathbb{Z} + m/n) = \mathbb{Z} + am/n = \mathbb{Z} - b + 1/n = \mathbb{Z} + 1/n$ . So  $\mathbb{Z} + 1/n \in K$ . Similarly  $\mathbb{Z} + 1/n' \in K$ . There are integers  $a', b'$  with  $a'n + b'n' = d$ . Hence  $b'(\mathbb{Z} + 1/n) + a'(\mathbb{Z} + 1/n') = \mathbb{Z} + b'/n + a'/n' = \mathbb{Z} + d/nn'$ . So  $\mathbb{Z} + d/nn' \in K$ . So  $nn'/d \leq n$  by the maximality of  $n$ . Hence  $n'/d \leq 1$ . So  $n' = d$  and  $n' | n$ . Therefore  $n'q = n$  and  $\mathbb{Z} + m'/n' = \mathbb{Z} + qm'/n = qm'(\mathbb{Z} + 1/n)$ . So  $K = \langle \mathbb{Z} + 1/n \rangle$ .
- (c)  $\langle \mathbb{Z} + 1/n \rangle$  is a subgroup of  $\mathbb{Q}/\mathbb{Z}$  having order  $n$ . Conversely let  $K$  be a subgroup of  $\mathbb{Q}/\mathbb{Z}$  with  $|K| = n$ . By (b) above  $K = \langle \mathbb{Z} + 1/n \rangle$ .
- (d) As  $(\mathbb{Z} + l/2^s) + (\mathbb{Z} + m/2^t) = \mathbb{Z} + (l2^t + m2^s)/2^{s+t} \in K$  we see that  $K$  is closed under addition.  $K$  contains  $\mathbb{Z} + 0$  (put  $l = 0$ ) the zero element of  $\mathbb{Q}/\mathbb{Z}$ , and  $K$  is closed under negation (replace  $l$  by  $-l$ ). So  $K$  is a subgroup of  $\mathbb{Q}/\mathbb{Z}$ . Each non-zero element of  $K$  is uniquely expressible  $\mathbb{Z} + l/2^s$ ,  $l$  odd,  $1 \leq l < 2^s$ ,  $s > 0$ . So  $K$  has an infinite number of elements, representatives being  $1, 1/2, 1/4, 3/4, 1/8, 3/8, 5/8, 7/8$ , etc.  $K$  is not cyclic as  $\mathbb{Q}/\mathbb{Z}$  contains no elements of infinite order. The finite subgroups of  $K$  are  $H_0 \subset H_1 \subset \dots \subset H_s \subset \dots$  where  $H_s = \langle \mathbb{Z} + 1/2^s \rangle$ . Let  $H$

be a subgroup of  $K$  and let  $S = \{s : \mathbb{Z} + 1/2^s \in H\}$ . If  $S$  is bounded above, then  $H = H_t$  where  $t = \max\{s : s \in S\}$ . If  $S$  is unbounded then  $S$  is the set of all non-negative integers and  $H = K$  is the only infinite subgroup of  $K$ .

#### Solution 4

(a) Omitting subscripts, the elements  $ng \in \mathbb{Z}_3 \oplus \mathbb{Z}_4$  for  $1 \leq n \leq 12$  are

$(\bar{1}, \bar{1}), (\bar{2}, \bar{2}), (\bar{0}, \bar{3}), (\bar{1}, \bar{0}), (\bar{2}, \bar{1}), (\bar{0}, \bar{2}), (\bar{1}, \bar{3}), (\bar{2}, \bar{0}), (\bar{0}, \bar{1}), (\bar{1}, \bar{2}), (\bar{2}, \bar{3}), (\bar{0}, \bar{0})$ , i.e. all elements of  $\mathbb{Z}_3 \oplus \mathbb{Z}_4$ . So  $g$  generates the additive abelian group  $\mathbb{Z}_3 \oplus \mathbb{Z}_4$  which is therefore cyclic of isomorphism type  $C_{12}$ .

(b) Write  $g_i = (\bar{1}_3, \bar{i}_3)$  for  $i = 0, 1, 2$  and let  $g_3 = (\bar{0}_3, \bar{1}_3)$ . The 8 non-zero elements of  $G$  are  $\pm g_0, \pm g_1, \pm g_2, \pm g_3$ . As  $g_i \neq 0, 2g_i = -g_i \neq 0, 3g_i = 0$  for  $0 \leq i \leq 3$  we see that  $g_i$  and  $-g_i$  have order 3. The 4 subgroups of order 3 are  $\langle g_i \rangle$  for  $0 \leq i \leq 3$ .  $G$  is a  $\mathbb{Z}$ -module, and as every non-zero element has order 3, we see that  $G$  is a  $\mathbb{Z}_3$ -module, i.e.  $G$  is a vector space over  $\mathbb{Z}_3$ .  $G = \langle g_0, g_3 \rangle$  has dimension 2 and is the internal direct sum of any two different 1-dimensional subspaces, i.e.

$$G = \langle g_0 \rangle \oplus \langle g_1 \rangle = \langle g_0 \rangle \oplus \langle g_2 \rangle = \langle g_0 \rangle \oplus \langle g_3 \rangle = \langle g_1 \rangle \oplus \langle g_2 \rangle = \langle g_1 \rangle \oplus \langle g_3 \rangle = \langle g_2 \rangle \oplus \langle g_3 \rangle.$$

The additive abelian group  $G$  has isomorphism type  $C_3 \oplus C_3$ .

(c) Write  $d = \gcd\{n_1, n_2\}$ . Then  $l(g_1, g_2) = (lg_1, lg_2) = ((n_2/d)n_1g_1, (n_1/d)n_2g_2) = (0_1, 0_2) = 0$ .

So  $(g_1, g_2)$  has finite order  $n$  say where  $n \mid l$ . Also  $n(g_1, g_2) = (0_1, 0_2)$  and so  $ng_1 = 0_1$  and  $ng_2 = 0_2$ . Hence  $n_1 \mid n$  and  $n_2 \mid n$ . So  $n = n_1q$  and  $(n_2/d) \mid (n_1/d)q$  on dividing  $n_2 \mid n_1q$  through by  $d$ . As  $\gcd\{n_1/d, n_2/d\} = 1$  we deduce  $(n_2/d) \mid q$ . So  $n_1(n_2/d) \mid n_1q$ , i.e.  $l \mid n$ . Hence  $l = n$ . So the order of  $(g_1, g_2)$  is  $l = \text{lcm}\{n_1, n_2\}$ .

(d) Write  $s' = \gcd\{s, m\}$ ,  $t' = \gcd\{t, n\}$ . By (2.7) the orders of  $\bar{s}_m = s(\bar{1}_m)$  and  $\bar{t}_n = t(\bar{1}_n)$  are  $m/s'$  and  $n/t'$  respectively. By (c) above  $(\bar{s}_m, \bar{t}_n)$  has order  $mn/s't'$  as  $\gcd\{m/s', n/t'\} = 1$ . But  $mn/s't' = mn \Leftrightarrow s't' = 1 \Leftrightarrow s' = t' = 1$ . In the case  $m = 7, n = 8$  there are  $6 = \phi(7)$  choices for  $\bar{s}_7$  and  $4 = \phi(8)$  choices for  $\bar{t}_8$ . Hence  $\mathbb{Z}_7 \oplus \mathbb{Z}_8$  has  $6 \times 4 = 24$  generators, i.e. there are 24 elements of order 56 in this group.

(e) As  $mn(g+h) = nmg + mnh = n0 + m0 = 0$  we see that  $g+h$  has finite order  $l$  where  $l \mid mn$ . Now  $l(g+h) = 0$  and so  $lg = -lh$ . Hence  $nlg = n(-lh) = -lnh = -l0 = 0$  showing that the order  $m$  of  $g$  is a divisor of  $nl$ , i.e.  $m \mid nl$ . As  $\gcd\{m, n\} = 1$  we deduce  $m \mid l$ . In the same way we obtain  $n \mid l$  and so  $mn \mid l$  using  $\gcd\{m, n\} = 1$  again. Therefore  $mn = l$ . Note that

$|G| = |K| \times |G/K| = mn$ . Replacing  $\varphi$  in Exercises 2.1, Question 4(b) by the natural

homomorphism  $\eta: G \rightarrow G/K$ , we see that the order  $s$  of  $h_0$  is a divisor of the order  $n$  of

$(h_0)\eta = K + h_0$ . So  $h = (s/n)h_0$  has order  $n$ . By the above  $g+h$  has order  $mn$ , as  $g$  has order  $m$  where  $K = \langle g \rangle$ . Therefore  $g+h$  generates  $G$ , i.e.  $G = \langle g+h \rangle$  is cyclic.

(f) Let  $g_1, g'_1, g''_1 \in G_1$  and  $g_2, g'_2, g''_2 \in G_2$ . Addition in  $G_1 \oplus G_2$  is associative as

$$\begin{aligned} ((g_1, g_2) + (g'_1, g'_2)) + (g''_1, g''_2) &= (g_1 + g'_1, g_2 + g'_2) + (g''_1, g''_2) = \\ ((g_1 + g'_1) + g''_1, (g_2 + g'_2) + g''_2) &= (g_1 + (g'_1 + g''_1), g_2 + (g'_2 + g''_2)) = \\ (g_1, g_2) + (g'_1 + g''_1, g'_2 + g''_2) &= (g_1, g_2) + ((g'_1, g'_2) + (g''_1, g''_2)). \end{aligned}$$

The zero element of  $G_1 \oplus G_2$  is  $(0_1, 0_2)$  since  $(0_1, 0_2) + (g_1, g_2) = (0_1 + g_1, 0_2 + g_2) = (g_1, g_2)$ .

The negative of  $(g_1, g_2)$  is  $(-g_1, -g_2)$  as  $(-g_1, -g_2) + (g_1, g_2) = (-g_1 + g_1, -g_2 + g_2) = (0_1, 0_2)$ .

Addition in  $G_1 \oplus G_2$  is commutative as

$$(g_1, g_2) + (g'_1, g'_2) = (g_1 + g'_1, g_2 + g'_2) = (g'_1 + g_1, g'_2 + g_2) = (g'_1, g'_2) + (g_1, g_2).$$

So  $G_1 \oplus G_2$  is an additive abelian group. Consider  $\alpha: G_1 \oplus G_2 \rightarrow G_2 \oplus G_1$  defined by

$$(g_1, g_2)\alpha = (g_2, g_1) \text{ for all } g_1 \in G_1, g_2 \in G_2. \text{ Then } \alpha: G_1 \oplus G_2 \cong G_2 \oplus G_1.$$

### Solution 5

(a) As  $r \equiv 7 \pmod{11}$  the possibilities for  $r$  with  $0 \leq r < 143$  are

7, 18, 29, 40, 51, 62, 73, 84, 95, 106, 117, 128, 139, whereas for  $r \equiv 6 \pmod{13}$  the list is

6, 19, 32, 45, 58, 71, 84, 97, 110, 123, 136. So  $\bar{r} = \overline{84}$ . Alternatively  $1 = 6 \times 11 - 5 \times 13$  and so

$$\bar{r} = \overline{6 \times 6 \times 11 - 7 \times 5 \times 13} = \overline{-59} = \overline{84}.$$

(b) In a field, the solutions of  $x^2 = x$  are 0 and 1. As  $\mathbb{Z}_{11}$  and  $\mathbb{Z}_{13}$  are fields, the solutions of  $x^2 = x$  in the ring  $\mathbb{Z}_{11} \oplus \mathbb{Z}_{13}$  are  $(\bar{0}, \bar{0}), (\bar{1}, \bar{0}), (\bar{0}, \bar{1}), (\bar{1}, \bar{1})$ . Using (2.11) the solutions of  $x^2 = x$  in  $\mathbb{Z}_{143}$  are  $\bar{0}, \bar{78}, \bar{66}, \bar{1}$ , as  $78 \equiv 1 \pmod{11}$ ,  $78 \equiv 0 \pmod{13}$  etc.

(c) The solutions of  $x^2 = 1$  in the ring  $\mathbb{Z}_{11} \oplus \mathbb{Z}_{13}$  are  $(\bar{1}, \bar{1}), (\bar{1}, -\bar{1}), (-\bar{1}, \bar{1}), (-\bar{1}, -\bar{1})$ . Using the ring isomorphism  $\alpha: \mathbb{Z}_{143} \cong \mathbb{Z}_{11} \oplus \mathbb{Z}_{13}$  of (2.11) we obtain the solutions  $\pm \bar{1}, \pm \bar{12}$  of  $x^2 = \bar{1}$  in  $\mathbb{Z}_{143}$ . The solutions of  $x^3 = x$  are  $-\bar{1}, \bar{0}, \bar{1}$  in  $\mathbb{Z}_{11}$  and in  $\mathbb{Z}_{13}$ . The pairs  $(\bar{0}, -\bar{1}), (-\bar{1}, \bar{0})$  in  $\mathbb{Z}_{11} \oplus \mathbb{Z}_{13}$  correspond to  $\bar{65}, \bar{77}$  in  $\mathbb{Z}_{143}$ . Using (b) above, the solutions of  $x^3 = x$  in  $\mathbb{Z}_{143}$  are  $\bar{0}, \pm \bar{1}, \pm \bar{12}, \pm \bar{65}, \pm \bar{66}$ .

(d) Using Exercises 2.1, Question 4(b) the number of  $\mathbb{Z}$ -linear  $\theta: \mathbb{Z}_3 \oplus \mathbb{Z}_5 \rightarrow \mathbb{Z}_{15}$  is 15, as for each  $\bar{r} \in \mathbb{Z}_{15}$  there is a unique  $\mathbb{Z}$ -linear  $\theta$  with  $(\bar{1}_3, \bar{1}_5)\theta = \bar{r}$  since the (additive) order of  $\bar{r}$  is a divisor of the order 15 of  $(\bar{1}_3, \bar{1}_5)$ . Of these 8  $\neq \phi(15)$  are group isomorphisms (those with  $\gcd\{r, 15\} = 1$ ) and just one ( $\bar{r} = \bar{1}$ ) is a ring isomorphism.

### Solution 6

(a) Suppose  $\mathbb{Z} = \langle m_1 \rangle + \langle m_2 \rangle + \dots + \langle m_t \rangle$ . There are integers  $a_1, a_2, \dots, a_t$  with

$1 = a_1 m_1 + a_2 m_2 + \dots + a_t m_t$ . Let  $d = \gcd\{m_1, m_2, \dots, m_t\}$ . Then  $d \mid m_i$  for all  $i$  with  $1 \leq i \leq t$ . So

$d \mid 1$  and so  $d = 1$ . Conversely suppose  $d = 1$ . There are  $t$  integers  $a_i$  as above and hence

$m = m a_1 m_1 + m a_2 m_2 + \dots + m a_t m_t$  showing  $\mathbb{Z} = \langle m_1 \rangle + \langle m_2 \rangle + \dots + \langle m_t \rangle$  as  $m a_i m_i \in \langle m_i \rangle$  for  $1 \leq i \leq t$ . As  $\gcd\{15, 36, 243\} = 3$  and  $\gcd\{15, 36, 80\} = 1$  the answers are 'No' and 'Yes'.

(b) Suppose to the contrary that the additive group  $\mathbb{Z}$  has non-trivial subgroups  $H_1$  and  $H_2$  such that  $\mathbb{Z} = H_1 \oplus H_2$ . As  $H_1$  and  $H_2$  are ideals of the ring  $\mathbb{Z}$ , by (1.15) there are positive integers  $n_1$  and  $n_2$  with  $H_1 = \langle n_1 \rangle$  and  $H_2 = \langle n_2 \rangle$ . But  $0 = 0 \times n_1 + 0 \times n_2 = n_2 \times n_1 + (-n_1) \times n_2$ , i.e. the integer zero is expressible in two different ways as a sum of integers from  $H_1$  and  $H_2$ . So  $\mathbb{Z}$  is indecomposable.

(c) Suppose  $h_1 + h_2 = 0$  where  $h_1 \in H_1, h_2 \in H_2$ . Then  $h_1 = -h_2$  showing  $h_1 \in H_2$  as  $H_2$  is closed under negation. So  $h_1 \in H_1 \cap H_2 = \{0\}$  and hence  $h_1 = 0$ . Therefore  $0 + h_2 = 0$  giving  $h_2 = 0$ . So  $H_1, H_2$  are independent submodules of  $G$ . Suppose given  $t$  submodules  $H_i$  of  $G$  as stated for

$1 \leq i \leq t$  and suppose  $h_1 + h_2 + \dots + h_{t-1} + h_t = 0$  where  $h_i \in H_i$ . Replacing  $H_1, H_2$  in the first part by  $H_1 + H_2 + \dots + H_{t-1}, H_t$  we deduce  $h_1 + h_2 + \dots + h_{t-1} = 0$  and  $h_t = 0$ . So

$h_1 = h_2 = \dots = h_{t-1} = 0$  by the independence of  $H_1, H_2, \dots, H_{t-1}$ . Hence the  $t$  submodules  $H_1, H_2, \dots, H_{t-1}, H_t$  are independent as  $h_i = 0$  for  $1 \leq i \leq t$ . Each element of  $H_1 \oplus H_2 \oplus \dots \oplus H_t$  can be expressed uniquely in the form  $h_1 + h_2 + \dots + h_t$  with  $h_i \in H_i$ . There are  $|H_i|$  choices for each  $h_i$  and so  $|H_1 \oplus H_2 \oplus \dots \oplus H_t| = |H_1| |H_2| \dots |H_t|$ .

(d) A typical element of  $G = \mathbb{Z}_3 \oplus \mathbb{Z}_9$  is  $(\bar{r}_3, \bar{s}_9)$  where  $1 \leq r \leq 3, 1 \leq s \leq 9$ . There are  $18 = 3 \times 6 = 3 \times \phi(9)$  elements  $(\bar{r}_3, \bar{s}_9)$  of order 9 as there are 3 choices for  $r$  and  $\phi(9)$  choices for  $s$  by 4(d) above. The remaining 8 non-trivial elements have order 3 as all elements of  $G$  have orders which are factors of 9. Each cyclic subgroup of order 9 contains  $\phi(9)$  elements of order 9 and so  $G$  contains  $18/\phi(9) = 18/6 = 3$  such subgroups, namely  $\langle(\bar{1}_3, \bar{1}_9)\rangle, \langle(\bar{2}_3, \bar{1}_9)\rangle, \langle(\bar{3}_3, \bar{1}_9)\rangle$ .

Similarly  $G$  has  $8/\phi(3) = 8/2 = 4$  (cyclic) subgroups of order 3 namely

$\langle(\bar{1}_3, \bar{3}_9)\rangle, \langle(\bar{1}_3, \bar{6}_9)\rangle, \langle(\bar{1}_3, \bar{9}_9)\rangle, \langle(\bar{3}_3, \bar{3}_9)\rangle$ . Each of the 3 cyclic subgroups  $H_2$  of order 9 contains just one subgroup of order 3 namely  $\langle(\bar{3}_3, \bar{3}_9)\rangle$ . So for each  $H_2$  there are 3 subgroups  $H_1$  of order 3 with  $H_1 \cap H_2 = \{0\}$ . There are  $3 \times 3 = 9$  such independent pairs  $H_1, H_2$  and for each  $G = H_1 \oplus H_2$  as  $|H_1 \oplus H_2| = |H_1| |H_2| = 3 \times 9 = 27 = |G|$ .

(e) Suppose  $k_1 + k_2 + \dots + k_t = 0$  where  $k_i \in K_i$  for  $1 \leq i \leq t$ . As  $k_i \in H_i$  for  $1 \leq i \leq t$  and  $H_1, H_2, \dots, H_t$  are independent, we see  $k_1 = k_2 = \dots = k_t = 0$ . By (2.14) the submodules  $K_1, K_2, \dots, K_t$  are also independent. So  $K = K_1 \oplus K_2 \oplus \dots \oplus K_t$ .

Suppose  $K + h = K + h'$  where  $h, h' \in H$ . There are unique elements  $h_i, h'_i \in H_i$  for  $1 \leq i \leq t$  with  $h = h_1 + h_2 + \dots + h_t$  and  $h' = h'_1 + h'_2 + \dots + h'_t$ . As  $h - h' \in K$  there are unique elements  $k_i \in K_i$  for  $1 \leq i \leq t$  with  $h - h' = k_1 + k_2 + \dots + k_t$ . But  $h - h' = (h_1 - h'_1) + (h_2 - h'_2) + \dots + (h_t - h'_t)$  which is the only way of expressing  $h - h'$  as a sum of elements, one from each  $H_i, 1 \leq i \leq t$ . As  $K_i \subseteq H_i$  we deduce  $h_i - h'_i = k_i$ , i.e.  $K_i + h_i = K_i + h'_i$  for  $1 \leq i \leq t$ . So  $\alpha$  is unambiguously defined.

Consider now any  $h, h' \in H$  and let  $h = h_1 + h_2 + \dots + h_t, h' = h'_1 + h'_2 + \dots + h'_t$  where  $h_i, h'_i \in H_i$  for  $1 \leq i \leq t$ . Then

$$\begin{aligned} ((K + h) + (K + h'))\alpha &= (K + (h + h'))\alpha = (K_1 + (h_1 + h'_1), \dots, K_t + (h_t + h'_t)) = \\ &= (K_1 + h_1, \dots, K_t + h_t) + (K_1 + h'_1, \dots, K_t + h'_t) = (K + h)\alpha + (K + h')\alpha \end{aligned}$$

showing  $\alpha$  to be additive.

Each  $t$ -tuple  $(K_1 + h_1, K_2 + h_2, \dots, K_t + h_t)$  can be written  $(K + h)\alpha$  where  $h = h_1 + h_2 + \dots + h_t$ . So  $\alpha$  is surjective. Suppose  $(K + h)\alpha = (K + h')\alpha$ . Then  $h_i - h'_i = k_i \in K_i$  for  $1 \leq i \leq t$ . Adding these  $t$  equations gives  $h - h' = (h_1 - h'_1) + (h_2 - h'_2) + \dots + (h_t - h'_t) = k_1 + k_2 + \dots + k_t \in K$  showing  $K + h = K + h'$ . So  $\alpha$  is injective. Therefore  $\alpha: H/K \cong (H_1/K_1) \oplus (H_2/K_2) \oplus \dots \oplus (H_t/K_t)$ .

### Solutions 2.3 (page 90)

#### Solution 1

(a) (i) Write  $K = \ker \theta$  and let  $k, k' \in K$ . Then  $(k + k')\theta = (k)\theta + (k')\theta = 0' + 0' = 0'$  showing that  $k + k' \in K$ , i.e.  $K$  is closed under addition. Also  $(-k)\theta = -(k)\theta = -0' = 0'$  and  $(0)\theta = 0'$  showing that  $-k \in K$  and  $0 \in K$ . As  $(mk)\theta = m((k)\theta) = m0' = 0'$  for  $m \in \mathbb{Z}$ , we conclude that  $mk \in K$  and so  $K$  is a submodule of the  $\mathbb{Z}$ -module  $G$ .

Suppose  $K = \{0\}$  and let  $g_1, g_2 \in G$  satisfy  $(g_1)\theta = (g_2)\theta$ . Then

$(g_1 - g_2)\theta = (g_1)\theta - (g_2)\theta = (g_1)\theta - (g_1)\theta = 0'$  showing  $g_1 - g_2 \in K$ . So  $g_1 - g_2 = 0$ , i.e.  $g_1 = g_2$  and  $\theta$  is injective. Conversely suppose that  $\theta$  is injective and let  $k \in K$ . Then  $(k)\theta = 0' = (0)\theta$ . So  $k = 0$  by the injectivity of  $\theta$  giving  $K = \{0\}$ .

(ii) Let  $g'_1, g'_2 \in \text{im } \theta$ . Then  $g'_1 = (g_1)\theta$  and  $g'_2 = (g_2)\theta$  for some  $g_1, g_2 \in G$ . Then

$g'_1 + g'_2 = (g_1)\theta + (g_2)\theta = (g_1 + g_2)\theta \in \text{im } \theta$  as  $g_1 + g_2 \in G$ . Also  $-g'_1 = -(g_1)\theta = (-g_1)\theta \in \text{im } \theta$  as  $-g_1 \in G$ . As  $0' = (0)\theta \in \text{im } \theta$  and  $mg'_1 = m((g_1)\theta) = (mg_1)\theta \in \text{im } \theta$  for all  $m \in \mathbb{Z}$ , we conclude that  $\text{im } \theta$  is a submodule of the  $\mathbb{Z}$ -module  $G'$ .

Yes,  $\text{im } \theta = G'$  is the same as  $\theta$  being surjective.

(b) As  $(1, 2)\theta = 4 \times 1 - 2 \times 2 = 0$  we see that  $(1, 2) \in \ker \theta$ . Yes,  $(-1, -2) = -(1, 2) \in \ker \theta$  and  $(2, 4) = 2(1, 2) \in \ker \theta$ . As  $m(1, 2) \in \ker \theta$  for all  $m \in \mathbb{Z}$  we see  $\langle (1, 2) \rangle \subseteq \ker \theta$ . Suppose  $(l, m) \in \ker \theta$ . Then  $4l - 2m = 0$  and so  $m = 2l$ . Hence  $(l, m) = (l, 2l) = l(1, 2) \in \langle (1, 2) \rangle$  and so  $\ker \theta \subseteq \langle (1, 2) \rangle$ . We conclude  $\ker \theta = \langle (1, 2) \rangle$ .

All even integers belong to  $\text{im } \theta$  as  $(0, -m)\theta = 2m$ . As  $(l, m)\theta = 2(2l - m)$  is even we see  $\text{im } \theta = \langle 2 \rangle$  which is infinite cyclic. Yes as  $(\ker \theta + (12, 20))\tilde{\theta} = (12, 20)\theta = 4 \times 12 - 2 \times 20 = 8$  and

$(\ker \theta + (17, 30))\tilde{\theta} = (17, 30)\theta = 4 \times 17 - 2 \times 30 = 8$ . By (2.16)  $\tilde{\theta}: \mathbb{Z} \oplus \mathbb{Z} / \ker \theta \cong \text{im } \theta$  and so  $\mathbb{Z} \oplus \mathbb{Z} / \ker \theta$  is infinite cyclic with generator  $\ker \theta + (0, 1)$ .

(c) (i) The additive group  $\mathbb{Z}_8 = \mathbb{Z} / \langle 8 \rangle$  has subgroups  $\langle \bar{1} \rangle, \langle \bar{2} \rangle, \langle \bar{4} \rangle, \langle \bar{8} \rangle$  and the homomorphic images of  $\mathbb{Z}_8$  are  $\mathbb{Z}_8 / \langle \bar{1} \rangle, \mathbb{Z}_8 / \langle \bar{2} \rangle, \mathbb{Z}_8 / \langle \bar{4} \rangle, \mathbb{Z}_8 / \langle \bar{8} \rangle$  which are cyclic of isomorphism types  $C_1, C_2, C_4, C_8$  respectively.

(ii) The additive group  $\mathbb{Z}_{12} = \mathbb{Z} / \langle 12 \rangle$  has subgroups  $\langle \bar{1} \rangle, \langle \bar{2} \rangle, \langle \bar{3} \rangle, \langle \bar{4} \rangle, \langle \bar{6} \rangle, \langle \bar{12} \rangle$  and the homomorphic images of  $\mathbb{Z}_{12}$  are  $\mathbb{Z}_{12} / \langle \bar{1} \rangle, \mathbb{Z}_{12} / \langle \bar{2} \rangle, \mathbb{Z}_{12} / \langle \bar{3} \rangle, \mathbb{Z}_{12} / \langle \bar{4} \rangle, \mathbb{Z}_{12} / \langle \bar{6} \rangle, \mathbb{Z}_{12} / \langle \bar{12} \rangle$  which are cyclic of isomorphism types  $C_1, C_2, C_3, C_4, C_6, C_{12}$  respectively.

(iii) The additive group  $\mathbb{Z}_n = \mathbb{Z} / \langle n \rangle$  where  $n > 0$  has subgroups  $\langle \bar{d} \rangle$  where  $d > 0$  and  $d \mid n$  by (2.2). So a typical homomorphic image of  $\mathbb{Z}_n$  is  $\mathbb{Z}_n / \langle \bar{d} \rangle$  which is cyclic of isomorphism type  $C_d$ .

(iv) The additive group  $\mathbb{Z}$  has subgroups  $\langle d \rangle$  where  $d \geq 0$  by (1.15). So  $\mathbb{Z} / \langle d \rangle$  is a typical homomorphic image of  $\mathbb{Z}$ .  $\mathbb{Z} / \langle d \rangle$  is cyclic of isomorphism type  $C_d$ .

(v) The Klein 4-group  $\mathbb{Z}_2 \oplus \mathbb{Z}_2 = \langle u, v \rangle$  has subgroups  $\{0\}, \langle u \rangle, \langle v \rangle, \langle u + v \rangle, \mathbb{Z}_2 \oplus \mathbb{Z}_2$ . The homomorphic images of  $\mathbb{Z}_2 \oplus \mathbb{Z}_2$  are

$$\mathbb{Z} \oplus \mathbb{Z} / \{0\}, \mathbb{Z} \oplus \mathbb{Z} / \langle u \rangle, \mathbb{Z} \oplus \mathbb{Z} / \langle v \rangle, \mathbb{Z} \oplus \mathbb{Z} / \langle u + v \rangle, \mathbb{Z} \oplus \mathbb{Z} / \mathbb{Z} \oplus \mathbb{Z}$$

and these are of isomorphism types  $C_2 \oplus C_2, C_2, C_2, C_2, C_1$  respectively.

(d) The identity homomorphism  $\iota: G_1 \rightarrow G_1$ , given by  $(g_1)\iota = g_1$  for all  $g_1$  in  $G_1$ , has  $\ker \iota = \{0\}$  and  $\text{im } \iota = G_1$ . So  $\tilde{\iota}: G_1/\{0\} \cong G_1$  by (2.16). The trivial homomorphism  $\mathcal{O}: G_1 \rightarrow G_1$ , given by  $(g_1)\mathcal{O} = 0$  for all  $g_1$  in  $G_1$ , has  $\ker \mathcal{O} = G_1$  and  $\text{im } \mathcal{O} = \{0\}$ . So  $G_1/G_1 \cong \{0\}$  by (2.16). The projection  $\pi_1: G_1 \oplus G_2 \rightarrow G_1$  given by  $(g_1, g_2)\pi_1 = g_1$  for all  $(g_1, g_2)$  in  $G_1 \oplus G_2$  is an homomorphism. As  $\text{im } \pi_1 = G_1$  we see directly that  $G_1$  is an homomorphic image of  $G_1 \oplus G_2$ . Similarly  $\pi_2: G_1 \oplus G_2 \rightarrow G_2$  given by  $(g_1, g_2)\pi_2 = g_2$  for all  $(g_1, g_2)$  in  $G_1 \oplus G_2$  is an homomorphism with  $\text{im } \pi_2 = G_2$  and so  $G_2$  is an homomorphic image of  $G_1 \oplus G_2$ . The mapping  $\theta: G_1 \oplus G_1 \rightarrow G_1$ , given by  $(g_1, g'_1)\theta = g_1 - g'_1$  for all  $g_1, g'_1 \in G_1$  is a homomorphism. As  $\ker \theta = K$  and  $\text{im } \theta = G_1$  we deduce  $(G_1 \oplus G_1)/K \cong G_1$  by (2.16) and so the answer is: Yes!

(e) As  $(g - (g)\theta)\theta = (g)\theta^2 - (g)\theta = (g)\theta - (g)\theta = 0$  we see  $g - (g)\theta \in \ker \theta$ . As  $(g)\theta \in \text{im } \theta$  the equation  $g = (g - (g)\theta) + (g)\theta$  shows  $G = \ker \theta + \text{im } \theta$ . To show that  $\ker \theta$  and  $\text{im } \theta$  are independent submodules of  $G$  suppose  $k + l = 0$  where  $k \in \ker \theta$ ,  $l \in \text{im } \theta$ . Then  $l = (g)\theta$  for some  $g \in G$ . Applying  $\theta$  to  $k + (g)\theta = 0$  gives  $(k + (g)\theta)\theta = (0)\theta = 0$  and so  $(k)\theta + (g)\theta^2 = 0$  which gives  $0 + (g)\theta = 0$ , i.e.  $l = 0$ . Hence  $k + 0 = 0$  and so  $k = 0$ . Therefore  $\ker \theta$  and  $\text{im } \theta$  are independent submodules of  $G$  and so  $G = \ker \theta \oplus \text{im } \theta$  by (2.15). For  $G = \mathbb{Z}_2 \oplus \mathbb{Z}_4$  we have  $(\bar{l}, \bar{m})\theta^2 = (\bar{m}, \overline{2l - m})\theta = (\overline{2l - m}, \overline{2m - (2l - m)}) = (\bar{m}, \overline{3m - 2l}) = (\bar{m}, \overline{2l - m}) \in \mathbb{Z}_2 \oplus \mathbb{Z}_4$  as  $\overline{2l} = \bar{0} \in \mathbb{Z}_2$  and  $\overline{2l} = -\overline{2l} \in \mathbb{Z}_4$ . So  $\theta$  is idempotent. By inspection  $\ker \theta = \langle (\bar{1}, \bar{2}) \rangle$  is cyclic of order 2 and  $\text{im } \theta = \langle (\bar{1}, \bar{1}) \rangle$  is cyclic of order 4.

## Solution 2

(a) (i) Consider  $h'_1, h'_2 \in H'$ . There are  $h_1, h_2 \in H$  with  $h'_1 = (h_1)\theta$ ,  $h'_2 = (h_2)\theta$ . Then  $h'_1 + h'_2 = (h_1)\theta + (h_2)\theta = (h_1 + h_2)\theta \in H'$  since  $h_1 + h_2 \in H$ . Also  $-h'_1 = -(h_1)\theta = (-h_1)\theta \in H'$  and  $0' = (0)\theta \in H'$  as  $-h_1, 0 \in H$ . So  $H'$  is a subgroup of  $G'$ . As  $\ker \theta \cap H$  is the kernel of  $\theta|_H$  and  $H'$  is the image of  $\theta|_H$ , applying (2.16) to  $\theta|_H$  gives  $H/(\ker \theta \cap H) \cong H'$ .

(ii) Let  $h_1, h_2 \in H$ . Then  $h_1 + h_2 \in H$  as  $(h_1 + h_2)\theta = (h_1)\theta + (h_2)\theta \in H'$  since  $(h_1)\theta, (h_2)\theta \in H'$ . Also  $-h_1, 0 \in H$  as  $(-h_1)\theta = -(h_1)\theta \in H'$  and  $(0)\theta = 0' \in H'$ . So  $H$  is a subgroup of  $G$ . In this case the kernel of  $\theta|_H$  is  $\ker \theta$  and the image of  $\theta|_H$  is  $H' \cap \text{im } \theta$ . Replacing  $\theta$  by  $\theta|_H$  in (2.16) now gives  $H/\ker \theta \cong H' \cap \text{im } \theta$ .

(b) As  $(1, 1)\theta = \bar{1} - \bar{1} = \bar{0}$  and  $(2, 0)\theta = \bar{2} = \bar{0}$  we see  $v_1 = (1, 1), v_2 = (2, 0) \in \ker \theta$ . Suppose  $m_1 v_1 + m_2 v_2 = 0 = (0, 0)$  for  $m_1, m_2 \in \mathbb{Z}$ . Then  $m_1 + 2m_2 = 0$  and  $m_1 = 0$ : so  $m_1 = m_2 = 0$  showing  $v_1, v_2$  to be  $\mathbb{Z}$ -independent. Let  $(l, m) \in \ker \theta$ . Then  $\overline{l - m} = \bar{0}$  in  $\mathbb{Z}_2$  and so there is  $k \in \mathbb{Z}$  with  $l - m = 2k$ . Hence  $(l, m) = mv_1 + kv_2$  showing that  $v_1, v_2$  generate  $\ker \theta$ . So  $v_1, v_2$  is a  $\mathbb{Z}$ -basis of  $\ker \theta$ . As  $\bar{0} = (0, 0)\theta$  and  $\bar{1} = (1, 0)\theta$  we see that  $\mathbb{Z}_2 = \text{im } \theta$ . By (2.15) we have  $\tilde{\theta}: \mathbb{Z} \oplus \mathbb{Z}/\ker \theta \cong \mathbb{Z}_2$  and so  $\mathbb{Z} \oplus \mathbb{Z}/\ker \theta$  has isomorphism type  $C_2$ . As  $(1, 0) \notin \ker \theta$  we see  $\ker \theta \neq \mathbb{Z} \oplus \mathbb{Z}$ . As 2 is prime, by Lagrange's theorem there are no subgroups  $H'$  with

$\{\bar{0}\} \subset H' \subset \mathbb{Z}_2$ . By (2.17) there are no subgroups  $H$  with  $\ker \theta \subset H \subset \mathbb{Z} \oplus \mathbb{Z}$ , i.e.  $\ker \theta$  is a maximal subgroup of  $\mathbb{Z} \oplus \mathbb{Z}$ .

(c) As  $(1, -1)\theta = \bar{1} + (-1) = \bar{0}$  and  $(0, 4)\theta = \bar{4} = \bar{0}$  we obtain  $v_1 = (1, -1), v_2 = (0, 4) \in \ker \theta$ . As in (b) above,  $v_1, v_2$  are  $\mathbb{Z}$ -independent. Let  $(l, m) \in \ker \theta$ . Then  $\overline{l+m} = \bar{0}$  in  $\mathbb{Z}_4$  and so there is  $k \in \mathbb{Z}$  with  $l+m = 4k$ . Hence  $(l, m) = lv_1 + kv_2$  showing  $\ker \theta = \langle v_1, v_2 \rangle$ . So  $\ker \theta$  has  $\mathbb{Z}$ -basis  $v_1, v_2$ . As  $(l, 0)\theta = \bar{l} \in \mathbb{Z}_4$  for  $l = 0, 1, 2, 3$  we see  $\text{im } \theta = \mathbb{Z}_4$ . By (2.15) we obtain  $(\mathbb{Z} \oplus \mathbb{Z})/\ker \theta \cong \mathbb{Z}_4$  and so  $(\mathbb{Z} \oplus \mathbb{Z})/\ker \theta$  has isomorphism type  $C_4$ . As  $\mathbb{Z}_4$  has exactly 3 subgroups  $H'$ , namely  $\langle \bar{4} \rangle, \langle \bar{2} \rangle, \langle \bar{1} \rangle$ , by (2.17) there are 3 corresponding subgroups of  $\mathbb{Z} \oplus \mathbb{Z}$  containing  $\ker \theta$ , namely  $\ker \theta, H, \mathbb{Z} \oplus \mathbb{Z}$ , where  $H = \{(l, m) \in \mathbb{Z} \oplus \mathbb{Z} : \overline{l+m} \in \langle \bar{2} \rangle\}$  has  $\mathbb{Z}$ -basis  $(1, -1), (0, 2)$ .

(d) As  $(l, m) \in \ker \theta \Leftrightarrow l = 2l', m = 4m' \ (l', m' \in \mathbb{Z}) \Leftrightarrow (l, m) = l'(2e_1) + m'(4e_2) \Leftrightarrow$

$(l, m) \in \langle 2e_1, 4e_2 \rangle$  we see  $\ker \theta = \langle 2e_1, 4e_2 \rangle$ , i.e.  $\ker \theta$  has  $\mathbb{Z}$ -basis  $2e_1, 4e_2$ . For  $H' = \langle (\bar{1}, \bar{0}) \rangle$  we have  $(l, m) \in H \Leftrightarrow (l, m)\theta \in \langle (\bar{1}, \bar{0}) \rangle \Leftrightarrow (\bar{l}, \bar{m}) \in \{(0, 0), (1, 0)\} \Leftrightarrow l$  arbitrary (any integer),

$m = 4m'$ . So  $\rho_1 = (1, 0), \rho_2 = (0, 4)$  is a  $\mathbb{Z}$ -basis of  $H$  and  $A = \begin{pmatrix} \rho_1 \\ \rho_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix}$  has invariant factors

$d_1 = 1, d_2 = 4$ .  $G'/H' = \langle H' + (\bar{0}, \bar{1}) \rangle$  is cyclic of order 4 and so of isomorphism type  $C_1 \oplus C_4 = C_4$ .

Similarly for  $H' = \langle (\bar{0}, \bar{2}) \rangle$  we have  $A = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$ ; so  $d_1 = d_2 = 2$  and

$G'/H' = \langle H' + (\bar{1}, \bar{0}), H' + (\bar{0}, \bar{1}) \rangle$  which has isomorphism type  $C_2 \oplus C_2$ . For  $H' = \langle (\bar{1}, \bar{1}) \rangle$  we see

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix};$$

so  $d_1 = 1, d_2 = 2$  and  $G'/H' = \langle H' + (\bar{0}, \bar{1}) \rangle$  which has isomorphism type  $C_1 \oplus C_2 = C_2$ .

(e) As  $\theta\eta$  is surjective, being the composition of two surjective mappings, we see  $\text{im } \theta\eta = G'/H'$ .

Also  $\ker \theta\eta = \{h \in G : (h)\theta\eta = H'\}$  as  $H'$  is the zero element of  $G'/H'$ . So

$\ker \theta\eta = \{h \in G : (h)\theta \in H'\} = H$ . Applying (2.16) to  $\theta\eta$  gives  $\widetilde{\theta\eta} : G/H \cong G'/H'$ .

### Solution 3

(a) There are  $k_1, k_2 \in K$  such that  $r_1 = r'_1 + k_1, r_2 = r'_1 + k_2$ . Therefore  $r_1 r_2 = (r'_1 + k_1)(r'_2 + k_2) = r'_1 r'_2 + k_3$  where  $k_3 = r'_1 k_2 + k_1 r'_2 + k_1 k_2$ . As  $K$  is an ideal of  $R$  we see that  $r'_1 k_2, k_1 r'_2, k_1 k_2 \in K$ , and so  $k_3 \in K$  as  $K$  is closed under addition. Hence  $r_1 r_2 \equiv r'_1 r'_2 \pmod{K}$  and so  $K + r_1 r_2 = K + r'_1 r'_2$  by (2.9), showing that coset multiplication is unambiguously defined. Write  $\bar{r} = K + r$  and then  $R/K$  has binary operations  $\bar{r}_1 + \bar{r}_2 = \overline{r_1 + r_2}$  and  $(\bar{r}_1)(\bar{r}_2) = \overline{r_1 r_2}$  where  $r_1, r_2 \in R$ . Now  $(R/K, +)$  is an abelian group by (2.10). Let  $r_1, r_2, r_3 \in R$ . Then  $((\bar{r}_1)(\bar{r}_2))(\bar{r}_3) = \overline{(r_1 r_2) r_3} = \overline{r_1 (r_2 r_3)} = \overline{r_1 (r_2 r_3)} = (\bar{r}_1)(\overline{r_2 r_3}) = (\bar{r}_1)((\bar{r}_2)(\bar{r}_3))$  showing that coset multiplication is associative. Coset multiplication is distributive because  $((\bar{r}_1) + (\bar{r}_2))(\bar{r}_3) = \overline{(r_1 + r_2) r_3} = \overline{r_1 r_3 + r_2 r_3} = \overline{r_1 r_3} + \overline{r_2 r_3} = (\bar{r}_1)(\bar{r}_3) + (\bar{r}_2)(\bar{r}_3)$  and similarly  $(\bar{r}_1)((\bar{r}_2) + (\bar{r}_3)) = (\bar{r}_1)(\overline{r_2 + r_3}) = \overline{r_1 (r_2 + r_3)} = \overline{r_1 r_2} + \overline{r_1 r_3} = (\bar{r}_1)(\bar{r}_2) + (\bar{r}_1)(\bar{r}_3)$ . Also  $(\bar{e})(\bar{r}) = \overline{er} = \bar{r} = \overline{re} = (\bar{r})(\bar{e})$  for all  $r \in R$ , and so  $R/K$  is a ring with 1-element  $\bar{e} = K + e$ .

By (2.10)  $\eta$  is additive. As  $(r_1 r_2)\eta = \overline{r_1 r_2} = (\overline{r_1})(\overline{r_2}) = (r_1)\eta (r_2)\eta$  for all  $r_1, r_2 \in R$  and  $(e)\eta = \bar{e}$  we see  $\eta$  is a ring homomorphism. Also  $\text{im } \eta = R/K$  and  $\ker \eta = K$ .

(b) By Question 1(a)(ii) above,  $\text{im } \theta$  is a subgroup of  $(R', +)$ . As  $(r_1)\theta (r_2)\theta = (r_1 r_2)\theta$  for all  $r_1, r_2 \in R$  we see  $\text{im } \theta$  is closed under multiplication. The 1-element  $e'$  of  $R'$  belongs to  $\text{im } \theta$  as  $e' = (e)\theta$ . So  $\text{im } \theta$  is a subring of  $R'$  and hence  $\text{im } \theta$  is itself a ring. By Question 1(a)(i) above,  $\ker \theta$  is a subgroup of  $(R, +)$ . Consider  $r \in R, k \in K$ ; then  $(rk)\theta = (r)\theta(k)\theta = (r)\theta \times 0 = 0$  and so  $rk \in K = \ker \theta$ .

Similarly  $kr \in K$  and so  $K$  is an ideal of  $R$ . Kernels of ring homomorphisms are ideals. By (2.16)

$\tilde{\theta}: R/K \cong \text{im } \theta$  is an isomorphism of additive abelian groups. Also

$((\overline{r_1})(\overline{r_2}))\tilde{\theta} = (\overline{r_1 r_2})\tilde{\theta} = (r_1 r_2)\theta = (r_1)\theta (r_2)\theta = (\overline{r_1})\tilde{\theta} (\overline{r_2})\tilde{\theta}$  for all  $r_1, r_2 \in R$ . So  $\tilde{\theta}$  is a ring isomorphism as  $(\bar{e})\tilde{\theta} = e'$ . Therefore  $\tilde{\theta}: R/K \cong \text{im } \theta$ .

(c) By (b) above  $\ker \theta$  is an ideal of the ring  $\mathbb{Z}$ . By (1.15) there is a non-negative integer  $d$  with  $\ker \theta = \langle d \rangle$ . By (b) above  $\tilde{\theta}: \mathbb{Z}/\langle d \rangle \cong \text{im } \theta$ , showing that the rings  $\mathbb{Z}_d = \mathbb{Z}/\langle d \rangle$  are, up to isomorphism, the (ring) homomorphic images of  $\mathbb{Z}$ .

(d) For  $r_1, r_2 \in R$  we see  $(r_1 + r_2)\theta\theta' = ((r_1)\theta + (r_2)\theta)\theta' = (r_1)\theta\theta' + (r_2)\theta\theta'$  and

$(r_1 r_2)\theta\theta' = ((r_1)\theta (r_2)\theta)\theta' = (r_1)\theta\theta' (r_2)\theta\theta'$  showing that  $\theta\theta'$  is additive and multiplicative. As

$(e)\theta = e'$  and  $(e')\theta' = e''$  we see  $(e)\theta\theta' = e''$  where  $e, e', e''$  are the 1-elements of  $R, R', R''$ . So

$\theta\theta': R \rightarrow R''$  is a ring homomorphism. Suppose that  $\theta$  is a ring isomorphism. Consider  $r'_1, r'_2 \in R'$  and write  $r_1 = (r'_1)\theta^{-1}, r_2 = (r'_2)\theta^{-1}$ . Then

$$((r'_1)\theta^{-1} + (r'_2)\theta^{-1})\theta = (r_1 + r_2)\theta = (r_1)\theta + (r_2)\theta = r'_1 + r'_2 = (r'_1 + r'_2)\theta^{-1}\theta$$

and so  $(r'_1)\theta^{-1} + (r'_2)\theta^{-1} = (r'_1 + r'_2)\theta^{-1}$  as  $\theta$  is injective, and  $\theta^{-1}$  is additive. Similarly

$$((r'_1)\theta^{-1} (r'_2)\theta^{-1})\theta = (r_1 r_2)\theta = (r_1)\theta (r_2)\theta = r'_1 r'_2 = (r'_1 r'_2)\theta^{-1}\theta \text{ and so } (r'_1)\theta^{-1} (r'_2)\theta^{-1} = (r'_1 r'_2)\theta^{-1}$$

as  $\theta$  is additive, and  $\theta^{-1}$  is multiplicative. As  $(e')\theta^{-1} = e$  we conclude  $\theta^{-1}: R' \rightarrow R$  is a ring

isomorphism. Take  $R = R' = R''$  and suppose  $\theta, \theta'$  are bijective, i.e. suppose  $\theta, \theta' \in \text{Aut } R$ . By the

above theory  $\theta\theta', \theta^{-1} \in \text{Aut } R$ . As the identity mapping  $\iota_R$  of  $R$  belongs to  $\text{Aut } R$  we see that  $\text{Aut } R$  is a group (it is a subgroup of the group of all bijections of  $R \rightarrow R$ ).

(e) By Exercises 2.2, Question 4(f) the direct sum  $R_1 \oplus R_2$  of the additive groups of  $R_1$  and  $R_2$  is itself an additive group. In order to verify the ring axioms involving multiplication, consider  $r_1, r'_1, r''_1 \in R_1$  and  $r_2, r'_2, r''_2 \in R_2$ . Then

$$\begin{aligned} (r_1, r_2)((r'_1, r'_2) + (r''_1, r''_2)) &= (r_1, r_2)(r'_1 + r''_1, r'_2 + r''_2) = (r_1(r'_1 + r''_1), r_2(r'_2 + r''_2)) = \\ (r_1 r'_1 + r_1 r''_1, r_2 r'_2 + r_2 r''_2) &= (r_1 r'_1, r_2 r'_2) + (r_1 r''_1, r_2 r''_2) = (r_1, r_2)((r'_1, r'_2) + (r''_1, r''_2)) \end{aligned}$$

which shows that one distributive law holds in  $R_1 \oplus R_2$ . The other distributive law holds in  $R_1 \oplus R_2$

and can be verified in the same way. The associative law of multiplication holds in  $R_1 \oplus R_2$  as

$$\begin{aligned} ((r_1, r_2)(r'_1, r'_2))(r''_1, r''_2) &= (r_1 r'_1, r_2 r'_2)(r''_1, r''_2) = ((r_1 r'_1) r''_1, (r_2 r'_2) r''_2) = \\ (r_1(r'_1 r''_1), r_2(r'_2 r''_2)) &= (r_1, r_2)(r'_1 r''_1, r'_2 r''_2) = (r_1, r_2)((r'_1, r'_2)(r''_1, r''_2)) \end{aligned}$$

using this law in the rings  $R_1$  and  $R_2$ . Let  $e_1$  and  $e_2$  denote the 1-elements of  $R_1$  and  $R_2$

respectively. Then  $(e_1, e_2)(r_1, r_2) = (e_1 r_1, e_2 r_2) = (r_1, r_2) = (r_1 e_1, r_2 e_2) = (r_1, r_2)(e_1, e_2)$  which shows that

$R_1 \oplus R_2$  has 1-element  $(e_1, e_2)$ . Therefore  $R_1 \oplus R_2$  is a ring.



(f) By Exercises 2.1, Question 7(a)(i) and (ii) both  $K \cap L$  and  $K + L$  are additive abelian groups. Consider  $r \in R$  and  $m \in K \cap L$ . Then  $m \in K$  and  $m \in L$ . As  $K$  is an ideal of  $R$  we see  $rm, mr \in K$ . As  $L$  is an ideal of  $R$  we see  $rm, mr \in L$ . So  $rm, mr \in K \cap L$  and  $K \cap L$  is an ideal of  $R$ . Consider  $r \in R, m \in K + L$ . Then  $m = k + l$  where  $k \in K$  and  $l \in L$ . So  $rm = r(k + l) = rk + rl \in K + L$  since  $rk \in K$  and  $rl \in L$  as before. Also  $mr = (k + l)r = kr + lr \in K + L$  since  $kr \in K$  and  $lr \in L$ . So  $K + L$  is an ideal of  $R$ . For  $r_1, r_2 \in R$  using addition and multiplication in the rings  $R/K$ ,  $R/L$  and  $R/K \oplus R/L$

$$\begin{aligned}(r_1 + r_2)\alpha &= (r_1 + r_2 + K, r_1 + r_2 + L) = ((r_1 + K) + (r_2 + K), (r_1 + L) + (r_2 + L)) = \\ &= (r_1 + K, r_1 + L) + (r_2 + K, r_2 + L) = (r_1)\alpha + (r_2)\alpha \text{ and} \\ (r_1 r_2)\alpha &= (r_1 r_2 + K, r_1 r_2 + L) = ((r_1 + K)(r_2 + K), (r_1 + L)(r_2 + L)) = \\ &= (r_1 + K, r_1 + L)(r_2 + K, r_2 + L) = (r_1)\alpha (r_2)\alpha.\end{aligned}$$

Let  $e$  be the 1-element of  $R$ . As  $(e)\alpha = (e + K, e + L)$  is the 1-element of  $R/K \oplus R/L$  we see that  $\alpha$  is a ring homomorphism. The 0-element of  $R/K \oplus R/L$  is  $(K, L)$ . As

$(r)\alpha = (K, L) \Leftrightarrow (r + K, r + L) \Leftrightarrow r \in K, r \in L \Leftrightarrow r \in K \cap L$  we see  $\ker \alpha = K \cap L$ . Now we use  $K + L = R$  to find  $\text{im } \alpha$ : there are elements  $k_0 \in K$  and  $l_0 \in L$  with  $k_0 + l_0 = e$ . Consider an arbitrary element  $(s + K, t + L)$  of  $R/K \oplus R/L$  and so  $s, t \in R$ . Write  $r = sl_0 + tk_0$ . Then

$r - s = r - se = s(l_0 - e) + tk_0 = s(-k_0) + tk_0 = (t - s)k_0 \in K$  and so  $r + K = s + K$ . Also  $r - t = r - te = sl_0 + t(k_0 - e) = sl_0 + t(-l_0) = (s - t)l_0 \in L$  and so  $r + L = t + L$ . Therefore  $(r)\alpha = (r + K, r + L) = (s + K, t + L)$  and  $\text{im } \alpha = R/K \oplus R/L$ . By (b) above

$\tilde{\alpha}: R/(K \cap L) \cong R/K \oplus R/L$  is a ring isomorphism where  $(r + K \cap L)\tilde{\alpha} = (r + K, r + L)$  for all  $r \in R$ .

#### Solution 4

(a) Suppose that  $K$  is normal in  $G$ . Let  $g \in G$  and consider  $kg \in Kg$  where  $k \in K$ . Then  $kg = g(g^{-1}kg) \in gK$  as  $g^{-1}kg \in K$ . So  $Kg \subseteq gK$ . Replacing  $g$  by  $g^{-1}$  in the normality condition gives  $gkg^{-1} \in K$  for all  $k \in K$ . Consider  $gk \in gK$ . Then  $gk = (gkg^{-1})g \in Kg$ . So  $gK \subseteq Kg$  and hence  $Kg = gK$ . Conversely suppose  $Kg = gK$  for all  $g \in G$ . For  $k \in K, g \in G$  we have  $kg \in Kg$  and so  $kg \in gK$ . There is  $k' \in K$  with  $kg = gk'$ . Hence  $g^{-1}kg = k' \in K$ , i.e.  $g^{-1}kg \in K$  for all  $k \in K, g \in G$ .

For  $g \in S_3$  the permutation  $g^{-1}\sigma^i g$  is even using the rules of parity, as  $g$  and  $g^{-1}$  have the same parity and  $\sigma^i$  is even. So  $g^{-1}\sigma^i g \in \langle \sigma \rangle$  showing that  $\langle \sigma \rangle$  is normal in  $S_3$ . As  $(3)\tau = 3, (3)\tau^2 = 3$  and  $(3)\sigma^{-1}\tau\sigma = (2)\tau\sigma = (1)\sigma = 2$  we see  $\sigma^{-1}\tau\sigma \notin \langle \tau \rangle$ . So  $\langle \tau \rangle$  is not a normal subgroup of  $S_3$ .

Suppose  $Kg_1 = Kg'_1$  and  $Kg_2 = Kg'_2$ . Using the above theory we obtain

$$Kg_1 g_2 = Kg'_1 g_2 = g'_1 g_2 K = g'_1 g'_2 K = Kg'_1 g'_2$$

showing that coset multiplication is unambiguously defined. So  $G/K$  is closed under coset multiplication. Let  $e$  denote the identity element of  $G$  and let  $g, g_1, g_2, g_3 \in G$ . Then

$(Kg_1 Kg_2)Kg_3 = K(g_1 g_2)g_3 = Kg_1(g_2 g_3) = Kg_1(Kg_2 Kg_3)$  showing that coset multiplication is associative. As  $KeKg = Keg = Kg = Kge = KgKe$  we see that  $K = Ke$  is the identity element of

$G/K$ . As  $Kg^{-1}Kg = Kg^{-1}g = Ke = Kgg^{-1} = KgKg^{-1}$  we see that  $Kg^{-1}$  is the inverse of  $Kg$ . So  $G/K$  is a group.

(b)  $G = \{\bar{1}, \bar{2}, \bar{4}, \bar{7}, \bar{8}, \bar{11}, \bar{13}, \bar{14}\}$ . For  $K = \{\bar{1}, \bar{4}\}$  we have

$K\bar{2} = \{\bar{2}, \bar{8}\}$ ,  $K\bar{7} = \{\bar{7}, \bar{13}\}$ ,  $K\bar{11} = \{\bar{11}, \bar{14}\}$  which partition  $G$ . The multiplication table of  $G/K$  is

$\times$	$K$	$K\bar{2}$	$K\bar{7}$	$K\bar{11}$
$K$	$K$	$K\bar{2}$	$K\bar{7}$	$K\bar{11}$
$K\bar{2}$	$K\bar{2}$	$K$	$K\bar{11}$	$K\bar{7}$
$K\bar{7}$	$K\bar{7}$	$K\bar{11}$	$K$	$K\bar{2}$
$K\bar{11}$	$K\bar{11}$	$K\bar{7}$	$K\bar{2}$	$K$

For example  $K\bar{2} \times K\bar{7} = K\bar{14} = K\bar{11}$ . The pattern in the table is the same as that in the addition table of the Klein 4-group  $\mathbb{Z}_2 \oplus \mathbb{Z}_2$ . So  $G/K$  has isomorphism type  $C_2 \oplus C_2$ .

For  $K = \{\bar{1}, \bar{14}\}$  we have  $K\bar{2} = \{\bar{2}, \bar{13}\}$ ,  $K\bar{4} = \{\bar{4}, \bar{11}\}$ ,  $K\bar{8} = \{\bar{7}, \bar{8}\}$  which partition  $G$ . The multiplication table of  $G/K$  is

$\times$	$K$	$K\bar{2}$	$K\bar{4}$	$K\bar{8}$
$K$	$K$	$K\bar{2}$	$K\bar{4}$	$K\bar{8}$
$K\bar{2}$	$K\bar{2}$	$K\bar{4}$	$K\bar{8}$	$K$
$K\bar{4}$	$K\bar{4}$	$K\bar{8}$	$K$	$K\bar{2}$
$K\bar{8}$	$K\bar{8}$	$K$	$K\bar{2}$	$K\bar{4}$

and so  $G/K$  is cyclic being generated by  $K\bar{2}$ . So  $G/K$  has isomorphism type  $C_4$ .

(i) Taking  $K_1 = \{\bar{1}, \bar{4}\}$  and  $K_2 = \{\bar{1}, \bar{14}\}$  we see that  $K_1 \cong K_2$  as both  $K_1$  and  $K_2$  are cyclic of order 2. However  $G/K_1$  and  $G/K_2$  are not isomorphic. So  $K_1 \cong K_2$  does not imply  $G/K_1 \cong G/K_2$ .

Let  $K = \{\bar{1}, \bar{4}, \bar{11}, \bar{14}\}$  which is a Klein 4-group. The two cosets  $K$  and  $K\bar{2} = \{\bar{2}, \bar{7}, \bar{8}, \bar{13}\}$  partition  $G$ . The multiplication table of  $G/K$  is

$\times$	$K$	$K\bar{2}$
$K$	$K$	$K\bar{2}$
$K\bar{2}$	$K\bar{2}$	$K$

and  $G/K$  is of isomorphism type  $C_2$ .

Let  $K = \{\bar{1}, \bar{2}, \bar{4}, \bar{8}\}$  which is cyclic of order 4. The two cosets  $K$  and  $K\bar{7} = \{\bar{7}, \bar{11}, \bar{13}, \bar{14}\}$  partition  $G$ . The multiplication table of  $G/K$  is

$\times$	$K$	$K\bar{7}$
$K$	$K$	$K\bar{7}$
$K\bar{7}$	$K\bar{7}$	$K$

and  $G/K$  has isomorphism type  $C_2$ .

(ii) Taking  $K_1 = \{\bar{1}, \bar{4}, \bar{11}, \bar{14}\}$  and  $K_2 = \{\bar{1}, \bar{2}, \bar{4}, \bar{8}\}$  we see that  $G/K_1$  and  $G/K_2$  are isomorphic although  $K_1$  and  $K_2$  are not isomorphic. So  $G/K_1 \cong G/K_2$  does not imply  $K_1 \cong K_2$ .

(c) Write  $x = (e)\theta$ . Then  $x^2 = (e)\theta(e)\theta = (e^2)\theta = (e)\theta = x$ . As  $x \in G'$  there is  $x^{-1} \in G'$  with  $xx^{-1} = e'$ . Hence  $e' = xx^{-1} = x^2x^{-1} = x$ , i.e.  $(e)\theta = e'$ . Applying  $\theta$  to  $g^{-1}g = e = gg^{-1}$  gives  $(g^{-1})\theta(g)\theta = e' = (g)\theta(g^{-1})\theta$  showing that  $(g^{-1})\theta$  is the inverse of  $(g)\theta$ , i.e.  $(g^{-1})\theta = ((g)\theta)^{-1}$  for all  $g \in G$ . As  $(e)\theta = e'$  we see  $e \in K$ . Let  $k_1, k_2 \in K$ . Then  $(k_1k_2)\theta = (k_1)\theta(k_2)\theta = e'e' = e'$  showing  $k_1k_2 \in K$ . As  $(k_1^{-1})\theta = ((k_1)\theta)^{-1} = e'^{-1} = e'$  we see  $k_1^{-1} \in K$ . Therefore  $K = \ker \theta$  is a subgroup of  $G$ . As  $(e)\theta = e'$  we see  $e' \in \text{im } \theta$ . As  $(g_1)\theta(g_2)\theta = (g_1g_2)\theta \in \text{im } \theta$  for  $g_1, g_2 \in G$  and so  $\text{im } \theta$  is closed under multiplication. As  $((g_1)\theta)^{-1} = (g_1^{-1})\theta \in \text{im } \theta$  for all  $g_1 \in G$  we see  $\text{im } \theta$  is a subgroup of  $G'$ . Let  $k \in K, g \in G$ . Then

$$(g^{-1}kg)\theta = (g^{-1})\theta(k)\theta(g)\theta = ((g)\theta)^{-1}e'(g)\theta = ((g)\theta)^{-1}(g)\theta = e'$$

showing that  $g^{-1}kg \in K$ . So  $K = \ker \theta$  is normal in  $G$ . Kernels of group homomorphisms are normal subgroups. As  $(kg)\theta = (k)\theta(g)\theta = e'(g)\theta = (g)\theta$  all elements of the coset  $Kg$  are mapped by  $\theta$  to the same element  $(g)\theta$ . So  $\tilde{\theta}: G/K \rightarrow \text{im } \theta$  defined by  $(Kg)\tilde{\theta} = (g)\theta$  is unambiguous and surjective. Suppose  $(Kg_1)\tilde{\theta} = (Kg_2)\tilde{\theta}$ . Then  $(g_1)\theta = (g_2)\theta$  and so

$$(g_1g_2^{-1})\theta = (g_1)\theta((g_2)\theta)^{-1} = (g_1)\theta((g_1)\theta)^{-1} = e'$$

showing  $g_1g_2^{-1} = k \in K$ . So  $g_1 = kg_2$  and hence  $Kg_1 = Kg_2$ , that is,  $\tilde{\theta}$  is injective. As

$$((Kg_1)(Kg_2))\tilde{\theta} = (Kg_1g_2)\tilde{\theta} = (g_1g_2)\theta = (g_1)\theta(g_2)\theta = (Kg_1)\tilde{\theta}(Kg_2)\tilde{\theta}$$

we see that  $\tilde{\theta}$  is a group isomorphism and so  $\tilde{\theta}: G/K \cong \text{im } \theta$ , the first isomorphism theorem for groups.

Let  $A_1, A_2 \in GL_t(R)$  where  $R$  is a non-trivial commutative ring. Then  $\det A_1, \det A_2 \in U(R)$  and

$(A_1A_2)\theta = \det A_1A_2 = \det A_1 \det A_2 = (A_1)\theta(A_2)\theta$  by the multiplicative property (1.18) of

determinants. So  $\theta: GL_t(R) \rightarrow U(R)$  is a group homomorphism. For  $u \in U(R)$  the diagonal  $t \times t$

matrix  $U = \text{diag}(u, e, e, \dots, e)$ , where  $e$  is the 1-element of  $R$ , is invertible over  $R$  and

$(U)\theta = \det U = u$ . So  $\theta$  is surjective, i.e.  $\text{im } \theta = U(R)$ . Also  $SL_t(R)$  is a normal subgroup of

$GL_t(R)$  with  $GL_t(R)/SL_t(R) \cong U(R)$  by the first isomorphism theorem above for groups. Taking

$R = \mathbb{Z}_p$  we have  $|U(\mathbb{Z}_p)| = |\mathbb{Z}_p^*| = p-1$  and so

$$|SL_t(\mathbb{Z}_p)| = |GL_t(\mathbb{Z}_p)| / (p-1) = (p^t - 1)(p^t - p) \cdots (p^t - p^{t-1}) / (p-1).$$

(d) Let  $g_1, g'_1, g''_1 \in G_1$  and  $g_2, g'_2, g''_2 \in G_2$ . Then

$$\begin{aligned} ((g_1, g_2)(g'_1, g'_2))(g''_1, g''_2) &= (g_1g'_1, g_2g'_2)(g''_1, g''_2) = ((g_1g'_1)g''_1, (g_2g'_2)g''_2) = \\ &= (g_1(g'_1g''_1), g_2(g'_2g''_2)) = (g_1, g_2)(g'_1g''_1, g'_2g''_2) = (g_1, g_2)((g'_1, g'_2)(g''_1, g''_2)) \end{aligned}$$

showing that componentwise multiplication on  $G_1 \times G_2$  is associative. The pair  $(e_1, e_2)$  consisting of the identity elements  $e_1$  of  $G_1$  and  $e_2$  of  $G_2$  is the identity element of  $G_1 \times G_2$  because

$$(g_1, g_2)(e_1, e_2) = (g_1e_1, g_2e_2) = (g_1, g_2) = (e_1g_1, e_2g_2) = (e_1, e_2)(g_1, g_2) \text{ for all } (g_1, g_2) \in G_1 \times G_2.$$

The inverse of  $(g_1, g_2)$  is  $(g_1^{-1}, g_2^{-1})$  as

$$(g_1, g_2)(g_1^{-1}, g_2^{-1}) = (g_1g_1^{-1}, g_2g_2^{-1}) = (e_1, e_2) = (g_1^{-1}g_1, g_2^{-1}g_2) = (g_1^{-1}, g_2^{-1})(g_1, g_2).$$

So  $G_1 \times G_2$  is a group, the external direct product of  $G_1$  and  $G_2$ .

Let  $K_1 = \ker \pi_2 = \{(g_1, e_2) : g_1 \in G_1\} \cong G_1$ . Then  $K_1$ , being the kernel of a homomorphism from  $G_1 \times G_2$  to  $G_2$ , is a normal subgroup of  $G_1 \times G_2$ . Similarly  $K_2 = \ker \pi_1 = \{(e_1, g_2) : g_2 \in G_2\} \cong G_2$  is also a normal subgroup of  $G_1 \times G_2$ . Then  $K_1 \cap K_2 = \{(e_1, e_2)\}$  is trivial. Let  $k_1 = (g_1, e_2) \in K_1$  and  $k_2 = (e_1, g_2) \in K_2$ . Then  $k_1 k_2 = (g_1, e_2)(e_1, g_2) = (g_1 e_1, e_2 g_2) = (g_1, g_2)$  and so  $K_1 K_2 = G_1 \times G_2$ .

### Solution 5

(a) Let  $m_1, m_2 \in \mathbb{Z}$ . Then  $(m_1 + m_2)\chi = (m_1 + m_2)e = m_1 e + m_2 e = (m_1)\chi + (m_2)\chi$  applying the result of Exercises 2.1, Question 8(c) to the additive group of  $F$ . Also

$(m_1 m_2)\chi = (m_1 m_2)e = (m_1 m_2)e^2 = (m_1 e)(m_2 e) = (m_1)\chi (m_2)\chi$ . So  $\chi$  is a ring homomorphism as  $(1)\chi = e$ . By (1.15) there is a unique non-negative integer  $d$  with  $\ker \chi = \langle d \rangle$ . By the first isomorphism theorem for rings  $\tilde{\chi} : \mathbb{Z}/\ker \chi \cong \text{im } \chi$ , i.e.  $\tilde{\chi} : \mathbb{Z}_d \cong \text{im } \chi$  defined by  $(\bar{i})\tilde{\chi} = ie$  for all  $\bar{i} \in \mathbb{Z}_d = \mathbb{Z}/\langle d \rangle$  is a ring isomorphism. Suppose  $d > 0$ . As  $e \neq 0$  we see  $d \neq 1$ . Suppose that  $d$  is not prime. Then  $d = d_1 d_2$  for positive integers  $d_1, d_2$ . Hence  $(d_1)\chi (d_2)\chi = (d_1 d_2)\chi = (d)\chi = 0$ . As  $F$  has no divisors of zero, either  $(d_1)\chi = 0$  or  $(d_2)\chi = 0$ , i.e. either  $d_1 \in \langle d \rangle$  or  $d_2 \in \langle d \rangle$  both of which are impossible as  $d$  is not a divisor of either  $d_1$  or  $d_2$ . So either  $d = 0$  or  $d = p$  a prime. For  $d = 0$  we have  $\tilde{\chi} : \mathbb{Z}_0 \cong \text{im } \chi$  and so  $\text{im } \chi$  has an infinite number of elements. So for each finite field  $F$  there is a prime  $p$  (the characteristic of  $F$ ) such that  $\tilde{\chi} : \mathbb{Z}_p \cong \text{im } \chi$ . As  $\mathbb{Z}_p$  is a field we see that  $\text{im } \chi = F_0$  is a subfield of  $F$ . Regard the elements  $v, v'$  of  $F$  as vectors and the elements  $a$  of  $F_0$  as scalars; then  $v + v' \in F$  and  $av \in F$  satisfy the vector space laws as these laws follow directly from the laws of a field. In short  $F$  is a vector space over  $F_0$ . As there are only a finite number of vectors, this vector space is finitely generated and so has a basis  $v_1, v_2, \dots, v_s$ . Each element of  $F$  can be uniquely expressed in the form  $a_1 v_1 + a_2 v_2 + \dots + a_s v_s$  where  $a_1, a_2, \dots, a_s \in F_0$ . As  $\mathbb{Z}_p \cong F_0$  there are  $p$  independent choices for each of the  $s$  scalars  $a_1, a_2, \dots, a_s$ . Hence  $|F| = p^s$ .

For  $0 < r < p$  the binomial coefficient  $\binom{p}{r}$  is divisible by the prime  $p$ . As  $pe = 0$  by the first

paragraph, we see  $\binom{p}{r} a^{p-r} b^r = \left( \binom{p}{r} / p \right) (pe) a^{p-r} b^r = 0$ . By the binomial theorem

$$(a+b)^p = \sum_{r=0}^p \binom{p}{r} a^{p-r} b^r = a^p + b^p \text{ as only the first and last terms contribute to the sum. Therefore}$$

$(a+b)\theta = (a+b)^p = a^p + b^p = (a)\theta + (b)\theta$  showing that  $\theta$  is additive. As

$\ker \theta = \{a \in F : a^p = 0\} = \{0\}$  we see that  $\theta$  is injective by Question 1(a)(i) above. As  $\theta : F \rightarrow F$  and  $F$  has only a finite number of elements we deduce that  $\theta$  is also surjective. Finally  $(ab)\theta = (ab)^p = a^p b^p = (a)\theta (b)\theta$  showing that  $\theta$  is multiplicative. As  $(e)\theta = e^p = e$  we conclude that  $\theta$  is an automorphism of the finite field  $F$ .

(b) Let  $A = (a_{ij})$  and  $B = (b_{ij})$  be  $t \times t$  matrices over  $\mathbb{Z}_n$ . Then

$$(A+B)\delta_i = ((a_{ij} + b_{ij})\delta_j) = ((a_{ij})\delta_j + (b_{ij})\delta_j) = ((a_{ij})\delta_j) + ((b_{ij})\delta_j) = (A)\delta_i + (B)\delta_i \text{ and}$$

$(AB)\delta_t = ((\sum_{j=1}^t a_{ij}b_{jk})\delta_1) = ((\sum_{j=1}^t (a_{ij})\delta_1(b_{jk})\delta_1)) = ((a_{ij})\delta_1)((b_{jk})\delta_1) = (A)\delta_t(B)\delta_t$ . As  $(\bar{1}_n)\delta_1 = \bar{1}_d$  and  $(\bar{0}_n)\delta_1 = \bar{0}_d$  we see that  $\delta_t$  maps the 1-element of the ring  $\mathfrak{M}_t(\mathbb{Z}_n)$  to the 1-element of  $\mathfrak{M}_t(\mathbb{Z}_d)$  and so  $\delta_t$  is a ring homomorphism. As  $\delta_1$  is surjective so also is  $\delta_t$ , i.e.  $\text{im } \delta_t = \mathfrak{M}_t(\mathbb{Z}_d)$ . Let  $A = (a_{ij}) \in \ker \delta_t$ . Then  $(a_{ij})\delta_1 = \bar{0}_d$  and so  $a_{ij} = \bar{m}_n$  where  $d \mid m$ . There are therefore  $n/d$  independent choices for each of the  $t^2$  entries in  $A$ . Hence  $|\ker \delta_t| = (n/d)^{t^2}$ . Using the multiplicative property of determinants  $A \in GL_t(\mathbb{Z}_n) \Leftrightarrow \det A = \bar{m}_n \in U(\mathbb{Z}_n) \Leftrightarrow \gcd\{m, p^s\} = 1 \Leftrightarrow \gcd\{m, p\} = 1 \Leftrightarrow \det(A)\delta_t = \bar{m}_p \in U(\mathbb{Z}_p) = \mathbb{Z}_p^* \Leftrightarrow (A)\delta_t \in GL_t(\mathbb{Z}_p)$ . Hence  $\delta_t| : GL_t(\mathbb{Z}_{p^s}) \rightarrow GL_t(\mathbb{Z}_p)$  makes sense and is surjective. As  $\delta_t$  respects products so also does the restriction  $\delta_t|$ , i.e. it is a homomorphism of multiplicative groups. Now  $A \in \ker \delta_t| \Leftrightarrow (A)\delta_t = (I)\delta_t \Leftrightarrow A \in \ker \delta_t + I$  and so  $\ker \delta_t| = \ker \delta_t + I$  is a normal subgroup of  $GL_t(\mathbb{Z}_{p^s})$  having  $|\ker \delta_t| = p^{(s-1)t^2}$  elements. By the first isomorphism theorem for groups  $GL_t(\mathbb{Z}_{p^s}) / \ker \delta_t| \cong GL_t(\mathbb{Z}_p)$ . So  $GL_t(\mathbb{Z}_{p^s})$  partitions into  $|GL_t(\mathbb{Z}_p)|$  cosets of  $\ker \delta_t|$ . Hence  $|GL_t(\mathbb{Z}_{p^s})| = p^{(s-1)t^2} (p^t - 1)(p^t - p) \cdots (p^t - p^{t-1})$  using the formula following (2.18). So  $|GL_3(\mathbb{Z}_4)| = 2^9 (2^3 - 1)(2^3 - 2)(2^3 - 2^2) = 86016$  and  $|GL_2(\mathbb{Z}_{17})| = 17^{0 \times 4} (17^2 - 1)(17^2 - 17) = 78336$ . As  $|GL_2(\mathbb{Z}_{125})| = |GL_2(\mathbb{Z}_{5^3})| = 5^8 (5^2 - 1)(5^2 - 5) = 187500000$  and  $|GL_2(\mathbb{Z}_{128})| = |GL_2(\mathbb{Z}_{2^7})| = 2^{24} (2^2 - 1)(2^2 - 2) = 100663296$ . So the answer is: No!

The 16 matrices in the kernel of  $\delta_2| : GL_2(\mathbb{Z}_4) \rightarrow GL_2(\mathbb{Z}_2)$  are:

$$\begin{pmatrix} \bar{1} & \bar{0} \\ \bar{0} & \bar{1} \end{pmatrix}, \begin{pmatrix} \bar{1} & \bar{2} \\ \bar{0} & \bar{1} \end{pmatrix}, \begin{pmatrix} \bar{1} & \bar{0} \\ \bar{2} & \bar{1} \end{pmatrix}, \begin{pmatrix} \bar{1} & \bar{2} \\ \bar{2} & \bar{1} \end{pmatrix}, \begin{pmatrix} \bar{3} & \bar{0} \\ \bar{0} & \bar{1} \end{pmatrix}, \begin{pmatrix} \bar{3} & \bar{2} \\ \bar{0} & \bar{1} \end{pmatrix}, \begin{pmatrix} \bar{3} & \bar{0} \\ \bar{2} & \bar{1} \end{pmatrix}, \begin{pmatrix} \bar{3} & \bar{2} \\ \bar{2} & \bar{1} \end{pmatrix}, \\
 \begin{pmatrix} \bar{1} & \bar{0} \\ \bar{0} & \bar{3} \end{pmatrix}, \begin{pmatrix} \bar{1} & \bar{2} \\ \bar{0} & \bar{3} \end{pmatrix}, \begin{pmatrix} \bar{1} & \bar{0} \\ \bar{2} & \bar{3} \end{pmatrix}, \begin{pmatrix} \bar{1} & \bar{2} \\ \bar{2} & \bar{3} \end{pmatrix}, \begin{pmatrix} \bar{3} & \bar{0} \\ \bar{0} & \bar{3} \end{pmatrix}, \begin{pmatrix} \bar{3} & \bar{2} \\ \bar{0} & \bar{3} \end{pmatrix}, \begin{pmatrix} \bar{3} & \bar{0} \\ \bar{2} & \bar{3} \end{pmatrix}, \begin{pmatrix} \bar{3} & \bar{2} \\ \bar{2} & \bar{3} \end{pmatrix}$$

each displaying the pattern  $\begin{pmatrix} \overline{\text{odd}} & \overline{\text{even}} \\ \overline{\text{even}} & \overline{\text{odd}} \end{pmatrix}$ .

$$(c) A_1 = \begin{pmatrix} \overline{41} & \overline{42} \\ \overline{43} & \overline{44} \end{pmatrix} \text{ as } 41 \equiv 1 \pmod{8}, 41 \equiv 5 \pmod{9} \text{ etc. and } A_2 = \begin{pmatrix} \overline{9} & \overline{64} \\ \overline{64} & \overline{9} \end{pmatrix} = \begin{pmatrix} \overline{9} & -\overline{8} \\ -\overline{8} & \overline{9} \end{pmatrix}.$$

$$(A_1 + A_2)\alpha = \begin{pmatrix} \overline{50} & \overline{34} \\ \overline{35} & \overline{53} \end{pmatrix} \alpha = \left( \begin{pmatrix} \overline{2} & \overline{2} \\ \overline{3} & \overline{5} \end{pmatrix}, \begin{pmatrix} \overline{5} & \overline{7} \\ \overline{8} & \overline{8} \end{pmatrix} \right) = (A_1)\alpha + (A_2)\alpha.$$

$$(A_1 A_2)\alpha = \begin{pmatrix} \overline{33} & \overline{50} \\ \overline{35} & \overline{52} \end{pmatrix} = \left( \begin{pmatrix} \overline{1} & \overline{2} \\ \overline{3} & \overline{4} \end{pmatrix}, \begin{pmatrix} \overline{6} & \overline{5} \\ \overline{8} & \overline{7} \end{pmatrix} \right) = (A_1)\alpha (A_2)\alpha.$$

Write  $p_i^{s_i} = q_i$  for  $1 \leq i \leq k$ . Then

$$\begin{aligned} |GL_t(\mathbb{Z}_n)| &= |GL_t(\mathbb{Z}_{q_1})| |GL_t(\mathbb{Z}_{q_2})| \cdots |GL_t(\mathbb{Z}_{q_k})| = \\ &= \prod_{i=1}^k (q_i/p_i)^{t^2} (p_i^t - 1)(p_i^t - p_i) \cdots (p_i^t - p_i^{t-1}). \end{aligned}$$

### Solution 6

(a) Suppose the invariant factors of the  $s \times t$  matrix  $A$  are not all positive. Then  $d_s = 0$ . There are invertible matrices  $P$  and  $Q$  over  $\mathbb{Z}$  with  $PA = DQ$  where  $D = \text{diag}(d_1, d_2, \dots, d_s)$ . The last row of  $D$  is zero and so the last row of  $DQ$  is zero. Let  $(m_1, m_2, \dots, m_s)$  denote the last row of  $P$ . The last row of  $PA$  is  $(m_1, m_2, \dots, m_s)A = m_1\rho_1 + m_2\rho_2 + \dots + m_s\rho_s = 0$ . As  $\det P \neq 0$  at least one of  $m_1, m_2, \dots, m_s$  is non-zero, showing that  $\rho_1, \rho_2, \dots, \rho_s$  are  $\mathbb{Z}$ -dependent. Hence  $\rho_1, \rho_2, \dots, \rho_s$  being  $\mathbb{Z}$ -independent implies  $d_1, d_2, \dots, d_s$  all positive. Conversely suppose that  $d_1, d_2, \dots, d_s$  are all positive. Let  $m_1, m_2, \dots, m_s$  be integers with  $m_1\rho_1 + m_2\rho_2 + \dots + m_s\rho_s = 0$ . Write  $\underline{m} = (m_1, m_2, \dots, m_s)$  and so  $\underline{m}A = 0$ . Hence  $\underline{m}P^{-1}DQ = 0$  where  $P, Q, D$  are as above. Hence  $\underline{m}P^{-1}D = 0$  which give the  $s$  equations  $l_i d_i = 0$  for  $1 \leq i \leq s$  where  $\underline{m}P^{-1} = (l_1, l_2, \dots, l_s)$ . So each  $l_i = 0$ , i.e.  $\underline{m}P^{-1} = 0$  and so  $\underline{m} = 0$  showing that  $\rho_1, \rho_2, \dots, \rho_s$  are  $\mathbb{Z}$ -independent.

Suppose there is a  $\mathbb{Z}$ -basis of  $\mathbb{Z}^t$  beginning with  $\rho_1, \rho_2, \dots, \rho_s$ . By (2.23) there is an invertible  $t \times t$  matrix  $Q$  over  $\mathbb{Z}$  of the type  $Q = \begin{pmatrix} A \\ B \end{pmatrix}$ , i.e. the first  $s$  rows of  $Q$  are  $\rho_1, \rho_2, \dots, \rho_s$ . Comparing the

first  $s$  rows in  $QQ^{-1} = I$  gives  $AQ^{-1} = (I \mid 0)$  where  $I$  is the  $s \times s$  identity matrix. So

$AQ^{-1} = S(A)$ , the Smith normal form of  $A$ , and  $d_1 = d_2 = \dots = d_s = 1$ . Conversely suppose  $d_1 = d_2 = \dots = d_s = 1$ . In the case  $s = t$  the  $t \times t$  matrix  $A$  is invertible over  $\mathbb{Z}$  and  $\rho_1, \rho_2, \dots, \rho_s$  is a  $\mathbb{Z}$ -basis of  $\mathbb{Z}^t$ . For  $s < t$ , by Exercises 1.3, Question 5(d) the  $s \times t$  matrix  $A$  can be changed into  $S(A) = (I \mid 0)$  using *ecos* only. By (1.4) there is an invertible  $t \times t$  matrix  $Q$  over  $\mathbb{Z}$  with

$AQ^{-1} = (I \mid 0)$ . Hence  $(I \mid 0)Q = A$  and so  $Q = \begin{pmatrix} A \\ B \end{pmatrix}$ . The rows of  $Q$  are therefore a  $\mathbb{Z}$ -basis of

$\mathbb{Z}^t$  beginning with  $\rho_1, \rho_2, \dots, \rho_s$ .

(b) Suppose  $\rho_1, \rho_2, \dots, \rho_s$  generate  $\mathbb{Z}^t$ . For each row  $e_i$  of the  $t \times t$  identity matrix  $I$  over  $\mathbb{Z}$  there are integers  $b_{ij}$  with  $b_{i1}\rho_1 + b_{i2}\rho_2 + \dots + b_{is}\rho_s = e_i$  for  $1 \leq i \leq t$ . So the  $t \times s$  matrix  $B = (b_{ij})$  satisfies  $BA = I$ . By (1.19) the non-negative gcd  $g_t(A)$  of the  $t$ -minors of  $A$  is a divisor of  $g_t(I) = 1$ . So  $g_t(A) = 1$  and as  $g_t(A) = d_1 d_2 \cdots d_t$  we see that  $d_1 = d_2 = \dots = d_t = 1$ .

Suppose  $d_1 = d_2 = \dots = d_t = 1$ . There are invertible matrices  $P$  and  $Q$  over  $\mathbb{Z}$  with  $PA = \begin{pmatrix} I \\ 0 \end{pmatrix} Q$ .

Consider  $(l_1, l_2, \dots, l_t) \in \mathbb{Z}^t$  and write  $(l_1, l_2, \dots, l_t)Q^{-1} = (l'_1, l'_2, \dots, l'_t)$ . There are unique integers

$m_1, m_2, \dots, m_s$  with  $(m_1, m_2, \dots, m_s)P^{-1} = (l'_1, l'_2, \dots, l'_t, 0, 0, \dots, 0) = \begin{pmatrix} I \\ 0 \end{pmatrix} (l'_1, l'_2, \dots, l'_t)$ . So

$(m_1, m_2, \dots, m_s)A = (l_1, l_2, \dots, l_t)$ , i.e.  $(l_1, l_2, \dots, l_t) = m_1\rho_1 + m_2\rho_2 + \dots + m_s\rho_s$ , showing that  $\rho_1, \rho_2, \dots, \rho_s$  generate  $\mathbb{Z}^t$ .

Suppose a  $\mathbb{Z}$ -basis  $\rho_{i_1}, \rho_{i_2}, \dots, \rho_{i_t}$  of  $\mathbb{Z}^t$  can be selected from  $\rho_1, \rho_2, \dots, \rho_s$ . The determinant  $\Delta$  of the  $t \times t$  submatrix of  $A$  made up of rows  $i_1, i_2, \dots, i_t$  has value  $\pm 1$ . The  $s \times s$  matrix  $P$  with rows  $\rho'_1, \rho'_2, \dots, \rho'_t$  together with the  $s-t$  rows  $e_j$  of the  $s \times s$  identity matrix  $I$  for  $j \neq i_1, i_2, \dots, i_t$  is invertible over  $\mathbb{Z}$  because expanding  $\det P$  along rows  $i_1, i_2, \dots, i_t$  gives  $\det P = \pm \Delta = \pm 1$ . By (2.23) the rows of  $P$  form a  $\mathbb{Z}$ -basis of  $\mathbb{Z}^s$ . So  $\rho'_1, \rho'_2, \dots, \rho'_t$  can be extended to a  $\mathbb{Z}$ -basis of  $\mathbb{Z}^s$  using  $s-t$  rows of the  $s \times s$  identity matrix. The converse is proved by reversing the above steps.

(c) (i) As  $\begin{vmatrix} 1 & 3 & 2 \\ 4 & 6 & 5 \\ 7 & 9 & 8 \end{vmatrix} = 0$  the given elements are  $\mathbb{Z}$ -dependent and so not contained in any  $\mathbb{Z}$ -basis of

$\mathbb{Z}^3$ . (ii) As  $S(A) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \end{pmatrix}$  where  $A = \begin{pmatrix} 1 & 3 & 7 \\ 3 & 5 & 9 \end{pmatrix}$ , the rows of  $A$  are  $\mathbb{Z}$ -independent but are not

contained in any  $\mathbb{Z}$ -basis of  $\mathbb{Z}^3$ . (iii) As  $S(A) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$  where  $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 3 & 5 \end{pmatrix}$ , the rows of  $A$  are  $\mathbb{Z}$ -independent and are contained in a  $\mathbb{Z}$ -basis of  $\mathbb{Z}^3$ . (iv) For

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \\ 1 & 1 & 4 \\ 4 & 1 & 1 \end{pmatrix}$$

we have  $d_1 = d_2 = 1, d_3 = 6$  and so the rows of  $A$  do not generate  $\mathbb{Z}^3$ . (v) As

$$\begin{vmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 4 \end{vmatrix} = -1,$$

we see that the given elements generate  $\mathbb{Z}^3$  and the first three of them form a  $\mathbb{Z}$ -basis of  $\mathbb{Z}^3$ . (vi) The 3-minors of

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 2 \\ 1 & 3 & 1 \\ 4 & 1 & 1 \end{pmatrix}$$

are  $-2, 3, -6, -17$  and so  $g_3(A) = 1$  giving  $d_1 = d_2 = d_3 = 1$ . So the rows of  $A$  generate  $\mathbb{Z}^3$  but a  $\mathbb{Z}$ -basis of  $\mathbb{Z}^3$  cannot be selected from them.

### Solution 7

(a) Suppose  $v'_1, v'_2, \dots, v'_t$  generate  $M'$ . Let  $v' \in M'$ . By (2.19)(i) there are  $r_1, r_2, \dots, r_t \in R$  with  $r_1 v'_1 + r_2 v'_2 + \dots + r_t v'_t = v'$ . Write  $v = r_1 v_1 + r_2 v_2 + \dots + r_t v_t \in M$ . Then

$$(v)\theta = (r_1 v_1 + r_2 v_2 + \dots + r_t v_t)\theta = r_1(v_1)\theta + r_2(v_2)\theta + \dots + r_t(v_t)\theta = r_1 v'_1 + r_2 v'_2 + \dots + r_t v'_t = v'$$

showing  $\theta$  to be surjective. Conversely suppose that  $\theta$  is surjective. Let  $v' \in M'$ . There is  $v \in M$  with  $(v)\theta = v'$ . As  $v_1, v_2, \dots, v_t$  generate  $M$  there are  $r_1, r_2, \dots, r_t \in R$  with  $v = r_1 v_1 + r_2 v_2 + \dots + r_t v_t$ . As  $\theta$  is  $R$ -linear we have

$$v' = (v)\theta = (r_1 v_1 + r_2 v_2 + \dots + r_t v_t)\theta = r_1(v_1)\theta + r_2(v_2)\theta + \dots + r_t(v_t)\theta = r_1 v'_1 + r_2 v'_2 + \dots + r_t v'_t$$

showing that  $v'_1, v'_2, \dots, v'_t$  generate  $M'$ . Suppose  $v'_1, v'_2, \dots, v'_t$  are  $R$ -independent elements of  $M'$ . Consider

$u \in \ker \theta$ . As  $v_1, v_2, \dots, v_t$  generate  $M$  there are  $r_1, r_2, \dots, r_t \in R$  with  $u = r_1 v_1 + r_2 v_2 + \dots + r_t v_t$ . As  $\theta$  is  $R$ -linear  $0 = (u)\theta = (r_1 v_1 + r_2 v_2 + \dots + r_t v_t)\theta = r_1(v_1)\theta + r_2(v_2)\theta + \dots + r_t(v_t)\theta = r_1 v'_1 + r_2 v'_2 + \dots + r_t v'_t$  and so  $r_1 = r_2 = \dots = r_t = 0$ . Hence  $u = 0v_1 + 0v_2 + \dots + 0v_t = 0$  showing  $\ker \theta = \{0\}$  and so  $\theta$  is

injective by Question 1(a)(i) above. Conversely suppose  $\theta$  is injective. Then  $\ker \theta = \{0\}$  by

Question 1(a)(i) above. Consider  $r_1 v'_1 + r_2 v'_2 + \dots + r_t v'_t = 0$  where  $r_1, r_2, \dots, r_t \in R$ . Then

$$(r_1 v_1 + r_2 v_2 + \dots + r_t v_t)\theta = 0, \text{ i.e. } r_1 v_1 + r_2 v_2 + \dots + r_t v_t \in \ker \theta. \text{ So } r_1 v_1 + r_2 v_2 + \dots + r_t v_t = 0. \text{ As}$$

$v_1, v_2, \dots, v_t$  are  $R$ -independent we conclude  $r_1 = r_2 = \dots = r_t = 0$  and so  $v'_1, v'_2, \dots, v'_t$  are

$R$ -independent elements of  $M'$ . Now suppose  $\theta: M \rightarrow M'$  is an isomorphism and  $M$  is free of rank  $t$ . Then  $M$  has  $R$ -basis  $v_1, v_2, \dots, v_t$ . Let  $v'_i = (v_i)\theta$  for  $1 \leq i \leq t$ . As  $\theta$  is surjective and injective

$v'_1, v'_2, \dots, v'_t$  generate  $M'$  and are  $R$ -independent using the above theory. So  $M'$  has  $R$ -basis

$v'_1, v'_2, \dots, v'_t$  and so is free of rank  $t = t'$ . Conversely suppose  $M$  and  $M'$  to be free  $R$ -modules of the same rank  $t$ . Let  $M$  have  $R$ -basis  $v_1, v_2, \dots, v_t$  and let  $M'$  have  $R$ -basis  $v'_1, v'_2, \dots, v'_t$ . Consider

$\theta: M \rightarrow M'$  defined by  $(r_1 v_1 + r_2 v_2 + \dots + r_t v_t)\theta = r_1 v'_1 + r_2 v'_2 + \dots + r_t v'_t$  for all  $r_1, r_2, \dots, r_t \in R$ . Let

$$u = r_1 v_1 + r_2 v_2 + \dots + r_t v_t, \quad v = s_1 v_1 + s_2 v_2 + \dots + s_t v_t \text{ where } s_i \in R \text{ for } 1 \leq i \leq t. \text{ Then}$$

$$(u+v)\theta = ((r_1 + s_1)v_1 + (r_2 + s_2)v_2 + \dots + (r_t + s_t)v_t)\theta =$$

$$(r_1 + s_1)v'_1 + (r_2 + s_2)v'_2 + \dots + (r_t + s_t)v'_t = (r_1 v'_1 + r_2 v'_2 + \dots + r_t v'_t) + (s_1 v'_1 + s_2 v'_2 + \dots + s_t v'_t) =$$

$$(r_1 v_1 + r_2 v_2 + \dots + r_t v_t)\theta + (s_1 v_1 + s_2 v_2 + \dots + s_t v_t)\theta = (u)\theta + (v)\theta$$

and so  $\theta$  is additive. Also for  $r \in R$  we see

$$(ru)\theta = (r(r_1 v_1 + r_2 v_2 + \dots + r_t v_t))\theta = (rr_1 v_1 + rr_2 v_2 + \dots + rr_t v_t)\theta =$$

$$rr_1 v'_1 + rr_2 v'_2 + \dots + rr_t v'_t = r(r_1 v'_1 + r_2 v'_2 + \dots + r_t v'_t) = r((u)\theta).$$

So  $\theta$  is  $R$ -linear and as  $(v_i)\theta = v'_i$  we can apply the above theory: since  $v'_1, v'_2, \dots, v'_t$  generate  $M'$  and are  $R$ -independent,  $\theta$  is surjective and injective, i.e.  $\theta: M \cong M'$ .



(b) By hypothesis  $M$  has  $R$ -basis  $v_1, v_2, \dots, v_t$ . There are  $t \times t$  matrices  $P = (p_{ij})$  and  $Q = (q_{jk})$  over  $R$  such that  $v_i = \sum_{j=1}^t p_{ij} u_j$  ( $1 \leq i \leq t$ ) and  $u_j = \sum_{k=1}^t q_{jk} v_k$  ( $1 \leq j \leq t$ ) as in the proof of (2.20). Then

$PQ = I$  on comparing coefficients in the equation  $v_i = \sum_{j=1}^t p_{ij} (\sum_{k=1}^t q_{jk} v_k) = \sum_{k=1}^t (\sum_{j=1}^t p_{ij} q_{jk}) v_k$ . So  $Q$  is invertible over  $R$  by (2.18). Hence  $u_1, u_2, \dots, u_t$  is an  $R$ -basis of  $M$  by (2.21).

(c) Regarding  $M$  and  $M'$  as  $\mathbb{Z}$ -modules, by Question 1(a)(i) and (ii) above,  $\ker \theta$  and  $\text{im } \theta$  are additive subgroups of  $M$  and  $M'$  respectively. For  $u \in \ker \theta$ ,  $r \in R$ , we have  $(ru)\theta = r((u)\theta) = r \times 0 = 0$  showing that  $ru \in \ker \theta$ . So  $\ker \theta$  is a submodule of the  $R$ -module  $M$  by (2.26). For  $u' \in \text{im } \theta$ ,  $r \in R$ , there is  $u \in M$  with  $(u)\theta = u'$ . So  $(ru)\theta = r((u)\theta) = ru'$  showing  $ru' \in \text{im } \theta$  as  $ru \in M$ . So  $\text{im } \theta$  is a submodule of the  $R$ -module  $M'$  by (2.26). Let  $\theta$  be bijective. Then  $\theta^{-1}: M' \rightarrow M$  is additive by Exercises 2.1, Question 4(d). Let  $r \in R$  and  $v' \in M'$  and write  $v = (v')\theta^{-1}$ . Then  $(rv)\theta = r((v)\theta) = rv'$ . Applying  $\theta^{-1}$  gives  $(rv')\theta^{-1} = rv = r((v')\theta^{-1})$  showing that  $\theta^{-1}$  is  $R$ -linear. Yes, the inverse of an isomorphism of  $R$ -modules is bijective and  $R$ -linear and so is itself an isomorphism of  $R$ -modules by (2.24).

(d) Consider  $r_1, r_2 \in R$  and  $v \in M$ . Using coset addition, before (2.10), and law 5 (part 2), before (2.19), we see  $(r_1 + r_2)(N + v) = N + (r_1 + r_2)v = N + r_1v + r_2v = (N + r_1) + (N + r_2)$  which shows that law 5 (part 2) holds in  $M/N$ . As law 6 holds in  $M$  we see  $(r_1r_2)(N + v) = N + (r_1r_2)v = N + r_1(r_2v) = r_1(N + r_2v) = r_1(r_2(N + v))$  showing that law 6 holds in  $M/N$ . The 1-element 1 of  $R$  satisfies  $1(N + v) = N + 1v = N + v$  and so law 7 holds in  $M/N$ . Therefore  $M/N$  is an  $R$ -module.

(e) There is an element  $u_0 \in N$ . Then  $z_0 = 0u_0 \in N$ . As  $0 + 1 = 1$  in  $R$  we see  $z_0 + u_0 = 0u_0 + 1u_0 = (0 + 1)u_0 = 1u_0 = u_0$  using the distributive law in  $M$ . Hence  $-u_0 + (z_0 + u_0) = -u_0 + u_0 = 0$  and so  $z_0 = 0$  (the 0-element of  $M$ ) on using the associative and commutative laws of addition in  $M$ . So  $N$  contains the 0-element of  $M$ . For  $u \in N$  we have  $(-1)u \in N$ . Then  $(-1)u + u = (-1)u + 1u = (-1 + 1)u = 0u = 0$  on replacing  $u_0$  by  $u$  in the above paragraph. So  $-u = (-1)u$  and so  $N$  is closed under negation.  $N$  is a subgroup of the additive group of  $M$ . So  $N$  is a submodule of  $M$ .

Consider  $v, v' \in N_1 + N_2$ . There are  $u_1, u'_1 \in N_1$  and  $u_2, u'_2 \in N_2$  with  $v = u_1 + u_2$ ,  $v' = u'_1 + u'_2$ . So  $v + v' = u_1 + u_2 + u'_1 + u'_2 = (u_1 + u'_1) + (u_2 + u'_2) \in N_1 + N_2$  since  $u_1 + u'_1 \in N_1$  and  $u_2 + u'_2 \in N_2$ . So  $N_1 + N_2$  is closed under addition. Also  $rv = r(u_1 + u_2) = ru_1 + ru_2 \in N_1 + N_2$  as  $ru_1 \in N_1$  and  $ru_2 \in N_2$  for all  $r \in R$ . By the above theory  $N_1 + N_2$  is a submodule of  $M$ . Consider  $u, u' \in N_1 \cap N_2$ . Then  $u, u' \in N_1$  and so  $u + u' \in N_1$ . Also  $u, u' \in N_2$  and so  $u + u' \in N_2$ . Therefore  $u + u' \in N_1 \cap N_2$  and so  $N_1 \cap N_2$  is closed under addition. For  $r \in R$  we have  $ru \in N_1$  as  $u \in N_1$ . Also  $ru \in N_2$  as  $u \in N_2$ . So  $ru \in N_1 \cap N_2$ . So  $N_1 \cap N_2$  is a submodule of  $M$ .

(f) By Exercises 2.2, Question 4(f) we know  $M_1 \oplus M_2$  is an additive abelian group. We next check that the  $R$ -module laws 5, 6 and 7 (before (2.19)) hold in  $M_1 \oplus M_2$  given that these laws hold in  $M_1$  and  $M_2$ . Consider  $v = (v_1, v_2) \in M_1 \oplus M_2$ ,  $v' = (v'_1, v'_2) \in M_1 \oplus M_2$  and  $r, r' \in R$ . Then

$$\begin{aligned} r(v + v') &= r((v_1, v_2) + (v'_1, v'_2)) = r(v_1 + v'_1, v_2 + v'_2) = (r(v_1 + v'_1), r(v_2 + v'_2)) = \\ &= (rv_1 + rv'_1, rv_2 + rv'_2) = (rv_1, rv_2) + (rv'_1, rv'_2) = r(v_1, v_2) + r(v'_1, v'_2) = rv + rv' \end{aligned}$$

and

$$\begin{aligned} (r + r')v &= (r + r')(v_1, v_2) = ((r + r')v_1, (r + r')v_2) = (rv_1 + r'v_1, rv_2 + r'v_2) = \\ &= r(v_1, v_2) + r'(v_1, v_2) = rv + r'v \end{aligned}$$

which shows that law 5 holds. Also

$(rr')v = (rr')(v_1, v_2) = ((rr')v_1, (rr')v_2) = (r(r'v_1), r(r'v_2)) = r(r'v_1, r'v_2) = r(r'(v_1, v_2)) = r(r'v)$  showing that law 6 holds. As  $1v = 1(v_1, v_2) = (1v_1, 1v_2) = (v_1, v_2) = v$  we see that law 7 holds and so  $M_1 \oplus M_2$  is an  $R$ -module.

Consider  $u, u' \in N_1$  and let  $r \in R$ . Then  $u = (v_1, 0), u' = (v'_1, 0)$  where  $v_1, v'_1 \in M_1$  and 0 is the 0-element of  $M_2$ . Then  $v_1 + v'_1 \in M_1$  and  $rv_1 \in M_1$  as  $M_1$  is an  $R$ -module. So

$u + u' = (v_1, 0) + (v'_1, 0) = (v_1 + v'_1, 0) \in N_1$  and  $ru = (rv_1, r0) = (rv_1, 0) \in N_1$ . Therefore by (e) above  $N_1$  is a submodule of  $M_1 \oplus M_2$ . In the same way  $N_2$  is a submodule of  $M_1 \oplus M_2$ .

We know  $N_1 + N_2$  to be a submodule of  $M_1 \oplus M_2$  by (e) above, and so  $N_1 + N_2 \subseteq M_1 \oplus M_2$ . On the other hand consider  $v \in M_1 \oplus M_2$ . Then  $v = (v_1, v_2) = (v_1, 0) + (0, v_2) \in N_1 + N_2$  showing

$M_1 \oplus M_2 \subseteq N_1 + N_2$ . So  $N_1 + N_2 = M_1 \oplus M_2$ .

As  $(0, 0) \in N_1$  and  $(0, 0) \in N_2$  we see  $\{(0, 0)\} \subseteq N_1 \cap N_2$ . Consider  $(v_1, v_2) \in N_1 \cap N_2$ . As  $(v_1, v_2) \in N_1$  we see  $v_2 = 0$ . As  $(v_1, v_2) \in N_2$  we see  $v_1 = 0$ . Therefore  $(v_1, v_2) = (0, 0)$  showing  $N_1 \cap N_2 \subseteq \{(0, 0)\}$ . So  $N_1 \cap N_2 = \{(0, 0)\}$ .

Consider  $\alpha_1 : N_1 \rightarrow M_1$  given by  $(v_1, 0)\alpha_1 = v_1$  for all  $v_1 \in M_1$ . Then  $\alpha_1$  is (clearly!) bijective and additive as  $((v_1, 0) + (v'_1, 0))\alpha_1 = (v_1 + v'_1, 0)\alpha_1 = v_1 + v'_1 = (v_1, 0)\alpha_1 + (v'_1, 0)\alpha_1$  for  $v_1, v'_1 \in M_1$ . Also  $\alpha_1$  is  $R$ -linear as  $(r(v_1, 0))\alpha_1 = (rv_1, 0)\alpha_1 = rv_1 = r((v_1, 0)\alpha_1)$  for all  $r \in R$  and  $v_1 \in M_1$ . So  $\alpha_1 : N_1 \cong M_1$  is an  $R$ -module isomorphism. In the same way  $\alpha_2 : N_2 \rightarrow M_2$  given by  $(0, v_2)\alpha_2 = v_2$  for all  $v_2 \in M_2$  is an  $R$ -module isomorphism  $\alpha_2 : N_2 \cong M_2$ .

They are equal, i.e.  $N_1 \oplus N_2 = M_1 \oplus M_2$ .

### Solutions 3.1 (page 113)

#### Solution 1

(a) The sequence  $c_1 - 2c_2, c_1 \leftrightarrow c_2, r_2 - r_1$  of elementary operations reduces  $A$  to its Smith normal form  $D = \text{diag}(2, 6)$ ; note that there are different sequences which have the same effect on  $A$ . Applying  $r_2 - r_1$  to  $I$  gives

$$P = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}$$

and applying the sequence  $r_2 + 2r_1, r_1 \leftrightarrow r_2$  of *eros* to  $I$  gives

$$Q = \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix}. \text{ Then } PA = \begin{pmatrix} 4 & 2 \\ 6 & 0 \end{pmatrix} = DQ.$$

The rows of the diagonal matrix  $D$  are non-zero and so are  $\mathbb{Z}$ -independent; hence the rows  $z_1, z_2$  of

$A = P^{-1}DQ$  are  $\mathbb{Z}$ -independent. By hypothesis  $z_1, z_2$  generate  $K = \ker \theta$  and so  $z_1, z_2$  form a  $\mathbb{Z}$ -basis of  $K$ . By (3.4) the rows  $2\rho_1 = (4, 2), 6\rho_2 = (6, 0)$  of  $DQ$  form a  $\mathbb{Z}$ -basis of  $K$ .

Rank  $K = 2$ .  $(\rho_1)\theta = (2e_1 + e_2)\theta = 2(e_1)\theta + (e_2)\theta = 2g_1 + g_2$  and  $(\rho_2)\theta = (e_1)\theta = g_1$ . Yes,  $G = \langle (\rho_1)\theta \rangle \oplus \langle (\rho_2)\theta \rangle = \langle 2g_1 + g_2 \rangle \oplus \langle g_1 \rangle$ . The orders of  $(\rho_1)\theta$  and  $(\rho_2)\theta$  are 2 and 6 respectively. The sequence of invariant factors of  $G$  is (2, 6) and  $G$  has isomorphism type  $C_2 \oplus C_6$ .

(b) The sequence  $c_2 - c_1, c_1 - c_2, c_2 - 2c_1, r_2 - 5r_1, -r_2$  of elementary operations reduces  $A$  to its Smith normal form  $D = \text{diag}(1, 8)$ . Applying the sequence of *eros*  $r_2 - 5r_1, -r_2$  to  $I$  gives  $P = \begin{pmatrix} 1 & 0 \\ 5 & -1 \end{pmatrix}$ .

Applying the conjugates of the above *ecos*, i.e.  $r_1 + r_2, r_2 + r_1, r_1 + 2r_2$ , to  $I$  gives  $Q = \begin{pmatrix} 3 & 5 \\ 1 & 2 \end{pmatrix}$ . Then

$PA = \begin{pmatrix} 3 & 5 \\ 8 & 16 \end{pmatrix} = DQ$ . As in (a) above the rows  $z_1, z_2$  of  $A$  form a  $\mathbb{Z}$ -basis of  $K$  as do the rows

$\rho_1, 8\rho_2$  of  $DQ$ . So rank  $K = 2$ .  $(\rho_1)\theta = (3, 5)\theta = 3g_1 + 5g_2$ ,  $(\rho_2)\theta = (1, 2)\theta = g_1 + 2g_2$ . Yes,  $G = \langle (\rho_1)\theta \rangle \oplus \langle (\rho_2)\theta \rangle = \langle 3g_1 + 5g_2 \rangle \oplus \langle g_1 + 2g_2 \rangle$ . The orders of  $(\rho_1)\theta$  and  $(\rho_2)\theta$  are 1 and 8 respectively. So  $G = \langle g_1 + 2g_2 \rangle$  is cyclic with invariant factor sequence (8) and isomorphism type  $C_8$ .

(c) The sequence  $c_2 - c_1, c_1 - 8c_2, c_1 \leftrightarrow c_2, -c_2, c_1 - c_2$  of *ecos* reduces  $A = \begin{pmatrix} 8 & 9 \\ 7 & 8 \end{pmatrix}$  to its Smith normal form  $D = I = \text{diag}(1, 1)$ . In this case  $P = I$  and  $Q = A$  is invertible over  $\mathbb{Z}$ . Also  $PA = A = Q = DQ$ . Yes, the rows  $z_1, z_2$  of  $A$  form a  $\mathbb{Z}$ -basis of  $K$  as do the rows  $\rho_1, \rho_2$  of  $DQ$ . Rank  $K = 2$  and  $K = \mathbb{Z}^2$ .  $(\rho_1)\theta = (8, 9)\theta = 8g_1 + 9g_2 = 0$ ,  $(\rho_2)\theta = (7, 8)\theta = 7g_1 + 8g_2 = 0$ . Yes,  $G = \langle (\rho_1)\theta \rangle \oplus \langle (\rho_2)\theta \rangle = \{0\}$  and so  $G$  is trivial. The orders of  $(\rho_1)\theta$  and  $(\rho_2)\theta$  are both 1 and  $G$  has invariant factor sequence  $\emptyset$  and isomorphism type  $C_1$ .

(d) The sequence  $c_2 - 2c_1, r_2 - 2r_1$  reduces  $A = \begin{pmatrix} 2 & 4 \\ 4 & 8 \end{pmatrix}$  to its Smith normal form  $D = \text{diag}(2, 0)$ .

$P = \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix}$  and  $Q = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$  are invertible over  $\mathbb{Z}$  and satisfy  $PA = \begin{pmatrix} 2 & 4 \\ 0 & 0 \end{pmatrix} = DQ$ . As  $2z_1 - z_2 = 0$

the rows  $z_1, z_2$  of  $A$  are  $\mathbb{Z}$ -dependent and so do not form a  $\mathbb{Z}$ -basis of  $K$ . The first row  $(2, 4)$  of  $DQ$  is a  $\mathbb{Z}$ -basis of  $K$  and  $\text{rank } K = 1$ .  $(\rho_1)\theta = (1, 2)\theta = g_1 + 2g_2$ ,  $(\rho_2)\theta = (0, 1)\theta = g_2$ . Yes  $G = \langle (\rho_1)\theta \rangle \oplus \langle (\rho_2)\theta \rangle = \langle g_1 + 2g_2 \rangle \oplus \langle g_2 \rangle$ . The order of  $(\rho_1)\theta$  is 2 and  $(\rho_2)\theta$  has infinite order. The invariant factor sequence of  $G$  is  $(2, 0)$  and  $G$  has isomorphism type  $C_2 \oplus C_0$ . The torsion-free rank of  $G$  is 1 and the order of its torsion submodule is 2.

(e) The sequence  $c_2 - c_1, c_1 - 3c_2, c_1 \leftrightarrow c_2, r_2 - 2r_1$  reduces  $A = \begin{pmatrix} 3 & 4 \\ 6 & 8 \end{pmatrix}$  to its Smith normal form

$$D = \text{diag}(1, 0). \quad P = \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix} \text{ and } Q = \begin{pmatrix} 3 & 4 \\ 1 & 1 \end{pmatrix} \text{ are invertible over } \mathbb{Z} \text{ and satisfy } PA = \begin{pmatrix} 3 & 4 \\ 0 & 0 \end{pmatrix} = DQ.$$

The rows of  $A$  are  $\mathbb{Z}$ -dependent and so do not form a  $\mathbb{Z}$ -basis of  $K$ . The first row  $(3, 4)$  of  $DQ$  is a  $\mathbb{Z}$ -basis of  $K$  and  $\text{rank } K = 1$ .  $(\rho_1)\theta = (3, 4)\theta = 3g_1 + 4g_2 = 0$ ,  $(\rho_2)\theta = (1, 1)\theta = g_1 + g_2$ . Yes  $G = \langle (\rho_1)\theta \rangle \oplus \langle (\rho_2)\theta \rangle = \langle 0 \rangle \oplus \langle g_1 + g_2 \rangle = \langle g_1 + g_2 \rangle$ . The order of  $(\rho_1)\theta$  is 1 and  $(\rho_2)\theta$  has infinite order. So  $G$  is infinite cyclic with invariant factor sequence  $(0)$  and isomorphism type  $C_0$ .

(f) The sequence  $c_2 - 2c_1, c_1 - c_2, c_2 - c_1, c_1 - 3c_2, c_1 \leftrightarrow c_2, r_2 + 2r_1, r_3 + 58r_1, r_3 - 21r_2$  reduces  $A$  to its

$$\text{Smith normal form } D = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 20 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \text{ Also } P = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 16 & -21 & 1 \end{pmatrix} \text{ and } Q = \begin{pmatrix} 7 & 18 \\ 2 & 5 \end{pmatrix} \text{ are invertible over } \mathbb{Z}$$

$$\text{such that } PA = \begin{pmatrix} 14 & 36 \\ 40 & 100 \\ 0 & 0 \end{pmatrix} = DQ. \text{ By (3.1) the 3 rows of } A \text{ cannot form a } \mathbb{Z}\text{-basis of } K \text{ as}$$

$\text{rank } K \leq 2$ . The non-zero rows of  $PA$  form a  $\mathbb{Z}$ -basis of  $K$ , i.e.  $K = \langle 2\rho_1, 20\rho_2 \rangle$  where  $\rho_1 = (7, 18)$ ,  $\rho_2 = (2, 5)$ . So  $\text{rank } K = 2$ .  $(\rho_1)\theta = 7g_1 + 18g_2$ ,  $(\rho_2)\theta = 2g_1 + 5g_2$ . Yes  $G = \langle (\rho_1)\theta \rangle \oplus \langle (\rho_2)\theta \rangle = \langle 7g_1 + 18g_2 \rangle \oplus \langle 2g_1 + 5g_2 \rangle$ . The orders of  $(\rho_1)\theta$  and  $(\rho_2)\theta$  are 2 and 20 respectively. The invariant factor sequence of  $G$  is  $(2, 20)$  and  $G$  has isomorphism type  $C_2 \oplus C_{20}$ .

(g) By (1.10) the  $2 \times 2$  matrix  $A = \text{diag}(n_1, n_2)$  has Smith normal form

$$D = \text{diag}(\gcd\{n_1, n_2\}, \text{lcm}\{n_1, n_2\}).$$

For  $\gcd\{n_1, n_2\} > 1$  the invariant factor sequence of  $G$  is  $(\gcd\{n_1, n_2\}, \text{lcm}\{n_1, n_2\})$  and  $G$  is not cyclic. For  $\gcd\{n_1, n_2\} = 1$  the invariant factor sequence of  $G$  is  $(n_1 n_2)$  or  $\emptyset$  according as  $n_1 n_2 > 1$  or  $n_1 = n_2 = 1$  and  $G$  is cyclic.

## Solution 2

(a) The sequence  $c_2 - 2c_1, c_3 - 2c_1, r_2 - 2r_1, r_3 - r_1, r_2 + r_3, c_2 - c_3$ , for example, reduces  $A$  to its Smith normal form  $D = \text{diag}(2, 2, 2)$ . Applying the *eros* in the above sequence to  $I$  gives

$$P = \begin{pmatrix} 1 & 0 & 0 \\ -3 & 1 & 1 \\ -1 & 0 & 1 \end{pmatrix}. \text{ Further } Q = \begin{pmatrix} 1 & 2 & 2 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \text{ results on applying the conjugates of the } \textit{ecos} \text{ in the above}$$

sequence to  $I$ . Then  $P$  and  $Q$  are invertible over  $\mathbb{Z}$  and

$$PA = \begin{pmatrix} 2 & 4 & 4 \\ 0 & 2 & 0 \\ 0 & 2 & 2 \end{pmatrix} = DQ.$$

Let  $\rho_1 = (1, 2, 2)$ ,  $\rho_2 = (0, 1, 0)$ ,  $\rho_3 = (0, 1, 1)$ . Then the rows  $\rho_1, \rho_2, \rho_3$  of  $Q$  constitute a  $\mathbb{Z}$ -basis of  $\mathbb{Z}^3$ . The rows  $2\rho_1, 2\rho_2, 2\rho_3$  of  $DQ$  form a  $\mathbb{Z}$ -basis of  $K = \ker \theta$ . By (3.4)

$$G = \langle (\rho_1)\theta \rangle \oplus \langle (\rho_2)\theta \rangle \oplus \langle (\rho_3)\theta \rangle = \langle g_1 + 2g_2 + 2g_3 \rangle \oplus \langle g_2 \rangle \oplus \langle g_2 + g_3 \rangle$$

has invariant factor sequence  $(2, 2, 2)$  and isomorphism type  $C_2 \oplus C_2 \oplus C_2$  as  $(\rho_i)\theta$  has order 2 ( $i=1, 2, 3$ ).

(b) The sequence  $c_1 - c_2, c_2 - 2c_1, c_3 - 5c_1, r_2 + 2r_1, r_3 + r_1, r_2 - r_3, c_3 - 3c_2, c_2 + c_3, -c_3$  reduces  $A$  to its Smith normal form  $D = \text{diag}(1, 2, 6)$  and gives rise to

$$P = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & -1 \\ 1 & 0 & 1 \end{pmatrix} \text{ and } Q = \begin{pmatrix} 3 & 2 & 5 \\ 1 & 1 & 3 \\ 1 & 1 & 2 \end{pmatrix}, \text{ both invertible over } \mathbb{Z}, \text{ satisfying } PA = \begin{pmatrix} 3 & 2 & 5 \\ 2 & 2 & 6 \\ 6 & 6 & 12 \end{pmatrix} = DQ.$$

Let  $\rho_1 = (3, 2, 5)$ ,  $\rho_2 = (1, 1, 3)$ ,  $\rho_3 = (1, 1, 2)$ . The rows  $\rho_1, \rho_2, \rho_3$  of  $Q$  form a  $\mathbb{Z}$ -basis of  $\mathbb{Z}^3$ . The rows  $\rho_1, 2\rho_2, 6\rho_3$  of  $DQ$  form a  $\mathbb{Z}$ -basis of  $K = \ker \theta$ . So  $(\rho_1)\theta = 0$  and the orders of  $(\rho_2)\theta$ ,  $(\rho_3)\theta$  are 2, 6 respectively. As  $G = \langle (\rho_1)\theta \rangle \oplus \langle (\rho_2)\theta \rangle \oplus \langle (\rho_3)\theta \rangle = \langle g_1 + g_2 + 3g_3 \rangle \oplus \langle g_1 + g_2 + 2g_3 \rangle$  we see that  $G$  has invariant factor sequence  $(2, 6)$  and isomorphism type  $C_2 \oplus C_6$ .

Yes, these  $\mathbb{Z}$ -modules are isomorphic as both have invariant factor sequence  $(2, 6)$ .

(c) The sequence  $c_2 - 2c_1, c_3 - c_1, r_2 - 2r_1, r_3 - r_1, r_3 + r_2, -r_2, c_3 - c_2$  reduces

$$A = \begin{pmatrix} 2 & 4 & 2 \\ 4 & 6 & 2 \\ 2 & 6 & 4 \end{pmatrix} \text{ to } D = \text{diag}(2, 2, 0) \text{ and gives rise to } P = \begin{pmatrix} 1 & 0 & 0 \\ 2 & -1 & 0 \\ -3 & 1 & 1 \end{pmatrix} \text{ and } Q = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \text{ satisfying}$$

$$PA = \begin{pmatrix} 2 & 4 & 2 \\ 0 & 2 & 2 \\ 0 & 0 & 0 \end{pmatrix} = DQ.$$

Write  $\rho_1 = (1, 2, 1)$ ,  $\rho_2 = (0, 1, 1)$ ,  $\rho_3 = (0, 0, 1)$ . Then the rows  $\rho_1, \rho_2, \rho_3$  of  $Q$  form a  $\mathbb{Z}$ -basis of  $\mathbb{Z}^3$ . Also  $2\rho_1, 2\rho_2$ , i.e. the non-zero rows of  $DQ$ , form a  $\mathbb{Z}$ -basis of  $K = \ker \theta$  and so  $\text{rank } K = 2$ . Then  $(\rho_1)\theta$ ,  $(\rho_2)\theta$  both have order 2 and  $(\rho_3)\theta$  has infinite order. So

$$G = \langle (\rho_1)\theta \rangle \oplus \langle (\rho_2)\theta \rangle \oplus \langle (\rho_3)\theta \rangle = \langle g_1 + 2g_2 + g_3 \rangle \oplus \langle g_2 + g_3 \rangle \oplus \langle g_3 \rangle$$

has invariant factor sequence  $(2, 2, 0)$  and isomorphism type  $C_2 \oplus C_2 \oplus C_0$ . The torsion submodule

$T = \langle (\rho_1)\theta, (\rho_2)\theta \rangle$  has invariant factor sequence  $(2, 2)$  and so  $T$  is a Klein 4-group. The torsion-free rank of  $G$  is 1, i.e.  $G/T = \langle T + (\rho_3)\theta \rangle$  is free of rank 1 and so is infinite cyclic.

(d) The sequence of elementary operations  $c_2 - c_1, c_1 - 2c_2, c_3 - 4c_2, c_2 - 2c_3, c_1 \leftrightarrow c_3$  reduces

$A = (8, 12, 18)$  to its Smith normal form  $D = (2, 0, 0)$  and produces the invertible matrix

$$Q = \begin{pmatrix} 4 & 6 & 9 \\ 2 & 3 & 4 \\ 1 & 1 & 0 \end{pmatrix}$$

over  $\mathbb{Z}$  satisfying  $A=DQ$  (in this case  $P$  is the  $1 \times 1$  identity matrix and is omitted). Let  $\rho_i$  denote row  $i$  of  $Q$  ( $i=1,2,3$ ). Then  $\mathbb{Z}^3 = \langle \rho_1, \rho_2, \rho_3 \rangle$  and  $K = \langle 2\rho_1 \rangle$ . So  $\text{rank } K = 1$ . Also  $(\rho_1)\theta$  has order 2 and both  $(\rho_2)\theta, (\rho_3)\theta$  have infinite order. Therefore

$$G = \langle (\rho_1)\theta \rangle \oplus \langle (\rho_2)\theta \rangle \oplus \langle (\rho_3)\theta \rangle = \langle 4g_1 + 6g_2 + 9g_3 \rangle \oplus \langle 2g_1 + 3g_2 + 4g_3 \rangle \oplus \langle g_1 + g_2 \rangle$$

has invariant factor sequence  $(2, 0, 0)$  and isomorphism type  $C_2 \oplus C_0 \oplus C_0$ . Also  $T$  has invariant factor sequence  $(2)$ . Also  $G$  has torsion-free rank 2 and  $G/T$  has invariant factor sequence  $(0, 0)$ .

### Solution 3

(a) Consider  $(m_1, m_2)$  and  $(m'_1, m'_2)$  in  $K$ . Then  $m_1 \equiv m_2 \pmod{2}$  and  $m'_1 \equiv m'_2 \pmod{2}$ . So

$$(m_1, m_2) + (m'_1, m'_2) = (m_1 + m'_1, m_2 + m'_2) \in K \text{ as } m_1 + m'_1 \equiv m_2 + m'_2 \pmod{2}, \text{ i.e.}$$

parity  $(m_1 + m'_1) = \text{parity}(m_2 + m'_2)$ , showing that  $K$  is closed under addition. Similarly

$$-(m_1, m_2) = (-m_1, -m_2) \in K. \text{ As } (0, 0) \in K \text{ we see that } K \text{ is a subgroup of the additive group } \mathbb{Z}^2,$$

i.e.  $K$  is a submodule of  $\mathbb{Z}^2$ . Clearly  $z_1 = (1, 1), z_2 = (0, 2)$  belong to  $K$  and are  $\mathbb{Z}$ -independent. As

$$m_2 = m_1 + 2q \text{ for } q \in \mathbb{Z}, \text{ we have } (m_1, m_2) = m_1(1, 1) + q(0, 2) \text{ showing that } z_1, z_2 \text{ generate } K. \text{ So}$$

$z_1, z_2$  form a  $\mathbb{Z}$ -basis of  $K$ . The single *eco*  $c_2 - c_1$  reduces  $A$  to its Smith normal form

$$D = \text{diag}(1, 2). \text{ As } (1, 1), (0, 1) \text{ is a } \mathbb{Z}\text{-basis of } \mathbb{Z}^2 \text{ and } (1, 1), (0, 2) \text{ is a } \mathbb{Z}\text{-basis of } K, \text{ we see that}$$

$$\mathbb{Z}^2/K \text{ has isomorphism type } C_1 \oplus C_2 = C_2.$$

(b) Consider  $z_1 = (1, 1, 1), z_2 = (0, 2, 0), z_3 = (0, 0, 2) \in K$ . Then  $z_1, z_2, z_3$  are  $\mathbb{Z}$ -independent as

$$\det A \neq 0. \text{ Let } (m_1, m_2, m_3) \in K. \text{ There are } q_2, q_3 \in \mathbb{Z} \text{ with } m_j = m_1 + 2q_j \text{ for } j = 2, 3. \text{ Hence}$$

$$(m_1, m_2, m_3) = m_1(1, 1, 1) + q_2(0, 2, 0) + q_3(0, 0, 2) \text{ showing that } z_1, z_2, z_3 \text{ generate } K. \text{ So } z_1, z_2, z_3 \text{ form a}$$

$\mathbb{Z}$ -basis of  $K$ . The *ecos*  $c_2 - c_1, c_3 - c_1$  reduce  $A$  to  $D = \text{diag}(1, 2, 2)$ . As  $(1, 1, 1), (0, 1, 0), (0, 0, 1)$  is a

$\mathbb{Z}$ -basis of  $\mathbb{Z}^3$  and  $(1, 1, 1), (0, 2, 0), (0, 0, 2)$  is a  $\mathbb{Z}$ -basis of  $K$ , we see that  $\mathbb{Z}^3/K$  has isomorphism

$$\text{type } C_1 \oplus C_2 \oplus C_2 = C_2 \oplus C_2.$$

(c) Consider  $z_1 = (1, 1, \dots, 1), z_2 = ne_2, z_3 = ne_3, \dots, z_t = ne_t \in K$ . Then  $z_1, z_2, \dots, z_t$  are  $\mathbb{Z}$ -independent

as  $\det A \neq 0$ . Let  $(m_1, m_2, \dots, m_t) \in K$ . There are  $q_j \in \mathbb{Z}$  with  $m_j = m_1 + nq_j$  for  $j = 2, 3, \dots, t$ . Hence

$$(m_1, m_2, \dots, m_t) = m_1(1, 1, \dots, 1) + nq_2e_2 + \dots + nq_te_t \text{ showing that } z_1, z_2, \dots, z_t \text{ generate } K. \text{ So}$$

$z_1, z_2, \dots, z_t$  form a  $\mathbb{Z}$ -basis of  $K$ . The *ecos*  $c_j - c_1$  ( $2 \leq j \leq t$ ) reduce  $A$  to  $D = \text{diag}(1, n, \dots, n)$ . As

$(1, 1, \dots, 1), e_2, \dots, e_t$  is a  $\mathbb{Z}$ -basis of  $\mathbb{Z}^t$  and  $(1, 1, \dots, 1), ne_2, \dots, ne_t$  is a  $\mathbb{Z}$ -basis of  $K$ , we see that

$\mathbb{Z}^t/K$  has isomorphism type

$$C_1 \oplus C_n \oplus \dots \oplus C_n = C_n \oplus C_n \oplus \dots \oplus C_n \text{ (} t-1 \text{ summands } C_n \text{)}.$$

### Solution 4

(a)  $24g = 24(\bar{1}, \bar{1}) = (3 \times \bar{8}, 2 \times \bar{12}) = (3 \times \bar{0}, 2 \times \bar{0}) = (\bar{0}, \bar{0}) = 0$ . So  $m \mid 24$  where  $m$  is the order of  $g$ .

Also  $mg = 0$ , i.e.  $(\bar{m}, \bar{m}) = (\bar{0}, \bar{0}) \in \mathbb{Z}_8 \oplus \mathbb{Z}_{12}$  and so  $8 \mid m$  and  $12 \mid m$ . Hence  $\text{lcm}\{8, 12\} \mid m$ , i.e.  $24 \mid m$ .

Hence  $m = 24$ . Since  $\bar{2}$  in  $\mathbb{Z}_8$  has order 4, we see, as above, that  $(\bar{2}, \bar{1})$  in  $\mathbb{Z}_8 \oplus \mathbb{Z}_{12}$  has order

$$\text{lcm}\{4, 12\} = 12. \text{ Similarly } (\bar{1}, \bar{2}) \text{ in } \mathbb{Z}_8 \oplus \mathbb{Z}_{12} \text{ has order } \text{lcm}\{8, 6\} = 24.$$

(b) Write  $m = \text{lcm}\{n_1, n_2\} = n_1 n_2 / d$  where  $d = \text{gcd}\{n_1, n_2\}$ . Then

$mg = m(\bar{1}, \bar{1}) = (\bar{m}, \bar{m}) = ((n_2/d)\bar{n}_1, (n_1/d)\bar{n}_2) = ((n_2/d)\bar{0}, (n_1/d)\bar{0}) = (\bar{0}, \bar{0}) = 0$  showing that  $g$  has finite order  $m'$  where  $m' \mid m$ . From  $m'g = 0$ , i.e.  $(\bar{m}', \bar{m}') = (\bar{0}, \bar{0})$  in  $\mathbb{Z}_{n_1} \oplus \mathbb{Z}_{n_2}$  we deduce  $n_1 \mid m'$  and  $n_2 \mid m'$ . So  $m'$  is a common multiple of  $n_1, n_2$  and so  $m \mid m'$  by Exercises 1.3, Question 1(c). Hence  $m' = m$ .

For each  $i$  with  $1 \leq i \leq t$  we have  $mh_i = 0$  as  $n_i \mid m$ . So

$$mg = mh_1 + mh_2 + \dots + mh_t = 0 + 0 + \dots + 0 = 0$$

showing that  $g$  has finite order  $m'$  where  $m' \mid m$ . From  $m'g = 0$ , i.e.  $m'h_1 + m'h_2 + \dots + m'h_t = 0$  and the independence (2.14) of the submodules  $H_i$  we deduce  $m'h_i = 0$  for  $1 \leq i \leq t$ . So  $n_i \mid m'$  for  $1 \leq i \leq t$  showing that  $m'$  is a common multiple of  $n_1, n_2, \dots, n_t$ . From Exercises 1.3, Question 1(c) we see  $m \mid m'$  and so  $m = m'$  as  $m, m'$  are positive.

(c) Suppose each  $n_i > 0$ . By (2.7) the element  $m_i h_i$  of  $H_i$  has order  $n_i / \text{gcd}\{m_i, n_i\}$  for  $1 \leq i \leq t$ .

By (b) above  $g = m_1 h_1 + m_2 h_2 + \dots + m_t h_t$  has order

$$\text{lcm}\{n_1 / \text{gcd}\{m_1, n_1\}, n_2 / \text{gcd}\{m_2, n_2\}, \dots, n_t / \text{gcd}\{m_t, n_t\}\}.$$

Suppose now that  $m_i \neq 0$  and  $n_i = 0$  for some  $i$  with  $1 \leq i \leq t$ . Then  $nm_i h_i \neq 0$  for all positive integers  $n$ . From  $ng = nm_1 h_1 + nm_2 h_2 + \dots + nm_t h_t \in H_1 \oplus H_2 \oplus \dots \oplus H_t$  we deduce  $ng \neq 0$ . So  $g$  has infinite order.

(d)

$$Q^{-1} = \begin{pmatrix} -5 & 18 \\ 2 & -7 \end{pmatrix} \text{ and so } \begin{aligned} g_1 &= -5h_1 + 18h_2 \\ g_2 &= 2h_1 - 7h_2 \end{aligned}$$

where  $h_1 = (\rho_1)\theta$  has order 2,  $h_2 = (\rho_2)\theta$  has order 20 and  $G = \langle h_1 \rangle \oplus \langle h_2 \rangle$ . Using (c) above,  $g_1$  has order  $\text{lcm}\{2 / \text{gcd}\{-5, 2\}, 20 / \text{gcd}\{18, 20\}\} = \text{lcm}\{2, 10\} = 10$ . In the same way,  $g_2$  has order

$$\text{lcm}\{2 / \text{gcd}\{2, 2\}, 20 / \text{gcd}\{-7, 20\}\} = \text{lcm}\{1, 20\} = 20.$$

(e) Applying the *ecos*  $c_1 - c_2, c_2 - 2c_1, c_3 - 5c_1, c_3 - 3c_2, c_2 + c_3, -c_3$  to the  $3 \times 3$  identity matrix  $I$  gives

$$Q^{-1} = \begin{pmatrix} 1 & -1 & -1 \\ -1 & -1 & 4 \\ 0 & 1 & -1 \end{pmatrix}. \text{ So } \begin{aligned} g_1 &= h_1 - h_2 - h_3 \\ g_2 &= -h_1 - h_2 + 4h_3 \\ g_3 &= h_2 - h_3 \end{aligned}$$

where  $h_1 = (\rho_1)\theta$  has order 1,  $h_2 = (\rho_2)\theta$  has order 2,  $h_3 = (\rho_3)\theta$  has order 6, and

$G = \langle h_1 \rangle \oplus \langle h_2 \rangle \oplus \langle h_3 \rangle$ . Using (b) above,

$$g_1 \text{ has order } \text{lcm}\{1 / \text{gcd}\{1, 1\}, 2 / \text{gcd}\{-1, 2\}, 6 / \text{gcd}\{-1, 6\}\} = \text{lcm}\{1, 2, 6\} = 6,$$

$$g_2 \text{ has order } \text{lcm}\{1 / \text{gcd}\{-1, 1\}, 2 / \text{gcd}\{-1, 2\}, 6 / \text{gcd}\{4, 6\}\} = \text{lcm}\{1, 2, 3\} = 6,$$

$$g_3 \text{ has order } \text{lcm}\{1 / \text{gcd}\{0, 1\}, 2 / \text{gcd}\{1, 2\}, 6 / \text{gcd}\{-1, 6\}\} = \text{lcm}\{1, 2, 6\} = 6.$$

### Solution 5

(a) Let  $\rho_i$  denote row  $i$  of  $Q$  ( $1 \leq i \leq t$ ) and write  $D = \text{diag}(d_1, d_2, \dots, d_{\min\{s, t\}})$ . Let

$r = \max\{0, i : d_i = 1\}$  and  $r + s' = \max\{0, i : d_i > 0\}$ . Then  $r + s' \leq \min\{s, t\}$  and

$d_1 \rho_1, d_2 \rho_2, \dots, d_{r+s'} \rho_{r+s'}$  are the non-zero rows of  $DQ$ . Write  $K' = \langle d_1 \rho_1, d_2 \rho_2, \dots, d_{r+s'} \rho_{r+s'} \rangle$ . As

the rows  $\rho_1, \rho_2, \dots, \rho_t$  of  $Q$  are  $\mathbb{Z}$ -independent we see that  $d_1\rho_1, d_2\rho_2, \dots, d_{r+s'}\rho_{r+s'}$  are also  $\mathbb{Z}$ -independent. Equating row  $i$  of  $PA = DQ$  gives  $p_{i1}z_1 + p_{i2}z_2 + \dots + p_{is}z_s = d_i\rho_i$  for  $1 \leq i \leq s$  where  $P = (p_{ij})$ , showing that  $d_1\rho_1, d_2\rho_2, \dots, d_{r+s'}\rho_{r+s'}$  belong to  $K = \ker \theta = \langle z_1, z_2, \dots, z_s \rangle$ . So  $K' \subseteq K$ . Equating row  $i$  of  $A = P^{-1}DQ$  gives  $z_i = p'_{i1}d_1\rho_1 + p'_{i2}d_2\rho_2 + \dots + p'_{ir+s'}d_{r+s'}\rho_{r+s'}$  for  $1 \leq i \leq s$  where  $P^{-1} = (p'_{ij})$ , showing  $K \subseteq K'$ . Hence  $K' = K$  and so  $d_1\rho_1, d_2\rho_2, \dots, d_{r+s'}\rho_{r+s'}$  generate  $K$ . Finally we see that  $\text{rank } K = r + s'$  and the non-zero rows of  $DQ$  form a  $\mathbb{Z}$ -basis of  $K$ .

(b) As  $G$  is generated by  $t$  say of its elements there is a surjective  $\mathbb{Z}$ -linear mapping  $\theta: \mathbb{Z}^t \rightarrow G$ . Consider  $k_1, k_2 \in K'$  and so  $(k_1)\theta = h_1 \in H$  and  $(k_2)\theta = h_2 \in H$ . Then  $(k_1 + k_2)\theta = (k_1)\theta + (k_2)\theta = h_1 + h_2 \in H$  showing  $k_1 + k_2 \in K'$ . Also  $(-k_1)\theta = -(k_1)\theta = -h_1 \in H$  showing  $-k_1 \in K'$ . As  $(0)\theta = 0 \in H$  we see  $0 \in K'$  and so  $K'$  is an additive subgroup of  $\mathbb{Z}^t$ , i.e.  $K'$  is a submodule of  $\mathbb{Z}^t$ . So  $K'$  is free with  $\mathbb{Z}$ -basis  $z_1, z_2, \dots, z_s$ ,  $s \leq t$  by (3.1). Let  $h \in H$ . There is  $k \in K'$  with  $(k)\theta = h$ . There are integers  $m_1, m_2, \dots, m_s$  with  $k = m_1z_1 + m_2z_2 + \dots + m_sz_s$ . As  $\theta$  is  $\mathbb{Z}$ -linear we obtain  $h = (k)\theta = (m_1z_1 + m_2z_2 + \dots + m_sz_s)\theta = m_1(z_1)\theta + m_2(z_2)\theta + \dots + m_s(z_s)\theta$  which shows that the  $s$  elements  $(z_1)\theta, (z_2)\theta, \dots, (z_s)\theta$  generate  $H$ . So  $H$  is finitely generated.

(c) Consider  $g = (g_1, g_2, \dots, g_s) \in G_1 \oplus G_2 \oplus \dots \oplus G_s$ . Then  $ng \in n(G_1 \oplus G_2 \oplus \dots \oplus G_s)$  and also  $ng = (ng_1, ng_2, \dots, ng_s) \in nG_1 \oplus nG_2 \oplus \dots \oplus nG_s$ . So the  $\mathbb{Z}$ -modules  $n(G_1 \oplus G_2 \oplus \dots \oplus G_s)$  and  $nG_1 \oplus nG_2 \oplus \dots \oplus nG_s$  are equal.

Consider  $g = (g_1, g_2, \dots, g_s) \in G_1 \oplus G_2 \oplus \dots \oplus G_s$ . Then  $g \in (G_1 \oplus G_2 \oplus \dots \oplus G_s)_{(n)} \Leftrightarrow ng = 0 \Leftrightarrow ng_i = 0$  for  $1 \leq i \leq s \Leftrightarrow g_i \in (G_i)_{(n)}$  for  $1 \leq i \leq s$ . So the  $\mathbb{Z}$ -modules  $(G_1 \oplus G_2 \oplus \dots \oplus G_s)_{(n)}$  and  $(G_1)_{(n)} \oplus (G_2)_{(n)} \oplus \dots \oplus (G_s)_{(n)}$  are equal.

(d) Consider  $m\bar{1}$  in  $(\mathbb{Z}_d)_{(n)}$ . Then  $mn\bar{1} = \bar{0}$ , the 0-element of  $\mathbb{Z}_d$ . As  $\bar{1}$  has order  $d$  in the additive group  $(\mathbb{Z}_d, +)$  we deduce  $d \mid mn$  from the discussion preceding (2.5). Therefore  $d/\gcd\{n, d\} \mid m(n/\gcd\{n, d\})$  and so  $d/\gcd\{n, d\} \mid m$  as  $d/\gcd\{n, d\}$  and  $n/\gcd\{n, d\}$  are coprime integers. So  $m\bar{1} = q(d/\gcd\{n, d\})\bar{1}$  where  $q \in \mathbb{Z}$  showing that  $(d/\gcd\{n, d\})\bar{1}$  generates the  $\mathbb{Z}$ -module  $(\mathbb{Z}_d)_{(n)}$ .

(e) As  $d_i(\mathbb{Z}_{d_j}) \cong \mathbb{Z}_{d_j/\gcd\{d_i, d_j\}}$  we see  $d_i(\mathbb{Z}_{d_j}) \cong \mathbb{Z}_1$  (trivial) for  $j \leq i$  and  $d_i(\mathbb{Z}_{d_j}) \cong \mathbb{Z}_{d_j/d_i}$  for  $j > i$ . Therefore  $d_i G \cong (\mathbb{Z}_1)^{m_1+m_2+\dots+m_i} \oplus (\mathbb{Z}_{d_{i+1}/d_i})^{m_{i+1}} \oplus \dots \oplus (\mathbb{Z}_{d_r/d_i})^{m_r}$  and so  $d_i G \cong (\mathbb{Z}_{d_{i+1}/d_i})^{m_{i+1}} \oplus \dots \oplus (\mathbb{Z}_{d_r/d_i})^{m_r}$ . As  $(\mathbb{Z}_{d_j})_{(d_i)} \cong \mathbb{Z}_{\gcd\{d_i, d_j\}}$  we see  $(\mathbb{Z}_{d_j})_{(d_i)} \cong \mathbb{Z}_{d_j}$  for  $j < i$  and  $(\mathbb{Z}_{d_j})_{(d_i)} \cong \mathbb{Z}_{d_i}$  for  $j \geq i$ . Therefore  $G_{(d_i)} \cong (\mathbb{Z}_{d_1})^{m_1} \oplus \dots \oplus (\mathbb{Z}_{d_{i-1}})^{m_{i-1}} \oplus (\mathbb{Z}_{d_i})^{m_i+m_{i+1}+\dots+m_r}$ . As  $(\mathbb{Z}_{d_j/d_i})_{(d_{i+1}/d_i)} \cong \mathbb{Z}_{\gcd\{d_{i+1}/d_i, d_j/d_i\}} = \mathbb{Z}_{d_{i+1}/d_i}$  for  $j > i$  we obtain  $(d_i G)_{(d_{i+1}/d_i)} \cong (\mathbb{Z}_{d_{i+1}/d_i})^{m_{i+1}+m_{i+2}+\dots+m_r}$ . So, yes  $(d_i G)_{(d_{i+1}/d_i)}$  is a free  $R$ -module with  $R = \mathbb{Z}_{d_{i+1}/d_i}$  of rank  $m_{i+1} + m_{i+2} + \dots + m_r$ .

## Solution 6



(a) Let  $g_1, g_2 \in G$ . Then

$$\begin{aligned} (\chi_1 + \chi_2)(g_1 + g_2) &= \chi_1(g_1 + g_2) + \chi_2(g_1 + g_2) = \chi_1(g_1) + \chi_1(g_2) + \chi_2(g_1) + \chi_2(g_2) = \\ &= \chi_1(g_1) + \chi_2(g_1) + \chi_1(g_2) + \chi_2(g_2) = (\chi_1 + \chi_2)(g_1) + (\chi_1 + \chi_2)(g_2) \end{aligned}$$

showing that  $\chi_1 + \chi_2$  is an additive mapping from  $G$  to  $\mathbb{Q}/\mathbb{Z}$ , i.e.  $\chi_1 + \chi_2$  is a character of  $G$ .

Therefore  $\chi_1, \chi_2 \in G^*$  implies  $\chi_1 + \chi_2 \in G^*$ , showing that  $G^*$  is closed under addition. We now verify that  $(G^*, +)$  obeys the laws of an abelian group. Consider  $\chi_1, \chi_2, \chi_3 \in G^*$ . Using the associative law of addition in  $\mathbb{Q}/\mathbb{Z}$  we see

$$\begin{aligned} ((\chi_1 + \chi_2) + \chi_3)(g) &= (\chi_1 + \chi_2)(g) + (\chi_3)(g) = ((\chi_1)(g) + (\chi_2)(g)) + (\chi_3)(g) = \\ &= (\chi_1)(g) + ((\chi_2)(g) + (\chi_3)(g)) = (\chi_1)(g) + (\chi_2 + \chi_3)(g) = (\chi_1 + (\chi_2 + \chi_3))(g) \end{aligned}$$

showing  $(\chi_1 + \chi_2) + \chi_3 = \chi_1 + (\chi_2 + \chi_3)$ , i.e.  $(G^*, +)$  obeys the associative law of addition. The mapping  $\chi_0: G \rightarrow \mathbb{Q}/\mathbb{Z}$  defined by  $\chi_0(g) = \mathbb{Z} + 0 = \mathbb{Z}$  for all  $g \in G$  is a character of  $G$ . We refer to  $\chi_0$  as the *zero character* of  $G$ . For  $\chi \in G^*$  we have  $\chi_0 + \chi = \chi$  since

$$(\chi_0 + \chi)(g) = \chi_0(g) + \chi(g) = \mathbb{Z} + \chi(g) = \chi(g) \in \mathbb{Q}/\mathbb{Z} \text{ for all } g \in G,$$

showing that  $\chi_0$  is the zero element of  $G^*$ . For each  $\chi \in G^*$  let  $-\chi: G \rightarrow \mathbb{Q}/\mathbb{Z}$  be given by

$$(-\chi)(g) = -\chi(g) \text{ for all } g \in G. \text{ Then}$$

$$(-\chi)(g_1 + g_2) = -\chi(g_1 + g_2) = -(\chi(g_1) + \chi(g_2)) = -\chi(g_1) - \chi(g_2) = (-\chi)(g_1) + (-\chi)(g_2)$$

which shows  $-\chi \in G^*$ . Now  $(-\chi + \chi)(g) = -\chi(g) + \chi(g) = \mathbb{Z}$ , the zero element of  $\mathbb{Q}/\mathbb{Z}$ , for all  $g \in G$  and so  $-\chi + \chi = \chi_0$ . Therefore each  $\chi \in G^*$  has a negative element in  $G^*$  namely  $-\chi$ .

Finally the commutative law is obeyed in  $(G^*, +)$  as, using the commutative law in  $\mathbb{Q}/\mathbb{Z}$ , we obtain

$$(\chi_1 + \chi_2)(g) = \chi_1(g) + \chi_2(g) = \chi_2(g) + \chi_1(g) = (\chi_2 + \chi_1)(g) \text{ for all } g \in G \text{ which shows}$$

$$\chi_1 + \chi_2 = \chi_2 + \chi_1 \text{ for all } \chi_1, \chi_2 \in G^*. \text{ We conclude: } G^* \text{ is an abelian group.}$$

Consider  $\chi_1, \chi_2 \in H^o$ . Then  $(\chi_1 + \chi_2)(h) = \chi_1(h) + \chi_2(h) = \mathbb{Z} + \mathbb{Z} = \mathbb{Z} \in \mathbb{Q}/\mathbb{Z}$  for all  $h \in H$  showing  $\chi_1 + \chi_2 \in H^o$ , i.e.  $H^o$  is closed under addition. As  $\chi_0(h) = \mathbb{Z}$  for all  $h \in H$  we see that the zero element  $\chi_0$  of  $G^*$  is in  $H^o$ . For  $\chi \in H^o$  we have  $(-\chi)(h) = -\chi(h) = -\mathbb{Z} = \mathbb{Z}$  for all  $h \in H$  which shows that  $-\chi \in H^o$ , i.e.  $H^o$  is closed under negation. So  $H^o$  is a subgroup of  $G^*$ .

The composition  $\eta\chi$  of the natural homomorphism  $\eta: G \rightarrow G/H$  and the homomorphism

$$\chi: G/H \rightarrow \mathbb{Q}/\mathbb{Z} \text{ is the homomorphism } (\chi)\alpha: G \rightarrow \mathbb{Q}/\mathbb{Z} \text{ as } (\eta\chi)(g) = \chi((g)\eta) = \chi(H + g) \text{ for all}$$

$g \in G$ , i.e.  $\eta\chi = (\chi)\alpha \in G^*$ . Also for  $h \in H$  we have  $((\chi)\alpha)(h) = \chi(H + h) = \chi(H) = \mathbb{Z}$  as  $\chi$  maps the zero element  $H$  of  $G/H$  to the zero element  $\mathbb{Z}$  of  $\mathbb{Q}/\mathbb{Z}$ , i.e.  $(\chi)\alpha \in H^o$ . So  $\alpha: (G/H)^* \rightarrow H^o$ .

Let  $\chi_1, \chi_2 \in (G/H)^*$ . Then

$$(\eta(\chi_1 + \chi_2))(g) = (\chi_1 + \chi_2)((g)\eta) = \chi_1((g)\eta) + \chi_2((g)\eta) = (g)(\eta\chi_1) + (g)(\eta\chi_2) \text{ for all } g \in G$$

giving  $\eta(\chi_1 + \chi_2) = \eta\chi_1 + \eta\chi_2$ . So  $(\chi_1 + \chi_2)\alpha = \eta(\chi_1 + \chi_2) = \eta\chi_1 + \eta\chi_2 = (\chi_1)\alpha + (\chi_2)\alpha$  showing  $\alpha$  to be additive, i.e.  $\alpha$  is a homomorphism. Consider  $\chi' \in H^o$ . As  $H \subseteq \ker \chi'$ , the mapping

$$\tilde{\chi}': G/H \rightarrow \mathbb{Q}/\mathbb{Z} \text{ given by } \tilde{\chi}'(H + g) = \chi'(g) \text{ for all } g \in G \text{ is an unambiguously defined}$$

homomorphism, i.e.  $\tilde{\chi}' \in (G/H)^*$  and  $(\tilde{\chi}')\alpha = \eta\tilde{\chi}' = \chi'$  showing  $H^o = \text{im } \alpha$ . Suppose

$$\chi: G/H \rightarrow \mathbb{Q}/\mathbb{Z} \text{ is in } \ker \alpha. \text{ Then } (\chi)\alpha = \eta\chi = \chi_0 \text{ the zero character of } G, \text{ i.e. } \chi((g)\eta) = \mathbb{Z} \text{ for all}$$

$g \in G$ . So  $\chi(H + g) = \mathbb{Z}$  for all  $H + g \in G/H$  which shows that  $\chi$  is the zero character of  $G/H$ , i.e.  $\chi$  is the zero element of  $(G/H)^*$ . Therefore  $\ker \alpha$  is trivial and so  $\alpha: (G/H)^* \cong H^o$ .

(b) For each element  $g$  of  $G$  there are integers  $m_i$  ( $1 \leq i \leq t$ ) with  $g = m_1 h_1 + m_2 h_2 + \dots + m_t h_t$  since  $G = \langle h_1, h_2, \dots, h_t \rangle$ . As  $G = \langle h_1 \rangle \oplus \langle h_2 \rangle \oplus \dots \oplus \langle h_t \rangle$  and  $h_i$  has order  $d_i$  ( $1 \leq i \leq t$ ), each  $m_i$  is unique modulo  $d_i$ , i.e. if also  $g = m'_1 h_1 + m'_2 h_2 + \dots + m'_t h_t$  where  $m'_i \in \mathbb{Z}$ , then  $m_i \equiv m'_i \pmod{d_i}$  for

$1 \leq i \leq t$ . Suppose there is a character  $\chi_i$  of  $G$  with the stated property. As  $\chi_i$  is  $\mathbb{Z}$ -linear we see  $\chi_i(g) = \chi_i(m_1 h_1 + m_2 h_2 + \dots + m_t h_t) = \mathbb{Z} + m_i/d_i = \mathbb{Z} + m'_i/d_i$  showing that there is at most one character  $\chi_i$  as stated. On the other hand let  $\chi_i: G \rightarrow \mathbb{Q}/\mathbb{Z}$  be defined by the preceding equation.

Consider  $g' \in G$  with  $g' = m''_1 h_1 + m''_2 h_2 + \dots + m''_t h_t$ . Then

$$g + g' = (m_1 + m''_1)h_1 + (m_2 + m''_2)h_2 + \dots + (m_t + m''_t)h_t \text{ and so}$$

$$(g + g')\chi_i = \mathbb{Z} + (m_i + m''_i)/d_i = (\mathbb{Z} + m_i/d_i) + (\mathbb{Z} + m''_i/d_i) = (g)\chi_i + (g')\chi_i$$

showing that  $\chi_i$  is a homomorphism. Replacing  $g$  by  $h_i$  gives  $m_i \equiv 1 \pmod{d_i}$ . So

$$(h_i)\chi_i = \mathbb{Z} + m_i/d_i = \mathbb{Z} + 1/d_i. \text{ Replacing } g \text{ by } h_j \text{ for } j \neq i \text{ gives } m_i \equiv 0 \pmod{d_i} \text{ and so}$$

$$(h_j)\chi_i = \mathbb{Z} + m_i/d_i = \mathbb{Z}, \text{ i.e. } \chi_i \text{ is a character of } G \text{ which acts on } h_1, h_2, \dots, h_t \text{ as stated.}$$

As  $(d_i \chi_i)(h_i) = d_i(\chi_i(h_i)) = d_i(\mathbb{Z} + 1/d_i) = \mathbb{Z} + d_i/d_i = \mathbb{Z} + 1 = \mathbb{Z}$  and

$$(d_i \chi_i)(h_j) = d_i(\chi_i(h_j)) = d_i \mathbb{Z} = \mathbb{Z} \text{ for } j \neq i. \text{ As } G = \langle h_1, h_2, \dots, h_t \rangle \text{ we deduce } d_i \chi_i = \chi_0 \text{ and so } \chi_i$$

has order  $n_i$  where  $n_i \mid d_i$ . Now  $n_i \chi_i = \chi_0$  and so  $(n_i \chi_i)(h_i) = \chi_0(h_i)$ , i.e.  $n_i(\mathbb{Z} + 1/d_i) = \mathbb{Z}$  showing  $d_i \mid n_i$ . As  $d_i$  and  $n_i$  are positive integers we see  $n_i = d_i$  showing that  $\chi_i$  has order  $d_i$  for  $1 \leq i \leq t$ .

Let  $\chi \in G^*$ . Suppose  $\chi(h_i) = \mathbb{Z} + a_i/b_i$  where  $a_i, b_i \in \mathbb{Z}$ . Then  $d_i(\chi(h_i)) = \chi(d_i h_i) = \chi(0) = \mathbb{Z}$ , i.e.  $\mathbb{Z} + d_i a_i/b_i = \mathbb{Z}$  and so  $d_i a_i/b_i = q_i \in \mathbb{Z}$ . Hence  $\chi(h_i) = \mathbb{Z} + q_i/d_i = q_i \chi_i(h_i)$  for  $1 \leq i \leq t$ .

Therefore the character  $q_1 \chi_1 + q_2 \chi_2 + \dots + q_t \chi_t$  agrees with  $\chi$  on each of the generators  $h_1, h_2, \dots, h_t$  of  $G$ . So  $\chi = q_1 \chi_1 + q_2 \chi_2 + \dots + q_t \chi_t$  showing  $G^* = \langle \chi_1 \rangle + \langle \chi_2 \rangle + \dots + \langle \chi_t \rangle$ . To show that this sum of cyclic subgroups is independent (2.14) suppose  $r_1 \chi_1 + r_2 \chi_2 + \dots + r_t \chi_t = \chi_0$  where  $r_i \in \mathbb{Z}$  for  $1 \leq i \leq t$ . Then  $(r_1 \chi_1 + r_2 \chi_2 + \dots + r_t \chi_t)(h_i) = \chi_0(h_i)$  which gives  $r_i \chi_i(h_i) = \chi_0(h_i)$  in  $\mathbb{Q}/\mathbb{Z}$  as  $\chi_j(h_i) = \mathbb{Z}$  for  $j \neq i$ . So  $\mathbb{Z} + r_i/d_i = \mathbb{Z}$  which means  $d_i \mid r_i$  and hence  $r_i \chi_i = \chi_0$  for  $1 \leq i \leq t$  as  $\chi_i$  has order  $d_i$  in  $G^*$ . So the above cyclic subgroups are independent and  $G^* = \langle \chi_1 \rangle \oplus \langle \chi_2 \rangle \oplus \dots \oplus \langle \chi_t \rangle$  by (2.15).

As above each  $g \in G$  is expressible as  $g = m_1 h_1 + m_2 h_2 + \dots + m_t h_t$  where each integer  $m_i$  is unique modulo  $d_i$  ( $1 \leq i \leq t$ ). In the same way each  $\chi \in G^*$  is expressible as  $\chi = m'_1 \chi_1 + m'_2 \chi_2 + \dots + m'_t \chi_t$  where each integer  $m'_i$  is unique modulo  $d_i$  ( $1 \leq i \leq t$ ). So  $G^*$  and  $G$  are abstractly identical, i.e. writing  $(\chi)\beta = g$  if and only if  $m'_i \equiv m_i \pmod{d_i}$  for  $1 \leq i \leq t$  defines an isomorphism  $\beta: G^* \cong G$  with  $(\chi_i)\beta = h_i$  for  $1 \leq i \leq t$ . Conversely let  $\beta': G^* \cong G$  with  $(\chi_i)\beta' = h_i$  for  $1 \leq i \leq t$ . As  $\beta'$  is  $\mathbb{Z}$ -linear we obtain

$$\begin{aligned} (\chi)\beta' &= (m'_1 \chi_1 + m'_2 \chi_2 + \dots + m'_t \chi_t)\beta' = m'_1((\chi_1)\beta') + m'_2((\chi_2)\beta') + \dots + m'_t((\chi_t)\beta') = \\ &= m'_1((\chi_1)\beta) + m'_2((\chi_2)\beta) + \dots + m'_t((\chi_t)\beta) = (m'_1 \chi_1 + m'_2 \chi_2 + \dots + m'_t \chi_t)\beta = (\chi)\beta \end{aligned}$$

for all  $\chi \in G^*$  showing  $\beta' = \beta$ .

(i) Consider  $h \in H$  and  $\chi \in H^o$ . There are integers  $m_i, m'_i$  ( $1 \leq i \leq t$ ) with  $h = m_1 h_1 + m_2 h_2 + \dots + m_t h_t$  and  $\chi = m'_1 \chi_1 + m'_2 \chi_2 + \dots + m'_t \chi_t$ . Then  $\chi(h) = \mathbb{Z}$  gives  $m'_1 m_1 / d_1 + m'_2 m_2 / d_2 + \dots + m'_t m_t / d_t \in \mathbb{Z}$ . Then  $(h)\beta^{-1} = m_1 \chi_1 + m_2 \chi_2 + \dots + m_t \chi_t$  and  $(\chi)\beta = m'_1 h_1 + m'_2 h_2 + \dots + m'_t h_t$ . Now

$$((h)\beta^{-1})(\chi)\beta = \mathbb{Z} + (m_1 m'_1 / d_1 + m_2 m'_2 / d_2 + \dots + m_t m'_t / d_t) = \mathbb{Z}$$

showing  $(h)\beta^{-1} \in ((H^o)\beta)^o$  as  $(\chi)\beta$  is a typical element of  $(H^o)\beta$ . As  $(h)\beta^{-1}$  is a typical element of  $(H)\beta^{-1}$  we obtain  $(H)\beta^{-1} \subseteq ((H^o)\beta)^o$ . However these subgroups of  $G^*$  have the same order:

$|(H)\beta^{-1}| = |H|$  and from (a) above we know  $|H^o| = |(G/H)^*|$ . Therefore

$|H^o| = |G/H| = |G|/|H|$  and so  $|(H^o)\beta| = |G|/|H|$ . Hence  $|((H^o)\beta)^o| = |G|/|H^o| = |H|$ . So

in fact  $(H)\beta^{-1} = ((H^o)\beta)^o$  and applying  $\beta$  to this set equality gives  $H = (((H^o)\beta)^o)\beta$ , i.e.

$$H = (H)\pi^2 \text{ for all } H \in \mathbb{L}(G).$$

(ii) Suppose  $H_1 \subseteq H_2$  where  $H_1$  and  $H_2$  are subgroups of  $G$ . Then  $H_1^o \supseteq H_2^o$  and so applying  $\beta$  gives  $(H_1^o)\beta \supseteq (H_2^o)\beta$ , i.e.  $(H_1)\pi \supseteq (H_2)\pi$ . Conversely suppose  $(H_1)\pi \supseteq (H_2)\pi$ , i.e.

$(H_2)\pi \subseteq (H_1)\pi$ . Applying  $\pi$  gives, on using the preceding theory,  $(H_2)\pi^2 \supseteq (H_1)\pi^2$ , i.e.  $H_1 \subseteq H_2$  by (i) above.

### Solution 7

(a)(i) Let  $e$  denote the 1-element of  $R$ . Then  $a \equiv a$  for all  $a \in R$  as  $a = ae$  and  $e \in U(R)$ . Suppose  $a \equiv b$ ; then  $a = bu$  for some  $u \in U(R)$ ; hence  $b \equiv a$  as  $u^{-1} \in U(R)$  and  $b = au^{-1}$ . Suppose  $a \equiv b$  and  $b \equiv c$  where  $a, b, c \in R$ ; there are  $u, v \in U(R)$  with  $a = bu$ ,  $b = cv$ ; hence  $a \equiv c$  as  $a = (cv)u = c(vu)$  and  $vu \in U(R)$ . So  $\equiv$  satisfies the reflexive, symmetric and transitive laws and so  $\equiv$  is an equivalence relation on  $R$ . Yes  $\{0\}$  and  $U(R)$  are associate classes for all commutative rings  $R$ . There are exactly two associate classes if and only if  $R$  is non-trivial and each non-zero element is invertible, i.e.  $R$  is a field. As  $U(\mathbb{Z}_{12}) = \{\bar{1}, \bar{5}, \bar{7}, \bar{11}\}$  we obtain the partition

$\mathbb{Z}_{12} = \{\bar{0}\} \cup \{\bar{1}, \bar{5}, \bar{7}, \bar{11}\} \cup \{\bar{2}, \bar{10}\} \cup \{\bar{3}, \bar{9}\} \cup \{\bar{4}, \bar{8}\} \cup \{\bar{6}\}$  into associate classes.

(ii) Suppose  $a \equiv b$ . Then  $a = bc$  for  $c \in U(R)$ . So  $b|a$  as  $U(R) \subseteq R$ . Also  $b = ac^{-1}$  and so  $a|b$  as  $c^{-1} \in R$ . Conversely suppose  $a|b$  and  $b|a$ . If  $a = 0$  then  $b = 0$  also as  $a|b$ , and so  $a \equiv b$  (equal elements are associate). So assume  $a \neq 0$ . There are  $c, d \in R$  with  $a = bc$ ,  $b = ad$ . Hence  $a = adc$  and so  $e = dc$  on cancelling  $a$  (cancellation of non-zero factors is legitimate in the integral domain  $R$ ). As  $R$  is commutative we see  $c \in U(R)$  and so  $a \equiv b$ .

Suppose  $\langle a \rangle = \langle b \rangle$ . As  $a \in \langle a \rangle$  we see  $a \in \langle b \rangle$  and so  $a = bc$  for  $c \in R$ . Therefore  $b|a$  and by the symmetry of the hypothesis  $a|b$  also. So  $a \equiv b$  using the above paragraph. Conversely suppose  $a \equiv b$ . Then  $a|b$  giving  $\langle b \rangle \subseteq \langle a \rangle$ . Also  $b|a$  giving  $\langle a \rangle \subseteq \langle b \rangle$ . Therefore  $\langle a \rangle = \langle b \rangle$ .

(b) As  $\langle a \rangle \subseteq \langle a \rangle + \langle b \rangle = \langle d \rangle$  we see  $d|a$  by (a)(ii) above. In the same way  $\langle b \rangle \subseteq \langle a \rangle + \langle b \rangle = \langle d \rangle$  gives  $d|b$ . Suppose  $d' \in R$  satisfies  $d'|a$  and  $d'|b$ . Then  $\langle a \rangle \subseteq \langle d' \rangle$  and  $\langle b \rangle \subseteq \langle d' \rangle$ . As  $\langle d' \rangle$  is closed under addition we see  $\langle d \rangle = \langle a \rangle + \langle b \rangle \subseteq \langle d' \rangle$  and so  $d'|d$ . So  $d$  is a gcd of  $a$  and  $b$ .

Conversely let  $d$  have these divisor properties. Then  $\langle a \rangle \subseteq \langle d \rangle$  and  $\langle b \rangle \subseteq \langle d \rangle$  and so  $\langle a \rangle + \langle b \rangle \subseteq \langle d \rangle$ .

Suppose  $\langle a \rangle + \langle b \rangle = \langle d' \rangle$ . Then  $d' \mid a$  and  $d' \mid b$ . So  $d' \mid d$  which means  $\langle d \rangle \subseteq \langle a \rangle + \langle b \rangle$ . Therefore  $\langle a \rangle + \langle b \rangle = \langle d \rangle$  showing that the gcd's of  $a$  and  $b$  are all associate, being precisely the generators  $d$  of the ideal  $\langle a \rangle + \langle b \rangle$ .

As  $\langle m \rangle = \langle a \rangle \cap \langle b \rangle \subseteq \langle a \rangle$  we see  $a \mid m$  by (a) (ii) above. In the same way  $\langle m \rangle = \langle a \rangle \cap \langle b \rangle \subseteq \langle b \rangle$  gives  $b \mid m$ . Suppose  $m' \in R$  satisfies  $a \mid m'$  and  $b \mid m'$ . Then  $\langle m' \rangle \subseteq \langle a \rangle$  and  $\langle m' \rangle \subseteq \langle b \rangle$ . So

$\langle m' \rangle \subseteq \langle a \rangle \cap \langle b \rangle = \langle m \rangle$  which gives  $m \mid m'$  by (a) (ii) above. So  $m$  is an lcm of  $a$  and  $b$ . Conversely let  $m$  have these divisor properties. Then  $a \mid m$  and  $b \mid m$  give  $\langle m \rangle \subseteq \langle a \rangle$  and  $\langle m \rangle \subseteq \langle b \rangle$ . So

$\langle m \rangle \subseteq \langle a \rangle \cap \langle b \rangle$ . Let  $\langle a \rangle \cap \langle b \rangle = \langle m' \rangle$ . Then  $a \mid m'$  and  $b \mid m'$  which gives  $m \mid m'$  by hypothesis. So  $\langle a \rangle \cap \langle b \rangle = \langle m' \rangle \subseteq \langle m \rangle$  which gives  $\langle m \rangle = \langle a \rangle \cap \langle b \rangle$ . So the lcms of  $a$  and  $b$  are all associate, being precisely the generators  $m$  of the ideal  $\langle a \rangle \cap \langle b \rangle$ .

As  $d \in \langle a \rangle + \langle b \rangle$  there are  $a', b \in R$  with  $d = a'a + b'b$ . Suppose  $d = 0$ . Then  $a = b = m = 0$  and so  $ab = 0 = dm$ . Suppose  $d \neq 0$ . Let  $m_0 = ab/d$ . Then  $a \mid m_0$  and  $b \mid m_0$  as  $m_0 = a(b/d)$  and  $m_0 = b(a/d)$ . Let  $m' \in R$  satisfy  $a \mid m'$  and  $b \mid m'$ . We wish to show  $m_0 \mid m'$ . This is certainly true in the case  $m' = 0$ . Suppose  $m' \neq 0$ . Then  $aa'' = m'$  and  $bb'' = m'$ . Substituting in  $m_0 da''b'' = aa''bb''$  gives

$$m_0(a'a + b'b)a''b'' = m_0(a'aa''b'' + b'bb''a'') = m_0(a'm'b'' + b'm'a'') = m_0(a'b'' + b'a'')m' = (m')^2$$

and so  $m_0(a'b'' + b'a'') = m'$  on cancelling  $m'$  from the last equation above. Therefore  $m_0 \mid m'$  showing that  $m_0$  is an lcm of  $a$  and  $b$ . So  $m_0 = ab/d \equiv m$  and hence  $ab \equiv dm$ .

Consider  $k_1, k_2 \in K = \bigcup_{i \geq 1} K_i$ . There are  $i_1, i_2$  with  $k_1 \in K_{i_1}$  and  $k_2 \in K_{i_2}$ . Write  $j = \max\{i_1, i_2\}$ . Then  $K_{i_1} \cup K_{i_2} = K_j$  as  $K_i \subseteq K_j$  for  $i \leq j$  by induction on  $j - i$ . So  $k_1, k_2 \in K_j$  and hence  $k_1 + k_2 \in K_j$  as  $K_j$  is an ideal of  $R$  and so closed under addition. As  $K_j \subseteq K$  we conclude  $k_1 + k_2 \in K$ , i.e.  $K$  is closed under addition. As  $0 \in K_1$  we see  $0 \in K$ . As  $-k_1, rk_1 \in K_{i_1}$  for all  $r \in R$  we see  $-k_1, rk_1 \in K$ . So  $K$  is an ideal of  $R$ . As  $R$  is a PID there is  $d \in R$  with  $K = \langle d \rangle$ . As  $d \in K$  there is  $l$  with  $d \in K_l$  and so  $K_i \subseteq K \subseteq K_l$  for all positive integers  $i$ . By hypothesis  $K_l \subseteq K_i$  for  $i \geq l$  and so  $K_i = K_l$  for  $i \geq l$ .

Write  $K_i = \langle b_i \rangle$  for  $i = 1, 2, \dots$ . As  $b_{i+1} \mid b_i$  we see  $K_i \subseteq K_{i+1}$  for  $i = 1, 2, \dots$ . By the preceding paragraph there is a positive integer  $l$  such that  $\langle b_i \rangle = K_i = K_l = \langle b_l \rangle$  for  $i \geq l$ . So  $b_i \equiv b_l$  for  $i \geq l$ .

(c) Yes, every  $1 \times 2$  diagonal matrix  $D = (d_1, 0)$  is in Snf as the condition  $d_i \mid d_j$  for

$1 \leq i \leq j \leq \min\{s, t\}$  is merely  $d_1 \mid d_1$  in this case.

(i) Consider  $A = (a, b)$  and let  $D = (d, 0)$  where  $\langle a \rangle + \langle b \rangle = \langle d \rangle$ . Then  $D$  is in Snf. Suppose first  $d = 0$ . Then  $a = b = 0$  and so  $A$  is in Snf and  $Q = I$ , the  $2 \times 2$  identity matrix over  $R$  satisfies

$A = DQ$ ,  $\det Q = e$ . Suppose  $d \neq 0$ . There are elements  $a', a'', b, b''$  of  $R$  with  $d = a'a + b'b$ ,  $a = a''d$ ,  $b = b''d$ . Therefore  $d = a'a''d + b'b''d = (a'a'' + b'b'')d$  and so  $a'a'' + b'b'' = e$

on cancelling  $d$ . Let  $Q = \begin{pmatrix} a'' & b'' \\ -b' & a' \end{pmatrix}$ . Then  $\det Q = a'a'' + b'b'' = e$  and  $DQ = (a''d, b''d) = (a, b) = A$ .

(ii) Let  $A = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$ . As above write  $\langle a \rangle + \langle b \rangle = \langle d \rangle$ . We may suppose  $a$  is not a divisor of  $b$  as otherwise  $A$  is in Snf and no elementary operations are needed. Then  $d \neq 0$  as  $d = 0 \Rightarrow a = b = 0 \Rightarrow a \mid b$ . There are  $a', b' \in R$  with  $a'a + b'b = d$ . The following sequence of *eros* and *ecos* of type (iii) reduces  $A$  to  $D$ :

$$\begin{aligned} A = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} &\equiv_{c_2 + a'c_1} \begin{pmatrix} a & a'a \\ 0 & b \end{pmatrix} \equiv_{r_1 + b'r_2} \begin{pmatrix} a & d \\ 0 & b \end{pmatrix} \equiv_{c_1 - (a/d-1)c_2} \begin{pmatrix} d & d \\ -ab/d + b & b \end{pmatrix} \\ &\equiv_{c_2 - c_1} \begin{pmatrix} d & 0 \\ -ab/d + b & ab/d \end{pmatrix} \equiv_{r_2 + (a/d-1)(b/d)r_1} \begin{pmatrix} d & 0 \\ 0 & ab/d \end{pmatrix} = D. \end{aligned}$$

As  $ab/d = d(a/d)(b/d)$  we see  $d \mid ab/d$  and so  $D$  is in Snf.

(d) Let the  $s \times s$  matrix  $P$  be specified by its rows:

$$e_i P = p_{11}e_i + p_{12}e_j, e_j P = p_{21}e_i + p_{22}e_j, e_k P = e_k \text{ for } k \neq i, j.$$

Suppose  $i < j$ . Apply to  $P$  the  $i-1$  row interchanges  $r_k \leftrightarrow r_{k-1}$  for  $k = i, i-1, \dots, 3, 2$ , followed by the  $j-2$  row interchanges  $r_k \leftrightarrow r_{k-1}$  for  $k = j, j-1, \dots, 4, 3$ , and also the  $i-1$  column interchanges  $c_k \leftrightarrow c_{k-1}$  for  $k = i, i-1, \dots, 3, 2$ , followed by the  $j-2$  column interchanges  $c_k \leftrightarrow c_{k-1}$  for  $k = j, j-1, \dots, 4, 3$ . The resulting matrix is  $P' \oplus I$  where  $I$  denotes the  $(s-2) \times (s-2)$  identity matrix. Comparing determinants gives  $\det P = (-1)^{2(i-1+j-2)} \det P' \det I = \det P' = e$ .

Suppose  $j < i$ . Carrying out the above row and column operations on  $P$  with  $i$  and  $j$  interchanged gives, on comparing determinants,  $\det P = (-1)^{2(j-1+i-2)} \det P'' \det I = \det P'' = e$  where

$$P'' = \begin{pmatrix} p_{22} & p_{21} \\ p_{12} & p_{11} \end{pmatrix} \text{ and so } \det P'' = \det P'. \text{ Therefore } \det P = e \text{ in all cases.}$$

Postmultiplying the above row equations by  $A$  gives  $e_i PA = p_{11}e_i A + p_{12}e_j A, e_j PA = p_{21}e_i A + p_{22}e_j A, e_k PA = e_k A$  for  $k \neq i, j$  which shows that  $PA$  is the result of applying this *nero* to  $A$ .

Let  $P' = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix}$  be a non-elementary matrix over  $R$  with  $\det P' = e$ . Note that the elementary

$2 \times 2$  matrices over  $R$  with determinant  $e$  are  $\begin{pmatrix} e & r \\ 0 & e \end{pmatrix}, \begin{pmatrix} e & 0 \\ r & e \end{pmatrix}$  for  $r \in R$  and in the case  $\chi(R) = 2$

$\begin{pmatrix} 0 & e \\ e & 0 \end{pmatrix}$ . Let  $A$  be an  $s \times t$  matrix over  $R$  with  $t \geq 2$  and let  $(i, j)$  be an ordered pair of distinct

integers with  $1 \leq i, j \leq t$ . Let the  $t \times t$  matrix  $Q$  be specified by its columns as follows:

$$Qe_i^T = p_{11}e_i^T + p_{21}e_j^T, Qe_j^T = p_{12}e_i^T + p_{22}e_j^T, Qe_k^T = e_k^T \text{ for } k \neq i, j. \text{ Mimicking the above theory (with rows instead of columns) gives } \det Q = \det(P')^T = e \text{ for } i < j \text{ and } \det Q = \det(P'')^T = e \text{ for } i > j.$$

Premultiplying the above column equations by  $A$  gives

$$AQe_i^T = p_{11}Ae_i^T + p_{21}Ae_j^T, AQe_j^T = p_{12}Ae_i^T + p_{22}Ae_j^T, AQe_k^T = Ae_k^T \text{ for } k \neq i, j$$

and these equations tell us that  $AQ$  is the matrix which results on applying this *neco* to  $A$ .

Using (c)(i) above, for  $i \neq j$  the  $(l, i)$ -entry  $a$  and the  $(l, j)$ -entry  $b$  in a matrix over a PID  $R$  can be replaced by  $\gcd\{a, b\}$  and 0 respectively on postmultiplying by a suitable matrix of determinant  $e$ .

Denote this **column operation** by  $c(l; i, j)$ .

In general  $c(l; i, j)$  is not elementary, but in the case  $R = \mathbb{Z}$  it can be carried out by a sequence of *ecos* as in (1.7). In an analogous way for  $i \neq j$  the  $(i, m)$ -entry  $a$  and the  $(j, m)$ -entry  $b$  in a matrix over a PID  $R$  can be replaced by  $\gcd\{a, b\}$  and 0 respectively on premultiplying by a suitable matrix of determinant  $e$ .

Denote this **row operation** by  $r(m; i, j)$ .

Notice that column  $k$  is unchanged by  $c(l; i, j)$  for  $k \neq i, j$  and row  $k$  is unchanged by  $r(m; i, j)$  for  $k \neq i, j$ . In the case  $a \mid b, a \neq 0$  these operations are elementary as  $c(l; i, j) = c_j - (b/a)c_i$  and  $r(m; i, j) = r_j - (b/a)r_i$ .

Yes (1.4) goes through without change and so elementary operations of type (iii) can be carried out using pre- and post-multiplication by elementary matrices of determinant  $e$ .

(e) Let  $A = (a_{ij})$  be an  $s \times t$  matrix over  $R$ . We generalise (1.9) by showing that there are square matrices  $P$  and  $Q$  over  $R$  with  $\det P = \det Q = e$  such that  $PAQ^{-1} = B = (b_{ij})$  where  $b_{1j} = b_{i1} = 0$  for  $1 < i \leq s, 1 < j \leq t$ . Should  $e_1 A \neq a_{11} e_1$  clear all off-diagonal entries in row 1 by applying the composite column operation  $c(1; 1, 2)c(1; 1, 3) \cdots c(1; 1, t)$  to  $A$ . So there is a  $t \times t$  matrix  $Q_1$  over  $R$  with  $e_1 A Q_1^{-1} = b_1 e_1$  and  $\det Q_1 = e$  (take  $Q_1 = I$  if  $e_1 A = a_{11} e_1$ ). Should  $A Q_1^{-1} e_1^T = b_1 e_1^T$  then the process terminates with  $A Q_1^{-1} = B$ . Otherwise clear all off-diagonal entries in column 1 of  $A Q_1^{-1}$  by applying the composite row operation  $r(1; 1, 2)r(1; 1, 3) \cdots r(1; 1, s)$  to  $A Q_1^{-1}$ . There is an  $s \times s$  matrix  $P_2$  with  $P_2 A Q_1^{-1} e_1^T = b_2 e_1^T$  and  $\det P_2 = e$ . So

$$b_1 e_1 e_1^T = e_1 A Q_1^{-1} e_1^T = e_1 P_2^{-1} P_2 A Q_1^{-1} e_1^T = e_1 P_2^{-1} (b_2 e_1^T) = b_2 (e_1 P_2^{-1} e_1^T)$$

which are equations relating  $1 \times 1$  matrices over  $R$ . Comparing entries gives  $b_2 \mid b_1$ . Should  $b_2 \equiv b_1$

then the process terminates with  $P_2 A Q_1^{-1} = B$  as  $r(1; 1, j)$  is an elementary row operation of type (iii) leaving row 1 unchanged for  $1 < j \leq s$ . Should  $b_2 \not\equiv b_1$  then the process is repeated starting with

$P_2 A Q_1^{-1}$  in place of  $A$ . Write  $B_0 = A, B_1 = A Q_1^{-1}$  and  $B_2 = P_2 A Q_1^{-1}$ . Is it possible for the process to be repeated an infinite number of times without terminating? If so there would be an infinite sequence  $B_1, B_2, \dots, B_i, \dots$  of  $s \times t$  matrices over  $R$  such that for each positive integer  $i$

$$e_1 B_i = b_i e_1, B_i = B_{i-1} Q_i^{-1}, \det Q_i = e \text{ (i odd) and } B_i e_1^T = b_i e_1^T, B_i = P_i B_{i-1}, \det P_i = e \text{ (i even)}.$$

For  $i$  odd  $b_i e_1 e_1^T = e_1 B_i e_1^T = e_1 P_{i+1}^{-1} B_{i+1} e_1^T = e_1 P_{i+1}^{-1} (b_{i+1} e_1^T) = b_{i+1} (e_1 P_{i+1}^{-1} e_1^T)$  and on comparing the entries in these  $1 \times 1$  matrices over  $R$  we see  $b_{i+1} \mid b_i$ . In the same way for  $i$  even

$b_i e_1 e_1^T = e_1 (b_i e_1^T) = e_1 B_i e_1^T = e_1 B_{i+1} Q_{i+1} e_1^T = b_{i+1} e_1 Q_{i+1} e_1^T$  and so we again obtain  $b_{i+1} \mid b_i$ . If the process does not terminate then  $b_{i+1} \not\equiv b_i$  for  $i \geq 1$  contrary to (b) above. Let  $l$  be the smallest integer with

$b_{l+1} \equiv b_l$ . Then the process terminates with  $B_{l+1} = B$ . For  $l = 2k$  we have  $PAQ^{-1} = B$  where

$P = P_{2k} \cdots P_4 P_2$  and  $Q = Q_{2k+1} \cdots Q_3 Q_1$ . For  $l = 2k - 1$  we have  $PAQ^{-1} = B$  where  $P = P_{2k} \cdots P_4 P_2$  and  $Q = Q_{2k-1} \cdots Q_3 Q_1$ . In both cases  $\det P = \det Q = e$  as  $P$  and  $Q$  are products of matrices all of which have determinant  $e$ .

Finally we show by induction on  $\min\{s, t\}$  that there are square matrices  $P$  and  $Q$  over  $R$  with  $\det P = \det Q = e$  such that  $PAQ^{-1}$  is in Snf.. Suppose  $\min\{s, t\} = 1$ . Then  $B$  as above is in Snf and so  $P$  and  $Q$  exist in this case. Suppose  $\min\{s, t\} > 1$ . By the preceding theory there are square matrices

$P_0$  and  $Q_0$  over  $R$  with  $\det P_0 = \det Q_0 = e$  such that  $P_0 A Q_0^{-1} = B = (b_{ij})$  where  $b_{1j} = b_{i1} = 0$  for  $1 < i \leq s, 1 < j \leq t$ . Let  $B'$  be the  $(s-1) \times (t-1)$  submatrix of  $B$  which remains on deleting row 1 and column 1. In fact  $B = (b_{11}) \oplus B'$  the direct sum (5.17) of the  $1 \times 1$  matrix  $(b_{11})$  and  $B'$ . As

$\min\{s-1, t-1\} = \min\{s, t\} - 1$  by the inductive hypothesis there is an  $(s-1) \times (s-1)$  matrix  $P'$  over  $R$  and a  $(t-1) \times (t-1)$  matrix  $Q'$  over  $R$  with  $\det P' = \det Q' = e$  and  $P'B'(Q')^{-1} = \text{diag}(d'_2, d'_3, \dots)$  in Snf. Let  $P_1 = (e) \oplus P'$  and  $Q_1 = (e) \oplus Q'$ . Then  $\det P_1 = \det Q_1 = e$  and

$$P_1 B Q_1^{-1} = ((e) \oplus P')((b_{11}) \oplus B')((e) \oplus (Q')^{-1}) = (b_{11}) \oplus P'B'(Q')^{-1} = \text{diag}(b_{11}, d'_2, d'_3, \dots) = D_1.$$

Applying (c)(ii) above to the leading  $2 \times 2$  matrix of  $D_1$ , there is an  $s \times s$  matrix  $P_2$  and a  $t \times t$  matrix  $Q_2$  over  $R$  with  $P_2 D_1 Q_2^{-1} = \text{diag}(d_1, m_2, d'_3, d'_4, \dots) = D_2$  where  $\det P_2 = \det Q_2 = e$  and

$d_1 = \gcd\{b_{11}, d'_2\}$ ,  $m_2 = \text{lcm}\{b_{11}, d'_2\}$ . In fact  $P_2$  and  $Q_2$  are products of elementary matrices of type (iii) over  $R$ . Write  $D'_2 = \text{diag}(m_2, d'_3, d'_4, \dots)$ . Applying the inductive hypothesis to the  $(s-1) \times (t-1)$  matrix  $D'_2$  there is an  $(s-1) \times (s-1)$  matrix  $P'_3$  over  $R$  and a  $(t-1) \times (t-1)$  matrix  $Q'_3$  over  $R$  with  $\det P'_3 = \det Q'_3 = e$  such that  $P'_3 D'_2 (Q'_3)^{-1} = D'$  in Snf. Let  $P_3 = (e) \oplus P'_3$  and

$Q_3 = (e) \oplus Q'_3$ . Then  $P_3 D_2 Q_3^{-1} = (d_1) \oplus D' = D$ . As  $d_1$  is a divisor of every entry in  $D'_2$  we see that  $d_1$  is a divisor of every entry in  $D'$  and so  $D$  is in Snf. Write  $P = P_3 P_2 P_1 P_0$  and  $Q = Q_3 Q_2 Q_1 Q_0$ . Then  $\det P = \det Q = e$  as  $P$  and  $Q$  are products of matrices having determinant  $e$ . Also

$$P A Q^{-1} = P_3 P_2 P_1 P_0 A Q_0^{-1} P_1^{-1} P_2^{-1} P_3^{-1} = D. \text{ The induction is now complete.}$$

(f) Let  $g_l(A)$  denote a gcd of the  $l$ -minors of  $A$ , i.e.  $\langle g_l(A) \rangle = \sum_{m_l} \langle m_l \rangle$  where  $m_l$  runs through the

$l$ -minors of  $A$ . So  $g_l(A)$  is a divisor of each  $m_l$  and  $g_l(A)$  is a linear combination of the  $\binom{s}{l} \times \binom{t}{l}$

elements  $m_l$  of  $R$ . Let  $B$  be an  $s \times r$  matrix over  $R$  and let  $C$  be a  $r \times t$  matrix over  $R$ . The proofs of (1.18) and (1.19) remain valid on replacing  $\mathbb{Z}$  by  $R$  and so  $g_l(B)$  and  $g_l(C)$  are divisors of

$g_l(BC)$  for  $1 \leq l \leq \min\{r, s, t\}$ . Suppose  $P D Q^{-1} = D'$ . For  $1 \leq l \leq \min\{s, t\}$  we deduce

$g_l(D) \mid g_l(PD)$  and  $g_l(PD) \mid g_l(PDQ^{-1}) = g_l(D')$ . Therefore  $g_l(D) \mid g_l(D')$ . As  $P^{-1}D'Q = D$  the roles of  $D$  and  $D'$  can be interchanged to give  $g_l(D') \mid g_l(D)$ . So  $g_l(D) \equiv g_l(D')$  by part (a)(ii)

above for  $1 \leq l \leq \min\{r, s, t\}$ . As in the proof of (1.20) the leading  $l$ -minor of  $D$  is  $d_1 d_2 \cdots d_l$  and this minor is a divisor of all  $l$ -minors of  $D$ . So  $g_l(D) \equiv d_1 d_2 \cdots d_l$  and in the same way

$g_l(D') \equiv d'_1 d'_2 \cdots d'_l$ . Taking  $l=1$  gives  $d_1 \equiv d'_1$ . Take  $l > 1$  and suppose inductively  $d_i \equiv d'_i$  for

$1 \leq i < l$ . We use  $d_1 d_2 \cdots d_l \equiv d'_1 d'_2 \cdots d'_l$ . In the case  $d_l \neq 0$  we have  $d_i \neq 0$  for  $1 \leq i < l$  as  $d_i \mid d_l$ , and cancelling these  $d_i$  gives  $d_l \equiv d'_l$ . Also  $d_l = 0$  implies  $d'_l = 0$  as  $d'_l \neq 0$  gives  $d'_1 d'_2 \cdots d'_l \neq 0$  and so  $d_l \equiv d'_l$  in this case also. The induction is now complete showing  $d_l \equiv d'_l$  for  $1 \leq l \leq \min\{s, t\}$ .

### Solution 8

(a) Consider  $u_1$  and  $u_2$  in  $T(M)$ . There are non-zero elements  $r_1$  and  $r_2$  of  $R$  with  $r_1 u_1 = 0$  and  $r_2 u_2 = 0$ . Then  $r_1 r_2 (u_1 + u_2) = r_1 r_2 u_1 + r_1 r_2 u_2 = r_2 (r_1 u_1) + r_1 (r_2 u_2) = r_2 \times 0 + r_1 \times 0 = 0$ . As  $R$  is an integral domain we know  $r_1 r_2 \neq 0$  and so the above equation shows  $u_1 + u_2 \in T(M)$ , i.e.  $T(M)$  is

closed under addition. The 1-element  $e$  of  $R$  is non-zero and satisfies  $e \times 0 = 0$  showing  $0 \in T(M)$ , i.e. the 0-element  $0$  of  $M$  belongs to  $T(M)$ . Also  $r_1(-u_1) = -r_1u_1 = -0 = 0$  shows  $-u_1 \in T(M)$ . So  $T(M)$  is a subgroup of the additive group of  $M$ . As  $r_1(ru_1) = r(r_1u_1) = r \times 0 = 0$  for all  $r \in R$  we see  $ru_1 \in T(M)$ , i.e.  $T(M)$  is closed under multiplication by  $r$ . Therefore  $T(M)$  is a submodule (2.26) of  $M$ .

Let  $\alpha: M \rightarrow M'$  be  $R$ -linear and let  $u \in T(M)$ . There is  $r \in R$  with  $ru = 0, r \neq 0$ . So  $r((u)\alpha) = (ru)\alpha = (0)\alpha = 0$  which shows  $(u)\alpha \in T(M')$ , i.e.  $(T(M))\alpha \subseteq T(M')$ . Suppose  $\alpha: M \cong M'$ , i.e. suppose  $\alpha$  is bijective (2.24). Then  $\alpha^{-1}: M' \cong M$  by Exercises 2.3, Question 7 (c). Replacing  $\alpha$  by  $\alpha^{-1}$  in the preceding theory gives  $(T(M'))\alpha^{-1} \subseteq T(M)$ . Applying  $\alpha$  to this set containment gives  $T(M') = (T(M')\alpha^{-1})\alpha \subseteq (T(M))\alpha$ . Therefore  $(T(M))\alpha = T(M')$  and so the restriction  $\alpha|$  of  $\alpha$  to  $T(M)$  satisfies  $\alpha|: T(M) \cong T(M')$ . The coset  $T(M) + v$  is a typical element of  $M/T(M)$  where  $v \in M$ . The  $R$ -linear mapping  $\tilde{\alpha}: M/T(M) \rightarrow M'/T(M')$  is unambiguously defined by  $(T(M) + v)\tilde{\alpha} = T(M') + (v)\alpha$  for all  $v \in M$ . Further  $\tilde{\alpha}$  is bijective as

$\tilde{\beta}: M'/T(M') \rightarrow M/T(M)$  is its inverse where  $\beta = \alpha^{-1}$ . So  $\tilde{\alpha}: M/T(M) \cong M'/T(M')$ , i.e. the isomorphism  $\alpha$  induces an isomorphism of the corresponding factor modules.

The torsion submodule of  $T(M)$  is  $T(T(M)) = T(M)$  directly from the above definition. Suppose  $T(M) + v$  belongs to the torsion submodule  $T(M/T(M))$  of  $M/T(M)$ . There is  $r \in R$  with  $r \neq 0$  and  $r(T(M) + v) = T(M)$ , i.e.  $T(M) + rv = T(M)$  giving  $rv \in T(M)$ . So there is  $r' \in R$  with  $r' \neq 0$  and  $r'(rv) = 0$ . As  $r'r \neq 0$  and  $(r'r)v = 0$  we see  $v \in T(M)$ . Therefore  $T(M) + v = T(M)$  showing  $M/T(M)$  torsion-free, i.e.  $T(M/T(M)) = \{T(M)\}$ , the torsion subgroup of  $M/T(M)$  is trivial.

(b) We use induction on  $t$ . In the case  $t = 0$  we see  $M$  is trivial and so  $N$  is also trivial. By convention  $N$  is free of rank 0. Suppose  $t = 1$ . Then  $M$  is cyclic with  $R$ -basis  $v_1$  say. We may suppose  $N$  is non-trivial by the above convention. Consider  $K = \{r_1 \in R: r_1v_1 \in N\}$ . Then  $K$  is a non-zero ideal of  $R$ . There is  $d \in R$  with  $d \neq 0$  such that  $K = \langle d \rangle$ . We claim that  $dv_1$  is an  $R$ -basis of  $N$ . As  $d \in K$  ( $d = de \in K$  where  $e$  is the 1-element of  $R$ ) we see  $dv_1 \in N$ . For  $r \in R$  suppose  $rdv_1 = 0$ . Then  $rd = 0$  as  $v_1$  is  $R$ -independent. So  $r = 0$  as  $R$  has no divisors of zero showing  $dv_1$  to be  $R$ -independent. Let  $u \in N$ . Then  $u \in M$  and so  $u = r_1v_1$  for some  $r_1 \in R$ . As  $r_1v_1 = u \in N$  we see  $r_1 \in K$  and so  $r_1 = r'd$  where  $r' \in R$ . Therefore  $u = r_1v_1 = r'dv_1$  showing that  $dv_1$  generates  $N$ . So  $dv_1$  is an  $R$ -basis of  $N$  which is therefore free of rank 1.

Now suppose  $t > 1$ . Then  $M$  has an  $R$ -basis  $v_1, v_2, \dots, v_t$ . Consider the free submodule  $M'$  generated by  $v_1, v_2, \dots, v_{t-1}$ . Then  $N \cap M'$  is a submodule of the free  $R$ -module  $M'$  of rank  $t-1$ . By inductive hypothesis we deduce that  $N \cap M'$  has  $R$ -basis  $u_1, u_2, \dots, u_{s-1}$  where  $0 \leq s-1 \leq t-1$ . Suppose  $N \subseteq M'$ . Then  $N = N \cap M'$  is free of rank  $s-1$ . So we suppose  $N \not\subseteq M'$ . For each  $u \in N$  there are unique elements  $r_1, r_2, \dots, r_t$  of  $R$  with  $u = r_1v_1 + r_2v_2 + \dots + r_tv_t$ . Let  $K$  denote the set of ring elements  $r_t$  which arise in this way, i.e.  $K = \{r_t: r_1v_1 + r_2v_2 + \dots + r_tv_t \in N\}$ . As  $N$  is a submodule of  $M$  it follows that  $K$  is an ideal of  $R$ . As  $N \not\subseteq M'$  we see that  $K$  is non-zero and so there is  $d \in R$  with  $K = \langle d \rangle$  and  $d \neq 0$ . As above  $d \in K$  and so  $d$  arises from an element of  $N$  which we call  $u_s$ , i.e.  $u_s \in N$  and  $u_s - dv_t \in M'$ . We show that  $u_1, u_2, \dots, u_{s-1}, u_s$  is an  $R$ -basis of  $N$ . Let  $u \in N$ . As above  $u = r_1v_1 + r_2v_2 + \dots + r_tv_t$  where  $r_1, r_2, \dots, r_t \in R$ . As  $r_t \in K$  there is  $r'_t \in R$  with  $r_t = r'_td$ .



Consider  $u - r'_t u_s \in N$ . As  $u - r'_t u_s = (r_1 v_1 + \dots + r_{t-1} v_{t-1}) - r'_t (u_s - dv_t) \in M'$  we see  $u - r'_t u_s \in N \cap M'$ . So  $u - r'_t u_s = r'_1 u_1 + \dots + r'_{s-1} u_{s-1}$  for  $r'_1, \dots, r'_{s-1} \in R$ . Therefore  $u = r'_1 u_1 + \dots + r'_{s-1} u_{s-1} + r'_t u_s$  showing that  $u_1, u_2, \dots, u_{s-1}, u_s$  generate  $N$ . Suppose  $r_1 u_1 + \dots + r_{s-1} u_{s-1} + r_s u_s = 0$  where  $r_1, r_2, \dots, r_s \in R$ . Then  $r_s u_s = -(r_1 u_1 + \dots + r_{s-1} u_{s-1}) \in M'$  and so  $r_s dv_t = r_s (dv_t - u_s) + r_s u_s \in M'$  also. So  $r_s dv_t$  is a linear combination of  $v_1, v_2, \dots, v_{t-1}$  which is impossible except in the case  $r_s d = 0$  as  $v_1, v_2, \dots, v_{t-1}, v_t$  are  $R$ -linearly independent. As  $d \neq 0$  we see  $r_s = 0$ . Therefore  $r_1 u_1 + \dots + r_{s-1} u_{s-1} = 0$  and so  $r_1 = r_2 = \dots = r_{s-1} = 0$  as  $u_1, u_2, \dots, u_{s-1}$  are  $R$ -linearly independent. We have shown that  $u_1, u_2, \dots, u_{s-1}, u_s$  are  $R$ -linearly independent. So  $u_1, u_2, \dots, u_{s-1}, u_s$  is an  $R$ -basis of  $N$  showing that  $N$  is free of rank  $s$  where  $s \leq t$ . The induction is now complete.

(c) Let  $k_1, k_2 \in K$ . Then  $k_1 v = 0$  and  $k_2 v = 0$ . So  $(k_1 + k_2)v = k_1 v + k_2 v = 0 + 0 = 0$  showing  $k_1 + k_2 \in K$ . As  $0 \times v = 0$  and  $(-k_1)v = -k_1 v = -0 = 0$  we see  $0, -k_1 \in K$  and so  $(K, +)$  is a subgroup of  $(R, +)$ . For  $r \in R, k \in K$  we have  $(rk)v = r(kv) = r \times 0 = 0$  showing  $rk \in K$ . Therefore  $K$  is an ideal of the commutative ring  $R$ . As  $R$  is a PID there is  $d \in R$  with  $K = \langle d \rangle$ .

Take  $t' = 0$  in the case  $M = 0$ . Suppose  $M \neq 0$  and let  $w_1, w_2, \dots, w_t$  generate  $M$ . So  $t \geq 1$ . Consider the  $R$ -linear mapping  $\theta: R^t \rightarrow M$  defined by  $(r_1, r_2, \dots, r_t)\theta = r_1 w_1 + r_2 w_2 + \dots + r_t w_t$  for all  $(r_1, r_2, \dots, r_t) \in R^t$ . Then  $\text{im } \theta = M$  and  $\ker \theta$  is a free submodule of rank  $s$  say where  $s \leq t$ . Suppose first  $s = 0$ . Then  $\theta: R^t \cong M$ . Take  $t' = t$  and  $v_j = (e_j)\theta = w_j$  for  $1 \leq j \leq t$ . Then  $rw_j = 0 \Rightarrow re_j = 0 \Rightarrow r = 0$  showing that  $w_j$  has order  $d_j = 0$  in  $M$ . As  $e_1, e_2, \dots, e_t$  form an  $R$ -basis of  $R^t$  we see  $M = N_1 \oplus N_2 \oplus \dots \oplus N_t$  is free of rank  $t$  and  $N_j$  is cyclic with generator  $v_j$  of order 0. The divisibility condition  $d_i \mid d_j$  is satisfied and  $T(M) = 0, N_0 = M$  in this case.

Suppose now  $s \geq 1$ . Construct the  $s \times t$  matrix  $A$  over  $R$  with  $z_i = e_i A$  for  $1 \leq i \leq s$  where  $z_1, z_2, \dots, z_s$  is an  $R$ -basis of  $\ker \theta$ . So the rows of  $A$  form an  $R$ -basis of  $\ker \theta$ . By Question 7 (e) above there are invertible matrices  $P$  and  $Q$  over  $R$  with  $PAQ^{-1} = D$  where  $D$  is in Smith normal form. The diagonal entries in  $D$  which are invertible elements of  $R$  are the leading  $s - s'$  entries where  $0 \leq s' \leq s$ . We may suppose that the first  $s - s'$  diagonal entries in  $D$  are  $e$  and so the remaining  $s'$  diagonal entries  $d_1, d_2, \dots, d_{s'}$  in  $D$  are not invertible elements of  $R$ . The rows  $\rho_1, \rho_2, \dots, \rho_t$  of  $Q$  constitute an  $R$ -basis of  $R^t$  by (2.23). As the rows of  $A$  form an  $R$ -basis of  $\ker \theta$  the same is true of the rows of  $PA$ . As  $PA = DQ$  the rows  $\rho_1, \dots, \rho_{s-s'}, d_1 \rho_{s-s'+1}, \dots, d_{s'} \rho_s$  of  $DQ$  also form an  $R$ -basis of  $\ker \theta$ . So  $d_1, d_2, \dots, d_{s'}$  are non-zero elements of  $R$ .

As  $\rho_1, \rho_2, \dots, \rho_t$  generate  $R^t$  and  $\theta: R^t \rightarrow M$  is surjective and  $R$ -linear we see that  $(\rho_1)\theta, (\rho_2)\theta, \dots, (\rho_t)\theta$  generate  $M$ . The first  $s - s'$  of these are zero as  $\rho_j \in \ker \theta$  for  $1 \leq j \leq s - s'$  and so the remaining  $t' = t - s + s'$ , i.e.  $(\rho_{s-s'+1})\theta, (\rho_{s-s'+2})\theta, \dots, (\rho_t)\theta$ , also generate  $M$ . Let  $N_i = \{r(\rho_{s-s'+i})\theta : r \in R\}$  and so  $N_i$  is the cyclic submodule of  $M$  generated by  $v_i = (\rho_{s-s'+i})\theta$  for  $1 \leq i \leq t'$ . As  $v_1, v_2, \dots, v_{t'}$  generate  $M$  we see

$$M = N_1 + N_2 + \dots + N_{t'}.$$

We show that this sum is direct using (2.14) and (2.15). So suppose  $u_1 + u_2 + \dots + u_{t'} = 0$  where  $u_i \in N_i$  for  $1 \leq i \leq t'$ . As  $u_i = r_{s-s'+i}(\rho_{s-s'+i})\theta$  where  $r_{s-s'+i} \in R$  for  $1 \leq i \leq t'$ , the above equation, on substituting for each  $u_i$ , gives

$$r_{s-s'+1}\rho_{s-s'+1} + r_{s-s'+2}\rho_{s-s'+2} + \dots + r_t\rho_t \in \ker \theta.$$

As  $\rho_1, \dots, \rho_{s-s'}, d_1\rho_{s-s'+1}, \dots, d_{s'}\rho_s$  is an  $R$ -basis of  $\ker \theta$ , there are  $r'_i \in R$  for  $1 \leq i \leq s$  such that  $r_{s-s'+1}\rho_{s-s'+1} + r_{s-s'+2}\rho_{s-s'+2} + \dots + r_t\rho_t = r'_1\rho_1 + \dots + r'_{s-s'}\rho_{s-s'} + r'_{s-s'+1}d_1\rho_{s-s'+1} + \dots + r'_s d_{s'}\rho_s$ . ♦

As  $\rho_1, \rho_2, \dots, \rho_t$  are  $R$ -independent the coefficients of each of  $\rho_1, \rho_2, \dots, \rho_t$  appearing on opposite sides of ♦ are equal. Therefore  $r'_i = 0$  for  $1 \leq i \leq s - s'$ . More to the point  $r_{s-s'+i} = r'_{s-s'+i}d_i$  and so  $u_i = r_{s-s'+i}(\rho_{s-s'+i})\theta = r'_{s-s'+i}d_i(\rho_{s-s'+i})\theta = r'_{s-s'+i}(d_i\rho_{s-s'+i})\theta = r_{s-s'+i} \times 0 = 0$  as  $d_i\rho_{s-s'+i} \in \ker \theta$  for  $1 \leq i \leq s'$ . Also  $r_i = 0$  for  $s < i \leq t$ , i.e.  $r_{s-s'+i} = 0$  for  $s' < i \leq t'$  giving  $u_i = r_{s-s'+i}(\rho_{s-s'+i})\theta = 0 \times (\rho_{s-s'+i})\theta = 0$  for  $s' < i \leq t'$ . So  $u_i = 0$  for  $1 \leq i \leq t'$  which shows that  $N_1, N_2, \dots, N_{t'}$  are independent submodules of  $M$  according to (2.14). Therefore

$$M = N_1 \oplus N_2 \oplus \dots \oplus N_{t'} \text{ (internal direct sum) by (2.15).}$$

Let  $K_i$  be the order ideal of  $v_i$  for  $1 \leq i \leq t'$ . For  $1 \leq i \leq s'$  we have

$d_i v_i = d_i(\rho_{s-s'+i})\theta = (d_i\rho_{s-s'+i})\theta = 0$  as  $d_i\rho_{s-s'+i} \in \ker \theta$  and so  $d_i \in K_i$ . On the other hand let  $k_i \in K_i$ . Then  $0 = k_i v_i = k_i(\rho_{s-s'+i})\theta = (k_i\rho_{s-s'+i})\theta$  showing  $k_i\rho_{s-s'+i} \in \ker \theta$ . As

$\rho_1, \dots, \rho_{s-s'}, d_1\rho_{s-s'+1}, \dots, d_{s'}\rho_s$  is an  $R$ -basis of  $\ker \theta$ , comparing coefficients as above gives  $k_i = r'_{s-s'+i}d_i$ . So  $K_i = \langle d_i \rangle$  showing that  $v_i$  has order  $d_i$  for  $1 \leq i \leq s'$ . Now take  $s' < i \leq t'$  and  $k_i \in K_i$ . As above  $k_i\rho_{s-s'+i} \in \ker \theta$  which is only possible in the case  $k_i = 0$  as  $\rho_1, \rho_2, \dots, \rho_t$  are  $R$ -independent. So  $K_i = \{0\}$  showing that  $v_i$  has order 0 for  $s' < i \leq t'$ . So  $d_i = 0$  for  $s' < i \leq t'$ .

Therefore the divisibility condition  $d_i \mid d_j$  for  $1 \leq i \leq j \leq t'$  is satisfied:  $d_i \mid d_j$  for  $1 \leq i \leq j \leq s'$  since  $D$  is in Snf whereas  $d_i \mid d_j$  for  $1 \leq i \leq j \leq t'$ ,  $s' < j$  since  $d_j = 0$  (all elements  $r$  of  $R$  satisfy  $r \mid 0$ ).

We show  $T(M) = N_1 \oplus N_2 \oplus \dots \oplus N_{s'}$ . Consider  $v \in T(M)$ . There is  $r \in R$  with  $rv = 0, r \neq 0$ . As  $v_1, v_2, \dots, v_{t'}$  generate  $M$  and  $T(M) \subseteq M$  there are  $r_1, r_2, \dots, r_{t'} \in R$  with  $v = r_1 v_1 + r_2 v_2 + \dots + r_{t'} v_{t'}$ . So  $0 = rv = rr_1 v_1 + rr_2 v_2 + \dots + rr_{t'} v_{t'}$ . As  $N_1, N_2, \dots, N_{t'}$  are independent (2.14) subgroups of  $(M, +)$  and  $rr_i v_i \in N_i$  for  $1 \leq i \leq t'$  we see  $rr_i v_i = 0$  for  $1 \leq i \leq t'$ . Since  $v_i$  has order  $d_i$  in  $M$  we obtain  $d_i \mid rr_i$  for  $1 \leq i \leq t'$ . So  $0 \mid rr_i$  for  $s' < i \leq t'$  and hence  $0 \mid r_i$  as  $r \neq 0$ , giving  $r_i = 0$  for  $s' < i \leq t'$ . Therefore  $v = r_1 v_1 + r_2 v_2 + \dots + r_{s'} v_{s'}$  showing  $T(M) \subseteq N_1 \oplus N_2 \oplus \dots \oplus N_{s'}$ . On the other hand  $d_{s'} v = 0$  for all  $v \in N_1 \oplus N_2 \oplus \dots \oplus N_{s'}$  showing  $N_1 \oplus N_2 \oplus \dots \oplus N_{s'} \subseteq T(M)$  as  $d_{s'} \neq 0$ . So we've shown  $T(M) = N_1 \oplus N_2 \oplus \dots \oplus N_{s'}$ . Write  $N_0 = N_{s'+1} \oplus N_{s'+2} \oplus \dots \oplus N_{t'}$ . Then  $N_0$  is free (2.22) of rank  $t' - s' = t - s$  having  $R$ -basis  $v_{s'+1}, v_{s'+2}, \dots, v_{t'}$ . Finally

$$M = (N_1 \oplus \dots \oplus N_{s'}) \oplus (N_{s'+1} \oplus \dots \oplus N_{t'}) = T(M) \oplus N_0.$$

(d) The order ideals of  $v$  and  $v'$  are  $\langle d \rangle$  and  $\langle d' \rangle$  respectively. Let  $\alpha: M \cong M'$ . Then  $(v)\alpha = u'$  is also a generator of  $M'$ . So there is  $r_0 \in R$  with  $v' = r_0 u'$ . Let  $K'$  be the order ideal of  $u'$ . Then  $k' \in K' \Rightarrow k'u' = 0 \Rightarrow k'v' = k'r_0 u' = r_0 k'u' = 0 \Rightarrow k' \in \langle d' \rangle$ , i.e.  $K' \subseteq \langle d' \rangle$ . As  $v'$  is a generator of  $M'$  there is  $r_1 \in R$  with  $u' = r_1 v'$ . Hence  $\langle d' \rangle \subseteq K'$  on interchanging the roles of  $u'$  and  $v'$  in the preceding

theory. So  $K' = \langle d' \rangle$ . Suppose  $kv = 0$ , i.e.  $k \in \langle d \rangle$ . Then  $ku' = k(v)\alpha = (kv)\alpha = (0)\alpha = 0$  showing  $k \in K' = \langle d' \rangle$ , i.e.  $\langle d \rangle \subseteq \langle d' \rangle$ . Using  $\alpha^{-1}: M' \cong M$  in place of  $\alpha$  gives  $\langle d' \rangle \subseteq \langle d \rangle$  and so  $\langle d \rangle = \langle d' \rangle$ . From Question 7 (a)(ii) above we deduce  $d \equiv d'$ .

Conversely suppose  $d \equiv d'$ . Then  $\langle d \rangle = \langle d' \rangle$  by Question 7 (a)(ii) above. The surjective  $R$ -linear mapping  $\beta: R \rightarrow M$  given by  $(r)\beta = rv$  for all  $r \in R$  has kernel  $\langle d \rangle$ . So  $\tilde{\beta}: R/\langle d \rangle \cong M$  by (2.28). In the same way  $\tilde{\beta}': R/\langle d' \rangle \cong M'$  where  $\beta': R \rightarrow M'$  is given by  $(r)\beta' = rv'$  for all  $r \in R$ . As  $\langle d \rangle = \langle d' \rangle$  we see  $\tilde{\beta}^{-1}\tilde{\beta}': M \cong M'$  showing that  $M$  and  $M'$  are isomorphic  $R$ -modules.

(e) Suppose  $\alpha: M \cong M'$ . By (a) above  $\alpha|_T: T(M) \cong T(M')$  where  $\alpha|_T$  denotes the restriction of  $\alpha$  to the torsion subgroup  $T(M)$  of  $M$ . By (c) above  $\alpha|_T: N_1 \oplus N_2 \oplus \dots \oplus N_{s'} \cong N'_1 \oplus N'_2 \oplus \dots \oplus N'_{s''}$  where  $s'$  and  $s''$  are such that  $d_i \neq 0$  for  $1 \leq i \leq s'$ ,  $d_i = 0$  for  $s' < i \leq t'$  and  $d'_i \neq 0$  for  $1 \leq i \leq s''$ ,  $d'_i = 0$  for  $s'' < i \leq t''$ . The  $R$ -modules  $M/T(M)$  and  $M'/T(M')$  are free of ranks  $t' - s'$  and  $t'' - s''$  respectively by part (c) above. As  $\tilde{\alpha}: M/T(M) \cong M'/T(M')$  is an isomorphism, where  $(T(M) + v)\tilde{\alpha} = T(M') + (v)\alpha$  for  $v \in M$ , we see  $t' - s' = t'' - s''$  by (2.25). It is therefore enough to prove  $s' = s''$  and  $d_i \equiv d'_i$  for  $1 \leq i \leq s'$  as then  $t' = t''$  and  $d_i = 0 = d'_i$  for  $s' < i \leq t'$  (and so  $d_i \equiv d'_i$  for  $1 \leq i \leq t'$ ). Let  $r$  denote the number of associate classes represented by  $d_1, d_2, \dots, d_{s'}$ . The proof is by induction on  $r$ . It's convenient to adopt notation analogous to that used in (3.7). For each  $R$ -module  $M$  and element  $d \in R$  write  $\mu_d: M \rightarrow M$  for the  $R$ -linear mapping defined by  $(v)\mu_d = dv$  for all  $v \in M$ . Write  $dM = \text{im } \mu_d$  and  $M_{(d)} = \ker \mu_d$ . As  $\mu_d\alpha = \alpha\mu_d$  we obtain  $\alpha|_T: dM \cong dM'$  and  $\alpha|_T: M_{(d)} \cong M'_{(d')}$ . As  $N_i$  is a cyclic  $R$ -module with generator of order  $d_i$  we see  $N_i \cong R/\langle d_i \rangle$  for  $1 \leq i \leq s'$ . Also  $dN_i \cong R/\langle d_i/\gcd\{d, d_i\} \rangle$  and  $(N_i)_{(d)} \cong R/\langle \gcd\{d, d_i\} \rangle$  for  $1 \leq i \leq s'$ . For the same reason  $N'_i \cong R/\langle d'_i \rangle$ ,  $dN'_i \cong R/\langle d'_i/\gcd\{d, d'_i\} \rangle$  and  $(N'_i)_{(d')}$  for  $1 \leq i \leq s''$ . The isomorphism  $\alpha|_T: T(M)_{(d_1)} \cong T(M')_{(d'_1)}$  gives

$\alpha|_T: (R/\langle d_1 \rangle)^{s'} \cong R/\langle \gcd\{d_1, d'_1\} \rangle \oplus R/\langle \gcd\{d_1, d'_2\} \rangle \oplus \dots \oplus R/\langle \gcd\{d_1, d'_{s''} \rangle$  which shows that the free  $R/\langle d_1 \rangle$ -module  $(R/\langle d_1 \rangle)^{s'}$  of rank  $s'$  is generated by  $s''$  of its elements, namely  $u_1, u_2, \dots, u_{s''}$  where  $(u_i)\alpha = e_i$  for  $1 \leq i \leq s''$ . So  $s'' \geq s'$  by (2.20). Using the isomorphism  $\alpha^{-1}|_T: T(M')_{(d'_1)} \cong T(M)_{(d_1)}$  we obtain  $s' \geq s''$  and so  $s' = s''$ . By Exercises 2.3, Question 7 (c) the  $s'$  elements  $u_1, u_2, \dots, u_{s'}$  form an  $R/\langle d_1 \rangle$ -basis of  $(R/\langle d_1 \rangle)^{s'}$  and so  $u_1$  has order  $d_1$  in the  $R$ -module  $T(M)_{(d_1)}$ . But  $e_1$  has order  $\gcd\{d_1, d'_1\}$  in the  $R$ -module  $T(M')_{(d'_1)}$ . As  $u_1$  and  $(u_1)\alpha = e_1$  have the same order we see  $d_1 \equiv \gcd\{d_1, d'_1\}$ , that is,  $d_1 | d'_1$ . Interchanging the roles of  $T(M)$  and  $T(M')$  gives  $d'_1 | d_1$  and so  $d_1 \equiv d'_1$  by Question 7 (c)(ii) above.

Let  $m_1$  denote the number of  $i$  with  $d_i \equiv d'_1$  and let  $m'_1$  denote the number of  $i$  with  $d'_i \equiv d_1$ . As the rings  $R/\langle d_i/d_1 \rangle = R/R$  are trivial for  $1 \leq i \leq m_1$  we see  $d_1 M \cong \sum_{m_1 < i \leq s} \oplus R/\langle d_i/d_1 \rangle$  showing that the  $R$ -module  $d_1 M$  is the direct sum of  $s - m_1$  non-trivial cyclic submodules. In the same way  $d_1 M' \cong \sum_{m'_1 < i \leq s} \oplus R/\langle d'_i/d_1 \rangle$  is the direct sum of  $s - m'_1$  non-trivial cyclic submodules.

In the case  $r=1$  the modules  $d_1M$  and  $d_1M'$  are trivial and so  $m_1=m'_1=s$  and  $d_i \equiv d'_i$  for  $1 \leq i \leq s$  as  $d_i \equiv d_1 \equiv d'_1 \equiv d'_i$ . Suppose  $r > 1$ . The elements  $d_i/d_1$  for  $i > m_1$  belong to  $r-1$  associate classes. Also  $d_i/d_1 \mid d_j/d_1$  for  $m_1 < i \leq j \leq s$  and  $d'_i/d_1 \mid d'_j/d_1$  for  $m'_1 < i \leq j \leq s$ . As  $d_1M \cong d_1M'$  the inductive hypothesis applies to  $d_1M$  and  $d_1M'$  giving  $s-m_1=s-m'_1$ . So  $m_1=m'_1$  and  $d_i/d_1 \equiv d'_i/d_1$  for  $m_1 < i \leq s$ . Multiplication of this equivalence by  $d_1$  gives  $d_i \equiv d'_i$  for  $m_1 < i \leq s$ . As  $d_i \equiv d'_i$  for  $1 \leq i \leq m_1$  we conclude  $d_i \equiv d'_i$  for  $1 \leq i \leq s$ . The induction is now complete.

Conversely suppose  $t'=t''$  and  $d_i \equiv d'_i$  for  $1 \leq i \leq t'$ . Then  $\langle d_i \rangle = \langle d'_i \rangle$  for  $1 \leq i \leq t'$ . Hence  $M \cong M'$  as both  $M$  and  $M'$  are isomorphic to  $R/\langle d_1 \rangle \oplus R/\langle d_2 \rangle \oplus \dots \oplus R/\langle d_{t'} \rangle$ .

### Solutions 3.2 (page 130)

#### Solution 1

(a)  $G_2 = \{\bar{0}, \bar{5}\}$ ,  $G_5 = \{\bar{0}, \bar{2}, \bar{4}, \bar{6}, \bar{8}\}$ . Each  $g \in G$  is expressible  $g = g_2 + g_5$  where  $g_2 \in G_2$ ,  $g_5 \in G_5$ :  $\bar{0} = \bar{0} + \bar{0}$ ,  $\bar{1} = \bar{5} + \bar{6}$ ,  $\bar{2} = \bar{0} + \bar{2}$ ,  $\bar{3} = \bar{5} + \bar{8}$ ,  $\bar{4} = \bar{0} + \bar{4}$ ,  $\bar{5} = \bar{5} + \bar{0}$ ,  $\bar{6} = \bar{0} + \bar{6}$ ,  $\bar{7} = \bar{5} + \bar{2}$ ,  $\bar{8} = \bar{0} + \bar{8}$ ,  $\bar{9} = \bar{5} + \bar{4}$ . So  $G = G_2 + G_5$ . By inspection  $G_2 \cap G_5 = \{\bar{0}\}$ . Hence  $G_2, G_5$  are independent submodules of  $G$  as  $g_2 + g_5 = \bar{0} \Rightarrow g_2 = -g_5 \in G_2 \cap G_5 = \{\bar{0}\} \Rightarrow g_2 = g_5 = \bar{0}$ . So  $G = G_2 \oplus G_5$  by (2.15).

$$(G_2)\mu_3 = 3G_2 = \{3 \times \bar{0}, 3 \times \bar{5}\} = \{\bar{0}, \bar{5}\} = G_2 \text{ and}$$

$$(G_5)\mu_3 = 3G_5 = \{3 \times \bar{0}, 3 \times \bar{2}, 3 \times \bar{4}, 3 \times \bar{6}, 3 \times \bar{8}\} = \{\bar{0}, \bar{6}, \bar{2}, \bar{8}, \bar{4}\} = G_5.$$

By Exercises 2.1, Question 4(a) we have  $\text{Aut } G = \{\mu_1, \mu_3, \mu_7, \mu_9\}$ . As  $\mu_3^2 = \mu_9$ ,  $\mu_3^3 = \mu_7$ ,  $\mu_3^4 = \mu_1$  we see  $\text{Aut } G$  is cyclic with generator  $\mu_3$ .

(b)  $G_2 = \{g \in G : 4g = 0\} = \{\bar{0}, \bar{3}, \bar{6}, \bar{9}\}$ ,  $G_3 = \{\bar{0}, \bar{4}, \bar{8}\}$ . As  $\bar{1} = \bar{9} + \bar{4} \in G_2 + G_3$ , on multiplying this equation by  $i$  for  $1 < i \leq 12$  we obtain  $\bar{i} \in G_2 + G_3$  for all  $\bar{i} \in \mathbb{Z}_{12} = G$  showing  $G = G_2 + G_3$ . As above  $G_2 \cap G_3 = \{\bar{0}\}$  and  $G = G_2 \oplus G_3$  by (2.15).

As  $G_2$  and  $G_3$  are of isomorphism type  $C_4$  and  $C_3$  respectively,  $G$  has elementary divisors 4, 3.

$$(G_2)\mu_5 = 5G_2 = \{5 \times \bar{0}, 5 \times \bar{3}, 5 \times \bar{6}, 5 \times \bar{9}\} = \{\bar{0}, \bar{3}, \bar{6}, \bar{9}\} = G_2 \text{ and}$$

$$(G_3)\mu_5 = 5G_3 = \{5 \times \bar{0}, 5 \times \bar{4}, 5 \times \bar{8}\} = \{\bar{0}, \bar{8}, \bar{4}\} = G_3.$$

By Exercises 2.1, Question 4(a) we see  $\text{Aut } G = \{\mu_1, \mu_5, \mu_7, \mu_{11}\}$  which is not cyclic as  $\mu_1^2 = \mu_5^2 = \mu_7^2 = \mu_{11}^2 = \mu_1$ : in fact  $\text{Aut } G$  is a Klein 4-group in multiplicative notation.

(c) By (2.2) and (3.10) the  $p_i$ -component  $G_{p_i}$  of the cyclic group  $G$  is cyclic of order  $p_i^{n_i}$  for  $1 \leq i \leq k$ . So  $G$  has  $k$  elementary divisors  $p_1^{n_1}, p_2^{n_2}, \dots, p_k^{n_k}$ .

(d) Let  $g, g' \in G_p$ . Then  $p^n g = 0$  and  $p^n g' = 0$ . Adding gives

$$p^n(g + g') = p^n g + p^n g' = 0 + 0 = 0 \text{ which shows } g + g' \in G_p. \text{ As } p^n 0 = 0 \text{ we see that } 0 \in G_p.$$

Also  $p^n(-g) = p^n(-g) + 0 = p^n(-g) + p^n g = p^n(-g + g) = p^n 0 = 0$  showing  $-g \in G_p$ . So  $G_p$  is a subgroup of  $G$ .

There are integers  $a, b$  with  $ap^n + bm = 1$  since  $\gcd\{p^n, m\} = 1$ . For  $g \in G_p$  we have

$$g = 1 \times g = (ap^n + bm)g = ap^n g + bmg = 0 + bmg = m(bg) \in mG \text{ showing } G_p \subseteq mG. \text{ Conversely}$$

consider  $g \in mG$ . Then  $g = mg'$  for some  $g' \in G$ . So  $p^n g = p^n mg' = |G|g' = 0$  showing  $mG \subseteq G_p$  and so  $mG = G_p$ .

Yes,  $G_p = mG_p$  as above  $bg \in G_p$  and so  $G_p \subseteq mG_p$ ; also  $mG_p$  is a subgroup of  $G_p$ .

#### Solution 2

(a) By the Chinese remainder theorem (2.11)  $\mathbb{Z}_{10} \cong \mathbb{Z}_2 \oplus \mathbb{Z}_5$ ,  $\mathbb{Z}_{12} \cong \mathbb{Z}_4 \oplus \mathbb{Z}_3$  and

$$\mathbb{Z}_{60} \cong \mathbb{Z}_5 \oplus \mathbb{Z}_4 \oplus \mathbb{Z}_3. \text{ Hence we obtain the ring isomorphisms}$$

$$\mathbb{Z}_{10} \oplus \mathbb{Z}_{12} \cong \mathbb{Z}_2 \oplus \mathbb{Z}_5 \oplus \mathbb{Z}_4 \oplus \mathbb{Z}_3 \cong \mathbb{Z}_2 \oplus \mathbb{Z}_{60}. \text{ Comparing additive subgroups we see that } G \text{ has}$$

invariant factors 2, 60. So  $G_2$  has invariant factors (and elementary divisors) 2, 4.  $G$  has elementary divisors 2, 4, 3, 5.

There are  $p(3) \times p(1) \times p(1) = 3 \times 1 \times 1 = 3$  isomorphism classes of abelian groups of order  $120 = 2^3 \times 3^1 \times 5^1$  their elementary divisors being  $2, 2, 2; 3; 5$  or  $2, 4; 3; 5$  or  $8; 3; 5$ . The corresponding invariant factor sequences are

$$(2, 2, 2 \times 3 \times 5) = (2, 2, 30), (2, 4 \times 3 \times 5) = (2, 60), (8 \times 3 \times 5) = (120).$$

(b) As  $200 = 2^3 \times 5^2$ , there are  $p(3) \times p(2) = 3 \times 2 = 6$  isomorphism classes of abelian groups of order 200. The elementary divisors of these classes are

$$2, 2, 2; 5, 5 \quad 2, 4; 5, 5 \quad 8; 5, 5 \quad 2, 2, 2; 25 \quad 2, 4; 25 \quad 8; 25$$

and the corresponding invariant factor sequences are

$$(2, 2 \times 5, 2 \times 5) = (2, 10, 10), (2 \times 5, 4 \times 5) = (10, 20),$$

$$(5, 8 \times 5) = (5, 40), (2, 2, 2 \times 25) = (2, 2, 50), (2, 4 \times 25) = (2, 100), (8 \times 25) = (200).$$

Their exponents are 10, 20, 40, 50, 100, 200 and so no two classes have the same exponent.

(c) The  $p(4) = 5$  isomorphism types of abelian groups of order  $16 = 2^4$  are

$$C_2 \oplus C_2 \oplus C_2 \oplus C_2, C_2 \oplus C_2 \oplus C_4, C_2 \oplus C_8, C_4 \oplus C_4, C_8.$$

The second and fourth have exponent 4 and so the answer is: Yes. As  $900^2 = 2^4 \times 3^4 \times 5^4$  the elementary divisors of the 8 isomorphism classes of abelian groups of order  $900^2$  having exponent 900 are

$$2, 2, 4; 3, 3, 9; 5, 5, 25 \quad 2, 2, 4; 3, 3, 9; 25, 25 \quad 2, 2, 4; 9, 9; 5, 5, 25 \quad 2, 2, 4; 9, 9; 25, 25 \\ 4, 4; 3, 3, 9; 5, 5, 25 \quad 4, 4; 3, 3, 9; 25, 25 \quad 4, 4; 9, 9; 5, 5, 25 \quad 4, 4; 9, 9; 25, 25$$

and the corresponding invariant factor sequences are

$$(2 \times 3 \times 5, 2 \times 3 \times 5, 4 \times 9 \times 25) = (30, 30, 900), (2 \times 3, 2 \times 3 \times 25, 4 \times 9 \times 25) = (6, 150, 900),$$

$$(2 \times 5, 2 \times 5 \times 9, 4 \times 9 \times 25) = (10, 90, 900), (2, 2 \times 9 \times 25, 4 \times 9 \times 25) = (2, 450, 900),$$

$$(3 \times 5, 4 \times 3 \times 5, 4 \times 9 \times 25) = (15, 60, 900), (3, 3 \times 4 \times 25, 4 \times 9 \times 25) = (3, 300, 900),$$

$$(5, 4 \times 9 \times 5, 4 \times 9 \times 25) = (5, 180, 900), (4 \times 9 \times 25, 4 \times 9 \times 25) = (900, 900).$$

(d) As  $400 = 2^4 \times 5^2$ , there are  $p(4) \times p(2) = 5 \times 2 = 10$  isomorphism classes of abelian groups of order 400. The elementary divisors of these classes are

$$2, 2, 2, 2; 5, 5 \quad 2, 2, 4; 5, 5 \quad 2, 8; 5, 5 \quad 4, 4; 5, 5 \quad 16; 5, 5 \quad 2, 2, 2, 2; 25 \quad 2, 2, 4; 25 \quad 2, 8; 25 \\ 4, 4; 25 \quad 16; 25 \text{ and the corresponding invariant factor sequences are}$$

$$(2, 2, 2 \times 5, 2 \times 5) = (2, 2, 10, 10), (2, 2 \times 5, 4 \times 5) = (2, 10, 20), (2 \times 5, 8 \times 5) = (10, 40),$$

$$(4 \times 5, 4 \times 5) = (20, 20), (5, 16 \times 5) = (5, 80), (2, 2, 2, 2 \times 25) = (2, 2, 2, 50),$$

$$(2, 2, 4 \times 25) = (2, 2, 100), (2, 8 \times 25) = (2, 200), (4, 4 \times 25) = (4, 100), (16 \times 8) = (400). \text{ As}$$

$144 = 2^4 \times 3^2$ , there are  $p(4) \times p(2) = 5 \times 2 = 10$  isomorphism classes of abelian groups of order 144.

(e) Among the  $p(6) = 11$  isomorphism classes of abelian groups of order  $64 = 2^6$  there are 3 having exponent 4, namely those with invariant factor sequences  $(2, 2, 2, 2, 4)$ ,  $(2, 2, 4, 4)$ ,  $(4, 4, 4)$ . So the additive groups of

$$\mathbb{Z}_2 \oplus \mathbb{Z}_2 \oplus \mathbb{Z}_2 \oplus \mathbb{Z}_2 \oplus \mathbb{Z}_4, \mathbb{Z}_2 \oplus \mathbb{Z}_2 \oplus \mathbb{Z}_4 \oplus \mathbb{Z}_4, \mathbb{Z}_4 \oplus \mathbb{Z}_4 \oplus \mathbb{Z}_4$$

provide an answer to the question. Another answer is: the additive groups of

$$\mathbb{Z}_2 \oplus \mathbb{Z}_2 \oplus \mathbb{Z}_2 \oplus \mathbb{Z}_8, \mathbb{Z}_2 \oplus \mathbb{Z}_4 \oplus \mathbb{Z}_8, \mathbb{Z}_8 \oplus \mathbb{Z}_8$$

each of which has order 64 and exponent 8.

### Solution 3

(a) The  $p(5) = 7$  partitions of 5 are

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

The  $p(9, 2) = 8$  partitions of 9 having all parts  $\geq 2$  are

$$(2, 2, 2, 3), (2, 2, 5), (2, 3, 4), (2, 7), (3, 3, 3), (3, 6), (4, 5), (9).$$

(b) As  $p(11) = p(10) + p(9, 2) + p(8, 3) + p(7, 4) + p(6, 5) + 1$  we see

$$p(11) = 42 + 8 + 3 + 1 + 1 + 1 = 56. \text{ Also } p(11, 2) = p(11) - p(10) = 56 - 42 = 14,$$

$$p(11, 3) = p(11, 2) - p(9, 2) = 14 - 8 = 6, \quad p(11, 4) = p(11, 3) - p(8, 3) = 6 - 3 = 3,$$

$p(11, 5) = p(11, 4) - p(7, 4) = 3 - 1 = 2$ . Finally  $p(11, j) = 1$  for  $6 \leq j \leq 11$  as (11) is the only partition of 11 with all parts  $\geq j$ , and  $p(11, j) = 0$  for  $j > 11$ .

As  $p(12) = p(11) + p(10, 2) + p(9, 3) + p(8, 4) + p(7, 5) + p(6, 6) + 1$  we see

$$p(12) = 56 + 12 + 4 + 2 + 1 + 1 + 1 = 77.$$

As  $p(13) = p(12) + p(10, 2) + p(9, 3) + p(8, 4) + p(7, 5) + p(6, 6) + 1$  we see

$$p(13) = 77 + 14 + 5 + 2 + 1 + 1 + 1 = 101.$$

As  $p(14) = p(13) + p(12, 2) + p(11, 3) + p(10, 4) + p(9, 5) + p(8, 6) + p(7, 7) + 1$  and

$$p(12, 2) = p(12) - p(11) = 21 \text{ we see } p(14) = 101 + 21 + 6 + 3 + 1 + 1 + 1 + 1 = 135.$$

(c) There are just 2 partitions of  $2n$  having all parts  $\geq n$  namely  $(n, n)$  and  $(2n)$ . So  $p(2n, n) = 2$ .

The partitions of  $3n$  with all parts  $\geq n$  have at most three parts. The partitions  $(n, n, n)$  and  $(3n)$  are the only such partitions having three parts and one part respectively. Such partitions with two parts are  $(n+k, 2n-k)$  for  $k \geq 0$  and  $n+k \leq 2n-k$ ; so  $0 \leq k \leq \lfloor n/2 \rfloor$ . Hence there are in all  $\lfloor n/2 \rfloor + 3$  partitions of  $3n$  with all parts  $\geq n$ .

(d) Let  $k$  be the number of parts 2 in a partition  $(1, 1, \dots, 1, 2, 2, \dots, 2)$  of  $n$  having all parts  $\leq 2$ .

Then  $0 \leq k \leq n/2$  and so the number of such partitions is  $\lfloor n/2 \rfloor + 1$ . The invariant factor sequence of an abelian group of order  $p^n$  and exponent  $p^2$  is  $(p, p, \dots, p, p^2, p^2, \dots, p^2)$ , the product of these invariant factors being  $p^n$  and the number  $k$  of invariant factors  $p^2$  being positive. So there are exactly  $\lfloor n/2 \rfloor$  isomorphism classes of abelian groups of order  $p^n$  having exponent  $p^2$ .

(e) Each partition of  $n$  having all parts  $\geq j$  and having exactly  $k$  parts  $j$  is of the form

$(j, j, \dots, j, t_1, t_2, \dots, t_s)$  where  $(t_1, t_2, \dots, t_s)$  is a partition of  $n - jk$  with all parts  $\geq j+1$ . So the

number of such partitions of  $n$  is  $p(n - jk, j+1)$ . Counting up the  $p(n, j)$  partitions of  $n$  having

all parts  $\geq j$ , by the number  $k$  of parts  $j$  each partition has, gives  $p(n, j) = \sum_{k=0}^{\lfloor n/j \rfloor} p(n - jk, j+1)$ .

So  $p(10) = p(10, 1) = \sum_{k=0}^{10} p(10 - k, 2)$ , i.e.  $42 = 12 + 8 + 7 + 4 + 4 + 2 + 2 + 1 + 1 + 0 + 1$ ,

$$p(10, 2) = \sum_{k=0}^5 p(10 - 2k, 3), \text{ i.e. } 12 = 5 + 3 + 2 + 1 + 0 + 1, \quad p(10, 3) = \sum_{k=0}^3 p(10 - 3k, 4), \text{ i.e.}$$

$$5 = 3 + 1 + 1 + 0, \quad p(10, 4) = \sum_{k=0}^2 p(10 - 4k, 5), \text{ i.e. } 3 = 2 + 1 + 0.$$

#### Solution 4

(a)  $G$  has 24 elements of order 5 namely its 24 non-zero elements.  $G$  has exponent 5 and so the answer is: Yes. Each non-zero element of  $G$  generates a subgroup of order 5. But each subgroup of order 5 is of the form  $\{0, g, 2g, 3g, 4g\} = \langle g \rangle = \langle 2g \rangle = \langle 3g \rangle = \langle 4g \rangle$ ,  $g \neq 0$ , which has exactly 4 generators. So  $G$  has  $24/4 = 6$  subgroups of order 5 and they can be expressed

$$\langle (\bar{1}, \bar{0}) \rangle, \langle (\bar{1}, \bar{1}) \rangle, \langle (\bar{1}, \bar{2}) \rangle, \langle (\bar{1}, \bar{3}) \rangle, \langle (\bar{1}, \bar{4}) \rangle, \langle (\bar{0}, \bar{1}) \rangle.$$

Any two  $H_1, H_2$  of these subgroups with  $H_1 \neq H_2$  satisfy  $H_1 \cap H_2 = \{(\bar{0}, \bar{0})\}$  and  $H_1 + H_2 = G$ , i.e.  $G = H_1 \oplus H_2$ . So there are 6 choices for  $H_1$  and 5 remaining choices for  $H_2$ . Hence there are  $6 \times 5 = 30$  ordered pairs of subgroups of order 5 such that  $G = H_1 \oplus H_2$ .

(b) Yes,  $G$  has exponent  $p$  and so is an elementary abelian  $p$ -group. As  $pG = \{0\}$  the  $|G| - 1 = p^2 - 1$  non-zero elements of  $G$  are each of order  $p$ . Each subgroup  $H$  of  $G$  with  $|H| = p$  is of the form  $H = \{g, 2g, \dots, (p-1)g, pg\} = \langle g \rangle = \langle 2g \rangle = \dots = \langle (p-1)g \rangle$  where  $g \neq 0$ . So each such subgroup  $H$  contains  $p-1$  non-zero elements of  $G$ . Hence  $G$  has exactly  $(p^2 - 1)/(p - 1) = p + 1$  subgroups  $H$  of order  $p$ . Each  $H$  has a generator  $g = (\bar{r}, \bar{s}) \in \mathbb{Z}_p \oplus \mathbb{Z}_p$  where either  $\bar{r} \neq \bar{0}$  or  $\bar{s} \neq \bar{0}$ . For  $\bar{r} \neq \bar{0}$  then  $H = \langle (\bar{r}, \bar{s}) \rangle = \langle (\bar{1}, \bar{t}) \rangle$  where  $\bar{t} = \bar{s}/\bar{r} \in \mathbb{Z}_p$ . For  $\bar{r} = \bar{0}$  then  $H = \langle (\bar{0}, \bar{s}) \rangle = \langle (\bar{0}, \bar{1}) \rangle$ . By Lagrange's theorem the subgroups of  $G$  have orders  $1, p, p^2$ . Let  $H_1$  and  $H_2$  be different subgroups of order  $p$ . Then  $H_1 \cap H_2 = \{0\}$  as  $|H_1 \cap H_2| < p$  and  $H_1 + H_2 = G$  as  $|H_1 + H_2| > p$ . So  $G = H_1 \oplus H_2$  and there are  $(p+1)p$  ordered pairs  $H_1, H_2$  of such subgroups, since there are  $p+1$  choices for  $H_1$  (any subgroup of order  $p$ ) and  $p$  remaining choices for  $H_2$  (any subgroup, except  $H_1$ , of  $G$  having order  $p$ ).

(c)  $H = \{(\bar{0}, \bar{0}), (\bar{0}, \bar{9}), (\bar{0}, \bar{18}), (\bar{3}, \bar{0}), (\bar{3}, \bar{9}), (\bar{3}, \bar{18}), (\bar{6}, \bar{0}), (\bar{6}, \bar{9}), (\bar{6}, \bar{18})\}$ . Yes,  $H = \langle (\bar{3}, \bar{0}), (\bar{0}, \bar{9}) \rangle$  is a subgroup of  $G$ . Yes,  $H$  is an elementary abelian 3-group. The invariant factor sequences of  $G$  and  $H$  are  $(9, 27)$  and  $(3, 3)$  respectively. As

$$G/H = (\mathbb{Z}_9 \oplus \mathbb{Z}_{27}) / (\langle (\bar{3}, \bar{0}) \rangle \oplus \langle (\bar{0}, \bar{9}) \rangle) \cong \mathbb{Z}_9 / \langle \bar{3} \rangle \oplus \mathbb{Z}_{27} / \langle 9 \rangle$$

has isomorphism class  $C_3 \oplus C_9$ , we see that  $G/H$  has invariant factor sequence  $(3, 9)$ .

(d) A typical element of  $p^{t-1}G$  is  $p^{t-1}g$  where  $g \in G$ . As  $p(p^{t-1}g) = p^t g = |G|g = 0$  we see  $p^{t-1}G \subseteq H$ . Let  $G = \langle g_0 \rangle$  and consider  $h \in H$ . Then  $h = mg_0$  for some  $m \in \mathbb{Z}$  as  $H \subseteq G$ . The order of  $g_0$  is  $p^t$  and the order of  $h$  is a divisor of  $p$  (either 1 or  $p$ ) since  $ph = 0$ . By (2.7)  $p^t / \gcd\{m, p^t\}$  is a divisor of  $p$  and so  $p^{t-1}$  is a divisor of  $\gcd\{m, p^t\}$ . Hence  $m = p^{t-1}m'$  where  $m' \in \mathbb{Z}$ . So  $h = p^{t-1}m'g_0 \in p^{t-1}G$  showing  $H \subseteq p^{t-1}G$  and so  $H = p^{t-1}G$ . In fact  $H = \langle p^{t-1}g_0 \rangle$  is cyclic of order  $p$  and  $|G/H| = |G|/|H| = p^t/p = p^{t-1}$ .

(e) Suppose  $h \in H$ . Then  $h = m_1h_1 + m_2h_2 + \dots + m_sh_s$  where  $m_i \in \mathbb{Z}$ ,  $h_i \in H_i$  for  $1 \leq i \leq s$ . As  $H_1, H_2, \dots, H_s$  are independent,  $ph = 0$  implies  $pm_ih_i = 0$  for  $1 \leq i \leq s$ . Using (d) above,  $m_i = p^{t_i-1}m'_i$  where  $m'_i \in \mathbb{Z}$  for  $1 \leq i \leq s$ . So  $m_ih_i = p^{t_i-1}m'_ih_i \in p^{t_i-1}H_i$  as  $m'_ih_i \in H_i$ . We've shown  $H \subseteq p^{t_1-1}H_1 \oplus p^{t_2-1}H_2 \oplus \dots \oplus p^{t_s-1}H_s$ . As  $p(p^{t_i-1}H_i) = \{0\}$  for  $1 \leq i \leq s$  we see  $p^{t_1-1}H_1 \oplus p^{t_2-1}H_2 \oplus \dots \oplus p^{t_s-1}H_s \subseteq H$  and so  $H = p^{t_1-1}H_1 \oplus p^{t_2-1}H_2 \oplus \dots \oplus p^{t_s-1}H_s$ . By (d) above each  $p^{t_i-1}H_i$  is cyclic of order  $p$  and so  $H$  is elementary abelian of order  $p^s$ . The invariant factor sequence of  $H$  is  $(p, p, \dots, p)$  having  $s$  terms. Suppose the first  $l$  parts of the partition  $(t_1, t_2, \dots, t_s)$  are 1 and  $t_{l+1} \geq 2$  ( $0 \leq l \leq s$ ). The invariant factor sequence of  $G/H$  is  $(p^{t_{l+1}-1}, p^{t_{l+2}-1}, \dots, p^{t_s-1})$  as by (3.2)

$$\frac{G}{H} = \frac{H_1 \oplus H_2 \oplus \dots \oplus H_s}{p^{t_1-1}H_1 \oplus p^{t_2-1}H_2 \oplus \dots \oplus p^{t_s-1}H_s} \cong \frac{H_1}{p^{t_1-1}H_1} \oplus \frac{H_2}{p^{t_2-1}H_2} \oplus \dots \oplus \frac{H_s}{p^{t_s-1}H_s}$$

(exceptionally using horizontal slashes for typographical convenience),  $H_i/p^{t_i-1}H_i$  being cyclic of order  $p^{t_i-1}$  by (d) above.  $G/H$  is an elementary abelian  $p$ -group if and only if  $G/H$  is non-trivial and all its invariant factors are  $p$ , which is so  $\Leftrightarrow t_s = 2$ .

(f) First consider an abelian group  $G$  of order  $p_1^{n_1}$  and let  $t_1$  be an integer with  $1 \leq t_1 \leq n_1$ . The additive group of  $\mathbb{Z}_{p_1} \oplus \mathbb{Z}_{p_1} \oplus \dots \oplus \mathbb{Z}_{p_1} \oplus \mathbb{Z}_{p_1^{t_1}}$  having  $n_1 - t_1$  summands  $\mathbb{Z}_{p_1}$  for  $t_1 \geq 2$  (and  $n_1$



summands  $\mathbb{Z}_{p_i}$  for  $t_i = 1$ ) has order  $p_1^{n_1}$  and exponent  $p_1^{t_1}$ . So there are  $n_1$  possibilities for the exponent of  $G$ , namely  $p, p^2, \dots, p^{n_1}$ . By (3.12) there are no further possibilities. In general  $|G| = p_1^{n_1} p_2^{n_2} \cdots p_k^{n_k}$  and there are  $n_j$  possibilities for the exponent of the  $p_j$ -component  $G_{p_j}$  of  $G$ . By (3.10)  $G = G_{p_1} \oplus G_{p_2} \oplus \dots \oplus G_{p_k}$  and so the largest invariant factor of  $G$  is the product of the largest invariant factors of the primary components  $G_{p_j}$ . By (3.12) the exponent of  $G$  is the product of the exponents of the primary components  $G_{p_j}$  for  $1 \leq j \leq k$ . As the exponents of the primary components are pair-wise coprime we see that there are  $n_1 n_2 \cdots n_k$  possibilities for the exponent of  $G$ , namely the positive integers  $e$  with  $p_1 p_2 \cdots p_k \mid e$  and  $e \mid p_1^{n_1} p_2^{n_2} \cdots p_k^{n_k}$ .

No,  $H$  is not necessarily cyclic; the subgroup  $H$  in (c) above is not cyclic. Suppose  $H$  is cyclic. Then each primary component  $H_{p_j}$  of  $H$  is cyclic. But  $H_{p_j} = \{h_j \in G_{p_j} : p_j h_j = 0\}$ . By (e) above  $G_{p_j}$  is cyclic for  $1 \leq j \leq k$  and so  $G$  is cyclic. So  $H$  cyclic implies  $G$  cyclic.

### Solution 5

(a) Let  $g, g' \in G = H_1 \oplus H_2$ . Then  $g = h_1 + h_2$ ,  $g' = h'_1 + h'_2$  where  $h_1, h'_1 \in H_1$  and  $h_2, h'_2 \in H_2$ . So

$$(g + g')\alpha = (h_1 + h_2 + h'_1 + h'_2)\alpha = ((h_1 + h'_1) + (h_2 + h'_2))\alpha = (h_1 + h'_1)\alpha_1 + (h_2 + h'_2)\alpha_2 = (h_1)\alpha_1 + (h'_1)\alpha_1 + (h_2)\alpha_2 + (h'_2)\alpha_2 = ((h_1)\alpha_1 + (h_2)\alpha_2) + ((h'_1)\alpha_1 + (h'_2)\alpha_2) = (g)\alpha + (g')\alpha$$

showing that  $\alpha$  is additive. Consider  $\gamma: G \rightarrow G$  defined by  $(h_1 + h_2)\gamma = (h_1)\alpha_1^{-1} + (h_2)\alpha_2^{-1}$  for all  $h_1 \in H_1, h_2 \in H_2$ . Then

$$(g)\alpha\gamma = (h_1 + h_2)\alpha\gamma = ((h_1)\alpha_1 + (h_2)\alpha_2)\gamma = (h_1)\alpha_1\alpha_1^{-1} + (h_2)\alpha_2\alpha_2^{-1} = h_1 + h_2 = g$$

and similarly  $(g)\gamma\alpha = g$  for all  $g \in G$ . So  $\gamma = \alpha^{-1}$  showing  $\alpha$  to be bijective. Hence  $\alpha$  is an automorphism of  $G$ . Also  $(h_1)\alpha = (h_1 + 0)\alpha = (h_1)\alpha_1 + (0)\alpha_2 = (h_1)\alpha_1 + 0 = (h_1)\alpha_1 \in H_1$  showing  $(H_1)\alpha \subseteq H_1$ . Replacing  $\alpha$  by  $\alpha^{-1}$  gives  $(H_1)\alpha^{-1} \subseteq H_1$ . Applying  $\alpha$  to the last set containment gives  $H_1 = (H_1)\alpha^{-1}\alpha \subseteq (H_1)\alpha$  and so  $(H_1)\alpha = H_1$ . Similarly  $(H_2)\alpha = H_2$ .

Define  $\beta_1: H_1 \rightarrow H_1$  by  $(h_1)\beta_1 = (h_1)\beta$  for all  $h_1 \in H_1$ . Then  $\beta_1$  (the restriction of  $\beta$  to  $H_1$ ) is additive and bijective. So  $\beta_1 \in \text{Aut } H_1$ . In the same way define  $\beta_2: H_2 \rightarrow H_2$  by  $(h_2)\beta_2 = (h_2)\beta$  for all  $h_2 \in H_2$ . Then  $\beta_2$  (the restriction of  $\beta$  to  $H_2$ ) is additive and bijective. So  $\beta_2 \in \text{Aut } H_2$ . Then  $(g)\beta = (h_1 + h_2)\beta = (h_1)\beta + (h_2)\beta = (h_1)\beta_1 + (h_2)\beta_2 = (g)(\beta_1 \oplus \beta_2)$  for all  $g \in G$ .

Therefore  $\beta = \beta_1 \oplus \beta_2$ .

Yes,  $L$ , being closed under composition and inversion and containing the identity mapping of  $G$ , is a subgroup of  $\text{Aut } G$ . Yes,  $L \cong \text{Aut } H_1 \times \text{Aut } H_2$  as the correspondence  $\beta \leftrightarrow (\beta_1, \beta_2)$  where  $\beta = \beta_1 \oplus \beta_2$  is a group isomorphism.

(b) Let  $g, g' \in G$ . By (3.10) there are unique elements  $g_j, g'_j \in G_{p_j}$  ( $1 \leq j \leq k$ ) with

$$g = g_1 + g_2 + \dots + g_k \text{ and } g' = g'_1 + g'_2 + \dots + g'_k. \text{ Then}$$

$$(g + g')\alpha = \left( \sum_{j=1}^k (g_j + g'_j) \right) \alpha = \sum_{j=1}^k (g_j + g'_j) \alpha_j = \sum_{j=1}^k ((g_j)\alpha_j + (g'_j)\alpha_j) = \sum_{j=1}^k (g_j)\alpha_j + \sum_{j=1}^k (g'_j)\alpha_j = (g)\alpha + (g')\alpha$$

showing that  $\alpha: G \rightarrow G$  is a homomorphism. As  $\alpha^{-1} = \alpha_1^{-1} \oplus \alpha_2^{-1} \oplus \dots \oplus \alpha_k^{-1}$  we see that  $\alpha$  is bijective and so  $\alpha \in \text{Aut } G$ . Consider  $\beta \in \text{Aut } G$ . As  $(G_{p_j})\beta = G_{p_j}$  the mapping  $\beta_j: G_{p_j} \rightarrow G_{p_j}$  defined by  $(g_j)\beta_j = (g_j)\beta$  for all  $g_j \in G_{p_j}$  is an automorphism of  $G_{p_j}$  for  $1 \leq j \leq k$ . Then

$$(g)\beta = \left(\sum_{j=1}^k g_j\right)\beta = \sum_{j=1}^k (g_j)\beta = \sum_{j=1}^k (g_j)\beta_j = (g)(\beta_1 \oplus \beta_2 \oplus \dots \oplus \beta_k)$$

for all  $g \in G$  showing  $\beta = \beta_1 \oplus \beta_2 \oplus \dots \oplus \beta_k$ . As  $\beta_j$  is the restriction of  $\beta$  to  $G_{p_j}$  for  $1 \leq j \leq k$  we see that the  $\beta_j$  are uniquely determined by  $\beta$ . Hence the mapping

$\text{Aut } G \rightarrow \text{Aut } G_{p_1} \times \text{Aut } G_{p_2} \times \dots \times \text{Aut } G_{p_k}$ , defined by  $\beta \rightarrow (\beta_1, \beta_2, \dots, \beta_k)$  for all  $\beta \in \text{Aut } G$ , is bijective. Let  $\beta, \beta' \in \text{Aut } G$ . There are  $\beta'_j \in \text{Aut } G_{p_j}$  ( $1 \leq j \leq k$ ) with

$\beta' = \beta'_1 \oplus \beta'_2 \oplus \dots \oplus \beta'_k$ . Then

$$(g)\beta\beta' = ((g)\beta)\beta' = \left(\sum_{j=1}^k (g_j)\beta_j\right)\beta' = \sum_{j=1}^k ((g_j)\beta_j)\beta'_j = \sum_{j=1}^k (g_j)\beta_j\beta'_j$$

showing that the correspondence  $\beta \leftrightarrow (\beta_1, \beta_2, \dots, \beta_k)$  is a group homomorphism. So

$$\text{Aut } G \cong \text{Aut } G_{p_1} \times \text{Aut } G_{p_2} \times \dots \times \text{Aut } G_{p_k}$$

is a group isomorphism.

### Solution 6

(a) Suppose  $H$  is indecomposable. Let  $H$  have  $t'$  invariant factors. Then  $t' = 1$  by (3.8) as otherwise  $H = H_1 \oplus (H_2 \oplus \dots \oplus H_{t'})$  with both  $H_1$  and  $H_2 \oplus \dots \oplus H_{t'}$  non-trivial. So  $H$  is cyclic of isomorphism type  $C_d$  where  $d \neq 1$ . Either  $d = 0$  or  $d \geq 2$ . In the latter case  $|H| = d$  is divisible by just one prime, as otherwise the primary decomposition (3.10) of  $H$  would be a non-trivial decomposition contradicting the fact that  $H$  is indecomposable. So  $d = p^n$  where  $p$  is prime. Conversely suppose  $H$  has isomorphism type  $C_d$  where either  $d = 0$  or  $d = p^n$  where  $p$  is prime. In the case  $d = 0$  we see  $H \cong \mathbb{Z}$ , that is,  $H$  is isomorphic to the additive group  $\mathbb{Z}$  of integers; but  $\mathbb{Z}$  (and hence  $H$ ) is indecomposable by Exercises 2.2, Question 6(b). By the discussion following (3.12), cyclic groups of prime power order  $d = p^n$  are indecomposable.

(b) The submodule  $H_i$  of the f.g.  $\mathbb{Z}$ -module  $G$  is itself f.g. by Exercises 3.1, Question 5(b) for  $1 \leq i \leq m$ . Let  $r_i$  be the torsion-free rank of  $H_i$  and let the torsion subgroup  $T_i$  of  $H_i$  have  $l_i$  elementary divisors. Then  $l_i + r_i \geq 1$  and, using (3.4), (3.5) and (3.10) we see  $H_i$  is the direct sum of  $l_i + r_i$  non-trivial indecomposable submodules for  $1 \leq i \leq m$ . Substituting for each  $H_i$  we obtain a decomposition  $G = H'_1 \oplus H'_2 \oplus \dots \oplus H'_{m'}$  where  $m \leq m'$  and each  $H'_{i'}$  is indecomposable for  $1 \leq i' \leq m'$ . By (a) above each  $H'_{i'}$  has isomorphism type  $C_0$  or  $C_{p^n}$ . The number of  $H'_{i'}$  having isomorphism type  $C_0$  is  $r$  the torsion-free rank of  $G$ . Now  $H'_{i'}$  has isomorphism type  $C_{p^n}$  if and only if  $p^n$  is an elementary divisor of the torsion subgroup  $T$  of  $G$ ; so the number of such  $H'_{i'}$  is  $l$ . Therefore  $m' = l + r$  and so  $m \leq l + r$ . From above we see  $l_1 + l_2 + \dots + l_m = l$  as  $T_1 \oplus T_2 \oplus \dots \oplus T_m = T$ . Also  $r_1 + r_2 + \dots + r_m = r$  on comparing torsion-free ranks in

$$H_1 \oplus H_2 \oplus \dots \oplus H_m = G.$$

Suppose  $m = l + r$ . Then  $m = m'$  which means  $l_i + r_i = 1$  for  $1 \leq i \leq m$ . So either  $r_i = 1$  or  $l_i = 1$ . In the first case  $H_i$  is indecomposable being of isomorphism type  $C_0$ . In the second case  $H_i$  is

indecomposable being of isomorphism type  $C_{p^n}$ . So  $m = l + r$  implies that  $H_i$  is indecomposable for  $1 \leq i \leq m$ . Conversely suppose  $H_i$  is indecomposable for  $1 \leq i \leq m$ . Then  $H'_i = H_i$  for  $1 \leq i \leq m$  and so  $m = m' = l + r$ .  
 Conversely suppose  $H_i$  is indecomposable for  $1 \leq i \leq m$ . Then  $H'_i = H_i$  for  $1 \leq i \leq m$  and so  $m = m' = l + r$ .

### Solutions 3.3 (page 146)

#### Solution 1

(a) For  $g_1, g_2$  in  $G$  by (2.3) we have

$$\begin{aligned}(g_1 + g_2)(\alpha + \alpha') &= (g_1 + g_2)\alpha + (g_1 + g_2)\alpha' = (g_1)\alpha + (g_2)\alpha + (g_1)\alpha' + (g_2)\alpha' = \\ &= (g_1)\alpha + (g_1)\alpha' + (g_2)\alpha + (g_2)\alpha' = (g_1)(\alpha + \alpha') + (g_2)(\alpha + \alpha')\end{aligned}$$

showing that  $\alpha + \alpha'$  is an endomorphism of  $G$ .

We verify the axioms of an additive abelian group (stated at the start of Chapter 2.1). Consider  $\alpha, \alpha', \alpha''$  in  $\text{End } G$ . Then  $(\alpha + \alpha') + \alpha'' = \alpha + (\alpha' + \alpha'')$  as

$$(g)((\alpha + \alpha') + \alpha'') = ((g)\alpha + (g)\alpha') + (g)\alpha'' = (g)\alpha + ((g)\alpha' + (g)\alpha'') = (g)(\alpha + (\alpha' + \alpha''))$$

for all  $g \in G$ . The zero endomorphism  $0$  satisfies  $0 + \alpha = \alpha$  as

$$(g)(0 + \alpha) = (g)0 + (g)\alpha = 0 + (g)\alpha = (g)\alpha \text{ for all } g \in G.$$

Write  $-\alpha: G \rightarrow G$  where  $(g)(-\alpha) = -(g)\alpha$  for all  $g \in G$ . Then  $-\alpha \in \text{End } G$  as

$$(g_1 + g_2)(-\alpha) = -(g_1 + g_2)\alpha = -((g_1)\alpha + (g_2)\alpha) = -(g_1)\alpha - (g_2)\alpha = (g_1)(-\alpha) + (g_2)(-\alpha)$$

for all  $g_1, g_2 \in G$ . Also  $-\alpha + \alpha = 0$  since

$$(g)(-\alpha + \alpha) = (g)(-\alpha) + (g)\alpha = -(g)\alpha + (g)\alpha = 0 = (g)0$$

for all  $g \in G$ . Finally  $\alpha + \alpha' = \alpha' + \alpha$  as

$$(g)(\alpha + \alpha') = (g)\alpha + (g)\alpha' = (g)\alpha' + (g)\alpha = (g)(\alpha' + \alpha)$$

for all  $g \in G$ . So  $(\text{End } G, +)$  is an abelian group.

As  $(g)(\alpha(\alpha' + \alpha'')) = ((g)\alpha)(\alpha' + \alpha'') = ((g)\alpha)\alpha' + ((g)\alpha)\alpha'' = (g)\alpha\alpha' + (g)\alpha\alpha'' = (g)(\alpha\alpha' + \alpha\alpha'')$  for all  $g \in G$ , the distributive law  $\alpha(\alpha' + \alpha'') = \alpha\alpha' + \alpha\alpha''$  holds.

(b) As  $\text{End } G$  is commutative for finite cyclic  $G$  by (3.18) and  $G$  is cyclic for  $|G| \leq 3$ , the smallest candidate with  $\text{End } G$  non-commutative is the non-cyclic group of order 4, that is, the Klein group. So let  $G$  be of isomorphism type  $C_2 \oplus C_2$ . By (3.16)  $\text{End } G \cong \mathfrak{M}_2(\mathbb{Z}_2)$  the ring of  $2 \times 2$  matrices over the field  $\mathbb{Z}_2$ . As

$$\begin{pmatrix} \bar{1} & \bar{0} \\ \bar{0} & \bar{0} \end{pmatrix} \begin{pmatrix} \bar{0} & \bar{1} \\ \bar{0} & \bar{0} \end{pmatrix} \neq \begin{pmatrix} \bar{0} & \bar{1} \\ \bar{0} & \bar{0} \end{pmatrix} \begin{pmatrix} \bar{1} & \bar{0} \\ \bar{0} & \bar{0} \end{pmatrix}$$

we see that  $\text{End } G$  is indeed non-commutative. By (3.14)  $\text{Aut } G \cong U(\mathfrak{M}_2(\mathbb{Z}_2)) = GL_2(\mathbb{Z}_2)$  and so

$|\text{Aut } G| = |GL_2(\mathbb{Z}_2)| = (2^2 - 1)(2^2 - 2) = 6$  using the theory following (2.18). In fact  $\text{Aut } G \cong S_3$  the symmetric group on 3 symbols  $\langle (\bar{1}, \bar{0}) \rangle, \langle (\bar{0}, \bar{1}) \rangle, \langle (\bar{1}, \bar{1}) \rangle$ .

(c) As  $(e_1)\alpha_0 = 0e_1 + 1e_2$ ,  $(e_2)\alpha_0 = 0e_1 + 0e_2$ , the matrix of  $\alpha_0$  relative to  $e_1, e_2$  is  $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ . By (3.15)

the matrices  $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$  relative to  $e_1, e_2$  of the endomorphisms  $\alpha$  in  $Z(\alpha_0)$  satisfy

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

which is true  $\Leftrightarrow a = d, c = 0$ . So  $\alpha$  in  $Z(\alpha_0)$  has matrix  $\begin{pmatrix} a & b \\ 0 & a \end{pmatrix}$  relative to  $e_1, e_2$ . As matrices of this type are the elements of a commutative subring of  $\mathfrak{M}_2(\mathbb{Z})$ , restricting the ring isomorphism  $\theta$  of (3.15) to  $Z(\alpha_0)$  gives

$$Z(\alpha_0) \cong \left\{ \begin{pmatrix} a & b \\ 0 & a \end{pmatrix} : a, b \in \mathbb{Z} \right\}$$

showing that  $Z(\alpha_0)$  is a commutative subring of  $\text{End } G$ .

The additive group of  $\text{End } G$  has isomorphism type  $C_0 \oplus C_0 \oplus C_0 \oplus C_0$ , as it is free of rank 4 having  $\mathbb{Z}$ -basis

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

As  $\begin{pmatrix} a & b \\ 0 & a \end{pmatrix} = a \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + b \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ , the additive group of  $Z(\alpha_0)$  has isomorphism type  $C_0 \oplus C_0$ , being free of rank 2 with  $\mathbb{Z}$ -basis

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

As  $\begin{pmatrix} a & b \\ 0 & a \end{pmatrix}$  is invertible  $\Leftrightarrow a = \pm 1$  and

$$\begin{pmatrix} -1 & b \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & -b \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}^2 \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}$$

we see  $U(Z(\alpha_0)) \cong H_1 \times H_2$ , that is,  $U(Z(\alpha_0))$  is the direct product of the multiplicative cyclic group

$H_1$  of order 2 with generator  $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$  and the multiplicative cyclic group  $H_2$  of infinite order with

generator  $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ . So the isomorphism type of  $U(Z(\alpha_0))$  is  $C_2 \oplus C_0$ .

(d) Let  $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$  be the matrix relative to  $e_1, e_2$  of  $\alpha \in Z(\alpha_0)$ . Then

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

which holds  $\Leftrightarrow a = d, b = -c$  on comparing entries. So

$$\begin{pmatrix} a & b \\ -b & a \end{pmatrix} = a \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + b \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

is the matrix of  $\alpha$  relative to  $e_1, e_2$ . So the additive groups of  $\text{End } G \cong \mathfrak{M}_2(\mathbb{Z}_3)$  and  $Z(\alpha_0)$  have isomorphism types  $C_3 \oplus C_3 \oplus C_3 \oplus C_3$  and  $C_3 \oplus C_3$  respectively.

Now  $\begin{vmatrix} a & b \\ -b & a \end{vmatrix} = a^2 + b^2$ . As  $\bar{0} = (\bar{0})^2$  and  $\bar{1} = (\bar{1})^2 = (\bar{2})^2$  in  $\mathbb{Z}_3$ , we see that  $a^2 + b^2 = \bar{0}, \bar{1}, \bar{2}$

according as none, one, or both of  $a, b$  is/are non-zero. So  $a^2 + b^2 = 0 \Leftrightarrow a = b = 0$  where  $a, b \in \mathbb{Z}_3$ .

Hence every non-zero element in the commutative ring  $Z(\alpha_0)$  is invertible, i.e.  $Z(\alpha_0)$  is a field. As

$|U(Z(\alpha_0))| = 9 - 1 = 8$ , by (3.17) we see  $U(Z(\alpha_0))$  has isomorphism type  $C_8$ ; in fact  $\begin{pmatrix} \bar{1} & \bar{1} \\ -\bar{1} & \bar{1} \end{pmatrix}$

generates the multiplicative group  $U(Z(\alpha_0)) = Z(\alpha_0)^*$ .

(e) As in (d) above  $\mathfrak{M}_2(\mathbb{Z}_5)$  is a vector space over  $\mathbb{Z}_5$  with basis

$$\begin{pmatrix} \bar{1} & \bar{0} \\ \bar{0} & \bar{0} \end{pmatrix}, \begin{pmatrix} \bar{0} & \bar{1} \\ \bar{0} & \bar{0} \end{pmatrix}, \begin{pmatrix} \bar{0} & \bar{0} \\ \bar{1} & \bar{0} \end{pmatrix}, \begin{pmatrix} \bar{0} & \bar{0} \\ \bar{0} & \bar{1} \end{pmatrix}.$$

By (3.16) with  $t = 2$ ,  $p = 5$ , the isomorphism type of the additive group of  $\text{End } G$  is

$C_5 \oplus C_5 \oplus C_5 \oplus C_5$ . By (3.16) the matrices  $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$  where  $a, b \in \mathbb{Z}_5$  are the elements of a subring

$R_0$  isomorphic to  $Z(\alpha_0)$ . As  $R_0$  is a vector space over  $\mathbb{Z}_5$  with basis  $\begin{pmatrix} \bar{1} & \bar{0} \\ \bar{0} & \bar{1} \end{pmatrix}, \begin{pmatrix} \bar{0} & \bar{1} \\ -\bar{1} & \bar{0} \end{pmatrix}$ , we see the

additive group of  $Z(\alpha_0)$  has isomorphism type  $C_5 \oplus C_5$ . Also  $|R_0| = 5 \times 5 = 25$ . How many of the 25

matrices  $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$  in  $R_0$  are *not* invertible over  $\mathbb{Z}_5$ ? In other words, how many ordered pairs  $(a, b)$

with  $a, b \in \mathbb{Z}_5$  satisfy  $a^2 + b^2 = 0$ ? As  $(\bar{0})^2 = \bar{0}, (\pm\bar{1})^2 = \bar{1}, (\pm\bar{2})^2 = -\bar{1}$  we see that the answer is 9, namely

$$(\bar{0}, \bar{0}), (\bar{1}, \bar{2}), (\bar{1}, -\bar{2}), (-\bar{1}, \bar{2}), (-\bar{1}, -\bar{2}), (\bar{2}, \bar{1}), (\bar{2}, -\bar{1}), (-\bar{2}, \bar{1}), (-\bar{2}, -\bar{1}).$$

Hence  $|U(R_0)| = 25 - 9 = 16$ .

Let  $H_1 = \left\{ \begin{pmatrix} \bar{2} & \bar{0} \\ \bar{0} & \bar{2} \end{pmatrix}, \begin{pmatrix} \bar{4} & \bar{0} \\ \bar{0} & \bar{4} \end{pmatrix}, \begin{pmatrix} \bar{3} & \bar{0} \\ \bar{0} & \bar{3} \end{pmatrix}, \begin{pmatrix} \bar{1} & \bar{0} \\ \bar{0} & \bar{1} \end{pmatrix} \right\}$  and let  $H_2 = \left\{ \begin{pmatrix} \bar{1} & \bar{1} \\ -\bar{1} & \bar{1} \end{pmatrix}, \begin{pmatrix} \bar{0} & \bar{2} \\ -\bar{2} & \bar{0} \end{pmatrix}, \begin{pmatrix} -\bar{2} & \bar{2} \\ -\bar{2} & -\bar{2} \end{pmatrix}, \begin{pmatrix} \bar{1} & \bar{0} \\ \bar{0} & \bar{1} \end{pmatrix} \right\}$ .

Then  $H_1$  and  $H_2$  are (multiplicative) cyclic subgroups (of  $U(R_0)$ ) having order 4. As  $|H_1 \cap H_2| = 1$

and  $|U(R_0)| = 16$  we see  $H_1 H_2 = U(R_0)$  as in Exercises 2.2, Question 3(f). So  $U(R_0) = H_1 \times H_2$

(direct product). Therefore  $U(Z(\alpha_0)) \cong U(R_0)$  has isomorphism type  $C_4 \oplus C_4$ .

(f) Let  $g, g' \in G$ . By (3.10) there are unique elements  $g_j, g'_j \in G_{p_j}$  ( $1 \leq j \leq k$ ) with

$$g = \sum_{j=1}^k g_j \quad \text{and} \quad g' = \sum_{j=1}^k g'_j. \quad \text{Then}$$

$$\begin{aligned} (g + g')\alpha &= \left( \sum_{j=1}^k (g_j + g'_j) \right) \alpha = \sum_{j=1}^k (g_j + g'_j) \alpha_j = \sum_{j=1}^k ((g_j) \alpha_j + (g'_j) \alpha_j) = \\ &= \sum_{j=1}^k (g_j) \alpha_j + \sum_{j=1}^k (g'_j) \alpha_j = (g) \alpha + (g') \alpha \end{aligned}$$

showing  $\alpha \in \text{End } G$ .

Consider  $\beta \in \text{End } G$  and let  $g_j \in G_{p_j}$ . Then  $p_j^{n_j}((g_j)\beta) = (p_j^{n_j}g_j)\beta = (0)\beta = 0$  showing  $(g_j)\beta \in G_{p_j}$ . So  $(G_{p_j})\beta \subseteq G_{p_j}$  and the mapping  $\beta_j : G_{p_j} \rightarrow G_{p_j}$  defined by  $(g_j)\beta_j = (g_j)\beta$  for all  $g_j \in G_{p_j}$  is an endomorphism of  $G_{p_j}$  for  $1 \leq j \leq k$ . Then

$$(g)\beta = \left(\sum_{j=1}^k g_j\right)\beta = \sum_{j=1}^k (g_j)\beta = \sum_{j=1}^k (g_j)\beta_j = (g)(\beta_1 \oplus \beta_2 \oplus \dots \oplus \beta_k)$$

for all  $g \in G$  showing  $\beta = \beta_1 \oplus \beta_2 \oplus \dots \oplus \beta_k$ . Also  $\beta_j$  is uniquely determined by  $\beta$  being the restriction of  $\beta$  to  $G_{p_j}$ . Hence the mapping

$$\text{End } G \rightarrow \text{End } G_{p_1} \oplus \text{End } G_{p_2} \oplus \dots \oplus \text{End } G_{p_k},$$

defined by  $\beta \rightarrow (\beta_1, \beta_2, \dots, \beta_k)$  for all  $\beta \in \text{End } G$ , is bijective. We show this mapping to be a ring isomorphism. Let  $\beta, \beta' \in \text{End } G$ . There are  $\beta'_j \in \text{End } G_{p_j}$  ( $1 \leq j \leq k$ ) with  $\beta' = \beta'_1 \oplus \beta'_2 \oplus \dots \oplus \beta'_k$ , i.e.  $\beta' \rightarrow (\beta'_1, \beta'_2, \dots, \beta'_k)$ . The restriction of  $\beta + \beta'$  to  $G_{p_j}$  is  $\beta_j + \beta'_j$  and the restriction of  $\beta\beta'$  to  $G_{p_j}$  is  $\beta_j\beta'_j$  for  $1 \leq j \leq k$ . So the correspondence  $\beta \leftrightarrow (\beta_1, \beta_2, \dots, \beta_k)$  respects addition, multiplication and 1-elements. Therefore

$$\text{End } G \cong \text{End } G_{p_1} \oplus \text{End } G_{p_2} \oplus \dots \oplus \text{End } G_{p_k},$$

i.e. these rings are isomorphic.

## Solution 2

(a) Taking  $r = 0$ , the powers of  $\bar{3}$  in  $\mathbb{Z}_{32}$  are

$$(\bar{3})^0 = \bar{1}, (\bar{3})^1 = \bar{3}, (\bar{3})^2 = \bar{9}, (\bar{3})^3 = \bar{27}, (\bar{3})^4 = \bar{17}, (\bar{3})^5 = \bar{19}, (\bar{3})^6 = \bar{25}, (\bar{3})^7 = \bar{11}.$$

Taking  $r = 1$ , the negatives of the powers of  $\bar{3}$  in  $\mathbb{Z}_{32}$  are

$$-(\bar{3})^0 = \bar{31}, -(\bar{3})^1 = \bar{29}, -(\bar{3})^2 = \bar{23}, -(\bar{3})^3 = \bar{5}, -(\bar{3})^4 = \bar{15}, -(\bar{3})^5 = \bar{13}, -(\bar{3})^6 = \bar{7}, -(\bar{3})^7 = \bar{21}.$$

Yes as each of the 16 elements of  $U(\mathbb{Z}_{32})$  is accounted for. Each element of  $U(\mathbb{Z}_{32})$  is uniquely expressible as  $(-\bar{1})^r \times (\bar{3})^s$  where  $0 \leq r < 2$ ,  $0 \leq s < 8$ . Hence  $U(\mathbb{Z}_{32}) = \langle -\bar{1} \rangle \times \langle \bar{3} \rangle$ , the direct product of cyclic groups of orders 2 and 8. So  $U(\mathbb{Z}_{32})$  has isomorphism type  $C_2 \oplus C_8$ .

(b) As  $\bar{2} \neq \bar{1}$  and  $(\bar{2})^2 = \bar{1}$  we see that  $\bar{2}$  in  $\mathbb{Z}_3$  has multiplicative order 2. The equation

$2^8 = 256 = 3 \times 81 + 13$  gives  $2^8 \equiv 13 \pmod{81}$ . Squaring produces  $2^{16} \equiv 7 \pmod{81}$  as  $13^2 = 169 = 2 \times 81 + 7$ . So  $2^{18} \equiv 28 \pmod{81}$  on multiplying by 4, i.e.  $2^{18} \not\equiv 1 \pmod{81}$ . Also  $2^{24} \equiv 10 \pmod{81}$  as  $7 \times 13 = 91 = 81 + 10$ . So  $2^{27} \equiv -1 \pmod{81}$  as  $8 \times 10 = 80 = 81 - 1$ . Therefore  $2^{27} \not\equiv 1 \pmod{81}$  and  $2^{54} \equiv 1 \pmod{81}$  on squaring  $2^{27} \equiv -1 \pmod{81}$ . So  $\bar{2}$  in  $U(\mathbb{Z}_{81})$  has

multiplicative order 54. The integer  $a = 2$  satisfies (3.19) with  $p = 3$  and so  $\bar{2}$  has order  $\phi(3^n) = 3^n - 3^{n-1}$  in  $\mathbb{Z}_{3^n}$  for  $n \geq 1$ . In particular  $\bar{2}$  has order  $3^5 - 3^4 = 162$  in  $\mathbb{Z}_{243}$  on taking  $n = 5$ .

(c) Try  $a = 3$ . Then  $a^2 \equiv 2 \pmod{7}$ ,  $a^3 \equiv -1 \pmod{7}$  and so  $\bar{a}$  in  $\mathbb{Z}_7$  has multiplicative order 6. Also  $a^6 = 3^6 = 243 \times 3 = (5 \times 49 - 2) \times 3 \equiv -6 \pmod{49}$  and so  $a^6 \not\equiv 1 \pmod{49}$ . The only other choice is  $a = 5$  in which case  $a^2 \equiv 4 \pmod{7}$ ,  $a^3 \equiv -1 \pmod{7}$ ,  $a^6 \equiv -6 \pmod{49}$ . In both cases  $a$  satisfies the conditions of (3.19). So  $\bar{a}$  in  $\mathbb{Z}_{49}$  has order  $\phi(49) = 49 - 7 = 42$  and  $\bar{a}$  in  $\mathbb{Z}_{(49)^2}$  has order

$$\phi((49)^2) = \phi(7^4) = 7^4 - 7^3 = 2058.$$

(d) We verify that  $\overline{14}$  in  $\mathbb{Z}_{29}$  has order 28. Notice  $14^2 \equiv -7 \pmod{29}$  as  $14^2 + 7 = 7 \times 29$ . Hence  $14^4 \equiv (-7)^2 \pmod{29}$ , i.e.  $14^4 \not\equiv 1 \pmod{29}$ . Also  $14^6 \equiv (-7)^3 \equiv 5 \pmod{29}$  since  $7^3 + 5 = 12 \times 29$ . Hence  $14^{14} \equiv (5 \times 14)^2 \equiv -1 \pmod{29}$  since  $70^2 + 1 = 169 \times 29$ , i.e.  $14^{14} \not\equiv 1 \pmod{29}$ . So  $\overline{14}$  in  $\mathbb{Z}_{29}$  has multiplicative order 28.

We show  $29^2$  is a divisor of  $14^{28} - 1$ . As  $14^{28} - 1 = (14^{14} + 1)(14^{14} - 1)$  it is enough to show  $29^2 = 841$  is a divisor of  $14^{14} + 1$ . Now  $14^3 \equiv 221 \pmod{841}$  as  $14^3 - 221 = 3 \times 841$ . Hence  $14^6 \equiv 63 \pmod{841}$  as  $221^2 - 63 = 58 \times 841$ . Hence  $14^{14} \equiv (63 \times 14)^2 \equiv 41^2 \equiv -1 \pmod{841}$  as  $63 \times 14 = 841 + 41$  and  $41^2 + 1 = 2 \times 841$ . The conclusion is:  $29^2$  is a divisor of  $14^{28} - 1$  and  $\overline{14}$  has order 28 in  $\mathbb{Z}_{841}$ . The statement in the question is therefore *not* true in general.

(e) Let  $\theta_2: \mathbb{Z}_{p^n} \rightarrow \mathbb{Z}_{p^2}$  denote the ring homomorphism given by  $(x_n)\theta_2 = x_2$  for all integers  $a$ , where  $x_n$  is the congruence class of  $a$  modulo  $p^n$ . As  $x_n \in U(\mathbb{Z}_{p^n}) \Leftrightarrow \gcd\{a, p\} = 1 \Leftrightarrow x_2 \in U(\mathbb{Z}_{p^2})$ , the restriction of  $\theta_2$  to  $U(\mathbb{Z}_{p^n})$  is the surjective group homomorphism  $\theta'_2: U(\mathbb{Z}_{p^n}) \rightarrow U(\mathbb{Z}_{p^2})$  given by  $(x_n)\theta'_2 = x_2$  for all integers  $a$  with  $\gcd\{a, p\} = 1$ . Similarly  $\theta'_1: U(\mathbb{Z}_{p^n}) \rightarrow U(\mathbb{Z}_p)$  given by  $(x_n)\theta'_1 = x_1$  for all integers  $a$  with  $\gcd\{a, p\} = 1$  is also a surjective group homomorphism (here  $x_1$  is the congruence class of  $a$  modulo  $p$ ). Let  $e_n$  denote the congruence class of 1 modulo  $p^n$  for  $n \geq 1$ . Suppose  $x_2$  generates  $U(\mathbb{Z}_{p^2})$ . Then  $x_2^{p-1} \neq e_2$  as  $x_2$  has order  $p(p-1)$  and so  $a \not\equiv 1 \pmod{p^2}$ . As  $\theta'_1: U(\mathbb{Z}_{p^2}) \rightarrow U(\mathbb{Z}_p)$  is surjective we see that  $\bar{a} = x_1 = (x_2)\theta'_1$  generates  $U(\mathbb{Z}_p) = \mathbb{Z}_p^*$ . From the second part of the proof of (3.19) we deduce that  $x_n$  generates  $U(\mathbb{Z}_{p^n})$  for all  $n \geq 2$ . Conversely suppose that  $x_n$  generates  $U(\mathbb{Z}_{p^n})$  for some integer  $n$  with  $n \geq 2$ . As  $\theta'_2$  is surjective we see that  $x_2 = (x_n)\theta'_2$  generates  $U(\mathbb{Z}_{p^2})$ .

(f) As  $(1+p)^{p^0} = 1+p$  we see  $k_0 = 1$ . So suppose  $i > 0$  and there is an integer  $k_{i-1}$  with  $\gcd\{k_{i-1}, p\} = 1$  such that  $(1+p)^{p^{i-1}} = 1 + k_{i-1}p^i$ . Raising this equation to the power  $p$  and using the binomial theorem gives  $(1+p)^{p^i} = ((1+p)^{p^{i-1}})^p = (1 + k_{i-1}p^i)^p = \sum_{r=0}^p \binom{p}{r} k_{i-1}^r p^{ir}$ . Each term in this sum with  $2 \leq r < p$  is divisible by  $p^{i+2}$  as  $\binom{p}{r}$  is divisible by  $p$  and  $1+ir \geq i+2$ . The last term  $k_{i-1}^p p^{ip}$  in the sum is also divisible by  $p^{i+2}$  as  $ip \geq i+2$  since  $i \geq 1, p \geq 3$ . Therefore there is an integer  $l_i$  with  $(1+p)^{p^i} = 1 + k_i p^{i+1} + l_i p^{i+2}$ . So  $k_i = k_{i-1} + l_i p$  satisfies  $\gcd\{k_i, p\} = \gcd\{k_{i-1}, p\} = 1$  and  $(1+p)^{p^i} = 1 + k_i p^{i+1}$ . The induction is now complete.

Taking  $i = n-2$  gives  $(1+p)^{p^{n-2}} = 1 + k_{n-2} p^{n-1} \not\equiv 1 \pmod{p^n}$  and so the order of  $g$ , as an element of the multiplicative group  $U(\mathbb{Z}_{p^n})$ , is not a divisor of  $p^{n-2}$ . Taking  $i = n-1$  gives

$(1+p)^{p^{n-1}} = 1 + k_{n-1} p^n \equiv 1 \pmod{p^n}$  showing that the order of  $g$  is a divisor of  $p^{n-1}$ . Hence  $g$  has order  $p^{n-1}$ . From (e) above the group homomorphism  $\theta'_1: U(\mathbb{Z}_{p^n}) \rightarrow U(\mathbb{Z}_p)$  is surjective. By (3.17)



$U(\mathbb{Z}_p)$  is cyclic of order  $p-1$  and so there is  $h' \in U(\mathbb{Z}_{p^n})$  such that  $(h')\theta'_1$  generates  $U(\mathbb{Z}_p)$ . By Exercises 2.1, Question 2(b) the order  $s$  of  $h'$  is divisible by  $p-1$ . So  $h = (h')^{s/(p-1)}$  has order  $p-1$ . As  $\gcd\{p^{n-1}, p-1\} = 1$  we conclude that  $gh$  has order  $p^{n-1}(p-1) = |U(\mathbb{Z}_{p^n})|$  using Exercises 2.2, Question 4(e) in multiplicative notation. Hence  $gh$  generates  $U(\mathbb{Z}_{p^n})$ , i.e.  $U(\mathbb{Z}_{p^n}) = \langle gh \rangle$ .

### Solution 3

(a) (i)  $|G| = 105 = 3 \times 5 \times 7$  and so  $U(\mathbb{Z}_{105}) \cong U(\mathbb{Z}_3) \times U(\mathbb{Z}_5) \times U(\mathbb{Z}_7)$  which is the direct product of cyclic groups of orders 2, 4, 6 respectively by (3.17). By the Chinese remainder theorem

$$C_2 \oplus C_4 \oplus C_6 = C_2 \oplus C_4 \oplus C_2 \oplus C_3 = C_2 \oplus C_2 \oplus C_4 \oplus C_3 = C_2 \oplus C_2 \oplus C_{12}$$

and so  $\text{Aut } G \cong U(\mathbb{Z}_{105})$  has invariant factor sequence (2, 2, 12).

(ii)  $|G| = 100 = 2^2 \times 5^2$  and so  $U(\mathbb{Z}_{100}) \cong U(\mathbb{Z}_4) \times U(\mathbb{Z}_{25})$  which is the direct product of cyclic groups of orders 2 and 20 respectively by (3.19). So  $\text{Aut } G \cong U(\mathbb{Z}_{100})$  has invariant factors 2, 20.

(iii)  $|G| = 98 = 2 \times 7^2$  and so  $U(\mathbb{Z}_{98}) \cong U(\mathbb{Z}_2) \times U(\mathbb{Z}_{49})$  which is the direct product of cyclic groups of orders 1 and 42 respectively by (3.19). So  $\text{Aut } G \cong U(\mathbb{Z}_{98})$  is cyclic with single invariant factor 42.

(iv)  $|G| = 96 = 2^5 \times 3$  and so  $U(\mathbb{Z}_{96}) \cong U(\mathbb{Z}_{32}) \times U(\mathbb{Z}_3)$  which is the direct product of cyclic groups of orders 2, 8, 2 by (3.19). Hence  $\text{Aut } G \cong U(\mathbb{Z}_{96})$  has invariant factor sequence (2, 2, 8).

(b) Yes,  $H$  is an elementary abelian 2-group being the subgroup of the abelian group  $\text{Aut } G$  consisting of elements having order 1 or 2. So  $|H| = 2^l$ . By (3.19) and the theory preceding it,

$\text{Aut } G$  is the direct product of finite cyclic groups of even order. Each such cyclic group has a unique element of order 2. Hence  $\text{Aut } G$  is the direct product of  $l$  finite cyclic groups of even order on comparing the numbers of elements of order 2 in the direct product decomposition of  $\text{Aut } G$ . For the same reason the invariant factor decomposition of  $\text{Aut } G$  is made up of  $l$  cyclic groups of even order, i.e. the invariant factors of  $\text{Aut } G$  are  $l$  in number and they are all even.

(c) By (3.19) we see  $\text{Aut } G \cong U(\mathbb{Z}_{p^n})$  is cyclic of order  $p^n - p^{n-1} = p^{n-1}(p-1)$  which is even. By (2.2) the group  $\text{Aut } G$  has a unique subgroup of order 2 and hence has a unique element of order 2. Conversely suppose  $\text{Aut } G$  has a unique element of order 2. Then  $|H| = 2$  where  $H$  is the subgroup of  $\text{Aut } G$  defined in (b) above. So  $\text{Aut } G$  is cyclic and non-trivial by (b) above, i.e.  $\text{Aut } G$  has exactly one invariant factor. Hence  $|G| = p^n$  or  $|G| = 2p^n$  where  $p$  is prime since  $\text{Aut } G$  is non-cyclic in all other cases. For  $|G| = 2^n$  by (3.19)  $\text{Aut } G$  is non-cyclic for  $n \geq 3$ ; as  $\text{Aut } G$  is trivial for  $n = 1$ , this leaves  $n = 2$ , i.e.  $G$  is cyclic of order 4. For  $|G| = p^n$  or  $|G| = 2p^n$  where  $p$  is an odd prime, by (3.17) and (3.19)  $\text{Aut } G$  is cyclic of order  $p^{n-1}(p-1)$ . So the other finite cyclic groups such that  $\text{Aut } G$  has a unique element of order 2 are of order 4 or  $2p^n$  where  $p$  is an odd prime.

(d) Let  $|G| = 2^n$  where  $n \geq 3$ ; then  $\text{Aut } G$  has invariant factor sequence  $(2, 2^{n-2})$  by (3.19). Let  $|G| = 4p^n$  where  $p$  is an odd prime; then  $\text{Aut } G$  has invariant factor sequence  $(2, p^{n-1}(p-1))$  by

(3.19). Let  $|G| = p_1^{n_1} p_2^{n_2}$  or  $|G| = 2 p_1^{n_1} p_2^{n_2}$  where  $p_1$  and  $p_2$  are different odd primes; then by the Chinese remainder theorem  $\text{Aut } G$  has invariant factor sequence

$$(\gcd\{p_1 - 1, p_2 - 1\}, p_1^{n_1-1} p_2^{n_2-1} \text{lcm}\{p_1 - 1, p_2 - 1\})$$

as  $\text{Aut } G$  is the direct product of two cyclic groups of order  $\phi(p_1^{n_1})$  and  $\phi(p_2^{n_2})$  by (3.19).

(e) Suppose that  $G$  is a finite cyclic group with  $\text{Aut } G$  being a non-trivial elementary abelian 2-group. Then all cyclic subgroups of  $\text{Aut } G$  have order 2. Now  $p^n - p^{n-1} = 2$ , where  $p$  is an odd prime, gives  $p = 3, n = 1$ . So (3.19) and the theory preceding it gives  $|G| = 2^{n_1} \times 3^{n_2}$  where  $n_1 \leq 3$  and  $n_2 \leq 1$ . As  $\text{Aut } G$  is non-trivial we see  $|G| \geq 3$ . Therefore  $|G| = 3, 4, 6, 8, 12, 24$ .

(f) As  $21 = 3 \times 7$ , using (3.19) and the theory preceding it,  $\text{Aut } \mathbb{Z}_{21}$  has isomorphism type  $C_2 \oplus C_6$ . As  $28 = 2^2 \times 7$  we see similarly that  $\text{Aut } \mathbb{Z}_{28}$  has isomorphism type  $C_2 \oplus C_6$ . As  $36 = 2^2 \times 3^2$  we see  $\text{Aut } \mathbb{Z}_{36}$  has isomorphism type  $C_2 \oplus C_6$ . As  $42 = 2 \times 3 \times 7$  we see  $\text{Aut } \mathbb{Z}_{42}$  has isomorphism type  $C_1 \oplus C_2 \oplus C_6 = C_2 \oplus C_6$ . Hence these four automorphism groups are isomorphic.

Suppose  $\text{Aut } \mathbb{Z}_{|G|}$  has isomorphism type  $C_2 \oplus C_{12}$  and let  $p^n$  be a divisor of  $|G|$  where  $p$  is prime.

By (3.19)  $p^n - p^{n-1}$  is a divisor of  $|\text{Aut } \mathbb{Z}_{|G|}| = 2 \times 12 = 24$ . So the possibilities are:

$p = 2$  and  $n \leq 4$ ,  $p = 3$  and  $n \leq 2$ ,  $p = 5$  and  $n = 1$ ,  $p = 7$  and  $n = 1$ ,  $p = 13$  and  $n = 1$ . By inspection we see  $|G| = 35, 39, 45, 52, 56, 70, 78, 90$ .

#### Solution 4

(a) The 9 subsequences of  $(2, 4, 4)$  are  $(1, 1, 1)$ ,  $(1, 1, 2)$ ,  $(1, 1, 4)$ ,  $(1, 2, 2)$ ,  $(1, 2, 4)$ ,  $(1, 4, 4)$ ,  $(2, 2, 2)$ ,  $(2, 2, 4)$ ,  $(2, 4, 4)$ . The 5 subsequences of  $(1, 3, 9)$  are  $(1, 1, 1)$ ,  $(1, 1, 3)$ ,  $(1, 1, 9)$ ,  $(1, 3, 3)$ ,  $(1, 3, 9)$ . As  $(2, 12, 36) = (2 \times 1, 4 \times 3, 4 \times 9)$  each subsequence of  $(2, 12, 36)$  corresponds to a pair of subsequences, one of  $(2, 4, 4)$  and one of  $(1, 3, 9)$ ; for example  $(1, 6, 12)$  corresponds to the pair  $(1, 2, 4)$ ,  $(1, 3, 3)$  on factorising each integer into a power of 2 multiplied by a power of 3. So there are  $9 \times 5 = 45$  subsequences of  $(2, 12, 36)$  using the paragraph below. By (3.22) there are 45 different isomorphism types among the subgroups of an abelian group of isomorphism type  $C_2 \oplus C_{12} \oplus C_{36}$ .

As  $\gcd\{d_s, \delta_s\} = 1$  each positive divisor of  $d_s \delta_s$  is uniquely expressible as  $d' \delta'$  where  $d'$  and  $\delta'$  are positive divisors of  $d_s$  and  $\delta_s$  respectively. In fact  $d' = \gcd\{d' \delta', d_s\}$  and  $\delta' = \gcd\{d' \delta', \delta_s\}$ . Further  $d' \delta' | d'' \delta'' \Leftrightarrow d' | \delta'$  and  $d'' | \delta''$  where  $d', d''$  and  $\delta', \delta''$  are positive divisors of  $d_s$  and  $\delta_s$  respectively. Hence each subsequence of  $(d_1 \delta_1, d_2 \delta_2, \dots, d_s \delta_s)$  is uniquely expressible as a 'product'  $(d'_1 \delta'_1, d'_2 \delta'_2, \dots, d'_s \delta'_s)$  where  $(d'_1, d'_2, \dots, d'_s)$  and  $(\delta'_1, \delta'_2, \dots, \delta'_s)$  are subsequences of  $(d_1, d_2, \dots, d_s)$  and  $(\delta_1, \delta_2, \dots, \delta_s)$  respectively. So the number of subsequences of  $(d_1 \delta_1, d_2 \delta_2, \dots, d_s \delta_s)$  is the product of the number of subsequences of  $(d_1, d_2, \dots, d_s)$  and the number of subsequences of  $(\delta_1, \delta_2, \dots, \delta_s)$ .

As  $(6, 60, 600) = (2 \times 3 \times 1, 2^2 \times 3 \times 5, 2^3 \times 3 \times 5^2)$  each subsequence of  $(6, 60, 600)$  is the 'product' of 3 subsequences, one of  $(2, 2^2, 2^3)$ , one of  $(3, 3, 3)$  and one of  $(1, 5, 5^2)$ . There are 14 subsequences of  $(2, 2^2, 2^3)$ , 4 subsequences of  $(3, 3, 3)$  and 5 subsequences of  $(1, 5, 5^2)$ . Hence  $(6, 60, 600)$  has  $14 \times 4 \times 5 = 280$  subsequences applying twice the conclusion of the above paragraph.

(b)  $G$  of isomorphism type  $C_2 \oplus C_{12} \oplus C_{84}$  (i) has a subgroup  $H$  of isomorphism type  $C_{14}$  as  $14 \mid 84$ , (ii) has a subgroup  $H$  of isomorphism type  $C_3 \oplus C_3 \oplus C_{14} = C_3 \oplus C_{42}$  as  $3 \mid 12$  and  $42 \mid 84$ , (iii) has no subgroup  $H$  of isomorphism type  $C_8$  as 8 is not a divisor of the exponent 84 of  $G$ .

(c) Suppose that  $H$  and  $G/H$  are isomorphic. Then  $|H| = |G/H| = |G|/|H|$  and so  $|H|^2 = |G| = 16$  giving  $|H| = 4$ . Hence  $H$  has isomorphism type either  $C_2 \oplus C_2$  or  $C_4$ . Now  $G$  has only one

subgroup of isomorphism type  $C_2 \oplus C_2$ , namely  $H_1 = \langle (\bar{1}, \bar{0}), (\bar{0}, \bar{4}) \rangle$  and  $G/H_1$  is cyclic of order 4 being generated by the coset  $H_1 + (\bar{0}, \bar{1})$ . So we are looking for a cyclic subgroup  $H$  of order 4 such that  $G/H$  is cyclic of order 4. Notice that  $H_2 = \langle (\bar{0}, \bar{2}) \rangle$  is cyclic of order 4 but

$G/H_2 = \langle H_2 + (\bar{1}, \bar{0}), H_2 + (\bar{0}, \bar{1}) \rangle$  is of isomorphism type  $C_2 \oplus C_2$ .

Consider  $H = \langle (\bar{1}, \bar{2}) \rangle = \{(\bar{1}, \bar{2}), (\bar{0}, \bar{4}), (\bar{1}, \bar{6}), (\bar{0}, \bar{0})\}$  which is cyclic of order 4. Then

$G/H = \langle H + (\bar{0}, \bar{1}) \rangle$  is also cyclic of order 4 and so  $H$  and  $G/H$  are isomorphic. In fact  $H$  is the only subgroup of  $G$  with this property.

(d) By (3.20) and (3.22) the subgroups of  $G$  are of isomorphism type  $C_1, C_2, C_4, C_8, C_2 \oplus C_2, C_2 \oplus C_4, C_2 \oplus C_8$  (all with torsion-free rank 0),  $C_0, C_2 \oplus C_0, C_4 \oplus C_0, C_8 \oplus C_0, C_2 \oplus C_2 \oplus C_0, C_2 \oplus C_4 \oplus C_0, C_2 \oplus C_4 \oplus C_0$  (all with torsion-free rank 1).

The elements of  $G$  are triples  $(x, y, m)$  where  $x \in \mathbb{Z}_2, y \in \mathbb{Z}_8, m \in \mathbb{Z}$ . Let

$H_1 = \langle (\bar{1}, \bar{0}, 0), (\bar{0}, \bar{2}, 0), (\bar{0}, \bar{0}, 12) \rangle$ . Then  $H_1$  has isomorphism type  $C_2 \oplus C_4 \oplus C_0$  and

$$G/H_1 \cong \frac{\langle (\bar{1}, \bar{0}, 0) \rangle}{\langle (\bar{1}, \bar{0}, 0) \rangle} \oplus \frac{\langle (\bar{0}, \bar{1}, 0) \rangle}{\langle (\bar{0}, \bar{2}, 0) \rangle} \oplus \frac{\langle (\bar{0}, \bar{0}, 1) \rangle}{\langle (\bar{0}, \bar{0}, 12) \rangle} \text{ has isomorphism type } C_1 \oplus C_2 \oplus C_{12} = C_2 \oplus C_{12}.$$

Let  $H_2 = \langle (\bar{0}, \bar{1}, 0), (\bar{0}, \bar{0}, 12) \rangle$  which has isomorphism type  $C_8 \oplus C_0$ . Then

$$G/H_2 \cong \frac{\langle (\bar{1}, \bar{0}, 0) \rangle}{\langle (\bar{0}, \bar{0}, 0) \rangle} \oplus \frac{\langle (\bar{0}, \bar{1}, 0) \rangle}{\langle (\bar{0}, \bar{1}, 0) \rangle} \oplus \frac{\langle (\bar{0}, \bar{0}, 1) \rangle}{\langle (\bar{0}, \bar{0}, 12) \rangle} \text{ has isomorphism type } C_2 \oplus C_1 \oplus C_{12} = C_2 \oplus C_{12}.$$

Let  $H_3 = \langle (\bar{0}, \bar{4}, 0), (\bar{0}, \bar{0}, 3) \rangle$  which has isomorphism type  $C_2 \oplus C_0$ . Then

$$G/H_3 \cong \frac{\langle (\bar{1}, \bar{0}, 0) \rangle}{\langle (\bar{0}, \bar{0}, 0) \rangle} \oplus \frac{\langle (\bar{0}, \bar{1}, 0) \rangle}{\langle (\bar{0}, \bar{4}, 0) \rangle} \oplus \frac{\langle (\bar{0}, \bar{0}, 1) \rangle}{\langle (\bar{0}, \bar{0}, 3) \rangle} \text{ has isomorphism type } C_2 \oplus C_4 \oplus C_3 = C_2 \oplus C_{12}.$$

$C_2 \oplus C_2, C_2 \oplus C_4, \dots, C_2 \oplus C_{50}$  (25 isomorphism types),

$C_2 \oplus C_2 \oplus C_2, C_2 \oplus C_2 \oplus C_4, \dots, C_2 \oplus C_2 \oplus C_{24}$  (12 isomorphism types),

$C_2 \oplus C_4 \oplus C_4, C_2 \oplus C_4 \oplus C_8, C_2 \oplus C_4 \oplus C_{12},$

$C_4 \oplus C_4, C_4 \oplus C_8, \dots, C_4 \oplus C_{24}$  (6 isomorphism types),  $C_8 \oplus C_8$ , giving 47 isomorphism types in all.

### Solution 5

(a) As  $d_i h_i = 0$  we see  $0 = (d_i h_i) \alpha = d_i ((h_i) \alpha) = d_i (\sum_{j=1}^s a_{ij} h_j) = \sum_{j=1}^s d_i a_{ij} h_j$ . As  $d_i a_{ij} h_j \in H_j$  and

$H_1, H_2, \dots, H_s$  are independent subgroups of  $G$  we deduce  $d_i a_{ij} h_j = 0$  for  $1 \leq i, j \leq s$ . As  $h_j$  has order  $d_j$  we conclude  $d_i a_{ij} \equiv 0 \pmod{d_j}$  for  $1 \leq i, j \leq s$ .

For  $i \geq j$  we have  $d_i a_{ij} = d_j (d_i/d_j) a_{ij} \equiv 0 \pmod{d_j}$  for arbitrary integers  $a_{ij}$ , as  $(d_i/d_j) \in \mathbb{Z}$ . For  $i < j$  we have  $(d_j/d_i) \in \mathbb{Z}$  and so  $d_i a_{ij} \equiv 0 \pmod{d_j} \Rightarrow a_{ij} \equiv 0 \pmod{d_j/d_i}$  on dividing through by  $d_i$ ; conversely  $a_{ij} \equiv 0 \pmod{d_j/d_i} \Rightarrow d_i a_{ij} \equiv 0 \pmod{d_j}$  on multiplying through by  $d_i$ .

Consider  $A = (a_{ij})$  and  $B = (b_{ij})$  in  $R_G$ . Adding  $d_i a_{ij} \equiv 0 \pmod{d_j}$  and  $d_i b_{ij} \equiv 0 \pmod{d_j}$  gives  $d_i(a_{ij} + b_{ij}) \equiv 0 \pmod{d_j}$  which shows  $A + B \in R_G$ . Also  $-A$  and the zero  $s \times s$  matrix over  $\mathbb{Z}$

belong to  $R_G$ . Consider the  $(i, k)$ -entry  $c_{ik} = \sum_{j=1}^s a_{ij} b_{jk}$  in  $AB$ . Then

$d_i a_{ij} b_{jk} = z_{ij} d_j b_{jk} = z_{ij} z'_{jk} d_k$  where  $z_{ij}, z'_{jk} \in \mathbb{Z}$ . So  $d_i a_{ij} b_{jk} \equiv 0 \pmod{d_k}$  for  $1 \leq j \leq s$ . Adding these congruences gives  $d_i c_{ik} \equiv 0 \pmod{d_k}$  which shows  $AB \in R_G$ . So  $R_G$  is a subring of  $\mathfrak{M}_s(\mathbb{Z})$ .

Suppose also  $g = x'_1 h_1 + x'_2 h_2 + \dots + x'_s h_s$  for  $x'_i \in \mathbb{Z}$ . Then

$$0 = g - g = (x_1 - x'_1)h_1 + (x_2 - x'_2)h_2 + \dots + (x_s - x'_s)h_s \in H_1 \oplus H_2 \oplus \dots \oplus H_s.$$

From the independence of  $H_1, H_2, \dots, H_s$  we deduce  $(x_i - x'_i)h_i = 0$  for  $1 \leq i \leq s$ . As  $h_i$  has order  $d_i$

we see  $x_i \equiv x'_i \pmod{d_i}$ , i.e. there are integers  $z_i$  with  $x_i - x'_i = z_i d_i$  ( $1 \leq i \leq s$ ). Let  $y'_j = \sum_{i=1}^s x'_i a_{ij}$  for

$1 \leq j \leq s$ . Hence  $y_j - y'_j = \sum_{i=1}^s (x_i - x'_i) a_{ij} = \sum_{i=1}^s z_i d_i a_{ij} \equiv 0 \pmod{d_j}$  since  $d_i a_{ij} \equiv 0 \pmod{d_j}$  for

$1 \leq i, j \leq s$  as  $A \in R_G$ . So  $y_j \equiv y'_j \pmod{d_j}$  giving  $y_j h_j = y'_j h_j$  for  $1 \leq j \leq s$  as  $h_j$  has order  $d_j$ .

Therefore  $\sum_{j=1}^s y_j h_j = \sum_{j=1}^s y'_j h_j$  showing that  $\alpha: G \rightarrow G$  is unambiguously defined.

Consider  $g' = x''_1 h_1 + x''_2 h_2 + \dots + x''_s h_s \in G$  where  $x''_i \in \mathbb{Z}$ . Write  $y''_j = \sum_{i=1}^s x''_i a_{ij}$  and so

$$(g')\alpha = \sum_{j=1}^s y''_j h_j. \text{ Now}$$

$$g + g' = (x_1 + x''_1)h_1 + (x_2 + x''_2)h_2 + \dots + (x_s + x''_s)h_s \text{ and } y_j + y''_j = \sum_{i=1}^s (x_i + x''_i) a_{ij} \text{ for } 1 \leq j \leq s.$$

Therefore

$$(g + g')\alpha = \sum_{j=1}^s (y_j + y''_j)h_j = \sum_{j=1}^s y_j h_j + \sum_{j=1}^s y''_j h_j = (g)\alpha + (g')\alpha$$

showing that  $\alpha = (A)\varphi$  is an endomorphism of  $G$ .

Consider  $A = (a_{ij}), B = (b_{ij}) \in R_G$ . Write  $y'''_j = \sum_{i=1}^s x_i b_{ij}$  and  $\beta = (B)\varphi$ . Then for all  $g \in G$  we have

$$(g)((A)\varphi + (B)\varphi) = (g)(\alpha + \beta) = (g)\alpha + (g)\beta = \sum_{j=1}^s y_j h_j + \sum_{j=1}^s y'''_j h_j = \sum_{j=1}^s (y_j + y'''_j)h_j = (g)(A+B)\varphi$$

since  $y_j + y'''_j = \sum_{i=1}^s x_i (a_{ij} + b_{ij})$  and  $a_{ij} + b_{ij}$  is the  $(i, j)$ -entry in  $A + B$ , i.e.  $(A)\varphi + (B)\varphi = (A+B)\varphi$ .

Write  $c_{ik} = \sum_{j=1}^s a_{ij} b_{jk}$  which is the  $(i, k)$ -entry in  $AB$ . Then

$$(g)((A)\varphi(B)\varphi) = (g)(\alpha\beta) = ((g)\alpha)\beta = \left(\sum_{j=1}^s y_j h_j\right)\beta = \sum_{k=1}^s z'_k h_k = (g)(AB)\varphi$$

where  $z'_k = \sum_{j=1}^s y_j b_{jk} = \sum_{j=1}^s (\sum_{i=1}^s x_i a_{ij}) y_j b_{jk} = \sum_{i=1}^s x_i c_{ik}$  and so  $(A)\varphi(B)\varphi = (AB)\varphi$ . Also  $x_i = y_i$  for  $1 \leq i \leq s$  in the case  $A = I$  giving  $(I)\varphi = \iota$ , the identity mapping of  $G$ , showing that  $\varphi$  maps the 1-element of  $R_G$  to the 1-element of  $\text{End } G$ . Therefore  $\varphi: R_G \rightarrow \text{End } G$  is a ring homomorphism.

Each  $\alpha \in \text{End } G$  gives rise to  $A = (a_{ij}) \in R_G$  where  $(h_i)\alpha = \sum_{j=1}^s a_{ij} h_j$  for  $1 \leq i \leq s$ . But

$$(h_i)(A)\varphi = \sum_{j=1}^s a_{ij} h_j \text{ as } (x_1, x_2, \dots, x_s) = e_i \text{ gives } g = h_i \text{ and } y_j = a_{ij} \text{ for } 1 \leq i, j \leq s. \text{ So } \alpha = (A)\varphi$$

showing  $\text{im } \varphi = \text{End } G$ .

Consider  $C = (c_{ij}) \in R_G$  where  $c_{ij} \equiv 0 \pmod{d_j}$  for  $1 \leq i, j \leq s$ . Then  $c_{ij} h_j = 0$  and so  $(h_i)(C)\varphi = 0$  for  $1 \leq i \leq s$ . Therefore  $(C)\varphi = 0$ , i.e.  $C \in \ker \varphi$ . Conversely suppose  $C \in \ker \varphi$ . Then  $(C)\varphi = 0$  and so  $(h_i)(C)\varphi = 0$  for  $1 \leq i \leq s$ . Therefore  $\sum_{j=1}^s c_{ij} h_j = 0$  and so  $c_{ij} h_j = 0$  for  $1 \leq i \leq s$ , i.e.

$c_{ij} \equiv 0 \pmod{d_j}$  for  $1 \leq i, j \leq s$ . We've shown  $\ker \varphi = \{C = (c_{ij}) : c_{ij} \equiv 0 \pmod{d_j}\}$ . By the first isomorphism theorem for rings (Exercises 2.3, Question 3(b))  $R_G / \ker \varphi \cong \text{End } G$ .

Consider a typical element  $B + \ker \varphi$  of  $R_G / \ker \varphi$ ; so  $B = (b_{ij}) \in R_G$ . By the division law for integers there are integers  $q_{ij}$  with  $b_{ij} = q_{ij} d_j + a_{ij}$  where  $0 \leq a_{ij} < d_j$  for  $1 \leq i, j \leq s$ . As  $C = -(q_{ij}) \in \ker \varphi$  we see  $A = (a_{ij}) \in B + \ker \varphi$ . Suppose also  $A' = (a'_{ij}) \in B + \ker \varphi$  where  $0 \leq a'_{ij} < d_j$  for  $1 \leq i, j \leq s$ . Then  $A - A' \in \ker \varphi$  showing  $d_j \mid (a_{ij} - a'_{ij})$ ; so  $a_{ij} = a'_{ij}$  for  $1 \leq i, j \leq s$  and so  $A = A'$  showing that each element of  $B + \ker \varphi$  of  $R_G / \ker \varphi$  contains a unique 'reduced' matrix  $A$ . For  $i < j$  the possible  $a_{ij}$  are  $d_i$  in number namely  $0, d_j/d_i, 2(d_j/d_i), \dots, (d_i-1)(d_j/d_i)$  as  $a_{ij} \equiv 0 \pmod{d_j/d_i}$ . For  $i \geq j$  all  $d_j$  integers  $a_{ij}$  with  $0 \leq a_{ij} < d_j$  are possible. The number of possibilities for the  $(i, j)$ -entry in a reduced matrix is therefore as shown in the symmetric matrix:

$$\begin{pmatrix} d_1 & d_1 & d_1 & \cdots & d_1 \\ d_1 & d_2 & d_2 & \cdots & d_2 \\ d_1 & d_2 & d_3 & \cdots & d_3 \\ \vdots & \vdots & \vdots & & \vdots \\ d_1 & d_2 & d_3 & \cdots & d_s \end{pmatrix}.$$

The number of  $d_i$  in this matrix is  $2s - 2i + 1$  for  $1 \leq i \leq s$ . The number of reduced matrices is the

product of the entries in the above matrix, i.e.  $|\text{End } G| = \prod_{i=1}^s d_i^{2s-2i+1}$  as the number of endomorphisms

of  $G$  equals the number of reduced matrices. (The invariant factors of  $\text{End } G$  are discussed in Question 6(d) below.)

(b) In the case  $i_0 = j_0$  we take  $k = 1$ . So suppose  $i_0 \neq j_0$ . Denote the (multiplicative) order of  $\pi$  by  $n$ .

Then  $n > 1$  and  $(i)\pi^n = i$  for all  $i \in S$ . As  $\pi^{n-1} = \pi^{-1}$  the positive integer  $k = n - 1$  satisfies

$(i_0)\pi^k = j_0$ . Also  $n \in K$  and so  $K$  is a non-zero ideal of  $\mathbb{Z}$ . By (1.15) there is a positive integer  $l_0$

with  $K = \langle l_0 \rangle$ . Therefore  $l_0$  is the smallest positive integer with  $(i_0)\pi^{l_0} = i_0$  and so

$i_0, (i_0)\pi, (i_0)\pi^2, \dots, (i_0)\pi^{l_0-1}$  are distinct. Let  $k$  be the remainder on dividing  $n-1$  by  $l_0$ . Then  $(i_0)\pi^k = (i_0)\pi^{n-1} = j_0$  and as  $0 < k < l_0$  we see that  $k$  is the smallest positive integer with  $(i_0)\pi^k = j_0$  and  $i_0, (i_0)\pi, (i_0)\pi^2, \dots, (i_0)\pi^{k-1}$  are distinct.

As  $A \in R_G$  taking  $i = i_0, j = (i_0)\pi$  we obtain  $d_{i_0} a_{i_0(i_0)\pi} \equiv 0 \pmod{d_{(i_0)\pi}}$  and so there is an integer  $z_1$  with  $d_{i_0} a_{i_0(i_0)\pi} = z_1 d_{(i_0)\pi}$ . In the same way there is an integer  $z_m$  with

$d_{(i_0)\pi^{m-1}} a_{(i_0)\pi^{m-1}(i_0)\pi^m} = z_m d_{(i_0)\pi^m}$  on taking  $i = (i_0)\pi^{m-1}, j = (i_0)\pi^m$  for  $1 \leq m < k$ . Substituting successively we obtain  $d_{i_0} a_{i_0(i_0)\pi} a_{(i_0)\pi(i_0)\pi^2} \cdots a_{(i_0)\pi^{k-1}j_0} = z_1 z_2 \cdots z_{k-1} d_{j_0}$  and so  $d_{i_0} t_\pi$  has factor  $d_{j_0}$ . As  $A_{j_0 i_0}$  is a sum of  $(s-1)!$  terms  $\pm t_\pi$  we conclude  $d_{i_0} A_{j_0 i_0} \equiv 0 \pmod{d_{j_0}}$  for  $1 \leq i_0, j_0 \leq s$  showing  $A \in R_G \Rightarrow \text{adj } A \in R_G$ .

(c) Suppose  $\alpha = (A)\varphi$  is an automorphism of  $G$ . Then  $\alpha^{-1}$  is also an automorphism of  $G$ . By (a) above there is an  $s \times s$  matrix  $B$  in  $R_G$  with  $\alpha^{-1} = (B)\varphi$ . So  $AB \equiv I \pmod{\ker \varphi}$ . All the entries in the matrices of  $\ker \varphi$  are divisible by  $p$  and so  $AB \equiv I \pmod{p}$ . Taking determinants gives  $|A| \not\equiv 0 \pmod{p}$ .

Conversely suppose  $|A| \not\equiv 0 \pmod{p}$ . Then  $\gcd\{|A|, p^{t_l}\} = 1$  and so there are integers  $b, b'$  with  $b|A| + b'p^{t_l} = 1$ . On multiplying  $A(\text{adj } A) = I = (\text{adj } A)A$  by  $b$  we see that the  $s \times s$  matrix  $B = b \text{adj } A$  over  $\mathbb{Z}$  satisfies  $AB \equiv I \pmod{p^{t_l}}$  and  $BA \equiv I \pmod{p^{t_l}}$ . Now  $Z \equiv 0 \pmod{p^{t_l}}$  implies  $Z \in \ker \varphi$  for all  $Z \in \mathfrak{M}_s(\mathbb{Z})$ . So  $AB \equiv I \pmod{\ker \varphi}$  and  $BA \equiv I \pmod{\ker \varphi}$ . By (b) above  $\text{adj } A$  and hence  $B$  belong to the ring  $R_G$  of (a) above. Let  $\beta = (B)\varphi$ . Then  $\alpha\beta = \iota$  and  $\beta\alpha = \iota$  showing that  $\beta = \alpha^{-1}$ . So  $\alpha \in \text{Aut } G$ .

Consider the entry  $a_{i'j'}$  in  $M_{ij}$  where  $i < j$ . As  $A \in R_G$  we see  $p^{t_i} a_{i'j'} \equiv 0 \pmod{p^{t_j}}$  which gives  $a_{i'j'} \equiv 0 \pmod{p}$  as  $t_i < t_j$ . So  $M_{ij} \equiv 0 \pmod{p}$  for  $i < j$ . Hence

$$|A| \equiv \begin{vmatrix} M_{11} & 0 & 0 & \cdots & 0 \\ M_{21} & M_{22} & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ M_{l-1,1} & M_{l-1,2} & \cdots & M_{l-1,l-1} & 0 \\ M_{l1} & M_{l2} & \cdots & M_{ll-1} & M_{ll} \end{vmatrix} \pmod{p}$$

and so  $|A| \equiv |M_{11}| |M_{22}| \cdots |M_{ll}| \pmod{p}$ . Hence  $|A| \not\equiv 0 \pmod{p}$  if and only if  $|M_{ii}| \not\equiv 0 \pmod{p}$  for  $1 \leq i \leq l$ , i.e.

$$\alpha = (A)\varphi \in \text{Aut } G \Leftrightarrow |M_{ii}| \not\equiv 0 \pmod{p} \text{ for } 1 \leq i \leq l.$$

The number of choices for each entry in  $M_{ij}$  ( $i \neq j$ ), so that  $\alpha = (A)\varphi \in \text{Aut } G$  remains as in (a) above, are  $p^{t_i}$  ( $i < j$ ) and  $p^{t_j}$  ( $i > j$ ). The number of integer solutions  $x$  with  $0 \leq x < p^{t_i}$  of

$x \equiv a \pmod{p}$  is  $p^{t_i-1}$ . Therefore the number of choices for the  $m_i \times m_i$  matrix  $M_{ii}$  so that  $\alpha = (A)\varphi \in \text{Aut } G$  is  $|GL_{m_i}(\mathbb{Z}_p)| \times (p^{t_i-1})^{m_i^2} = r_{m_i} (p^{t_i})^{m_i^2}$ , as each  $M_{ii}$  corresponds to a matrix  $X$  in  $GL_{m_i}(\mathbb{Z}_p)$  and each of the  $m_i^2$  entries  $\bar{a}$  in  $X$  contains  $p^{t_i-1}$  integers  $x$  in the range  $0 \leq x < p^{t_i}$  where  $1 \leq i \leq l$ . As the number of choices for  $M_{ii}$  so that  $\alpha = (A)\varphi \in \text{End } G$  is  $(p^{t_i})^{m_i^2}$  for  $1 \leq i \leq l$ , we conclude  $|\text{Aut } G| = r_{m_1} r_{m_2} \cdots r_{m_l} |\text{End } G|$ .

(d) Using the above partition into submatrices  $M_{ij}$  of the reduced matrix  $A$ , there are  $p^{t_i}$  choices for each of the  $m_i m_j$  entries in  $M_{ij}$  for  $i < j$ , namely any integer  $a = zp^{t_j-t_i}$  where  $0 \leq z < p^{t_i}$  as  $p^{t_i}a \equiv 0 \pmod{p^{t_j}}$  and  $0 \leq a < p^{t_j}$  since  $A \in R_G$  and  $A$  is reduced. So the number of possible matrices  $M_{ij}$  for  $i < j$  is  $p^{t_i m_i m_j}$ . For  $i \geq j$  there are  $p^{t_j}$  choices for each of the  $m_i m_j$  entries in  $M_{ij}$ , namely any integer  $z$  with  $0 \leq z < p^{t_j}$ . So the number of possible matrices  $M_{ij}$  for  $i \geq j$  is  $p^{t_j m_i m_j}$ . The total number of reduced matrices  $A \in R_G$  is  $p$  to the power of the sum of entries in the  $l \times l$  symmetric matrix

$$\begin{pmatrix} t_1 m_1^2 & t_1 m_1 m_2 & t_1 m_1 m_3 & \cdots & t_1 m_1 m_l \\ t_1 m_2 m_1 & t_2 m_2^2 & t_2 m_2 m_3 & \cdots & t_2 m_2 m_l \\ t_1 m_3 m_1 & t_2 m_3 m_2 & t_3 m_3^2 & \cdots & t_3 m_3 m_l \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ t_1 m_l m_1 & t_2 m_l m_2 & t_3 m_l m_3 & \cdots & t_l m_l^2 \end{pmatrix}.$$

The number of terms involving  $t_i m_i$  is  $m_i + 2(m_{i+1} + m_{i+2} + \dots + m_l)$ . Therefore  $|\text{End } G| = p^e$  where

$$e = \sum_{i=1}^l t_i m_i (m_i + 2(m_{i+1} + m_{i+2} + \dots + m_l)).$$

(e) Here  $d_1 = 3, d_2 = 9, d_3 = 9$ . The endomorphism condition  $d_i a_{ij} \equiv 0 \pmod{d_j}$  is automatically satisfied for  $i \geq j$  as  $d_j | d_i$ . As  $3 \times 3 \equiv 0 \pmod{9}$  we see  $d_1 a_{12} \equiv 0 \pmod{d_2}$ . Also  $3 \times 6 \equiv 0 \pmod{9}$  gives  $d_1 a_{13} \equiv 0 \pmod{d_3}$  and  $9 \times 4 \equiv 0 \pmod{9}$  gives  $d_2 a_{23} \equiv 0 \pmod{d_3}$  accounting for all cases with  $i < j$ . So  $A$  satisfies the endomorphism condition.

The entries  $a_{i1}$  in col 1 of  $A$  satisfy  $0 \leq a_{i1} < 3$ , the entries  $a_{i2}$  in col 2 of  $A$  satisfy  $0 \leq a_{i2} < 9$  and the entries  $a_{i3}$  in col 3 of  $A$  satisfy  $0 \leq a_{i3} < 9$ ; therefore  $A$  is reduced.

Now  $\text{adj } A = \begin{pmatrix} 31 & -9 & -30 \\ -5 & 10 & -2 \\ 1 & -2 & 11 \end{pmatrix}$  and  $|A| = 53 \equiv -1 \pmod{9}$ . So  $(-\text{adj } A)A \equiv I \pmod{9}$ . Now

$-\text{adj } A \equiv B \pmod{\ker \varphi}$  where  $B = \begin{pmatrix} 2 & 0 & 3 \\ 2 & 8 & 2 \\ 2 & 2 & 7 \end{pmatrix}$  is reduced. Then  $AB \equiv I \pmod{\ker \varphi}$ .

$$A^2 \equiv \begin{pmatrix} 1 & 6 & 0 \\ 0 & 2 & 0 \\ 1 & 3 & 2 \end{pmatrix} \pmod{\ker \varphi} \text{ and } A^4 \equiv \begin{pmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{pmatrix} \pmod{\ker \varphi}. \text{ Now } 4^3 \equiv 1 \pmod{9} \text{ and so}$$

$A^{12} \equiv I \pmod{\ker \varphi}$ . As  $A^3 \not\equiv I \pmod{\ker \varphi}$  and  $A^6 \not\equiv I \pmod{\ker \varphi}$  we conclude that  $\alpha$  has multiplicative order 12.

Yes,  $\alpha^{-1} = (B)\varphi$  as  $AB \equiv I \pmod{\ker \varphi}$ .

(f)(i) Here  $d_1 = 4, d_2 = 8, d_3 = 8, s = 3$ . By the analogue of Frobenius' theorem in (a) above

$|\text{End } G| = 4^5 \times 8^3 \times 8 = 2^{22}$ . Alternatively using the formula of (d) above with

$p = 2, t_1 = 2, t_2 = 3, m_1 = 1, m_2 = 2, l = 2$  we obtain  $|\text{End } G| = 2^{2 \times 5 + 6 \times 2} = 2^{22}$ . We have

$r_1 = |GL_1(\mathbb{Z}_2)|/2 = (2-1)/2 = 1/2$  and  $r_2 = |GL_2(\mathbb{Z}_2)|/2^4 = (2^2 - 2)(2^2 - 1)/2^4 = 3/8$ . Using (c)

above  $|\text{Aut } G| = r_1 r_2 |\text{End } G| = (1/2)(3/8)2^{22} = 3 \times 2^{18}$ .

(ii) In this case  $d_1 = 3, d_2 = 3, d_3 = 9, s = 3$ . By the analogue of Frobenius' theorem

$|\text{End } G| = 3^5 \times 3^3 \times 9 = 3^{10}$ . Alternatively using the formula of (d) above with

$p = 3, t_1 = 1, t_2 = 2, m_1 = 2, m_2 = 1, l = 2$  we obtain  $|\text{End } G| = 3^{2 \times 4 + 2 \times 1} = 3^{10}$ . We have

$r_2 = |GL_2(\mathbb{Z}_3)|/3^2 = (3^2 - 1)(3^2 - 3)/3^2 = 16/3$  and  $r_1 = |GL_1(\mathbb{Z}_3)|/3 = (3-1)/3 = 2/3$ . Using (c)

above  $|\text{Aut } G| = r_2 r_1 |\text{End } G| = (16/3)(2/3)3^{10} = 2^5 \times 3^8$ .

(iii) The 2-component and 3-component of a group  $G$  of isomorphism type  $C_{12} \oplus C_{24} \oplus C_{72}$  have isomorphism types  $C_4 \oplus C_8 \oplus C_8$  (as in (i) above) and  $C_3 \oplus C_3 \oplus C_9$  (as in (ii) above) respectively.

As  $\text{End } G$  is the direct sum of the endomorphism rings of the prime components of  $G$  we see

$|\text{End } G| = 2^{22} \times 3^{10}$ . Alternatively the analogue of Frobenius' theorem gives

$|\text{End } G| = 12^5 \times 24^3 \times 72 = 2^{22} \times 3^{10}$  directly. As  $\text{Aut } G$  is the direct product of the automorphism

groups of the prime components of  $G$  we obtain  $|\text{Aut } G| = (3 \times 2^{18}) \times (2^5 \times 3^8) = 2^{23} \times 3^9$ .

(g) Both  $A$  and  $B$  are invertible over  $\mathbb{Z}$  and belong to the ring  $R_G$ . So  $(A)\varphi$  and  $(B)\varphi$  are in

$\text{Aut } G$ . Also

$$AB = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & d_2/d_1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & d_2/d_1 \\ 1 & (d_2/d_1) + 1 \end{pmatrix} \text{ and } BA = \begin{pmatrix} 1 + (d_2/d_1) & d_2/d_1 \\ 1 & 1 \end{pmatrix}.$$

As  $(d_2/d_1) + 1 \not\equiv 1 \pmod{d_2}$  we see  $AB \not\equiv BA \pmod{\ker \varphi}$  on comparing  $(2, 2)$ -entries. So  $(A)\varphi$

and  $(B)\varphi$  do not commute and  $\text{Aut } G$  is non-abelian.

Suppose  $d_1 = 0$ . Then  $G$  is free of rank 2 and  $\text{Aut } G \cong GL_2(\mathbb{Z})$ , the group of invertible  $2 \times 2$  matrices

over  $\mathbb{Z}$ . This group is non-abelian as  $C = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$  satisfies  $AC = \begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix} \neq \begin{pmatrix} 1 & 0 \\ -1 & -1 \end{pmatrix} = CA$ . So

$d_1 \neq 0$  and therefore  $d_1 \geq 2$ , as 1 is never an invariant factor. In this case  $R_G$  consists of matrices

$\begin{pmatrix} a_{11} & 0 \\ a_{21} & a_{22} \end{pmatrix}$  as  $d_1 a_{12} \equiv 0 \pmod{0} \Leftrightarrow d_1 a_{12} = 0 \Leftrightarrow a_{12} = 0$ . Also  $\text{Aut } G$  abelian implies  $(A)\varphi$  and

$(C)\varphi$  commute and so  $AC \equiv CA \pmod{\ker \varphi}$  where



$$\ker \varphi = \left\{ \begin{pmatrix} d_1 a'_{11} & 0 \\ d_1 a'_{12} & 0 \end{pmatrix} : a_{11}, a_{12} \in \mathbb{Z} \right\}.$$

So  $1 \equiv -1 \pmod{d_1}$  showing  $d_1 \mid 2$  on comparing  $(2,1)$ -entries. Therefore  $d_1 = 2$ . In this case  $\text{Aut } G$  is abelian as  $G$  has just 4 automorphisms, namely those corresponding to

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix}.$$

### Solution 6

(a) For  $g_1, g_2 \in G$  we have

$$\begin{aligned} (g_1 + g_2)(\alpha + \beta) &= (g_1 + g_2)\alpha + (g_1 + g_2)\beta = (g_1)\alpha + (g_2)\alpha + (g_1)\beta + (g_2)\beta = \\ &= (g_1)\alpha + (g_1)\beta + (g_2)\alpha + (g_2)\beta = (g_1)(\alpha + \beta) + (g_2)(\alpha + \beta) \end{aligned}$$

showing that  $\alpha + \beta : G \rightarrow G'$  is a group homomorphism.

We verify the axioms for an additive abelian group in the case of  $(\text{Hom}(G, G'), +)$ . Consider

$\alpha, \beta, \gamma \in \text{Hom}(G, G')$ . Then for  $g \in G$  we have

$$(g)((\alpha + \beta) + \gamma) = (g)(\alpha + \beta) + (g)\gamma = (g)\alpha + (g)\beta + (g)\gamma = (g)\alpha + (g)(\beta + \gamma) = (g)(\alpha + (\beta + \gamma))$$

showing  $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$ . The zero mapping  $0 : G \rightarrow G'$ , given by  $(g)0 = 0$  (the zero element of  $G'$ ) for all  $g \in G$ , belongs to  $\text{Hom}(G, G')$  and satisfies  $0 + \alpha = \alpha$  for all  $\alpha \in \text{Hom}(G, G')$ .

For each  $\alpha \in \text{Hom}(G, G')$  the mapping  $-\alpha : G \rightarrow G'$ , given by  $(g)(-\alpha) = -g$  for all  $g \in G$ , belongs to  $\text{Hom}(G, G')$  and satisfies  $-\alpha + \alpha = 0$ . Also  $\alpha + \beta = \beta + \alpha$  for all  $\alpha, \beta \in \text{Hom}(G, G')$ . Therefore  $(\text{Hom}(G, G'), +)$  is an additive abelian group.

(b) Let  $\iota_1 : G_1 \rightarrow G_1 \oplus G_2$  be defined by  $(g_1)\iota_1 = (g_1, 0)$  for all  $g_1 \in G_1$ . Also let  $\iota_2 : G_2 \rightarrow G_1 \oplus G_2$  be defined by  $(g_2)\iota_2 = (0, g_2)$  for all  $g_2 \in G_2$ . Then  $\iota_1$  and  $\iota_2$  are injective group homomorphisms. For

$\alpha \in \text{Hom}(G_1 \oplus G_2, G')$  write  $(\alpha)\theta = (\iota_1\alpha, \iota_2\alpha)$ . As  $\iota_i\alpha \in \text{Hom}(G_i, G')$  for  $i=1, 2$ , we see

$\theta : \text{Hom}(G_1 \oplus G_2, G') \rightarrow \text{Hom}(G_1, G') \oplus \text{Hom}(G_2, G')$ . Consider  $\alpha, \beta \in \text{Hom}(G_1 \oplus G_2, G')$ . As

$$(g_i)\iota_i(\alpha + \beta) = (g_i)\iota_i\alpha + (g_i)\iota_i\beta \text{ for all } g_i \in G_i \text{ we see } \iota_i(\alpha + \beta) = \iota_i\alpha + \iota_i\beta \text{ for } i=1, 2. \text{ Hence}$$

$$(\alpha + \beta)\theta = (\iota_1(\alpha + \beta), \iota_2(\alpha + \beta)) = (\iota_1\alpha + \iota_1\beta, \iota_2\alpha + \iota_2\beta) = (\iota_1\alpha, \iota_2\alpha) + (\iota_1\beta, \iota_2\beta) = (\alpha)\theta + (\beta)\theta$$

showing that  $\theta$  is additive. Suppose  $(\alpha)\theta = (\iota_1\alpha, \iota_2\alpha) = 0$ . Then  $\iota_1\alpha = 0$  and  $\iota_2\alpha = 0$ . As

$$(g_1, g_2)\alpha = ((g_1, 0) + (0, g_2))\alpha = ((g_1)\iota_1 + (g_2)\iota_2)\alpha = (g_1)\iota_1\alpha + (g_2)\iota_2\alpha = (g_1)0 + (g_2)0 = 0$$

we see  $\alpha = 0$  and so  $\ker \theta = 0$ . Consider  $\alpha_i \in \text{Hom}(G_i, G')$  for  $i=1, 2$ . We define  $\alpha : G_1 \oplus G_2 \rightarrow G'$

by  $(g_1, g_2)\alpha = (g_1)\alpha_1 + (g_2)\alpha_2$  for all  $(g_1, g_2) \in G_1 \oplus G_2$ . Then  $\alpha$  is additive as

$$((g_1, g_2) + (g'_1, g'_2))\alpha = (g_1 + g'_1, g_2 + g'_2)\alpha = (g_1 + g'_1)\alpha_1 + (g_2 + g'_2)\alpha_2 =$$

$$(g_1)\alpha_1 + (g'_1)\alpha_1 + (g_2)\alpha_2 + (g'_2)\alpha_2 = (g_1)\alpha_1 + (g_2)\alpha_2 + (g'_1)\alpha_1 + (g'_2)\alpha_2 = (g_1, g_2)\alpha + (g'_1, g'_2)\alpha$$

for all  $g_i, g'_i \in G_i$  for  $i=1, 2$  on commuting the elements  $(g'_1)\alpha_1$  and  $(g_2)\alpha_2$  of the abelian group  $G'$ .

So  $\alpha \in \text{Hom}(G_1 \oplus G_2, G')$ . As  $(g_1)\iota_1\alpha = (g_1, 0)\alpha = (g_1)\alpha_1 + (0)\alpha_2 = (g_1)\alpha_1 + 0 = (g_1)\alpha_1$  for all

$g_1 \in G_1$  we see  $\iota_1\alpha = \alpha_1$  and in the same way  $\iota_2\alpha = \alpha_2$ . Therefore  $(\alpha)\theta = (\alpha_1, \alpha_2)$  which shows

$\text{im } \theta = \text{Hom}(G_1, G') \oplus \text{Hom}(G_2, G')$ . The conclusion is:  $\theta$  is a group isomorphism, i.e.

$$\theta : \text{Hom}(G_1 \oplus G_2, G') \cong \text{Hom}(G_1, G') \oplus \text{Hom}(G_2, G').$$

Let  $\pi_1 : G'_1 \oplus G'_2 \rightarrow G'_1$  and  $\pi_2 : G'_1 \oplus G'_2 \rightarrow G'_2$  be the projections defined by  $(g'_1, g'_2)\pi_1 = g'_1$  and  $(g'_1, g'_2)\pi_2 = g'_2$  for all  $(g'_1, g'_2) \in G'_1 \oplus G'_2$ . Then  $\pi_1$  and  $\pi_2$  are surjective group homomorphisms. For  $\alpha \in \text{Hom}(G, G'_1 \oplus G'_2)$  write  $(\alpha)\varphi = (\alpha\pi_1, \alpha\pi_2)$ . As  $\alpha\pi_i \in \text{Hom}(G, G'_i)$  for  $i=1,2$  we see  $\varphi : \text{Hom}(G, G'_1 \oplus G'_2) \rightarrow \text{Hom}(G, G'_1) \oplus \text{Hom}(G, G'_2)$ . Consider  $\alpha, \beta \in \text{Hom}(G, G'_1 \oplus G'_2)$ . As  $(g)(\alpha + \beta)\pi_i = ((g)\alpha + (g)\beta)\pi_i = (g)\alpha\pi_i + (g)\beta\pi_i$  for all  $g \in G$  we see  $(\alpha + \beta)\pi_i = \alpha\pi_i + \beta\pi_i$  for  $i=1,2$ . Hence

$$(\alpha + \beta)\varphi = ((\alpha + \beta)\pi_1, (\alpha + \beta)\pi_2) = (\alpha\pi_1 + \beta\pi_1, \alpha\pi_2 + \beta\pi_2) = (\alpha\pi_1, \alpha\pi_2) + (\beta\pi_1, \beta\pi_2) = (\alpha)\varphi + (\beta)\varphi$$

showing that  $\varphi$  is additive. Suppose  $(\alpha)\varphi = (\alpha\pi_1, \alpha\pi_2) = 0$ , i.e.  $\alpha\pi_1 = 0$  and  $\alpha\pi_2 = 0$ . Write  $(g)\alpha = (g'_1, g'_2)$  where  $g \in G$ . Then  $g'_i = (g)\alpha\pi_i = (g)0 = 0$  for  $i=1,2$ . So  $(g)\alpha = (0,0)$  for all  $g \in G$  showing  $\alpha = 0$ . Therefore  $\ker \varphi = 0$ . Consider  $\alpha_i \in \text{Hom}(G, G'_i)$  for  $i=1,2$ . We define  $\alpha : G \rightarrow G'_1 \oplus G'_2$  by  $(g)\alpha = ((g)\alpha_1, (g)\alpha_2)$  for all  $g \in G$ . Then  $\alpha$  is additive as

$$\begin{aligned} (g + g')\alpha &= ((g + g')\alpha_1, (g + g')\alpha_2) = ((g)\alpha_1 + (g')\alpha_1, (g)\alpha_2 + (g')\alpha_2) = \\ &= ((g)\alpha_1, (g)\alpha_2) + ((g')\alpha_1, (g')\alpha_2) = (g)\alpha + (g')\alpha \end{aligned}$$

for  $g, g' \in G$ . So  $\alpha \in \text{Hom}(G, G'_1 \oplus G'_2)$ . As  $(g)\alpha\pi_i = (g)\alpha_i$  for all  $g \in G$  we see  $\alpha\pi_i = \alpha_i$  for  $i=1,2$ . Therefore  $\text{im } \varphi = \text{Hom}(G, G'_1) \oplus \text{Hom}(G, G'_2)$ . So  $\varphi$  is a group isomorphism, i.e.

$$\varphi : \text{Hom}(G, G'_1 \oplus G'_2) \cong \text{Hom}(G, G'_1) \oplus \text{Hom}(G, G'_2).$$

(c) Write  $d = \gcd\{m, n\}$ . Then  $\overline{n/d} \in \mathbb{Z}_n$  has order ideal  $\langle d \rangle = \{r \in \mathbb{Z} : r(\overline{n/d}) = \bar{0}\}$ . As  $\bar{1} \in \mathbb{Z}_m$  has order ideal  $\langle m \rangle$  and  $d \mid m$  by Exercises 2.1, Question 4(b) there is a unique  $\alpha_0 \in \text{Hom}(\mathbb{Z}_m, \mathbb{Z}_n)$  with  $(\bar{1})\alpha_0 = \overline{n/d}$ . Also each  $\alpha \in \text{Hom}(\mathbb{Z}_m, \mathbb{Z}_n)$  is of the form  $\alpha = r\alpha_0$  for  $1 \leq r \leq d$ . So  $\text{Hom}(\mathbb{Z}_m, \mathbb{Z}_n)$  is cyclic of isomorphism type  $C_d$  being generated by  $\alpha_0$ .

Let  $\alpha_0 : \mathbb{Z} \rightarrow \mathbb{Z}/\langle n \rangle = \mathbb{Z}_n$  denote the natural mapping. Then  $\alpha_0 \in \text{Hom}(\mathbb{Z}, \mathbb{Z}_n)$  has additive order  $n$ . Also  $\text{Hom}(\mathbb{Z}, \mathbb{Z}_n)$  has isomorphism type  $C_n$  and generator  $\alpha_0$ .

Consider  $\alpha \in \text{Hom}(\mathbb{Z}_m, \mathbb{Z})$ . Then  $m(\bar{1})\alpha = (\bar{m})\alpha = (\bar{0})\alpha = 0$ . As  $\mathbb{Z}$  has no divisors of zero we see  $(\bar{1})\alpha = 0$  and so  $\alpha = 0$ . Therefore  $\text{Hom}(\mathbb{Z}_m, \mathbb{Z})$  is trivial, i.e. cyclic of isomorphism type  $C_1$ .

The identity mapping  $\iota$  of  $\mathbb{Z}$  generates  $\text{Hom}(\mathbb{Z}, \mathbb{Z})$  by (3.15) with  $t=1$ . So  $\text{Hom}(\mathbb{Z}, \mathbb{Z})$  has isomorphism type  $C_0$ .

By (3.4) the f.g. abelian group  $G$  decomposes as an internal direct sum  $G = H_1 \oplus H_2 \oplus \dots \oplus H_s$  where  $H_i$  is a cyclic subgroup of  $G$  for  $1 \leq i \leq s$ . In the same way  $G' = H'_1 \oplus H'_2 \oplus \dots \oplus H'_{s'}$  where  $H'_j$  is a cyclic subgroup of  $G'$  for  $1 \leq j \leq s'$ . By (b) above and induction

$$\text{Hom}(G, G') \cong \sum_{i,j} \oplus \text{Hom}(H_i, H'_j)$$

where the direct summation has  $ss'$  terms  $\text{Hom}(H_i, H'_j)$  each being cyclic by the preceding part for  $1 \leq i \leq s$ ,  $1 \leq j \leq s'$ . Therefore  $\text{Hom}(G, G')$  is generated by  $ss'$  of its elements, one for each cyclic group  $\text{Hom}(H_i, H'_j)$ . So  $\text{Hom}(G, G')$  is finitely generated.

(d) By (3.4) the finite abelian group  $G$  decomposes  $G = H_1 \oplus H_2 \oplus \dots \oplus H_s$  where  $H_i$  is a cyclic subgroup of  $G$  with  $H_i \cong \mathbb{Z}_{d_i}$  for  $1 \leq i \leq s$ . In the same way  $G' = H'_1 \oplus H'_2 \oplus \dots \oplus H'_{s'}$  where  $H'_j$  is a cyclic subgroup of  $G'$  with  $H'_j \cong \mathbb{Z}_{d'_j}$  for  $1 \leq j \leq s'$ . By (c) above  $\text{Hom}(H_i, H'_j) \cong \mathbb{Z}_{\gcd\{d_i, d'_j\}}$  and

$$\text{Hom}(G, G') \cong \sum_{i,j} \oplus \text{Hom}(H_i, H'_j) \cong \sum_{i,j} \oplus \mathbb{Z}_{\gcd\{d_i, d'_j\}}$$

which expresses  $\text{Hom}(G, G')$  as a direct sum of finite cyclic groups. As  $\gcd\{d_i, d'_j\} = \gcd\{d'_j, d_i\}$  we see  $\text{Hom}(H_i, H'_j) \cong \text{Hom}(H'_j, H_i)$  for  $1 \leq i \leq s$  and  $1 \leq j \leq s'$ . So yes  $\text{Hom}(G, G')$  and  $\text{Hom}(G', G)$ , where  $G$  and  $G'$  are finite abelian groups, are isomorphic as

$$\begin{aligned} \text{Hom}(G, G') &\cong \sum_{i,j} \oplus \text{Hom}(H_i, H'_j) \cong \sum_{i,j} \oplus \mathbb{Z}_{\gcd\{d_i, d'_j\}} = \\ &\sum_{i,j} \oplus \mathbb{Z}_{\gcd\{d'_j, d_i\}} \cong \sum_{j,i} \oplus \text{Hom}(H'_j, H_i) \cong \text{Hom}(G', G). \end{aligned}$$

As  $\gcd\{d_i, d_j\} = d_i = \gcd\{d_j, d_i\}$  for  $1 \leq i \leq j \leq s$  the invariant factors of  $\text{End } G = \text{Hom}(G, G)$  are the entries in the  $s \times s$  matrix

$$\begin{pmatrix} d_1 & d_1 & d_1 & \cdots & d_1 \\ d_1 & d_2 & d_2 & \cdots & d_2 \\ d_1 & d_2 & d_3 & \cdots & d_3 \\ \vdots & \vdots & \vdots & & \vdots \\ d_1 & d_2 & d_3 & \cdots & d_s \end{pmatrix}$$

which we used in the solution of Question 5(a) above. Therefore  $\text{End } G$  has  $2(s-i)+1$  invariant factors  $d_i$  for  $i=1, 2, \dots, s$ .

The table having  $(i, j)$ -entry  $\gcd\{d_i, d'_j\}$  is:

gcd	3	3	6	24
2	1	1	2	2
6	3	3	6	6
12	3	3	6	12

in this case. So  $\text{Hom}(G, G')$  has isomorphism type

$$\begin{aligned} C_1 \oplus C_1 \oplus C_2 \oplus C_2 \oplus C_3 \oplus C_3 \oplus C_3 \oplus C_3 \oplus C_6 \oplus C_6 \oplus C_6 \oplus C_{12} = \\ C_3 \oplus C_3 \oplus C_6 \oplus C_6 \oplus C_6 \oplus C_6 \oplus C_6 \oplus C_{12} \end{aligned}$$

on using  $C_2 \oplus C_3 = C_6$  twice and rearranging. Therefore  $(3, 3, 6, 6, 6, 6, 6, 12)$  is the invariant factor sequence of  $\text{Hom}(G, G')$ .

As  $G \oplus G'$  has isomorphism type

$$C_2 \oplus C_6 \oplus C_{12} \oplus C_3 \oplus C_3 \oplus C_6 \oplus C_{24} = C_3 \oplus C_6 \oplus C_6 \oplus C_6 \oplus C_{12} \oplus C_{24}$$

the table as above in the case of  $\text{Hom}(G, G \oplus G')$  is

gcd	3	6	6	6	12	24
2	1	2	2	2	2	2
6	3	6	6	6	6	6
12	3	6	6	6	12	12

and so  $\text{Hom}(G, G \oplus G')$  has invariant factor sequence  $(2, 2, 2, 6, 6, 6, 6, 6, 6, 6, 12, 12)$ .

The gcd table in the case of  $\text{Hom}(G', G \oplus G') \cong \text{Hom}(G \oplus G', G')$  is

gcd	3	6	6	6	12	24
3	3	3	3	3	3	3
3	3	3	3	3	3	3
6	3	6	6	6	6	6
24	3	6	6	6	12	24

and so  $\text{Hom}(G \oplus G', G')$  has invariant factor sequence

$(3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 6, 6, 6, 6, 6, 6, 6, 6, 12, 24)$ . In the case of  $\text{End } G$  the gcd table is

gcd	2	6	12
2	2	2	2
6	2	6	6
12	2	6	12

and so  $(2, 2, 2, 2, 2, 6, 6, 6, 12)$  is its invariant factor sequence.

(e) As  $T(G)$  is finite and  $M(G')$  is free we see  $\text{Hom}(T(G), M(G')) = 0$  using (b) above, (3.4) and  $\text{Hom}(\mathbb{Z}_m, \mathbb{Z}) = 0$  from (c) above. Also from (c) above we know  $\text{Hom}(\mathbb{Z}, \mathbb{Z}_m) \cong \mathbb{Z}_m$  and so  $\text{Hom}(M(G), T(G')) \cong T(G') \oplus T(G') \oplus \dots \oplus T(G')$  the direct sum of  $r = \text{rank } M(G)$  copies of  $T(G')$ . By (b) above we have

$$\begin{aligned} \text{Hom}(G, G') &= \text{Hom}(T(G) \oplus M(G), T(G') \oplus M(G')) \cong \\ &\text{Hom}(T(G), T(G')) \oplus \text{Hom}(M(G), T(G')) \oplus \text{Hom}(M(G), M(G')). \end{aligned}$$

the first two summands (terms) in the preceding decomposition are finite and the third is free, i.e.  $\text{Hom}(T(G), T(G')) \oplus \text{Hom}(M(G), T(G'))$  is isomorphic to the torsion subgroup of  $\text{Hom}(G, G')$  and  $\text{Hom}(M(G), M(G'))$  is free of rank  $rr'$  using  $\text{Hom}(\mathbb{Z}, \mathbb{Z}) \cong \mathbb{Z}$ , (b) above and induction. So  $\text{Hom}(G, G')$  has torsion-free rank  $rr'$ . Taking  $G' = G$  we see

$$\mathrm{End}T(G) \oplus T(G) \oplus T(G) \oplus \dots \oplus T(G)$$

where there are  $r$  summands  $T(G)$ , is isomorphic to the torsion subgroup of  $\text{End } G$  and  $r^2$  is the torsion-free rank of  $\text{End } G$ .

As  $T(G)$  and  $T(G')$  have invariant factor sequences  $(2, 4)$  and  $(2, 2, 4, 4)$  respectively we construct the table

gcd	2	2	4	4
2	2	2	2	2
4	2	2	4	4

from which we see that  $\text{Hom}(T(G), T(G'))$  has invariant factor sequence  $(2, 2, 2, 2, 2, 2, 4, 4)$ .

Therefore  $\text{Hom}(G, G')$  has invariant factor sequence  $(2, 2, 2, 2, 2, 2, 2, 2, 4, 4, 4, 4, 0, 0)$  as

$\text{Hom}(T(G), T(G')) \oplus T(G')$  is isomorphic to  $T(\text{Hom}(G, G'))$  and 2 is the torsion-free rank of  $\text{Hom}(G, G')$ .

Interchanging  $G$  and  $G'$  we see  $T(\text{Hom}(G', G)) \cong \text{Hom}(T(G'), T(G)) \oplus T(G) \oplus T(G)$  has invariant factor sequence  $(2, 2, 2, 2, 2, 2, 2, 4, 4, 4, 4)$ . Therefore  $\text{Hom}(G', G)$  and  $\text{Hom}(G, G')$  have the same invariant factor sequence and so are isomorphic in this case.

[illegible]

## Solutions (page 163)

### Solution 1

$$\begin{aligned}
 xI - A &= \begin{pmatrix} x-3 & -1 & -1 \\ 8 & x+3 & 4 \\ -4 & -2 & x-3 \end{pmatrix} \begin{matrix} \equiv \\ c_1 + (x-3)c_2 \\ c_1 \leftrightarrow c_2 \end{matrix} \begin{pmatrix} -1 & 0 & -1 \\ x+3 & x^2-1 & 4 \\ -2 & -2x+2 & x-3 \end{pmatrix} \begin{matrix} \equiv \\ -c_1 \\ c_3 + c_1 \end{matrix} \\
 &\begin{pmatrix} 1 & 0 & 0 \\ -x-3 & x^2-1 & -x+1 \\ 2 & -2x+2 & x-1 \end{pmatrix} \begin{matrix} \equiv \\ r_2 + (x+3)r_1 \\ r_3 - 2r_1 \end{matrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & x^2-1 & -x+1 \\ 0 & -2x+2 & x-1 \end{pmatrix} \begin{matrix} \equiv \\ c_2 + (x+1)c_3 \\ c_2 \leftrightarrow c_3 \end{matrix} \\
 &\begin{pmatrix} 1 & 0 & 0 \\ 0 & -x+1 & 0 \\ 0 & x-1 & (x-1)^2 \end{pmatrix} \begin{matrix} \equiv \\ -c_2 \\ r_3 + r_2 \end{matrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & x-1 & 0 \\ 0 & 0 & (x-1)^2 \end{pmatrix} = D(x).
 \end{aligned}$$

Therefore the rcf of  $A$  is  $C(x-1) \oplus C((x-1)^2) = \left( \begin{array}{c|cc} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & -1 & 2 \end{array} \right) = C$ .

Applying the *eros* in the above sequence to the  $3 \times 3$  identity matrix  $I$  gives

$$I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{matrix} \equiv \\ r_2 + (x+3)r_1 \\ r_3 - 2r_1 \end{matrix} \begin{pmatrix} 1 & 0 & 0 \\ x+3 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix} \begin{matrix} \equiv \\ r_3 + r_2 \\ \end{matrix} \begin{pmatrix} 1 & 0 & 0 \\ x+3 & 1 & 0 \\ x+1 & 1 & 1 \end{pmatrix} = P(x).$$

Applying the conjugates of the *ecos* in the above sequence to the  $3 \times 3$  identity matrix  $I$  gives

$$\begin{aligned}
 I &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{matrix} \equiv \\ r_2 - (x-3)r_1 \\ r_1 \leftrightarrow r_2 \end{matrix} \begin{pmatrix} -x+3 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{matrix} \equiv \\ -r_1 \\ r_1 + r_3 \end{matrix} \begin{pmatrix} x-3 & -1 & -1 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{matrix} \equiv \\ r_3 - (x+1)r_2 \\ r_2 \leftrightarrow r_3 \end{matrix} \\
 &\begin{pmatrix} x-3 & -1 & -1 \\ -x-1 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \begin{matrix} \equiv \\ -r_2 \\ \end{matrix} \begin{pmatrix} x-3 & -1 & -1 \\ x+1 & 0 & -1 \\ 1 & 0 & 0 \end{pmatrix} = Q(x).
 \end{aligned}$$

Then  $P(x)$  and  $Q(x)$  are invertible over  $\mathbb{Q}[x]$  and satisfy  $P(x)(xI - A) = D(x)Q(x)$ . Note that  $P(x)$  and  $Q(x)$  are not unique – the chances are that you will have found a valid but different pair  $P(x), Q(x)$  having reduced  $xI - A$  to  $D(x)$  using a different sequence of elementary operations.

Now  $(\rho_1(x))\theta_A = ((x-3)e_1 - e_2 - e_3)\theta_A = e_1A - (3,1,1) = 0$  which has order 1 in  $M(A)$ . Write

$v_1 = (\rho_2(x))\theta_A = (x+1, 0, -1)\theta_A = (xe_1)\theta_A + (e_1 - e_3)\theta_A = e_1A + e_1 - e_3 = (3,1,1) + (1,0,-1) = (4,1,0)$ .

Then  $v_1$  has order  $x-1$  in  $M(A)$  as  $v_1 \neq 0$  but  $(x-1)v_1 = (4,1,0)(A - I) = (0,0,0)$ . Write

$v_2 = (\rho_3(x))\theta_A = (1,0,0)$ . As  $v_2 = (1,0,0)$  and  $xv_2 = (3,1,1)$  are linearly independent and

$x^2v_2 = xv_2A = (3,1,1)A = (5,2,2)$  we see  $x^2v_2 = 2xv_2 - v_2$ , i.e.  $(x-1)^2v_2 = 0$ . So  $v_2$  has order  $(x-1)^2$

in  $M(A)$ . The matrix

$$X = \begin{pmatrix} v_1 \\ v_2 \\ xv_2 \end{pmatrix} = \begin{pmatrix} 4 & 1 & 0 \\ 1 & 0 & 0 \\ 3 & 1 & 1 \end{pmatrix}$$

satisfies  $\det X = -1 \neq 0$  and  $XA = CX$ . Hence  $X$  is invertible over  $\mathbb{Q}$  and  $XAX^{-1} = C$ .

Replace  $xv_2$  in  $X$  by  $(x-1)v_2$  which has order  $x-1$  in  $M(A)$ . Let

$$X_1 = \begin{pmatrix} v_1 \\ v_2 \\ (x-1)v_2 \end{pmatrix} = \begin{pmatrix} 4 & 1 & 0 \\ 1 & 0 & 0 \\ 2 & 1 & 1 \end{pmatrix}.$$

Then  $\det X_1 = -1 \neq 0$  and  $X_1 A = J X_1$  where

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

is the Jnf of  $A$ . So  $X_1$  is invertible over  $\mathbb{Q}$  and  $X_1 A X_1^{-1} = J$  is in Jnf. Rows 1 and 3 of  $X_1$ , i.e.  $(4,1,0)$  and  $(2,1,1)$ , are linearly independent eigenvectors of  $A$  with eigenvalue 1.

### Solution 2

$$\begin{aligned} xI - A &= \begin{pmatrix} x-2 & 2 & -1 \\ -2 & x+3 & -2 \\ -1 & 2 & x-2 \end{pmatrix} \begin{matrix} \equiv \\ r_1 \leftrightarrow r_3 \\ -c_1 \end{matrix} \begin{pmatrix} 1 & 2 & x-2 \\ 2 & x+3 & -2 \\ 2-x & 2 & -1 \end{pmatrix} \begin{matrix} \equiv \\ c_2 - 2c_1 \\ c_3 - (x-2)c_1 \end{matrix} \begin{pmatrix} 1 & 0 & 0 \\ 2 & x-1 & -2x+2 \\ 2-x & 2x-2 & (x-3)(x-1) \end{pmatrix} \\ &\equiv \begin{pmatrix} 1 & 0 & 0 \\ 0 & x-1 & 2-2x \\ 0 & 2x-2 & (x-3)(x-1) \end{pmatrix} \begin{matrix} \equiv \\ r_2 - 2r_1 \\ r_3 + (x-2)r_1 \end{matrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & x-1 & 0 \\ 0 & 0 & (x+1)(x-1) \end{pmatrix} \begin{matrix} \equiv \\ c_3 + 2c_2 \\ r_3 - 2r_2 \end{matrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & x-1 & 0 \\ 0 & 0 & (x+1)(x-1) \end{pmatrix} = D(x). \end{aligned}$$

Applying the *eros* in the above sequence to  $I$ :

$$I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{matrix} \equiv \\ r_1 \leftrightarrow r_3 \\ r_2 - 2r_1 \end{matrix} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & -2 \\ 1 & 0 & 0 \end{pmatrix} \begin{matrix} \equiv \\ r_3 + (x-2)r_1 \\ r_3 - 2r_2 \end{matrix} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & -2 \\ 1 & -2 & x+2 \end{pmatrix} = P(x).$$

Applying the conjugates of the *ecos* in the above sequence to  $I$ :

$$I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{matrix} \equiv \\ -r_1 \\ r_1 + 2r_2 \end{matrix} \begin{pmatrix} -1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{matrix} \equiv \\ r_1 + (x-2)r_3 \\ r_2 - 2r_3 \end{matrix} \begin{pmatrix} -1 & 2 & x-2 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \rho_1(x) \\ \rho_2(x) \\ \rho_3(x) \end{pmatrix} = Q(x).$$

Then  $P(x)$  and  $Q(x)$  are invertible over  $\mathbb{Q}[x]$  as  $\det P(x) = \det Q(x) = -1$  and satisfy

$P(x)(xI - A) = D(x)Q(x)$ . The rcf of  $A$  is

$$C(x-1) \oplus C((x+1)(x-1)) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix} = C.$$

Evaluating the rows of  $Q(x)$  in  $M(A)$ :  $(\rho_1(x))\theta_A = (-1, 2, x-2)\theta_A = (-1, 2, -2) + (0, 0, 1)A = (0, 0, 0)$  which has order 1 in  $M(A)$ . Also  $v_1 = (\rho_2(x))\theta_A = (0, 1, -2)\theta_A = (0, 1, -2)$  has order  $x-1$  in  $M(A)$  as  $v_1 \neq 0$  and  $xv_1 = v_1 A = (0, 1, -2) = v_1$ , i.e.  $(x-1)v_1 = 0$ . Also  $v_2 = (\rho_3(x))\theta_A = (0, 0, 1)$  has order  $x^2 - 1$  in  $M(A)$  as  $v_2 = (0, 0, 1)$ ,  $xv_2 = (1, -2, 2)$  are linearly independent and  $x^2 v_2 = (1, -2, 2)A = (0, 0, 1) = v_2$ , i.e.  $(x^2 - 1)v_2 = 0$ . We conclude that the order of  $(\rho_i(x))\theta_A$  in  $M(A)$  is the  $(i, i)$ -entry in  $D(x)$  for  $i = 1, 2, 3$ . The matrix

$$X = \begin{pmatrix} \frac{v_1}{v_2} \\ \frac{v_1}{v_2} \\ xv_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 & -2 \\ 0 & 0 & 1 \\ 1 & -2 & 2 \end{pmatrix}$$

has  $\det X = 1$  and satisfies  $XA = CX$ . Hence  $X$  is invertible over  $\mathbb{Q}$  and  $XAX^{-1} = C$ .

As  $v_2$  has order  $x^2 - 1 = (x+1)(x-1)$  in  $M(A)$ , we see that  $(x+1)v_2$  and  $(x-1)v_2$  have orders  $x-1$  and  $x+1$  in  $M(A)$  respectively. So

$$X_1 = \begin{pmatrix} \frac{v_1}{(x+1)v_2} \\ \frac{v_1}{(x+1)v_2} \\ (x-1)v_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 & -2 \\ 1 & -2 & 3 \\ 1 & -2 & 1 \end{pmatrix}$$

has  $\det X_1 = 2$  and satisfies  $X_1A = \text{diag}(1, 1, -1)X_1$ , i.e.  $X_1AX_1^{-1}$  is diagonal.

### Solution 3

$$\begin{aligned} xI - A &= \begin{pmatrix} x-3 & -2 & 2 \\ 2 & x & -3 \\ -1 & -1 & x \end{pmatrix} \begin{matrix} r_1 \leftrightarrow r_3 \\ -r_1 \end{matrix} \begin{pmatrix} 1 & 1 & -x \\ 2 & x & -3 \\ x-3 & -2 & 2 \end{pmatrix} \begin{matrix} c_2 - c_1 \\ c_3 + xc_1 \end{matrix} \begin{pmatrix} 1 & 0 & 0 \\ 2 & x-2 & 2x-3 \\ x-3 & 1-x & x^2-3x+2 \end{pmatrix} \\ &\equiv \begin{pmatrix} 1 & 0 & 0 \\ 0 & x-2 & 2x-3 \\ 0 & 1-x & x^2-3x+2 \end{pmatrix} \begin{matrix} r_2 - 2r_1 \\ r_3 - (x-3)r_1 \end{matrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -x^2+x+1 \\ 0 & 1-x & (x-1)(x-2) \end{pmatrix} \begin{matrix} r_2 + r_3 \\ -r_3 \end{matrix} \\ &\equiv \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -x^2+x+1 \\ 0 & 0 & (x-1)^3 \end{pmatrix} \begin{matrix} c_3 + (x^2-x-1)c_2 \end{matrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & (x-1)^3 \end{pmatrix} = D(x). \end{aligned}$$

The rcf of  $A$  is

$$C((x-1)^3) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -3 & 3 \end{pmatrix}$$

as  $(x-1)^3 = x^3 - 3x^2 + 3x - 1$ .

Applying the *eros* in the above sequence to  $I$ :

$$\begin{aligned} I &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{matrix} r_1 \leftrightarrow r_3 \\ -r_1 \end{matrix} \begin{pmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{matrix} r_2 - 2r_1 \\ r_3 - (x-3)r_1 \end{matrix} \begin{pmatrix} 0 & 0 & -1 \\ 0 & 1 & 2 \\ 1 & 0 & x-3 \end{pmatrix} \begin{matrix} r_2 + r_3 \\ -r_2 \end{matrix} \begin{pmatrix} 0 & 0 & -1 \\ -1 & -1 & 1-x \\ 1 & 0 & x-3 \end{pmatrix} \\ &\equiv \begin{pmatrix} 0 & 0 & -1 \\ -1 & -1 & 1-x \\ x-2 & x-1 & x^2-3x+4 \end{pmatrix} \begin{matrix} r_3 + (x-1)r_2 \\ -r_3 \end{matrix} = P(x). \end{aligned}$$

Applying the *conjugates* of the *ecos* in the above sequence to  $I$ :

$$I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{matrix} r_1 + r_2 \\ r_1 - xr_3 \end{matrix} \begin{pmatrix} 1 & 1 & -x \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{matrix} r_2 - (x^2-x-1)r_3 \end{matrix} \begin{pmatrix} 1 & 1 & -x \\ 0 & 1 & -x^2+x+1 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \rho_1(x) \\ \rho_2(x) \\ \rho_3(x) \end{pmatrix} = Q(x).$$

Then  $P(x)$  and  $Q(x)$  are invertible over  $\mathbb{Q}[x]$  and satisfy

$$P(x)(xI - A) = D(x)Q(x).$$

As  $(\rho_1(x))\theta_A = (1, 1, -x)\theta_A = ((1, 1, 0) - x(0, 0, 1))\theta_A = (1, 1, 0) - (0, 0, 1)A = (0, 0, 0)$  we see that

$(\rho_1(x))\theta_A$  has order 1, the  $(1, 1)$ -entry in  $D(x)$ . In the same way

$$\begin{aligned} (\rho_2)\theta_A &= (0, 1, -x^2 + x + 1) = ((0, 1, 1) - x^2(0, 0, 1) + x(0, 0, 1))\theta_A = (0, 1, 1) - (0, 0, 1)A^2 + (0, 0, 1)A = \\ &= (0, 1, 1) - (1, 1, 0)A + (1, 1, 0) = (1, 2, 1) - (1, 2, 1) = (0, 0, 0) \end{aligned}$$

and so  $(\rho_2(x))\theta_A$  being zero has order 1, the  $(2, 2)$ -entry in  $D(x)$ . Let  $v_1 = (\rho_3(x))\theta_A = (0, 0, 1)$ .

Then  $xv_1 = (0, 0, 1)A = (1, 1, 0)$  and  $x^2v_1 = (1, 1, 0)A = (1, 2, 1)$ . Now  $\begin{vmatrix} v_1 \\ xv_1 \\ x^2v_1 \end{vmatrix} = \begin{vmatrix} 0 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 2 & 1 \end{vmatrix} = 1 \neq 0$  showing

that  $v_1, xv_1, x^2v_1$  are linearly independent and so the order of  $v_1$  in  $M(A)$  is of degree at least 3. But

$$x^3v_1 = x(x^2v_1) = (1, 2, 1)A = (0, 3, 4) = (0, 0, 1) - 3(1, 1, 0) + 3(1, 2, 1) = v_1 - 3xv_1 + 3x^2v_1$$

showing  $(x^3 - 3x^2 + 3x - 1)v_1 = 0$ , i.e.  $(x-1)^3v_1 = 0$ . So the order of  $v_1$  in  $M(A)$  is  $(x-1)^3$ , as expected, which is the  $(3, 3)$ -entry in  $M(A)$ . The matrix

$$X = \begin{pmatrix} v_1 \\ xv_1 \\ x^2v_1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 2 & 1 \end{pmatrix}$$

is invertible over  $\mathbb{Q}$  and satisfies  $XA = CX$ . So  $XAX^{-1} = C$ . The module  $M(A)$  is cyclic with generator  $v_1$  of order  $(x-1)^3$ . There are exactly four submodules of  $M(A)$  namely

$$M(A) = \langle v_1 \rangle, \langle (x-1)v_1 \rangle, \langle (x-1)^2v_1 \rangle, \langle (x-1)^3v_1 \rangle = \{(0, 0, 0)\}$$

corresponding to the four monic divisors of  $(x-1)^3$ . As  $(x-1)v_1 = (1, 1, -1)$  has order  $(x-1)^2$  in

$M(A)$ , we see that  $(x-1)v_1$  is not an eigenvector of  $A$ . But  $(x-1)^2v_1 = (-1, 0, 2)$  is an eigenvector of  $A$  as its order in  $M(A)$  is  $x-1$  which has degree 1. In fact all eigenvectors of  $A$  are proportional to  $(-1, 0, 2)$ .



## Solutions 4.1 (page 180)

### Solution 1

(a) (i) First  $r_1(x) = xr_2(x) + x^2 - x + 1$  giving  $q_2(x) = x$ ,  $r_3(x) = x^2 - x + 1$ . Next  $r_2(x) = (x+1)r_3(x)$  giving  $q_3(x) = x+1$ ,  $r_4(x) = 0(x)$ . So  $k=3$ ,  $c=1$  and  $d(x) = r_3(x)$ . Also

$$T = T_2 T_3 = \begin{pmatrix} x & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x+1 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} x^2 + x + 1 & x \\ x+1 & 1 \end{pmatrix}.$$

Col 1 of  $T$  gives  $r_1(x) = (x^2 + x + 1)d(x)$ ,  $r_2(x) = (x+1)d(x)$ , i.e.  $r_1'(x) = x^2 + x + 1$ ,  $r_2'(x) = x + 1$ .

Col 2 of  $T$  gives  $a_1(x) = 1$ ,  $a_2(x) = -x$ .

(ii) First  $r_1(x) = (x+1)r_2(x) - 2x^2 + 2$  giving  $q_2(x) = x+1$ ,  $r_3(x) = -2x^2 + 2$ . Next  $r_2(x) = (-x/2)r_3(x) + x - 1$  giving  $q_3(x) = -x/2$ ,  $r_4(x) = x - 1$ . Finally  $r_3(x) = -2(x+1)r_4(x)$  giving  $q_4(x) = -2(x+1)$ ,  $r_5(x) = 0(x)$ . So  $k=4$ ,  $c=1$  and  $d(x) = r_4(x)$ . In this case

$$T = T_2 T_3 T_4 = \begin{pmatrix} x+1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} -x/2 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} -2(x+1) & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} x^3 + 2x^2 - 1 & -x^2/2 - x/2 + 1 \\ x^2 + x + 1 & -x/2 \end{pmatrix}.$$

From col 1 of  $T$  we obtain  $r_1'(x) = x^3 + 2x^2 - 1$ ,  $r_2'(x) = x^2 + x + 1$ . From col 2 of  $T$  we obtain

$$a_1(x) = x/2, a_2(x) = -(x^2/2) - (x/2) + 1.$$

(iii) First  $r_1(x) = (x+1)r_2(x) + x + 2$  showing  $q_2(x) = x+1$ ,  $r_3(x) = x+2$ . Next  $r_2(x) = (x-2)r_3(x) + 3$  showing  $q_3(x) = x-2$ ,  $r_4(x) = 3$ . Lastly  $r_3(x) = ((x+2)/3)r_4(x)$  showing  $q_4(x) = (x+2)/3$ ,  $r_5(x) = 0(x)$ . So  $k=4$ ,  $c=3$  and  $d(x) = (1/3)r_4(x) = 1$ . Here

$$T = T_2 T_3 T_4 = \begin{pmatrix} x+1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x-2 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} (x+2)/3 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} (x^3 + x^2 + 1)/3 & x^2 - x - 1 \\ (x^2 - 1)/3 & x - 2 \end{pmatrix}.$$

So  $r_i(x) = r_i'(x)$  for  $i=1, 2$  and from col 2 of  $T$  we see  $a_1(x) = -(x-2)/3$ ,  $a_2(x) = (x^2 - x - 1)/3$ .

(iv)  $46 = 1 \times 32 + 14$ ,  $32 = 2 \times 14 + 4$ ,  $14 = 3 \times 4 + 2$ ,  $4 = 2 \times 2$  showing  $\gcd\{46, 32\} = 2$ . This sequence of divisions corresponds (see (c) below) to:  $x^{46} - 1 = (x^{14})(x^{32} - 1) + x^{14} - 1$ ,  $x^{32} - 1 = (x^{18} + x^4)(x^{14} - 1) + x^4 - 1$ ,  $x^{14} - 1 = (x^{10} + x^6 + x^2)(x^4 - 1) + x^2 - 1$ ,  $x^4 - 1 = (x^2 + 1)(x^2 - 1)$ .

So  $\gcd\{x^{46} - 1, x^{32} - 1\} = x^2 - 1$ . Here  $r_1'(x) = (x^{46} - 1)/(x^2 - 1) = x^{44} + x^{42} + x^{40} + \dots + x^2 + 1$  and  $r_2'(x) = (x^{32} - 1)/(x^2 - 1) = x^{30} + x^{28} + x^{26} + \dots + x^2 + 1$ . From col 2 of  $T_2 T_3 T_4 T_5$  which equals col 1 of

$$T_2 T_3 T_4 = \begin{pmatrix} x^{14} & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x^{18} + x^4 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x^{10} + x^6 + x^2 & 1 \\ 1 & 0 \end{pmatrix}$$

we get  $a_1(x) = q_3(x)q_4(x) + 1 = (x^{18} + x^4)(x^{10} + x^6 + x^2) + 1 = x^{28} + x^{24} + x^{20} + x^{14} + x^{10} + x^6 + 1$  and

$$a_2(x) = -(q_2(x)q_3(x) + 1)q_4(x) - q_2(x) = -(x^{32} + x^{18} + 1)(x^{10} + x^6 + x^2) - x^{14} = -(x^{42} + x^{38} + x^{34} + x^{28} + x^{24} + x^{20} + x^{14} + x^{10} + x^6 + x^2).$$

(b) (i) Working over  $\mathbb{Z}_2$ ,  $r_1(x) = r_2(x) + x^4 + x^3 + x^2 + x$  showing  $q_2(x) = \bar{1}$ ,  $r_3(x) = x^4 + x^3 + x^2 + x$ .

Dividing  $r_2(x)$  by  $r_3(x)$  gives  $r_2(x) = (x + \bar{1})r_3(x) + x^2 + \bar{1}$  and so  $q_3(x) = x + \bar{1}$ ,  $r_4(x) = x^2 + \bar{1}$ .

Dividing  $r_3(x)$  by  $r_4(x)$  gives  $r_3(x) = (x^2 + x)r_4(x)$  and so  $q_4(x) = x^2 + x$ ,  $r_5(x) = 0(x)$ . Therefore

$d(x) = \gcd\{r_1(x), r_2(x)\} = r_4(x) = x^2 + \bar{1}$ . Dividing  $r_1(x)$  by  $r_4(x)$  gives  $r_1(x) = (x^3 + x^2 + \bar{1})r_4(x)$  and so  $r'_1(x) = x^3 + x^2 + \bar{1}$ . Similarly  $r_2(x) = (x^3 + x + \bar{1})r_4(x)$  giving  $r'_2(x) = x^3 + x + \bar{1}$ .

(ii) Working over  $\mathbb{Z}_3$ ,  $r_1(x) = r_2(x) + x^3 + x^2 + \bar{1}$  and so  $q_2(x) = \bar{1}$ ,  $r_3(x) = x^3 + x^2 + \bar{1}$ . Dividing  $r_2(x)$  by  $r_3(x)$  gives  $r_2(x) = (x - \bar{1})r_3(x) + x^2 - x - \bar{1}$  giving  $q_3(x) = x - \bar{1}$ ,  $r_4(x) = x^2 - x - \bar{1}$ . As  $r_3(x) = (x - \bar{1})r_4(x)$  we obtain  $q_4(x) = x - \bar{1}$ ,  $r_5(x) = 0(x)$ . So

$d(x) = \gcd\{r_1(x), r_2(x)\} = r_4(x) = x^2 - x - \bar{1}$ . As  $r_1(x) = (x^2 - x + \bar{1})r_4(x)$  and  $r_2(x) = (x^2 + x - \bar{1})r_4(x)$  we see  $r'_1(x) = x^2 - x + \bar{1}$  and  $r'_2(x) = x^2 + x - \bar{1}$ .

(iii) Working over  $\mathbb{Z}_5$ ,  $r_1(x) = \bar{4}r_2(x)$  showing  $q_2(x) = \bar{4}$ ,  $r_3(x) = 0(x)$ . In this case  $k = 2$ ,  $c = \bar{3}$  and  $d(x) = \bar{2}r_2(x) = x^3 + \bar{3}x^2 + \bar{4}x + \bar{2}$ . Also  $r'_1(x) = \bar{2}$ ,  $r'_2(x) = \bar{3}$ .

(c) For  $q = 1$  the equation  $x^m - e = x^{m-n}(x^n - e) + x^{m-n} - e$  gives  $q(x) = x^{m-n}$ ,  $r(x) = x^{m-n} - e$  as  $\deg r(x) = m - n = m - qn = r < n$ . For  $q \geq 1$  the equation

$x^m - e = (x^{m-n} + x^{m-2n} + \dots + x^{m-qn})(x^n - e) + x^{m-qn} - e$  gives  $q(x) = x^{m-n} + x^{m-2n} + \dots + x^{m-qn}$  and  $r(x) = x^{m-qn} - e$  as  $\deg r(x) = m - qn = r < n = \deg(x^n - e)$ . Now  $(x^m - e) \mid (x^n - e) \Leftrightarrow r(x) = 0(x)$

and  $x^{m-qn} - e = 0(x) \Leftrightarrow m - qn = 0 \Leftrightarrow n \mid m$ . Write  $r_1 = m$ ,  $r_2 = n$  and let  $r_3, r_4, \dots, r_k$  be the positive integer remainders produced by the Euclidean algorithm. So  $r_k = d = \gcd\{m, n\}$ . By the first part,

applying the Euclidean algorithm to  $x^{r_1} - e$  and  $x^{r_2} - e$  produces the monic polynomial remainders  $x^{r_3} - e, x^{r_4} - e, \dots, x^{r_k} - e$ . Hence  $\gcd\{x^m - e, x^n - e\} = \gcd\{x^{r_1} - e, x^{r_2} - e\} = x^{r_k} - e = x^d - e$ .

(d) Using Exercises 2.3, Question 3(f) we see that  $\langle f_1(x) \rangle \cap \langle f_2(x) \rangle$  is a non-zero ideal of  $F[x]$  as the non-zero polynomial  $f_1(x)f_2(x)$  belongs to it. By (4.4) there is a unique monic polynomial  $l(x)$  with  $\langle l(x) \rangle = \langle f_1(x) \rangle \cap \langle f_2(x) \rangle$ . As  $\langle l(x) \rangle = \langle f_1(x) \rangle \cap \langle f_2(x) \rangle \subseteq \langle f_1(x) \rangle$  we deduce  $f_1(x) \mid l(x)$  and in the same way  $f_2(x) \mid l(x)$  showing that (i) holds. Suppose  $l'(x) \in F[x]$  is a common multiple of  $f_1(x)$  and  $f_2(x)$ , i.e.  $f_1(x) \mid l'(x)$  and  $f_2(x) \mid l'(x)$ . Then  $\langle l'(x) \rangle \subseteq \langle f_1(x) \rangle$  and  $\langle l'(x) \rangle \subseteq \langle f_2(x) \rangle$ . Therefore  $\langle l'(x) \rangle \subseteq \langle f_1(x) \rangle \cap \langle f_2(x) \rangle = \langle l(x) \rangle$  which gives  $l(x) \mid l'(x)$  and so (ii) holds. Suppose that  $m(x)$  is also a monic polynomial over  $F$  satisfying the same conditions (i) and (ii) as  $l(x)$ . As  $m(x)$  is a common multiple of  $f_1(x)$  and  $f_2(x)$  we see  $l(x) \mid m(x)$ . Interchanging the roles of  $l(x)$  and  $m(x)$  gives

$m(x) \mid l(x)$ . As  $l(x)$  and  $m(x)$  are divisors of each other and both are monic we conclude  $l(x) = m(x)$ .

So the lcm  $l(x)$  of  $f_1(x)$  and  $f_2(x)$  is unique.

Suppose  $f_1(x)$  and  $f_2(x)$  are both monic and write  $d(x) = \gcd\{f_1(x), f_2(x)\}$ . Then

$l(x) = (f_1(x)f_2(x))/d(x) = (f_1(x)/d(x))f_2(x) = (f_2(x)/d(x))f_1(x)$  showing that  $l(x)$  satisfies (i) as  $f_1(x)/d(x)$  and  $f_2(x)/d(x)$  belong to  $F[x]$ , i.e.  $l(x)$  is a common multiple of  $f_1(x)$  and  $f_2(x)$ .

Suppose  $l'(x)$  is a common multiple of  $f_1(x)$  and  $f_2(x)$  and so there are  $q_1(x), q_2(x) \in F[x]$  with  $l'(x) = q_1(x)f_1(x) = q_2(x)f_2(x)$ . By (4.6) there are  $a_1(x), a_2(x) \in F[x]$  with

$a_1(x)f_1(x) + a_2(x)f_2(x) = d(x)$ . Multiplying this equation by  $l'(x)/d(x)$  and substituting for  $l'(x)$  on the l.h.s. gives  $(a_1(x)q_2(x) + a_2(x)q_1(x))l(x) = l'(x)$  showing that  $l(x) \mid l'(x)$ , i.e. (ii) holds and  $l(x)$ , being monic, is the lcm of  $f_1(x)$  and  $f_2(x)$ .

(e) Suppose  $f(x) = p(x)^n$  where  $p(x)$  is irreducible over  $F$ . Then  $t = \deg f(x) = n \deg p(x)$  and so  $n \leq t$  as  $\deg p(x) \geq 1$ . The monic divisors over  $F$  of  $f(x)$  are  $p(x)^i$  for  $0 \leq i \leq n$  and so there are  $n+1$  of them. The binomial expansion of  $(1+1)^n$  consists of  $n+1$  terms each of which is a positive integer. So  $2^n \geq n+1$ . Hence  $n+1 \leq 2^n \leq 2^t$ . In the general case  $f(x) = p_1(x)^{n_1} p_2(x)^{n_2} \cdots p_k(x)^{n_k}$  where  $p_1(x), p_2(x), \dots, p_k(x)$  are distinct monic polynomials over  $F$ . A typical monic divisor over  $F$  of  $f(x)$  is  $p_1(x)^{i_1} p_2(x)^{i_2} \cdots p_k(x)^{i_k}$  where  $0 \leq i_l \leq n_l$  for  $1 \leq l \leq k$ , and so the number of such divisors is  $(n_1+1)(n_2+1) \cdots (n_k+1)$ . By the above  $n_l+1 \leq 2^{n_l}$  and so  $n_l+1 \leq 2^{n_l \deg p_l(x)}$  for  $1 \leq l \leq k$ . The product of these  $k$  inequalities gives  $(n_1+1)(n_2+1) \cdots (n_k+1) \leq 2^t$  as  $t = \sum_{l=1}^k n_l \deg p_l(x)$ .

### Solution 2

(a) Consider  $f(x) = \sum_{i \geq 0} a_i x^i$  and  $g(x) = \sum_{i \geq 0} b_i x^i$  in  $F[x]$ . Then

$$(f(x) + g(x))\mathcal{E}_a = \left( \sum_{i \geq 0} (a_i + b_i) x^i \right) \mathcal{E}_a = \sum_{i \geq 0} (a_i + b_i) a^i.$$

As  $a_i, b_i, a^i$  belong to the commutative ring  $F$  we obtain

$$\sum_{i \geq 0} (a_i + b_i) a^i = \sum_{i \geq 0} a_i a^i + \sum_{i \geq 0} b_i a^i$$

from the commutative laws and the distributive law. As

$$\sum_{i \geq 0} a_i a^i + \sum_{i \geq 0} b_i a^i = f(a) + g(a) = (f(x))\mathcal{E}_a + (g(x))\mathcal{E}_a$$

we obtain

$$(f(x) + g(x))\mathcal{E}_a = (f(x))\mathcal{E}_a + (g(x))\mathcal{E}_a,$$

showing that  $\mathcal{E}_a$  is additive. Similarly

$$(f(x)g(x))\mathcal{E}_a = \left( \sum_{i \geq 0} (a_0 b_i + a_1 b_{i-1} + \dots + a_i b_0) x^i \right) \mathcal{E}_a = \sum_{i \geq 0} (a_0 b_i + a_1 b_{i-1} + \dots + a_i b_0) a^i$$

which is the result of collecting together terms involving  $a^j a^k$  where  $j+k=i$  in the product

$$\left( \sum_{j \geq 0} a_j a^j \right) \left( \sum_{k \geq 0} b_k a^k \right) = f(a)g(a) = (f(x))\mathcal{E}_a (g(x))\mathcal{E}_a. \text{ So } (f(x)g(x))\mathcal{E}_a = (f(x))\mathcal{E}_a (g(x))\mathcal{E}_a$$

showing that  $\mathcal{E}_a$  is multiplicative. For all  $c \in F$  we have  $(c)\mathcal{E}_a = c$  showing that  $\mathcal{E}_a : F[x] \rightarrow F$  is surjective and  $\mathcal{E}_a$  maps the 1-element of  $F[x]$  to the 1-element of  $F$ . Therefore  $\mathcal{E}_a$  is a surjective ring homomorphism.

(b) Suppose  $h(x) \in \langle f(x), g(x) \rangle$ . There are  $a(x), b(x) \in F[x]$  with  $h(x) = a(x)f(x) + b(x)g(x)$ .

Hence  $h(x) = (a(x)q(x) + b(x))g(x) + a(x)r(x) \in \langle g(x), r(x) \rangle$  showing  $\langle f(x), g(x) \rangle \subseteq \langle g(x), r(x) \rangle$ .

As  $r(x) = f(x) - q(x)g(x)$  we obtain  $\langle g(x), r(x) \rangle \subseteq \langle f(x), g(x) \rangle$  and so  $\langle f(x), g(x) \rangle = \langle g(x), r(x) \rangle$ .

On comparing the monic generators of these ideals of  $F[x]$  we see  $\gcd\{f(x), g(x)\} = \gcd\{g(x), r(x)\}$  by (4.4). Write  $d(x) = \gcd\{d_1(x), d_2(x), \dots, d_t(x)\}$  and  $d'(x) = \gcd\{d_1(x), \gcd\{d_2(x), \dots, d_t(x)\}\}$ .

Then  $d(x) \mid d_1(x)$  and  $d(x) \mid d_i(x)$  for  $2 \leq i \leq t$ . So  $d(x) \mid d_1(x)$  and  $d(x) \mid \gcd\{d_2(x), \dots, d_t(x)\}$  which combine to give  $d(x) \mid d'(x)$ . Conversely  $d'(x) \mid d_1(x)$  and  $d'(x) \mid \gcd\{d_2(x), \dots, d_t(x)\}$  which combine to give  $d'(x) \mid d_1(x)$  and  $d'(x) \mid d_i(x)$  for  $2 \leq i \leq t$ . So  $d'(x) \mid d(x)$ . Therefore  $d(x) = d'(x)$ .

(c) Consider  $(a_i)$  and  $(b_i)$  in  $P(R)$ . There are non-negative integers  $m$  and  $n$  with  $a_i = 0$  for all  $i > m$  and  $b_i = 0$  for all  $i > n$ ; the least such integers are denoted  $\deg(a_i)$  and  $\deg(b_i)$ , the degrees of the non-zero sequences  $(a_i)$  and  $(b_i)$  respectively. Then  $a_i + b_i = 0$  for  $i > \max\{m, n\}$  and

$$\sum_{j+k=i} a_j b_k = a_0 b_i + a_1 b_{i-1} + \dots + a_i b_0 = 0 \text{ for } i > m+n$$

as each term in the sum is zero. Therefore only a finite number of the entries in  $(a_i) + (b_i)$  and  $(a_i) \times (b_i)$  are non-zero, and so these sequences belong to  $P(R)$ , i.e.,  $P(R)$  is closed under sum and product of sequences.

As the elements of  $R$  form an additive abelian group and addition of sequences is carried out entrywise, the elements of  $P(R)$  also form an additive abelian group: e.g.

$$(a_i) + (b_i) = (a_i + b_i) = (b_i + a_i) = (b_i) + (a_i)$$

shows addition of sequences to be commutative. The zero sequence  $(0)$  having all entries zero is the 0-element of  $P(R)$ . Consider  $(a_i), (b_i), (c_i) \in P(R)$ . As the elements of  $R$  obey the ring laws we obtain

$$\sum_{l+k=t} \left( \sum_{i+j=l} a_i b_j \right) c_k = \sum_{i+j+k=t} (a_i b_j) c_k = \sum_{i+j+k=t} a_i (b_j c_k) = \sum_{i+s=t} a_i \left( \sum_{j+k=s} b_j c_k \right)$$

showing that  $((a_i)(b_i))(c_i) = (a_i)((b_i)(c_i))$  as these sequences have the same entry  $t$ , i.e. multiplication in  $P(R)$  is associative. Similarly

$$\sum_{j+k=i} (a_j + b_j) c_k = \sum_{j+k=i} a_j c_k + \sum_{j+k=i} b_j c_k$$

shows that entry  $i$  in the sequence  $((a_i) + (b_i))(c_i)$  is equal to entry  $i$  in the sequence  $(a_i)(c_i) + (b_i)(c_i)$  for all  $i \geq 0$ . So these sequences are equal, i.e.,  $((a_i) + (b_i))(c_i) = (a_i)(c_i) + (b_i)(c_i)$  showing that the right distributive law holds in  $P(R)$ . In the same way the left distributive law

$$(a_i)((b_i) + (c_i)) = (a_i)(b_i) + (a_i)(c_i)$$

holds in  $P(R)$ . The sequence  $e_0 = (1, 0, 0, \dots, 0, \dots)$  is the 1-element of  $P(R)$  as

$$e_0(a_i) = (a_i) = (a_i)e_0. \text{ So } P(R) \text{ is a ring.}$$

Let  $a_0, b_0 \in R$ . Then  $(a_0 + b_0)t' = (a_0 + b_0, 0, 0, \dots) = (a_0, 0, 0, \dots) + (b_0, 0, 0, \dots) = (a_0)t' + (b_0)t'$ ,  $(a_0 b_0)t' = (a_0 b_0, 0, 0, \dots) = (a_0, 0, 0, \dots)(b_0, 0, 0, \dots) = (a_0)t'(b_0)t'$  and  $(1)t' = e_0$  showing that  $t': R \rightarrow P(R)$  is a ring homomorphism. As  $(a_0)t' = (b_0)t'$ , i.e.  $(a_0, 0, 0, \dots) = (b_0, 0, 0, \dots)$  implies  $a_0 = b_0$  we see that  $t'$  is injective. Hence  $a_0 \rightarrow (a_0)t'$  is a ring isomorphism between  $R$  and  $\text{im } t' = R'$ , showing that  $R'$ , being the image of a ring isomorphism, is a subring of  $P(R)$  (see Exercises 2.3, Question 3(b)) and  $R'$  is isomorphic to  $R$ .

Let  $R$  be an integral domain, i.e.  $R$  is commutative, non-trivial and has no zero-divisors. Consider  $(a_i), (b_i) \in P(R)$ . Then  $\sum_{j+k=i} a_j b_k = \sum_{k+j=i} b_k a_j$  showing  $(a_i)(b_i) = (b_i)(a_i)$ , i.e. the ring  $P(R)$  is

commutative. As  $R'$  is non-trivial, being isomorphic to  $R$ , we see that  $P(R)$  is also non-trivial, as it contains the subring  $R'$ . Suppose  $(a_i) \neq (0), (b_i) \neq (0)$ . Then  $\sum_{j+k=i} a_j b_k = 0$  for  $i > m+n$  where

$$m = \deg(a_i), n = \deg(b_i) \text{ as each term in the sum is zero. On the other hand } \sum_{j+k=m+n} a_j b_k = a_m b_n \neq 0,$$

only one term in the sum being non-zero as  $R$  has no zero-divisors. So  $(a_i)(b_i) \neq 0$ , showing that  $P(R)$

has no zero-divisors and in fact  $\deg(a_i)(b_i) = m + n = \deg(a_i) + \deg(b_i)$ . So  $P(R)$  is an integral domain. Conversely  $P(R)$  an integral domain implies that its subring  $R'$  is an integral domain and hence  $R$  is an integral domain as  $R \cong R'$ .

Using the multiplication rule in  $P(R)$  we obtain

$$(a_0)t'x = (a_0, 0, 0, \dots)(0, 1, 0, 0, \dots) = (0, a_0, 0, 0, \dots) = (0, 1, 0, 0, \dots)(a_0, 0, 0, \dots) = x(a_0)t'$$

for all  $a_0 \in R$ . By convention  $x^0 = e_0$  the 1-element of  $P(R)$ , and  $x^1 = x = e_1$  by definition. Suppose  $x^{i-1} = e_{i-1}$  for some  $i > 1$ . Then  $x^i = x x^{i-1} = e_i e_{i-1}$  and using the multiplication rule in  $P(R)$  we see  $e_i e_{i-1} = e_i$ , showing  $x^i = e_i$  and completing the induction. Hence

$$(0, 0, \dots, 0, a_i, 0, 0, \dots) = (a_i, 0, 0, \dots)e_i = (a_i)t'x^i \text{ and so } (a_0, a_1, \dots, a_i, \dots) = \sum_{i \geq 0} (a_i)t'x^i$$

which is a polynomial in the indeterminate  $x$  over  $R'$ .

### Solution 3

(a)(i) Suppose  $x^2 + \bar{1}$  is reducible over  $F$ . By (4.8)(ii) there is  $c \in F$  with  $c^2 = -\bar{1}$ . So  $c$  generates a cyclic subgroup of order 4, namely  $\{c, c^2, c^3, c^4\} = \{c, -\bar{1}, -c, \bar{1}\}$ , of the multiplicative group  $F^*$  of non-zero elements of  $F$ . As  $|F^*| = |F| - 1$ , by Lagrange's theorem  $4 \mid (|F| - 1)$ , i.e.  $|F| \equiv 1 \pmod{4}$ .

Conversely suppose  $|F| \equiv 1 \pmod{4}$ . By (3.17) the group  $F^*$  is cyclic with generator  $g$  of order

$|F| - 1$ . Write  $c = g^{(|F|-1)/4}$ . Then  $c^2$  is a zero of  $x^2 - \bar{1}$  over  $F$ . As  $c^2 \neq \bar{1}$  and

$x^2 - \bar{1} = (x - \bar{1})(x + \bar{1})$  we see  $c^2 = -\bar{1}$ . So  $c$  is a zero in  $F$  of  $x^2 + \bar{1} = (x - c)(x + c)$  which is therefore reducible over  $F$ .

(ii) Write  $f(x) = x^2 + x + \bar{1}$ . Then  $f(x)$  is irreducible over  $\mathbb{Z}_2$  by (4.8)(ii) as  $f(\bar{0}) \neq \bar{0}, f(\bar{1}) \neq \bar{0}$ .

Over  $\mathbb{Z}_3$  we see  $f(x) = (x - \bar{1})^2$  is reducible. Over  $\mathbb{Z}_5$  the quadratic  $f(x)$  has no zeros:

$f(\bar{0}) = \bar{1}, f(\bar{1}) = \bar{3}, f(\bar{2}) = \bar{2}, f(\bar{3}) = \bar{3}, f(\bar{4}) = \bar{1}$ , and so by (4.8)(ii) we see that  $f(x)$  is irreducible.

Over  $\mathbb{Z}_7$  we have  $f(x) = (x - \bar{2})(x - \bar{4})$  showing that  $f(x)$  is reducible.

Suppose  $|F| \equiv 0 \pmod{3}$ ; then  $|F| = 3^s$  by Exercises 2.3, Question 5(a) and the characteristic of  $F$  is 3. So  $f(x) = (x - \bar{1})^2$  is reducible over  $F$  as  $-\bar{1} - \bar{1} = \bar{1}$ .

Suppose  $|F| \equiv 1 \pmod{3}$ ; by (3.17) there is  $a \in F^*$  having multiplicative order  $|F^*| = |F| - 1$  and so

$c = a^{|F^*|/3}$  has order 3. As  $\bar{0} = c^3 - \bar{1} = (c - \bar{1})(c^2 + c + \bar{1})$  and  $c \neq \bar{1}$  we obtain  $c^2 + c + \bar{1} = \bar{0}$

showing that  $c$  is a zero of  $f(x)$  over  $F$ . So  $f(x) = (x - c)(x - c^2)$  is reducible over  $F$ .

Suppose  $|F| \equiv -1 \pmod{3}$  and suppose  $f(x)$  is reducible over  $F$ . By (4.8)(ii) there is  $c \in F$  satisfying  $c^2 + c + \bar{1} = \bar{0}$  and so  $c^3 = \bar{1}$  as above. As  $|F| \not\equiv 0 \pmod{3}$  we see  $c \neq \bar{1}$  and so  $c$  has order 3 in the group  $F^*$  of order  $|F| - 1$ . By Lagrange's theorem  $3 \mid (|F| - 1)$ , i.e.  $|F| \equiv 1 \pmod{3}$ .

This contradiction shows that  $f(x)$  is irreducible over  $F$ .

So  $x^2 + x + \bar{1}$  is irreducible over  $F$  if and only if  $|F| \equiv -1 \pmod{3}$ .

(b) In  $\mathbb{Z}_7$  we see  $(\bar{0})^3 = \bar{0}, (\bar{1})^3 = \bar{1}, (\bar{2})^3 = \bar{1}, (\bar{3})^3 = \bar{6}, (\bar{4})^3 = \bar{1}, (\bar{5})^3 = \bar{6}, (\bar{6})^3 = \bar{6}$ . By inspection there is no  $c \in \mathbb{Z}_7$  with  $c^3 = \bar{2}$ , i.e. the cubic  $x^3 - \bar{2}$  has no zeros in  $\mathbb{Z}_7$ . So  $x^3 - \bar{2}$  is irreducible over  $\mathbb{Z}_7$  by (4.8)(ii). For the same reason  $x^3 - \bar{3}, x^3 - \bar{4}, x^3 - \bar{5}$  are irreducible over  $\mathbb{Z}_7$ .

Suppose  $|F| \not\equiv 1 \pmod{3}$  and so  $|F^*| = |F| - 1 \not\equiv 0 \pmod{3}$ . Now

$\ker \theta = \{b \in F^* : (b)\theta = 1\} = \{b \in F^* : b^3 = 1\} = \{1\}$  as  $F^*$  contains no elements of order 3 by Lagrange's theorem. So  $\theta : F^* \rightarrow F^*$  is injective by Exercises 2.3, Question 1(a)(i). As  $F^*$  is a finite set,  $\theta$  is also surjective. Therefore there is  $b \in F$  with  $b^3 = a$  showing that  $x^3 - a$  has factor  $x - b$  and so is reducible over  $F$ .

Suppose  $|F| \equiv 1 \pmod{3}$  and so  $|F^*| = |F| - 1 \equiv 0 \pmod{3}$ . By (3.17) the group  $F^*$  is cyclic with

generator  $c$  say. Then  $F^*$  has a unique subgroup of order 3, namely  $\{1, b, b^2\}$  where  $b = c^{|F^*|/3}$ . So

$\ker \theta = \{1, b, b^2\}$  as all non-trivial elements of  $\ker \theta$  have order 3. Hence  $|\operatorname{im} \theta| = |F^*|/3$ . Now

$x^3 - a$  is reducible over  $F$  if and only if either  $a = 0$  or  $a \in \operatorname{im} \theta$ , and so there are

$1 + |F^*|/3 = (|F| + 2)/3$  such polynomials as  $|F^*| = |F| - 1$ .

(c) First note that  $b = 0$  is a zero of  $x^{q^n} - x$ . For  $b \neq 0$  we have that  $b$  belongs to the multiplicative group  $E^*$  of non-zero elements of the field  $E$ . As  $E$  is an  $n$ -dimensional vector space over  $F$  we see that  $|E| = |F|^n = q^n$  and so  $|E^*| = q^n - 1$ . By the  $|G|$ -lemma in multiplicative notation,  $b^{q^n-1} = 1$  and so  $b^{q^n} - b = 0$  showing that  $b$  is a zero of  $x^{q^n} - x$ . The  $q^n$  elements of  $E$  are therefore the  $q^n$  zeros of  $x^{q^n} - x$ . So  $x - b$  is a monic, irreducible over  $E$ , factor of  $x^{q^n} - x$  for each  $b \in E$ . Hence

$$x^{q^n} - x = \prod_{b \in E} (x - b)$$

is the factorisation of  $x^{q^n} - x$  into monic irreducible polynomials over  $E$ . So  $x^{q^n} - x$  splits over  $E$  into *distinct* monic factors.

Write  $d(x) = \gcd\{p(x), x^{q^n} - x\}$ . As  $p(x)$  and  $x^{q^n} - x$  are polynomials over  $F$ ,  $d(x)$  is also a polynomial over  $F$ . Either  $d(x) = p(x)$  or  $d(x) = 1$  as  $p(x)$  is monic and irreducible over  $F$  and  $d(x) \mid p(x)$ . Let  $c = \langle p(x) \rangle + x \in E$ . The discussion following (4.9) shows  $p(c) = 0$  and the paragraph above shows  $c^{q^n} - c = 0$ . Suppose  $d(x) = 1$ . By (4.6) there are  $a_1(x), a_2(x) \in F[x]$  with  $a_1(x)p(x) + a_2(x)(x^{q^n} - x) = 1$ . Applying the evaluation at  $c$  ring homomorphism  $\varepsilon_c$  gives

$$0 = a_1(c)0 + a_2(c)0 = a_1(c)p(c) + a_2(c)(c^{q^n} - c) = 1$$

which is a contradiction, showing  $d(x) = p(x)$  and so  $p(x) \mid (x^{q^n} - x)$ .

As  $x^{q^n} - x$  splits over  $E$  into a product of distinct factors so also does its divisor  $p(x)$ .

Substituting  $p'(x)$  for  $p(x)$  in the preceding part of the question we see that  $p'(x) \mid (x^{q^n} - x)$  and so there is  $c' \in E$  with  $p'(c') = 0$ . By (4.4) the evaluation homomorphism  $\varepsilon_{c'} : F[x] \rightarrow E$  has kernel  $\langle p'(x) \rangle$ . As  $1, c', (c')^2, \dots, (c')^{n-1}$  is a basis of the  $n$ -dimensional vector space  $E$  over  $F$  we see  $\varepsilon_{c'}$  is surjective. Hence  $\tilde{\varepsilon}_{c'} : F[x]/\langle p'(x) \rangle \cong E$  where  $(\langle p'(x) \rangle + f(x))\tilde{\varepsilon}_{c'} = f(c')$  for all  $f(x) \in F[x]$ .

Let  $p(x)$  be irreducible over  $F$  and have zero  $c$  in an extension field  $E$  of  $F$ . As above

$p(x) \mid (x^{q^n} - x)$  and so  $p(x)$  has no squared factor of degree  $> 0$ , i.e. it is impossible for an irreducible polynomial over a finite field to have a repeated zero in an extension field.

(d) As  $p(\bar{0}) = \bar{1}$  by (4.2)(i) we see that  $x$  is not a divisor of  $p(x)$ . As  $x^3 + \bar{1} = (x^2 + x + \bar{1})(x + \bar{1})$  and  $\gcd\{p(x), x^3 + \bar{1}\} = \bar{1}$  neither  $x + \bar{1}$  nor  $x^2 + x + \bar{1}$  is a factor of  $p(x)$ . As  $x^2 + x + \bar{1}$  is the only irreducible quadratic over  $\mathbb{Z}_2$  we see that  $p(x)$  has no factors of degree 1 or 2 over  $\mathbb{Z}_2$ . As

$\deg p(x) = 5 = 1 + 4 = 2 + 3$  are the only partitions of 5 into two parts, we conclude that  $p(x)$  is irreducible over  $\mathbb{Z}_2$  (any factorisation  $p(x) = p'(x)p''(x)$  with  $1 \leq \deg p'(x) \leq \deg p''(x)$  leads to the partition  $(\deg p'(x), \deg p''(x))$  of 5 which is impossible for the above reasons). As

$(x + \bar{1})(x^4 + x^3 + \bar{1}) = x^5 + x^3 + x + \bar{1}$  and  $(x^2 + x + \bar{1})(x^3 + x + \bar{1}) = x^5 + x^4 + \bar{1}$  these polynomials are reducible over  $\mathbb{Z}_2$ . Write  $p(x) = x^5 + x^2 + \bar{1}$ . Then  $p(\bar{0}) = \bar{1}$  and  $p(x) = (x^3 + \bar{1})x^2 + \bar{1}$  showing

$\gcd\{p(x), x^3 + \bar{1}\} = \bar{1}$ . So  $p(x)$  is irreducible over  $\mathbb{Z}_2$ . Also  $x^5 + x + \bar{1} = (x^2 + x + 1)(x^3 + x^2 + 1)$  is reducible over  $\mathbb{Z}_2$ . From (c) above  $p(x) \mid x^{32} - x$  and  $x^{32} - x = (x^{31} - \bar{1})x$ . As  $\gcd\{p(x), x\} = \bar{1}$  we deduce  $p(x) \mid x^{31} - \bar{1}$  and so  $\gcd\{p(x), x^{31} - \bar{1}\} = p(x)$ . Alternatively, dividing  $x^{31} - \bar{1}$  by  $p(x)$  gives  $x^{31} - \bar{1} = (x^{26} + x^{23} + x^{21} + x^{20} + x^{17} + x^{16} + x^{15} + x^{14} + x^{13} + x^9 + x^8 + x^6 + x^5 + x^4 + x^2 + \bar{1})p(x)$  and so we obtain  $\gcd\{p(x), x^{31} - \bar{1}\} = p(x)$  as before.

#### Solution 4

(a) Suppose the integer  $m$  is positive. Then  $me = e + e + \dots + e$  ( $m$  terms). As  $\theta$  is additive and  $(e)\theta = e$  we obtain  $(me)\theta = (e + e + \dots + e)\theta = (e)\theta + (e)\theta + \dots + (e)\theta = e + e + \dots + e = me$ . Hence  $((-m)e)\theta = (-me)\theta = -(me)\theta = -me = (-m)e$  as  $\theta$  respects negation. So  $(me)\theta = me$  for all integers  $m$  as  $(0e)\theta = (0)\theta = 0 = 0e$ . Let  $a_0 = me/ne = (me)(ne)^{-1}$ . As  $\theta$  is multiplicative and so respects inversion we see  $(a_0)\theta = ((me)(ne)^{-1})\theta = (me)\theta((ne)\theta)^{-1} = (me)(ne)^{-1} = a_0$ , i.e.  $\theta$  fixes all elements  $a_0 \in F_0$ .

Write  $L = \{a \in F : (a)\theta = a\}$ . Consider  $a, a' \in L$ . As  $\theta$  is additive we have

$(a + a')\theta = (a)\theta + (a')\theta = a + a'$  showing  $a + a' \in L$ . Also  $(-a)\theta = -(a)\theta = -a$  and  $(0)\theta = 0$

showing  $-a, 0 \in L$ . So  $L$  is closed under addition and is a subgroup of the additive group of  $F$ . As  $\theta$

respects multiplication we have  $(aa')\theta = (a)\theta(a')\theta = aa'$  showing  $aa' \in L$ . For  $a \neq 0$  we have

$(a^{-1})\theta = ((a)\theta)^{-1} = a^{-1}$  and  $(e)\theta = e$  showing  $a^{-1}, e \in L$ . So  $L$  is a subfield of  $F$  and  $F_0 \subseteq L \subseteq F$ .

Write  $f(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n$  where  $a_0, a_1, \dots, a_n \in F_0$ . Then

$$(f(c))\theta = (a_0 + a_1c + a_2c^2 + \dots + a_nc^n)\theta = (a_0)\theta + (a_1c)\theta + (a_2c^2)\theta + \dots + (a_nc^n)\theta =$$

$$(a_0)\theta + (a_1)\theta(c)\theta + (a_2)\theta((c)\theta)^2 + \dots + (a_n)\theta((c)\theta)^n = a_0 + a_1(c)\theta + a_2((c)\theta)^2 + \dots + a_n((c)\theta)^n = f((c)\theta)$$

as  $\theta$  respects addition and multiplication in  $F$ . Let  $c$  be a zero of  $f(x)$ ,  $c \in F$ . Then  $f(c) = 0$  and

so by the foregoing theory  $f((c)\theta) = (f(c))\theta = (0)\theta = 0$  showing that  $(c)\theta$  is also a zero of  $f(x)$ .

Conversely suppose that  $(c)\theta$  is a zero of  $f(x)$ . By Exercises 2.3, Question 3(d) the inverse  $\theta^{-1}$  of  $\theta$  is also an automorphism of  $F$ . Hence  $((c)\theta)\theta^{-1} = c$  is a zero of  $f(x)$ .

(b) Consider the elements  $x = a + b\sqrt{2}, y = c + d\sqrt{2}$  of  $\mathbb{Q}(\sqrt{2})$ . Then

$$(x+y)\theta = ((a+c) + (b+d)\sqrt{2})\theta = (a+c) - (b+d)\sqrt{2} = (a-b\sqrt{2}) + (c-d\sqrt{2}) = (a+b\sqrt{2})\theta + (c+d\sqrt{2})\theta = (x)\theta + (y)\theta$$

showing that  $\theta$  is additive. Similarly

$$(xy)\theta = ((a+b\sqrt{2})(c+d\sqrt{2}))\theta = ((ac+2bd) + (ad+bc)\sqrt{2})\theta = (ac+2bd) - (ad+bc)\sqrt{2} = (a-b\sqrt{2})(c-d\sqrt{2}) = (a+b\sqrt{2})\theta(c+d\sqrt{2})\theta = (x)\theta(y)\theta$$

showing that  $\theta$  is multiplicative. Also  $(1)\theta = (1+0\sqrt{2})\theta = 1-0\sqrt{2} = 1$ . As

$$(x)\theta^2 = ((a+b\sqrt{2})\theta)\theta = (a-b\sqrt{2})\theta = a-(-b)\sqrt{2} = a+b\sqrt{2} = x$$

for all  $x \in \mathbb{Q}(\sqrt{2})$  we see that  $\theta^{-1} = \theta$  and so  $\theta$  is bijective being self-inverse. We conclude that  $\theta$  is an automorphism of  $\mathbb{Q}(\sqrt{2})$ .

Let  $\mathbb{Q}(\sqrt{2})$  have automorphism  $\varphi$ . As  $\mathbb{Q}$  is the prime subfield of  $\mathbb{Q}(\sqrt{2})$ , by (a) above  $(a)\varphi = a$  for all  $a \in \mathbb{Q}$  and  $(\sqrt{2})\varphi = \pm\sqrt{2}$  as the zeros  $\sqrt{2}, -\sqrt{2}$  of  $x^2 - 2$  over  $\mathbb{Q}$  are fixed or interchanged by  $\varphi$ . Suppose  $(\sqrt{2})\varphi = \sqrt{2}$ . Then  $(a+b\sqrt{2})\varphi = (a)\varphi + (b)\varphi(\sqrt{2})\varphi = a+b\sqrt{2}$  showing that  $\varphi$  is the identity automorphism of  $\mathbb{Q}(\sqrt{2})$ . Suppose  $(\sqrt{2})\varphi = -\sqrt{2}$ . Then

$$(a+b\sqrt{2})\varphi = (a)\varphi + (b)\varphi(\sqrt{2})\varphi = a-b\sqrt{2}$$

showing that  $\varphi = \theta$ . We conclude that  $\theta$  is the only non-identity automorphism of  $\mathbb{Q}(\sqrt{2})$ .

(c) Consider  $e = a+bc, e' = a'+b'c \in F(c)$ . Then

$$\begin{aligned} (e+e')\theta &= ((a+a') + (b+b')c)\theta = a+a' - (b+b')a_1 - (b+b')c = \\ &= (a-ba_1-bc) + (a'-b'a_1-b'c) = (a+bc)\theta + (a'+b'c)\theta = (e)\theta + (e')\theta \end{aligned}$$

showing that  $\theta$  respects addition. Now  $p(c) = 0$  gives  $c^2 = -a_0 - a_1c$ . Therefore

$$\begin{aligned} (ee')\theta &= ((a+bc)(a'+b'c))\theta = (aa' - bb'a_0 + (ab' + ba' - bb'a_1)c)\theta = \\ &= aa' - bb'a_0 - (ab' + ba' - bb'a_1)a_1 - (ab' + ba' - bb'a_1)c. \end{aligned}$$

But

$$\begin{aligned} (e)\theta(e')\theta &= (a-ba_1-bc)(a'-b'a_1-b'c) = \\ &= (a-ba_1)(a'-b'a_1) - bb'a_0 - ((a-ba_1)b' + b(a'-b'a_1) + bb'a_1)c. \end{aligned}$$

Comparison of the above expressions gives  $(ee')\theta = (e)\theta(e')\theta$  showing that  $\theta$  respects multiplication. As  $(a)\theta = a$  for all  $a \in F$  we see  $(1)\theta = 1$  showing that  $\theta$  respects the 1-element of  $F$  which is also the 1-element of  $F(c)$ . Finally

$$(e)\theta^2 = ((e)\theta)\theta = (a+bc)\theta = (a-ba_1-bc)\theta = a-ba_1 - (-b)a_1 - (-b)c = a+bc = e \text{ for all } e \in F(c)$$

and so  $\theta$  is self-inverse. Hence  $\theta$  is bijective. So  $\theta$  is an automorphism of  $F(c)$ .

The fixed field  $L$  of  $\theta$  contains  $F$ . Suppose  $c \neq -a_1 - c$ , that is,  $a_1 + 2c \neq 0$ . Consider  $e \in L$ . Then  $(e)\theta = e$  gives  $a-ba_1-bc = a+bc$ , that is,  $b(2c+a_1) = 0$  giving  $b = 0$ . So  $e = a \in F$  and so  $L = F$ . Conversely suppose  $L = F$ . As  $p(x)$  is irreducible of degree 2 over  $F$  and  $p(c) = 0$  we see  $c \notin F$ . Therefore  $(c)\theta \neq c$ , that is,  $-a_1 - c \neq c$ .

Incidentally it is possible for  $L \neq F$ : take  $p(x)$  inseparable (see (6.19)) in which case  $-a_1 - c = c$  and  $L = F(c)$ .

## Solution 5



(a) Notice first  $F_0 = \{e, 2e, 3e, \dots, pe\}$  where  $e$  is the 1-element of  $F$  (Exercises 2.3, Question 5(a)).

So  $a_0 = ne$  where  $1 \leq n \leq p$ . As  $\theta$  is additive and  $(e)\theta = e$  we deduce

$(a_0)\theta = (ne)\theta = n(e)\theta = ne = a_0$ . From the solution of Question 4(a) above,  $L$  is a subfield of  $F$  with  $F_0 \subseteq L \subseteq F$ .

Let  $n_0$  denote the multiplicative order of  $\theta$ . Using Question 3(c) above with  $p$  in place of  $q$  we see

$a^{p^n} = a$  for all  $a \in F$ . So  $(a)\theta^n = a^{p^n} = a$  showing  $n_0 \leq n$ . Suppose  $n_0 < n$ . Then

$a^{p^{n_0}} = (a)\theta^{n_0} = a$  for all  $a \in F$  showing that the polynomial  $x^{p^{n_0}} = x$  of degree  $p^{n_0}$  over  $F$  has

$|F| = p^n > p^{n_0}$  zeros in  $F$  contrary to (4.2)(ii). Therefore  $n_0 = n$ .

(b) As  $p(\bar{0}) = \bar{1}, p(\bar{1}) = \bar{1}$  we see that  $p(x)$  has no zeros in  $\mathbb{Z}_2$ . Hence  $p(x)$  is irreducible over  $\mathbb{Z}_2$  by (4.8)(ii) as  $\deg p(x) = 3$ . The addition and multiplication tables of the field

$\mathbb{Z}_2(c) = \{\bar{0}, \bar{1}, c, \bar{1}+c, c^2, \bar{1}+c^2, c+c^2, \bar{1}+c+c^2\}$  are:

+	$\bar{0}$	$\bar{1}$	$c$	$\bar{1}+c$	$c^2$	$\bar{1}+c^2$	$c+c^2$	$\bar{1}+c+c^2$
$\bar{0}$	$\bar{0}$	$\bar{1}$	$c$	$\bar{1}+c$	$c^2$	$\bar{1}+c^2$	$c+c^2$	$\bar{1}+c+c^2$
$\bar{1}$	$\bar{1}$	$\bar{0}$	$\bar{1}+c$	$c$	$\bar{1}+c^2$	$c^2$	$\bar{1}+c+c^2$	$c+c^2$
$c$	$c$	$\bar{1}+c$	$\bar{0}$	$\bar{1}$	$c+c^2$	$\bar{1}+c+c^2$	$c^2$	$\bar{1}+c^2$
$\bar{1}+c$	$\bar{1}+c$	$c$	$\bar{1}$	$\bar{0}$	$\bar{1}+c+c^2$	$c+c^2$	$\bar{1}+c^2$	$c^2$
$c^2$	$c^2$	$\bar{1}+c^2$	$c+c^2$	$\bar{1}+c+c^2$	$\bar{0}$	$\bar{1}$	$c$	$\bar{1}+c$
$\bar{1}+c^2$	$\bar{1}+c^2$	$c^2$	$\bar{1}+c+c^2$	$c+c^2$	$\bar{1}$	$\bar{0}$	$\bar{1}+c$	$c$
$c+c^2$	$c+c^2$	$\bar{1}+c+c^2$	$c^2$	$\bar{1}+c^2$	$c$	$\bar{1}+c$	$\bar{0}$	$\bar{1}$
$\bar{1}+c+c^2$	$\bar{1}+c+c^2$	$c+c^2$	$\bar{1}+c^2$	$c^2$	$\bar{1}+c$	$c$	$\bar{1}$	$\bar{0}$

and (using  $p(c) = \bar{0}$ , i.e.  $c^3 = \bar{1}+c$ )

$\times$	$\bar{0}$	$\bar{1}$	$c$	$\bar{1}+c$	$c^2$	$\bar{1}+c^2$	$c+c^2$	$\bar{1}+c+c^2$
$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$
$\bar{1}$	$\bar{0}$	$\bar{1}$	$c$	$\bar{1}+c$	$c^2$	$\bar{1}+c^2$	$c+c^2$	$\bar{1}+c+c^2$
$c$	$\bar{0}$	$c$	$c^2$	$c+c^2$	$\bar{1}+c$	$\bar{1}$	$\bar{1}+c+c^2$	$\bar{1}+c^2$
$\bar{1}+c$	$\bar{0}$	$\bar{1}+c$	$c+c^2$	$\bar{1}+c^2$	$\bar{1}+c+c^2$	$c^2$	$\bar{1}$	$c$
$c^2$	$\bar{0}$	$c^2$	$\bar{1}+c$	$\bar{1}+c+c^2$	$c+c^2$	$c$	$\bar{1}+c^2$	$\bar{1}$
$\bar{1}+c^2$	$\bar{0}$	$\bar{1}+c^2$	$\bar{1}$	$c^2$	$c$	$\bar{1}+c+c^2$	$\bar{1}+c$	$c+c^2$
$c+c^2$	$\bar{0}$	$c+c^2$	$\bar{1}+c+c^2$	$\bar{1}$	$\bar{1}+c^2$	$\bar{1}+c$	$c$	$c^2$
$\bar{1}+c+c^2$	$\bar{0}$	$\bar{1}+c+c^2$	$\bar{1}+c^2$	$c$	$\bar{1}$	$c+c^2$	$c^2$	$\bar{1}+c$

The powers of  $c$  are  $c, c^2, c^3 = \bar{1}+c, c^4 = c+c^2, c^5 = \bar{1}+c+c^2, c^6 = \bar{1}+c^2, c^7 = \bar{1}$  accounting for all the seven elements of  $\mathbb{Z}_2(c)^*$ . By Question 4(a) above

$p(x) = (x-c)(x-(c)\theta)(x-(c)\theta^2) = (x-c)(x-c^2)(x-c^4) = (x-c)(x-c^2)(x-c-c^2)$  which is the factorisation of  $p(x)$  over  $\mathbb{Z}_2(c)$ . Since  $p'(\bar{0}) = \bar{1}, p'(\bar{1}) = \bar{1}$ , by (4.8)(ii)  $p'(x)$  is irreducible over  $\mathbb{Z}_2$ . As

$$(\bar{1}+c)^3 + (\bar{1}+c)^2 + \bar{1} = c^9 + c^6 + \bar{1} = c^2 + (\bar{1}+c^2) + \bar{1} = \bar{0}$$

we see  $p'(\bar{1}+c) = \bar{0}$  showing that  $\bar{1}+c$  is a zero of  $p'(x)$ . Hence

$p'(x) = (x-c^3)(x-(c^3)\theta)(x-(c^3)\theta^2) = (x-c^3)(x-c^6)(x-c^{12}) = (x-\bar{1}-c)(x-\bar{1}-c^2)(x-\bar{1}-c-c^2)$  showing how  $p'(x)$  splits over  $\mathbb{Z}_2(c)$ . As in Question 3(c) above

$$x^8 - x = x(x-\bar{1})\{(x-c)(x-c^2)(x-c^4)\}\{(x-c^3)(x-c^6)(x-c^5)\}$$

is the splitting factorisation of  $x^8 - x$  over  $\mathbb{Z}_2(c)$ . Hence  $x^8 - x = x(x-\bar{1})p(x)p'(x)$  is the factorisation of  $x^8 - x$  into irreducible polynomials over  $\mathbb{Z}_2$ . By Question 3(c) above every irreducible polynomial of degree 3 over  $\mathbb{Z}_2$  is a divisor of  $x^8 - x$ . So  $p(x)$  and  $p'(x)$  are the only such polynomials.

(c) Dividing  $p(x)$  by  $x^3 + \bar{1}$  gives  $p(x) = x^5 + x^3 + \bar{1} = (x^2 + \bar{1})(x^3 + \bar{1}) + x^2$ . Dividing  $x^3 + \bar{1}$  by  $x^2$  gives  $x^3 + \bar{1} = x(x^2) + \bar{1}$  and so  $\gcd\{p(x), x^3 + \bar{1}\} = \bar{1}$ . As  $p(\bar{0}) = \bar{1}$  we conclude that  $p(x)$  is irreducible over  $\mathbb{Z}_2$  by Question 3(d) above. We know  $p(c) = \bar{0}$ , i.e.  $c^5 = c^3 + \bar{1}$ . By Question 4(b) above  $(c)\theta^i$  is a zero of  $p(x)$  for  $1 \leq i \leq 5$ . So  $c, c^2, c^4$  are zeros of  $p(x)$  as is

$c^8 = c^3c^5 = c^3(c^3 + \bar{1}) = cc^5 + c^3 = c(c^3 + \bar{1}) + c^3 = c^4 + c^3 + c$ . Squaring this equation, i.e. applying  $\theta$ , gives  $c^{16} = c^8 + c^6 + c^2 = c^4 + c^3 + c + c(c^3 + \bar{1}) + c^2 = c^3 + c^2$  which is also a zero of  $p(x)$ . The five zeros of  $p(x)$  in  $\mathbb{Z}_2(c)$  are therefore  $c, c^2, c^4, c^4 + c^3 + c, c^3 + c^2$ .

As  $f(c) = f((c+a)-a)$  we see that  $c$  is a zero of  $f(x) \Leftrightarrow c+a$  is a zero of  $f(x-a)$ . Also  $f(x) = f_1(x)f_2(x) \Leftrightarrow f(x-a) = f_1(x-a)f_2(x-a)$ . As  $\deg f(x) = \deg f(x-a)$  for all  $f(x) \in F[x]$ ,  $a \in F$  it follows that  $f(x)$  is irreducible over  $F$  if and only if  $f(x-a)$  is irreducible over  $F$ . Hence

$$p(x-\bar{1}) = (x-\bar{1})^5 + (x-\bar{1})^3 + \bar{1} = (x-\bar{1})(x^4 - \bar{1}) + (x-\bar{1})(x^2 - \bar{1}) + \bar{1} = x^5 + x^4 + x^3 + x^2 + \bar{1}$$

has zero  $c + \bar{1}$ . So

$p'(x) = x^5 + x^4 + x^3 + x^2 + \bar{1} = (x+c+\bar{1})(x+c^2+\bar{1})(x+c^4+\bar{1})(x+c^4+c^3+c+\bar{1})(x+c^3+c^2+\bar{1})$  is the factorisation of  $p'(x)$  into irreducible polynomials over  $\mathbb{Z}_2(c)$ .

### Solution 6

(a) As none of  $(\bar{0})^2 + \bar{1} = \bar{1}, (\bar{1})^2 + \bar{1} = -\bar{1}, (-\bar{1})^2 + \bar{1} = -\bar{1}$  is zero, we see that  $x^2 + \bar{1}$  has no zeros in  $\mathbb{Z}_3 = \{\bar{0}, \bar{1}, \bar{2}\} = \{\bar{0}, \bar{1}, -\bar{1}\}$ . As  $x^2 + \bar{1}$  has degree 2, by (4.8)(ii)  $x^2 + \bar{1}$  is irreducible over  $\mathbb{Z}_3$ .

Similarly  $(\bar{0})^2 + \bar{0} - \bar{1} = -\bar{1}, (\bar{1})^2 + \bar{1} - \bar{1} = \bar{1}, (-\bar{1})^2 + (-\bar{1}) - \bar{1} = -\bar{1}$  showing that  $x^2 + x - \bar{1}$  has no zeros in  $\mathbb{Z}_3$  and so is irreducible over  $\mathbb{Z}_3$  by (4.8)(ii). Replacing  $x$  by  $-x$  we deduce that

$(-x)^2 + (-x) - \bar{1} = x^2 - x - \bar{1}$  is also irreducible over  $\mathbb{Z}_3$ . Multiplying out gives

$$(x^2 + x - \bar{1})(x^2 - x - \bar{1}) = x^4 + \bar{1} \text{ as } -x^2 - x^2 - x^2 = \bar{0}. \text{ Hence}$$

$$(x^2 + \bar{1})(x^2 + x - \bar{1})(x^2 - x - \bar{1}) = (x^2 + \bar{1})(x^4 + \bar{1}) = x^6 + x^4 + x^2 + \bar{1}. \text{ As}$$

$$x^8 - 1 = (x^2 - 1)(x^6 + x^4 + x^2 + 1) \text{ over all rings and fields, we see}$$

$x^9 - x = x(x^8 - \bar{1}) = x(x - \bar{1})(x + \bar{1})(x^2 + \bar{1})(x^2 + x - \bar{1})(x^2 - x - \bar{1})$  which is the factorisation of  $x^9 - x$  into monic irreducible polynomials over  $\mathbb{Z}_3$ . There are no further monic irreducible quadratic

polynomials  $p(x)$  over  $\mathbb{Z}_3$ , as  $p(x) \mid (x^9 - x)$  by Question 3(c) above, and so  $p(x)$  is present in the above factorisation of  $x^9 - x$ .

Just as with complex numbers  $x^2 + \bar{1} = (x - i)(x + i)$  and multiplying pairs of ‘conjugate’ factors together gives  $(x + i - \bar{1})(x - i - \bar{1}) = x^2 + x - \bar{1}$ ,  $(x + i + \bar{1})(x - i + \bar{1}) = x^2 - x - \bar{1}$  which are the factorisations of these polynomials over  $\mathbb{Z}_3$  into monic irreducible polynomials over  $\mathbb{Z}_3(i)$ .

$(a + bi)^3 = a^3 + 3a^2bi + 3a(bi)^2 + (bi)^3 = a^3 + b^3i^3 = a^3 - b^3i$  as  $3a^2bi = 3a(bi)^2 = \bar{0}$  and  $i^3 = i^2 \times i = (-\bar{1})i = -i$ . But  $a^3 = a, b^3 = b$  for  $a, b \in \mathbb{Z}_3$ . So  $(a + bi)\theta = (a + bi)^3 = a - bi$  showing that the Frobenius automorphism  $\theta$  of  $\mathbb{Z}_3(i)$  coincides with the ‘conjugation’ mapping  $a + bi \rightarrow a - bi$ . The element  $g = \bar{1} + i$  generates  $\mathbb{Z}_3(i)^*$  as  $g^2 = (\bar{1} + i)(\bar{1} + i) = -i$  and so  $g^4 = -\bar{1}$  showing that  $g$  has order  $8 = |\mathbb{Z}_3(i)^*|$ . In fact  $\mathbb{Z}_3(i)^*$  is generated by any one of  $\bar{1} + i, \bar{1} - i, -\bar{1} + i, -\bar{1} - i$ .

The addition and multiplication tables of  $\mathbb{Z}_3(i)$  are

+	$\bar{0}$	$\bar{1}$	$-\bar{1}$	$i$	$\bar{1} + i$	$-\bar{1} + i$	$-i$	$\bar{1} - i$	$-\bar{1} - i$
$\bar{0}$	$\bar{0}$	$\bar{1}$	$-\bar{1}$	$i$	$\bar{1} + i$	$-\bar{1} + i$	$-i$	$\bar{1} - i$	$-\bar{1} - i$
$\bar{1}$	$\bar{1}$	$-\bar{1}$	$\bar{0}$	$\bar{1} + i$	$-\bar{1} + i$	$i$	$\bar{1} - i$	$-\bar{1} - i$	$-i$
$-\bar{1}$	$-\bar{1}$	$\bar{0}$	$\bar{1}$	$-\bar{1} + i$	$i$	$\bar{1} + i$	$-\bar{1} - i$	$-i$	$\bar{1} - i$
$i$	$i$	$\bar{1} + i$	$-\bar{1} + i$	$-i$	$\bar{1} - i$	$-\bar{1} - i$	$\bar{0}$	$\bar{1}$	$-\bar{1}$
$\bar{1} + i$	$\bar{1} + i$	$-\bar{1} + i$	$i$	$\bar{1} - i$	$-\bar{1} - i$	$-i$	$\bar{1}$	$-\bar{1}$	$\bar{0}$
$-\bar{1} + i$	$-\bar{1} + i$	$i$	$\bar{1} + i$	$-\bar{1} - i$	$-i$	$\bar{1} - i$	$-\bar{1}$	$\bar{0}$	$\bar{1}$
$-i$	$-i$	$\bar{1} - i$	$-\bar{1} - i$	$\bar{0}$	$\bar{1}$	$-\bar{1}$	$i$	$\bar{1} + i$	$-\bar{1} + i$
$\bar{1} - i$	$\bar{1} - i$	$-\bar{1} - i$	$-i$	$\bar{1}$	$-\bar{1}$	$\bar{0}$	$\bar{1} + i$	$-\bar{1} + i$	$i$
$-\bar{1} - i$	$-\bar{1} - i$	$-i$	$\bar{1} - i$	$-\bar{1}$	$\bar{0}$	$\bar{1}$	$-\bar{1} + i$	$i$	$\bar{1} + i$

$\times$	$\bar{0}$	$\bar{1}$	$-\bar{1}$	$i$	$\bar{1} + i$	$-\bar{1} + i$	$-i$	$\bar{1} - i$	$-\bar{1} - i$
$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$
$\bar{1}$	$\bar{0}$	$\bar{1}$	$-\bar{1}$	$i$	$\bar{1} + i$	$-\bar{1} + i$	$-i$	$\bar{1} - i$	$-\bar{1} - i$
$-\bar{1}$	$\bar{0}$	$-\bar{1}$	$\bar{1}$	$-i$	$-\bar{1} - i$	$\bar{1} - i$	$i$	$-\bar{1} + i$	$\bar{1} + i$
$i$	$\bar{0}$	$i$	$-i$	$-\bar{1}$	$-\bar{1} + i$	$-\bar{1} - i$	$\bar{1}$	$\bar{1} + i$	$\bar{1} - i$
$\bar{1} + i$	$\bar{0}$	$\bar{1} + i$	$-\bar{1} - i$	$-\bar{1} + i$	$-i$	$\bar{1}$	$\bar{1} - i$	$-\bar{1}$	$i$
$-\bar{1} + i$	$\bar{0}$	$-\bar{1} + i$	$\bar{1} - i$	$-\bar{1} - i$	$\bar{1}$	$i$	$\bar{1} + i$	$-i$	$-\bar{1}$
$-i$	$\bar{0}$	$-i$	$i$	$\bar{1}$	$\bar{1} - i$	$\bar{1} + i$	$-\bar{1}$	$-\bar{1} - i$	$-\bar{1} + i$
$\bar{1} - i$	$\bar{0}$	$\bar{1} - i$	$-\bar{1} + i$	$\bar{1} + i$	$-\bar{1}$	$-i$	$-\bar{1} - i$	$i$	$\bar{1}$
$-\bar{1} - i$	$\bar{0}$	$-\bar{1} - i$	$\bar{1} + i$	$\bar{1} - i$	$i$	$-\bar{1}$	$-\bar{1} + i$	$\bar{1}$	$-i$

(b) As  $a^3 = a$  for all  $a \in \mathbb{Z}_3$  we see  $p(a) = a^3 - a - \bar{1} = -\bar{1}$  showing that  $p(x)$  has no zeros in  $\mathbb{Z}_3$ . As  $\deg p(x) = 3$  we deduce from (4.8)(ii) that  $p(x)$  is irreducible over  $\mathbb{Z}_3$ . Cubing  $c^3 = c + \bar{1}$  gives  $c^9 = c^3 + \bar{1} = c + \bar{1} + \bar{1} = c - \bar{1}$ . Hence  $c^{12} = c^3 \times c^9 = (c + \bar{1})(c - \bar{1}) = c - \bar{1}$  and so  $c^{13} = c(c^2 - \bar{1}) = c^3 - c = \bar{1}$ . As  $c \notin \mathbb{Z}_3$  we see that  $c$  has order 13. So  $c$  does not generate  $\mathbb{Z}_3(c)^*$  which is a group of order 26. As  $p(x)$  is monic and irreducible over  $\mathbb{Z}_3$  either  $\gcd\{p(x), x^{13} - \bar{1}\} = \bar{1}$  or  $\gcd\{p(x), x^{13} - \bar{1}\} = p(x)$ . Suppose the former is true. By (4.6) there are polynomials  $a_1(x), a_2(x)$  over  $\mathbb{Z}_3$  with  $a_1(x)p(x) + a_2(x)(x^{13} - \bar{1}) = \bar{1}$ . Using the evaluation homomorphism  $\mathcal{E}_c$  gives  $\bar{0} = a_1(c) \times \bar{0} + a_2(c) \times \bar{0} = a_1(c)p(c) + a_2(c)(c^{13} - \bar{1}) = \bar{1}$ , i.e.  $\bar{0} = \bar{1}$ . This contradiction shows that  $\gcd\{p(x), x^{13} - \bar{1}\} = p(x)$ , i.e.  $p(x)$  is a divisor of  $x^{13} - \bar{1}$ .

As  $(-c)^{13} = -c^{13} = -\bar{1}$  and  $(-c)^2 \neq \bar{1}$  we see that  $-c$  has order 26. So  $-c$  does generate  $\mathbb{Z}_3(c)^*$ . As  $-p(-x)$  is monic irreducible over  $\mathbb{Z}_3$  and  $-c$  is a zero of both  $-p(-x)$  and  $x^{13} + \bar{1}$  we obtain, as above, that  $\gcd\{-p(-x), x^{13} + \bar{1}\} = -p(-x)$ , i.e.  $-p(-x)$  is a divisor of  $x^{13} + \bar{1}$ .

The zeros of  $x^3 - x$  over  $\mathbb{Z}_3$  are  $\bar{0}, \bar{1}, -\bar{1}$ , i.e. the three elements  $a \in \mathbb{Z}_3$ . By (4.2)(ii) this polynomial of degree 3 over  $\mathbb{Z}_3$  has no further zeros in  $\mathbb{Z}_3(c)$ , i.e. for  $a \in \mathbb{Z}_3(c), a \notin \mathbb{Z}_3$  we have  $a^3 \neq a$ . So  $a^3 = a \Leftrightarrow a \in \mathbb{Z}_3$ . In terms of the Frobenius automorphism  $\theta: \mathbb{Z}_3(c) \rightarrow \mathbb{Z}_3(c)$  we therefore obtain  $(a)\theta = a$  if and only if  $a \in \mathbb{Z}_3$ , i.e.  $\mathbb{Z}_3$  is the fixed field of  $\theta$ . Write

$p_a(x) = (x - a)(x - a^3)(x - a^9) = x^3 + b_2x^2 + b_1x + b_0$  and so, on multiplying out and equating coefficients we obtain  $b_2 = -(a + a^3 + a^9)$ ,  $b_1 = a \times a^3 + a^3 \times a^9 + a^9 \times a$ ,  $b_0 = -a \times a^3 \times a^9$ . The coefficients  $b_0, b_1, b_2$  belong to  $\mathbb{Z}_3(c)$  and we use the above property of  $\theta$  to show that they in fact belong to  $\mathbb{Z}_3$ . Notice that  $(a)\theta = a^3, (a^3)\theta = a^9, (a^9)\theta = a^{27} = a$  by Qu. 3 (c) above, i.e.  $\theta$  cyclically permutes the three zeros of  $p_a(x)$ . Hence  $(b_2)\theta = -(a + a^3 + a^9)\theta = -(a^3 + a^9 + a) = b_2$  and so  $b_2 \in \mathbb{Z}_3$ . Similarly  $(b_1)\theta = (a \times a^3 + a^3 \times a^9 + a^9 \times a)\theta = a^3 \times a^9 + a^9 \times a + a \times a^3 = b_1$  showing  $b_1 \in \mathbb{Z}_3$ . Also  $(b_0)\theta = (-a \times a^3 \times a^9)\theta = -a^3 \times a^9 \times a = b_0$  and so  $b_0 \in \mathbb{Z}_3$ . We conclude that  $p_a(x)$  is a polynomial over  $\mathbb{Z}_3$ . By (4.8)(ii)  $p_a(x)$  is irreducible over  $\mathbb{Z}_3$  where  $a \in \mathbb{Z}_3(c), a \notin \mathbb{Z}_3$  as  $\deg p_a(x) = 3$  and  $p_a(x)$  has no zeros in  $\mathbb{Z}_3$ . Each subset  $\{a, a^3, a^9\}$  of  $\mathbb{Z}_3(c)$  with  $a \notin \mathbb{Z}_3$  gives rise to a monic irreducible polynomial  $p_a(x) = p_{a^3}(x) = p_{a^9}(x)$  of degree 3 over  $\mathbb{Z}_3$ . There are  $(27 - 3)/3 = 8$  such subsets  $\{a, a^3, a^9\}$ . Combining Question 3(c) and Question 4(a) above we see that each monic irreducible polynomial of degree 3 over  $\mathbb{Z}_3$  is of the type  $p_a(x)$ , and so there are 8 such polynomials.

The 13 powers of  $c$ , i.e. the elements of the multiplicative group generated by  $c$ , are the zeros in  $\mathbb{Z}_3(c)$  of  $x^{13} - \bar{1}$  over  $\mathbb{Z}_3$ . So  $x^3 - x - \bar{1} = (x - c)(x - c^3)(x - c^9)$  is a divisor of  $x^{13} - \bar{1}$ . The monic polynomial having the inverses of  $c, c^3, c^9$  as its zeros is  $((1/x)^3 - (1/x) - \bar{1})(-x^3) = x^3 + x^2 - \bar{1} = (x - c^{-1})(x - c^{-3})(x - c^{-9}) = (x - c^{12})(x - c^{10})(x - c^4) = p_{c^4}(x)$  which is a divisor of  $x^{13} - \bar{1}$ . The polynomial  $p_{c^2}(x) = (x - c^2)(x - c^6)(x - c^{18})$  is a divisor of  $x^{13} - \bar{1}$ . Note  $c^6 = (c^3)^2 = c^2 - c + \bar{1}$  and

$$c^{18} = (c^6)^3 = (c^2 - c + \bar{1})^3 = c^6 - c^3 + \bar{1} = c^2 - c + \bar{1} - c - \bar{1} + \bar{1} = c^2 + c + \bar{1}.$$

Hence  $p_{c^2}(x) = (x - c^2)(x - c^2 + c - \bar{1})(x - c^2 - c - \bar{1})$ . The coefficients of  $x^2, x^1, x^0$  in  $p_{c^2}(x)$  are

$$-c^2 - c^2 + c - \bar{1} - c^2 - c - \bar{1} = \bar{1}, \quad c^2(c^2 - c + \bar{1}) + c^2(c^2 + c + \bar{1}) + (c^2 - c + \bar{1})(c^2 + c + \bar{1}) = 3c^4 + 0c^3 + 3c^2 + 0c + \bar{1} = \bar{1}, \quad -c^2 \times c^6 \times c^{18} = -c^{26} = -\bar{1}$$

respectively and so  $p_{c^2}(x) = x^3 + x^2 + x - \bar{1}$ .

The monic polynomial having zeros the reciprocals (inverses) of  $c^2, c^4, c^{18}$  is

$$((1/x)^3 + (1/x)^2 + (1/x) - \bar{1})(-x^3) = x^3 - x^2 - x - \bar{1} = (x - c^{-2})(x - c^{-6})(x - c^{-18}) = p_{c^7}(x)$$

as  $c^{-6} = c^7$  and this polynomial is also a divisor of  $x^{13} - \bar{1}$ . Therefore

$$x^{13} - \bar{1} = (x - \bar{1})(x^3 - x - \bar{1})(x^3 + x^2 - \bar{1})(x^3 + x^2 + x - \bar{1})(x^3 - x^2 - x - \bar{1})$$

is the factorisation of  $x^{13} - \bar{1}$  into monic irreducible polynomials over  $\mathbb{Z}_3$ .

Replacing  $x$  by  $-x$  in the above factorisation and changing the sign of each factor, thereby making all factors monic, produces the factorisation of  $x^{13} + \bar{1}$  into monic irreducible polynomials over  $\mathbb{Z}_3$ , namely

$$x^{13} + \bar{1} = (x + \bar{1})(x^3 - x - \bar{1})(x^3 + x^2 - \bar{1})(x^3 - x^2 + x + \bar{1})(x^3 + x^2 - x + \bar{1}).$$

(c) As  $p(x) = x^4 + x^2 - \bar{1} = x^2(x^2 + \bar{1}) - \bar{1}$  we see that  $p(x)$  is not divisible by  $x^2 + \bar{1}$  over  $\mathbb{Z}_3$ . As  $x^4 + x^2 - \bar{1} = (x^2 - x)(x^2 + x - \bar{1}) - x - \bar{1}$  and  $x^4 + x^2 - \bar{1} = (x^2 + x)(x^2 - x - \bar{1}) + x - \bar{1}$  we see that  $p(x)$  is not divisible by either  $x^2 + x - \bar{1}$  or  $x^2 - x - \bar{1}$ . So  $p(x)$  has no irreducible factor of degree 2 over  $\mathbb{Z}_3$ . As  $p(\bar{0}) = -\bar{1}$ ,  $p(\bar{1}) = p(-\bar{1}) = \bar{1}$  we see that  $p(x)$  has no factor of degree 1 over  $\mathbb{Z}_3$ .

Therefore  $p(x)$  is irreducible over  $\mathbb{Z}_3$ . As  $c^4 + c^2 = \bar{1}$  we have

$$i^2 = (c^2 - \bar{1})^2 = c^4 + c^2 + \bar{1} = \bar{1} + \bar{1} = -\bar{1}. \quad \mathbb{Z}_3(i) = \{\bar{0}, \bar{1}, -\bar{1}, i, i + \bar{1}, i - \bar{1}, -i, -i + \bar{1}, -i + \bar{1}\}$$

is closed under addition and multiplication. Also  $\mathbb{Z}_3(i) \cong \mathbb{Z}_3[x]/\langle x^2 + \bar{1} \rangle$  is a field by (4.9), since  $i$  is a zero of  $x^2 + \bar{1}$  which is irreducible over  $\mathbb{Z}_3$ .

Using the Frobenius automorphism  $\theta$  of  $\mathbb{Z}_3(c)$  given by  $(a)\theta = a^3$  for all  $a \in \mathbb{Z}_3(c)$ , we obtain

$p(x) = (x - c)(x - c^3)(x - c^9)(x - c^{27})$  which is the factorisation of  $p(x)$  into monic irreducible polynomials over  $\mathbb{Z}_3(c)$ . Now  $(a)\theta^2 = a^9$  for all  $a \in \mathbb{Z}_3(c)$ . Hence  $a \in \mathbb{Z}_3(i) \Leftrightarrow (a)\theta^2 = a$  showing that  $\mathbb{Z}_3(i)$  is the fixed field of  $\theta^2$ . As  $c^2 = i + \bar{1}$  on squaring we get  $c^4 = (i + \bar{1}) = i^2 - i + \bar{1} = -i$  and so incidentally  $c$  has multiplicative order 16. Also  $c^8 = -\bar{1}$  and so  $c + c^9 = \bar{0}$ ,

$$c^{10} = c^8 \times c^2 = -c^2 = -i - \bar{1}. \quad \text{Therefore } (x - c)(x - c^9) = x^2 - i - \bar{1}. \quad \text{Similarly}$$

$$(x - c^3)(x - c^{27}) = x^2 + i - \bar{1} \quad \text{and so } p(x) = (x^2 - i - \bar{1})(x^2 + i - \bar{1}) \text{ is the factorisation of } p(x) \text{ into monic irreducible quadratics over } \mathbb{Z}_3(i).$$

The polynomial  $x^{81} - x$  splits over  $\mathbb{Z}_3(c)$  each subset  $\{a, a^3, a^9, a^{27}\}$  of  $\mathbb{Z}_3(c)$  for  $a \notin \mathbb{Z}_3(i)$  being the zeros of a monic irreducible polynomial of degree 4 over  $\mathbb{Z}_3(c)$ . There are

$$(|\mathbb{Z}_3(c)| - |\mathbb{Z}_3(i)|)/4 = (81 - 9)/4 = 18 \text{ such subsets and so there are 18 monic irreducible polynomials of degree 4 over } \mathbb{Z}_3(c).$$

## Solution 7

(a) Write  $n = [E : F]$ . The  $n+1$  elements  $c^0, c^1, c^2, \dots, c^n$  of the  $n$ -dimensional vector space  $E$  over  $F$  are linearly dependent. So there are 'scalars'  $a_0, a_1, a_2, \dots, a_n \in F$ , not all zero, satisfying  $a_0 + a_1 c + a_2 c^2 + \dots + a_n c^n = 0$ . Therefore  $(f_0(x))\mathcal{E}_c = 0$  where  $f_0(x) = a_0 + a_1 x + a_2 x^2 + \dots + a_n x^n \in F[x]$ , i.e.  $f_0(x) \in \ker \mathcal{E}_c = K$ . Hence  $K$  is non-zero as  $f_0(x)$  is non-zero. By (4.4) there is a monic polynomial  $m_c(x) \in F[x]$  with  $K = \langle m_c(x) \rangle$ . So  $m_c(c) = 0$ ,  $\deg m_c(x) \geq 1$  and  $m_c(x)$  is the monic polynomial of smallest degree over  $F$  having  $c$  as its zero, i.e.  $m_c(x)$  is **the minimum polynomial** of  $c$  over  $F$ . To show that  $m_c(x)$  is irreducible over  $F$  suppose that there are  $g(x), h(x) \in F[x]$  with  $m_c(x) = g(x)h(x)$ . Applying  $\mathcal{E}_c$  gives  $0 = m_c(c) = (m_c(x))\mathcal{E}_c = (g(x)h(x))\mathcal{E}_c = (g(x))\mathcal{E}_c (h(x))\mathcal{E}_c = g(c)h(c)$  and so either  $g(c) = 0$  or  $h(c) = 0$  (or both) as the field  $E$  has no zero-divisors. Hence either  $g(x) \in K$  or  $h(x) \in K$ , i.e.  $m_c(x) \mid g(x)$  or  $m_c(x) \mid h(x)$ . So either  $\deg g(x) \geq \deg m_c(x)$  or  $\deg h(x) \geq \deg m_c(x)$ . Therefore  $m_c(x)$  is irreducible over  $F$  by (4.7). Taking  $\theta = \mathcal{E}_c$  in (4.9) we see that  $F(c) = \{f(c) : f(x) \in F[x]\} = \text{im } \mathcal{E}_c$  is a field, i.e.  $F(c)$  is a subfield of  $E$ . In fact  $F(c)$  is the smallest subfield of  $E$  which contains  $F$  and  $c$ . From (4.9) we deduce  $\tilde{\mathcal{E}}_c : F[x]/K \cong F[c]$  as  $K = \ker \mathcal{E}_c$ ,  $F(c) = \text{im } \mathcal{E}_c$ . We show that  $c^0, c^1, c^2, \dots, c^{m-1}$  where  $m = \deg m_c(x)$  is a basis of the vector space  $F(c)$  over  $F$ . By the minimality of  $m$  we see that  $c^0, c^1, c^2, \dots, c^{m-1}$  are linearly independent. On dividing  $f(x)$  by  $m_c(x)$  there are  $q(x), r(x) \in F[x]$  with  $f(x) = q(x)m_c(x) + r(x)$  where  $\deg r(x) < m$ . Hence  $f(c) = r(c)$ , i.e. a typical element  $f(c)$  of  $F(c)$  is expressible as  $r(c)$  which is a linear combination of  $c^0, c^1, c^2, \dots, c^{m-1}$ . So  $c^0, c^1, c^2, \dots, c^{m-1}$  span  $F(c)$ . Therefore  $[F(c) : F] = m = \deg m_c(x)$ .

(b) Consider  $w \in E$ . There are  $l_1, l_2, \dots, l_n \in L$  with  $w = l_1 v_1 + l_2 v_2 + \dots + l_n v_n$  since  $v_1, v_2, \dots, v_n$  span the vector space  $E$  over the field  $L$ . For each  $l_j$  ( $1 \leq j \leq n$ ) there are  $a_{1j}, a_{2j}, \dots, a_{mj} \in F$  with  $l_j = a_{1j} u_1 + a_{2j} u_2 + \dots + a_{mj} u_m$  since  $u_1, u_2, \dots, u_m$  span the vector space  $L$  over the field  $F$ . On substituting for each  $l_j$  we obtain

$$w = \sum_{j=1}^n \left( \sum_{i=1}^m a_{ij} u_i \right) v_j = \sum_{i=1}^m \sum_{j=1}^n a_{ij} u_i v_j$$

showing that the elements  $u_i v_j$  ( $1 \leq i \leq m, 1 \leq j \leq n$ ) span the vector space  $E$  over the field  $F$ .

To show that the  $u_i v_j$  ( $1 \leq i \leq m, 1 \leq j \leq n$ ) are linearly independent elements of the vector space  $E$

over  $F$  suppose  $\sum_{i=1}^m \sum_{j=1}^n a_{ij} u_i v_j = 0$  where each  $a_{ij} \in F$ . Then  $\sum_{j=1}^n \left( \sum_{i=1}^m a_{ij} u_i \right) v_j = 0$ . As

$l_j = \sum_{i=1}^m a_{ij} u_i \in L$  for  $1 \leq j \leq n$  and  $v_1, v_2, \dots, v_n$  are linearly independent elements of the vector space  $E$

over the field  $L$  we deduce  $l_j = 0$  for  $1 \leq j \leq n$ . From  $l_j = \sum_{i=1}^m a_{ij} u_i = 0$  we deduce  $a_{ij} = 0$  for

$1 \leq i \leq m$  as  $u_1, u_2, \dots, u_m$  are linearly independent elements of the vector space  $L$  over  $F$ . Therefore the  $u_i v_j$  ( $1 \leq i \leq m, 1 \leq j \leq n$ ) are indeed linearly independent elements of  $E$  over  $F$  and so these  $mn$  elements constitute a basis of  $E$  over  $F$ . We have proved  $[E : F] = mn = nm = [E : L][L : F]$ .

(c) (i) From Exercises 2.3, Question 5(a) we see  $|E| = q^n$  and  $[L:F] = d$ . Hence from (b) above  $n = [E:L]d$  showing  $d \mid n$ .

Conversely suppose  $d \mid n$ . Then  $q^n - 1 = (q^d - 1)(q^{n-d} + q^{n-2d} + q^{n-3d} + \dots + q^d + 1)$  showing

$(q^d - 1) \mid (q^n - 1)$ . So  $x^{q^n} - x = x(x^{q^{n-1}} - 1) = x(x^{q^{d-1}} - 1)g(x) = (x^{q^d} - x)g(x)$  where

$g(x) = (x^{q^{n-1}} - 1)/(x^{q^{d-1}} - 1) = x^{q^{n-1} - (q^d - 1)} + x^{q^{n-1} - 2(q^d - 1)} + x^{q^{n-1} - 3(q^d - 1)} + \dots + x^{q^{d-1}} + 1$  showing  $(x^{q^d} - x) \mid (x^{q^n} - x)$ . By Question 3(c) above  $E$  consists of the  $q^n$  zeros of  $x^{q^n} - x$ . Hence  $L$

consists of the  $q^d$  zeros of  $x^{q^d} - x$  in  $E$ . As  $L$  is the fixed field of  $\theta^{dt}$  where  $\theta$  is the Frobenius automorphism of  $E$  and  $q = p^t$  ( $p$  prime), we see that  $L$  is a subfield of  $E$  with  $|L| = q^d$  and  $F \subseteq L$ .

So  $L$  exists given the existence of  $E$ . By Question 3(c) above every subfield of  $E$  having order  $q^d$  consists of the zeros of  $x^{q^d} - x$  and so is unique.

(ii) As  $L \cap M$  is a subfield of  $E$  containing  $F$  by (i) above there is a positive integer  $m$  with

$|L \cap M| = q^m$  and  $m \mid n$ . As  $L \cap M$  is a subfield of  $L$  and a subfield of  $M$  we see  $m \mid d$  and  $m \mid e$  by

(i) above. So  $m \mid \gcd\{d, e\}$ . By (i) above  $E$  has a unique subfield  $N$  with  $|N| = q^{\gcd\{d, e\}}$ . As

$\gcd\{d, e\} \mid d$  and  $\gcd\{d, e\} \mid e$  we see  $N \subseteq L$  and  $N \subseteq M$  also by (i) above. Therefore  $N \subseteq L \cap M$

and so  $\gcd\{d, e\} \mid m$  by (i) above. So  $m = \gcd\{d, e\}$  as each of these positive integers is a divisor of the other.

(iii) By convention  $\bigcap_{j \in \emptyset} L_j = E$  and so the equation  $\left| \bigcap_{j \in X} L_j \right| = q^{n/\pi_X}$  is valid for  $X = \emptyset$  as

$|E| = q^n = q^{n/\pi_\emptyset}$ . Suppose  $|X| = s \geq 1$  and write  $X = \{j_1, j_2, \dots, j_s\}$ ,  $X_0 = \{j_1, j_2, \dots, j_{s-1}\}$ . We

assume inductively that  $\left| \bigcap_{j \in X_0} L_j \right| = q^{n/\pi_{X_0}}$ . From (ii) above  $\left| \left( \bigcap_{j \in X_0} L_j \right) \cap L_{j_s} \right| = q^{\gcd\{n/\pi_{X_0}, n/p_{j_s}\}}$ . As

$$\bigcap_{j \in X} L_j = \left( \bigcap_{j \in X_0} L_j \right) \cap L_{j_s}$$

and

$$\gcd\{n/\pi_{X_0}, n/p_{j_s}\} = \gcd\left\{ \prod_{j \in X_0} p_j^{n_j-1} \prod_{j \notin X_0} p_j^{n_j}, p_{j_s}^{n_{j_s}-1} \prod_{j \neq j_s} p_j^{n_j} \right\} = \left( \prod_{j \in X_0} p_j^{n_j-1} \right) p_{j_s}^{n_{j_s}-1} \left( \prod_{j \notin X} p_j^{n_j} \right) = n / \pi_X$$

we see  $\left| \bigcap_{j \in X} L_j \right| = q^{n/\pi_X}$  completing the inductive step.

By (a) above the minimum polynomial of each of the  $r$  elements of  $E$  not in any subfield  $L$  of  $E$  ( $L \neq E$ ) is monic and irreducible of degree  $n$  over  $F$ . By Question 3(c) above all such polynomials arise there being  $n$  elements of  $E$  having the same minimum polynomial (its  $n$  distinct zeros in  $E$ ). So

$$\left( \sum_X (-1)^{|X|} q^{n/\pi_X} \right) / n$$

is the number of monic irreducible polynomials of degree  $n$  over  $F \cong \mathbb{F}_q$ .

Taking  $q = 2$ ,  $n = 12 = 2^2 \times 3$  gives  $(2^{12} - 2^6 - 2^4 - 2^2)/12 = 4020/12 = 335$  monic irreducible polynomials of degree 12 over  $\mathbb{Z}_2$ . Taking  $q = 4$ ,  $n = 6 = 2 \times 3$  gives  $(4^6 - 4^3 - 4^2 + 4)/6 = 670$  monic irreducible polynomials of degree 6 over  $\mathbb{F}_4$ .

Taking  $q = 3$ ,  $n = 12 = 2^2 \times 3$  gives  $(3^{12} - 3^6 - 3^4 + 3^2)/12 = 44220$  monic irreducible polynomials of degree 12 over  $\mathbb{Z}_3$ . Taking  $q = 9$ ,  $n = 6 = 2 \times 3$  gives  $(9^6 - 9^3 - 9^2 + 9)/6 = 88440$  monic irreducible polynomials of degree 6 over  $\mathbb{F}_9$ .

### Solution 8

(a) The elements of the ring  $R_m$  are cosets  $K + f(x)$  where  $K = \langle p(x)^m \rangle$  and  $f(x) \in F[x]$ . The 1-element of  $R_m$  is  $K + 1$ . Suppose  $K + f(x) \in G_m = U(R_m)$ . There is  $g(x) \in F[x]$  with  $(K + g(x))(K + f(x)) = K + 1$ . So  $g(x)f(x) - 1 \in K$  which gives  $g(x)f(x) - 1 = q(x)p(x)^m$ . Therefore  $p(x)$  is not a divisor of  $f(x)$ , as otherwise  $p(x) \mid 1$  which is not true. So  $\gcd\{f(x), p(x)\} = 1$ . Conversely suppose  $\gcd\{f(x), p(x)\} = 1$ . Then  $\gcd\{f(x), p(x)^m\} = 1$  also as 1 is the only monic divisor of  $p(x)^m$  which is not divisible by  $p(x)$ . By (4.6) there are  $g(x), q(x) \in F[x]$  with  $g(x)f(x) - q(x)p(x)^m = 1$  which gives  $g(x)f(x) - 1 \in K$ . So  $(K + g(x))(K + f(x)) = K + 1$ , showing that  $K + f(x)$  is an invertible element of  $R_m$  (its inverse is  $K + g(x)$ ), i.e.  $K + f(x) \in G_m$ .

Consider  $K + f_1(x), K + f_2(x) \in H_m$ . Write  $L = \langle p(x) \rangle$ . Then  $f_1(x) \equiv 1 \pmod{L}$  and  $f_2(x) \equiv 1 \pmod{L}$ . Multiplying these congruences together gives  $f_1(x)f_2(x) \equiv 1 \pmod{L}$ . So  $(K + f_1(x))(K + f_2(x)) = K + f_1(x)f_2(x) \in H_m$  showing that  $H_m$  is closed under multiplication. As  $1 \equiv 1 \pmod{L}$  we see that  $H_m$  contains the 1-element  $K + 1$  of  $G_m$ . Suppose  $(K + g(x))(K + f(x)) = K + 1$  where  $K + f(x) \in H_m$ . Then  $f(x) \equiv 1 \pmod{L}$  and hence  $g(x) \equiv g(x)f(x) \pmod{L}$ . As  $K \subseteq L$  we see  $g(x)f(x) \equiv 1 \pmod{L}$  and so  $g(x) \equiv 1 \pmod{L}$ . So  $K + g(x) \in H_m$ . Therefore  $H_m$  is closed under inversion. So  $H_m$  is a subgroup of  $G_m$ . Suppose  $F$  is a finite field with  $|F| = q$ . Each element of  $R_m$  is uniquely expressible as  $K + f(x)$  where  $\deg f(x) < \deg p(x)^m = mn$ . Therefore  $|R_m| = q^{mn}$  as there are  $q^{mn}$  polynomials over  $F$  of degree less than  $mn$ . The number of these polynomials which are divisible by  $p(x)$  is  $q^{mn-n} = q^{(m-1)n}$  as such a polynomial is  $h(x)p(x)$  where  $\deg h(x) < mn - n$ . By the preceding part of the question  $|G_m| = q^{mn} - q^{(m-1)n} = q^{(m-1)n}(q^n - 1)$ . The elements of  $H_m$  are  $K + h(x)p(x) + 1$  where as above  $\deg h(x) < mn - n$ . So  $|H_m| = q^{(m-1)n}$ .

As  $K \subseteq L$  the natural mapping  $\eta: R_m \rightarrow F[x]/L$ , given by  $(K + f(x))\eta = L + f(x)$  for all  $f(x) \in F[x]$ , is a surjective ring homomorphism from  $R_m$  to the finite field  $E = F[x]/L$  of order  $q^n$ . Also  $\ker \eta = \{K + f(x) : p(x) \mid f(x)\} = \langle K + p(x) \rangle$ . By (3.17) there is an element  $g_1 \in G_m$  such that  $(g_1)\eta$  generates the multiplicative group  $E^*$  of  $q^n - 1$  non-zero elements of  $E$ . Now  $(g_1)\eta$  has order  $q^n - 1$  in  $E^*$  and so  $(g_1^{q^n-1})\eta = ((g_1)\eta)^{q^n-1} = L + 1$ . Therefore  $g_1^{q^n-1} \in H_m$  and further  $g_1^l \in H_m \Leftrightarrow (q^n - 1) \mid l$ . Let  $d$  be the order of  $g_1^{q^n-1}$  and so  $g_1$  has order  $d(q^n - 1)$  as



$\gcd\{d, q^n - 1\} = 1$ . Write  $g_0 = g_1^d$ . Then  $g_0$  has order  $q^n - 1$ . Let  $H_0$  be the cyclic subgroup of  $G_m$  generated by  $g_0$ . Then  $|H_m \cap H_0| = 1$  as  $\gcd\{|H_m|, |H_0|\} = \gcd\{q^{(m-1)n}, q^n - 1\} = 1$ . Also  $G_m = H_m H_0$  as  $|H_m H_0| = |H_m| |H_0| = q^{(m-1)n} (q^n - 1) = |G_m|$ . Therefore  $G_m \cong H_m \times H_0$  using the multiplicative version of (2.15) with  $t = 2$ .

Let  $d_1, d_2, \dots, d_s$  be the invariant factors of  $H_m$ . As  $H_0$  has the single invariant factor  $q^n - 1$  and  $\gcd\{d_s, q^n - 1\} = 1$ , the invariant factors of  $G_m \cong H_m \times H_0$  are  $d_1, d_2, \dots, d_{s-1}, d_s(q^n - 1)$ .

(b) Taking  $q = 2, n = 1$  in (a) above we obtain  $|H_0| = 1$  and so  $G_m = H_m$  and  $|G_m| = 2^{m-1}$ . As  $|G_1| = 1, |G_2| = 2$  the invariant factors of  $G_1$  and  $G_2$  are  $\emptyset$  (the empty set) and 2 respectively.

The group  $G_3$  has four elements  $K+1, K+1+x, K+1+x^2, K+1+x+x^2$  where  $K = \langle x^3 \rangle$  and  $1+1=0$ . As  $(K+1+x)^2 = K+1+x^2, (K+1+x)^3 = K+1+x+x^2, (K+1+x)^4 = K+1$  since  $x^4 \in K$ , we see that  $G_3$  is cyclic of order 4 and so  $G_3$  has the single invariant factor 4.

$G_4$  is a group of order 8. With  $K = \langle x^4 \rangle$  the element  $K+1+x$  has order 4 as  $(K+1+x)^2 = K+1+x^2, (K+1+x)^4 = K+1$ . Also  $K+1+x^3$  has order 2 as  $x^6 \in K$ . Further  $K+1+x^3$  is not contained in the subgroup generated by  $K+1+x$ . So  $K+1+x^3$  and  $K+1+x$  generate independent subgroups of  $G_4$  having orders 2 and 4 respectively. Therefore  $G_4$  is the internal direct product of these subgroups and so  $G_4$  has invariant factors 2, 4.

$G_5$  is a group of order 16. With  $K = \langle x^5 \rangle$  the element  $K+1+x$  has order 8 as  $(K+1+x)^4 = K+1+x^4 \neq K+1$  but  $(K+1+x)^8 = K+1$ . The element  $K+1+x^3$  has order 2 and is different from the unique element  $K+1+x^4$  of order 2 in the subgroup generated by  $K+1+x$ . Therefore  $G_5$  has invariant factors 2, 8.

$G_6$  is a group of order 32. With  $K = \langle x^6 \rangle$  the element  $K+1+x$  has order 8. The elements  $K+1+x^3$  and  $K+1+x^5$  have order 2 and together with  $(K+1+x)^4 = K+1+x^4$  generate an elementary abelian 2-group with invariant factors 2, 2, 2. Therefore  $G_6$  has invariant factors 2, 2, 8. So  ${}^2G_6$  has invariant factors 2, 2, 2. Also  ${}^4G_6$  has invariant factors 2, 2, 4 and  ${}^8G_6 = G_6$  has invariant factors 2, 2, 8 as above.

(c) Let  $\lfloor (j-1)m/j \rfloor = k$ . Then  $(j-1)m/j = k + \varepsilon$  where  $0 \leq \varepsilon < 1$ . So

$m/j = (1 - (l-1)/l)m = m - k - \varepsilon$ . Let  $i$  be an integer. Then  $m/j \leq i \Leftrightarrow m - k - \varepsilon \leq i \Leftrightarrow m - k \leq i$ .

So  $m/j \leq i < m \Leftrightarrow m - k \leq i < m$ . There are  $k$  integers  $i$  in the latter range. So there are  $\lfloor (j-1)m/j \rfloor$  integers  $i$  with  $m/j \leq i < m$ .

Consider  $g = K + f(x) \in G_m$  where  $K = \langle x^m \rangle \subseteq \mathbb{Z}_2[x]$ . We may assume  $\deg f(x) < m$  and so

$f(x) = \sum_{i=0}^{m-1} a_i x^i$  where  $a_0 = 1$ . Then

$$g \in {}^{2^j}G_m \Leftrightarrow g^{2^j} = K + 1 \Leftrightarrow f(x)^{2^j} \equiv 1 \pmod{K} \Leftrightarrow \sum_{i=0}^{m-1} a_i^{2^j} x^{i2^j} \equiv 1 \pmod{K} \Leftrightarrow a_i = 0 \text{ for } 0 < i2^j < m.$$

The coefficients  $a_i$  for  $m \leq i2^j$  do not appear above and so are arbitrary elements of  $\mathbb{Z}_2$ . The number

$t_j$  of these coefficients is the number of integers  $i$  with  $m/2^j \leq i < m$ . So  $|2^j G_m| = 2^{t_j}$  where

$t_j = \lfloor (2^j - 1)m/2^j \rfloor$  by the first paragraph of (c).

As  $2^{r-1} < m \leq 2^r$  we see  $1/2 < m/2^r \leq 1$  and so  $(2^r - 1)m/2^r = m - m/2^r \geq m - 1$  showing  $t_r = m - 1$ .

So  $2^r G_m = G_m$  and hence  $2^j G_m = G_m$  for  $j \geq r$ , i.e.  $t_j = m - 1$  for  $j \geq r$ .

Comparing orders using Exercises 3.1, Question 5(c) gives  $t_0 = 0$  and

$t_j = \sum_{1 \leq i < j} i s_i + j(s_j + s_{j+1} + \dots + s_{m-1})$  for  $j \geq 1$ . Solving these linear equations we obtain

$s_j = -t_{j-1} + 2t_j - t_{j+1}$  for  $j \geq 1$ . Hence  $s_j = 0$  for  $j > r$  as  $t_r = t_{r+1} = \dots = m - 1$ . Also

$$s_j = -t_{j-1} + 2t_j - t_{j+1} = -\lfloor (2^{j-1} - 1)m/2^{j-1} \rfloor + 2\lfloor (2^j - 1)m/2^j \rfloor - \lfloor (2^{j+1} - 1)m/2^{j+1} \rfloor \text{ for } j \geq 1.$$

Take  $m = 25$ . So  $r = 5$  and

$$t_1 = \lfloor 25/2 \rfloor = 12, t_2 = \lfloor 3 \times 25/4 \rfloor = 18, t_3 = \lfloor 7 \times 25/8 \rfloor = 21, t_4 = \lfloor 15 \times 25/16 \rfloor = 23, t_5 = \lfloor 31 \times 25/32 \rfloor = 24.$$

Therefore  $s_1 = 6, s_2 = 3, s_3 = 1, s_4 = 1, s_5 = 1, s_j = 0$  for  $j \geq 6$ , showing that  $G_{25}$  has invariant factor sequence  $(2, 2, 2, 2, 2, 2, 4, 4, 4, 8, 16, 32)$ .

Take  $m = 32$ . So  $r = 5$  and  $t_j = \lfloor (2^j - 1)2^5/2^j \rfloor = 2^5 - 2^{5-j}$  for  $0 \leq j \leq 5$ . Therefore

$s_1 = 8, s_2 = 4, s_3 = 2, s_4 = 1, s_5 = 1, s_j = 0$  for  $j \geq 6$ , showing that  $G_{32}$  has invariant factor sequence  $(2, 2, 2, 2, 2, 2, 2, 2, 4, 4, 4, 8, 16, 32)$ .

(d) To show  $\alpha$  is unambiguously defined (u.d.), consider  $f_1(x), f_2(x) \in F[x]$  such that

$\langle g(x)h(x) \rangle + f_1(x) = \langle g(x)h(x) \rangle + f_2(x)$ . Then  $f_1(x) - f_2(x) \in \langle g(x)h(x) \rangle$  and so

$g(x)h(x) \mid (f_1(x) - f_2(x))$ . Hence  $g(x) \mid (f_1(x) - f_2(x))$  and  $h(x) \mid (f_1(x) - f_2(x))$ , i.e.

$f_1(x) - f_2(x) \in \langle g(x) \rangle$  and  $f_1(x) - f_2(x) \in \langle h(x) \rangle$ . Therefore  $\langle g(x) \rangle + f_1(x) = \langle g(x) \rangle + f_2(x)$

and  $\langle h(x) \rangle + f_1(x) = \langle h(x) \rangle + f_2(x)$  showing that  $\alpha$  is u.d. as we set out to prove.

To show that  $\alpha$  respects addition and multiplication, consider elements  $a, a' \in F[x]/\langle g(x)h(x) \rangle$ . There are  $f(x), f'(x) \in F[x]$  with  $a = \langle g(x)h(x) \rangle + f(x)$  and  $a' = \langle g(x)h(x) \rangle + f'(x)$ . Then

$$(a + a')\alpha = (\langle g(x)h(x) \rangle + f(x) + f'(x))\alpha = (\langle g(x) \rangle + f(x) + f'(x), \langle h(x) \rangle + f(x) + f'(x)) = (\langle g(x) \rangle + f(x), \langle h(x) \rangle + f(x)) + (\langle g(x) \rangle + f'(x), \langle h(x) \rangle + f'(x)) = (a)\alpha + (a')\alpha$$

and

$$(aa')\alpha = (\langle g(x)h(x) \rangle + f(x)f'(x))\alpha = (\langle g(x) \rangle + f(x)f'(x), \langle h(x) \rangle + f(x)f'(x)) = (\langle g(x) \rangle + f(x), \langle h(x) \rangle + f(x))(\langle g(x) \rangle + f'(x), \langle h(x) \rangle + f'(x)) = (a)\alpha (a')\alpha$$

showing that  $\alpha$  does respect addition and multiplication. As

$$(\langle g(x)h(x) \rangle + 1)\alpha = (\langle g(x) \rangle + 1, \langle h(x) \rangle + 1)$$

we see that  $\alpha$  respects 1-elements. So  $\alpha$  is a ring homomorphism.

Suppose  $\gcd\{g(x), h(x)\} = 1$ . To show that  $\alpha$  is injective suppose  $(a)\alpha = (a')\alpha$ . Using the above notation we obtain  $(\langle g(x) \rangle + f(x), \langle h(x) \rangle + f(x)) = (\langle g(x) \rangle + f'(x), \langle h(x) \rangle + f'(x))$  and this equality between ordered pairs gives  $\langle g(x) \rangle + f(x) = \langle g(x) \rangle + f'(x)$  and  $\langle h(x) \rangle + f(x) = \langle h(x) \rangle + f'(x)$ .

Therefore  $f(x) - f'(x)$  is divisible by  $g(x)$  and by  $h(x)$ . So  $f(x) - f'(x)$  is divisible by  $g(x)h(x)$  giving  $a = \langle g(x)h(x) \rangle + f(x) = \langle g(x)h(x) \rangle + f'(x) = a'$ . So  $\alpha$  is injective. To show that  $\alpha$  is

surjective consider a typical element  $(\langle g(x) \rangle + s(x), \langle h(x) \rangle + t(x))$  of  $(F[x]/\langle g(x) \rangle) \oplus (F[x]/\langle h(x) \rangle)$ .

By (4.6) there are  $a(x), b(x) \in F[x]$  with  $a(x)g(x) + b(x)h(x) = 1$ . Let

$r(x) = t(x)a(x)g(x) + s(x)b(x)h(x)$ . Then  $r(x) \equiv s(x) \pmod{g(x)}$  and  $r(x) \equiv t(x) \pmod{h(x)}$  and so

$(\langle g(x) \rangle + s(x), \langle h(x) \rangle + t(x)) = (\langle g(x)h(x) \rangle + r(x))\alpha$  showing that  $\alpha$  is surjective. The conclusion is:

$$\alpha: F[x]/\langle g(x)h(x) \rangle \cong (F[x]/\langle g(x) \rangle) \oplus (F[x]/\langle h(x) \rangle), \text{ i.e. } \alpha \text{ is a ring isomorphism.}$$

Let  $\sigma: \mathbb{Z}_2[x] \rightarrow \mathbb{Z}_2[x]$  be given by  $(f(x))\sigma = f(x-1)$  for all  $f(x) \in F[x]$ . Then  $\sigma$  is self-inverse

$(\sigma^{-1} = \sigma)$  and  $\sigma$  is an automorphism of the ring  $\mathbb{Z}_2[x]$ . Also  $(\langle x^m \rangle)\sigma = \langle (x-1)^m \rangle$  and so

$\tilde{\sigma}: R_m \cong R'_m$  is a ring isomorphism where  $(\langle x^m \rangle + f(x))\tilde{\sigma} = \langle (x-1)^m \rangle + f(x-1)$ .

Restricting  $\tilde{\sigma}$  to the group  $U(R_m)=G_m$  gives the group isomorphism  $\tilde{\sigma}|_{G_m}:G_m\cong U(R'_m)$ .

By the above polynomial version of the Chinese remainder theorem  $\alpha: \mathbb{Z}_2[x]/\langle x^l(x-1)^m \rangle \cong R_l \oplus R'_m$

is a ring isomorphism. The restriction of  $\alpha$  to the group of invertible elements of  $\mathbb{Z}_2[x]/\langle x^l(x-1)^m \rangle$  is

an isomorphism  $U(\mathbb{Z}_2[x]/\langle x^l(x-1)^m \rangle) \cong U(R_l \oplus R'_m)$ . As  $U(R \oplus R') = U(R) \times U(R')$  for all rings

$R$  and  $R'$  we conclude  $U(\mathbb{Z}_2[x]/\langle x^l(x-1)^m \rangle) \cong U(R_l) \times U(R'_m) \cong G_l \times G_m$ .

The invariant factor sequence of  $U(\mathbb{Z}_2[x]/\langle x^{25}(x-1)^{32} \rangle) \cong G_{25} \times G_{32}$  is

(2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 2, 4, 4, 4, 4, 4, 4, 8, 8, 16, 16, 32, 32) on combining the invariant factors of  $G_{25}$  and  $G_{32}$  from (c) above.

(e) Express  $f(x)$  in the scale of  $p(x)$ : let  $r_0(x)$  be the remainder on dividing  $f(x)$  by  $p(x)$ . Then

$\deg r_0(x) < n$  and  $f(x) - r_0(x) \equiv 0 \pmod{p(x)}$ . Let the integer  $j$  satisfy  $1 \leq j < m$  and suppose

inductively that there are polynomials  $r_i(x)$  with  $\deg r_i(x) < n$  for  $0 \leq i < j$  such that

$$f(x) - \sum_{i=0}^{j-1} r_i(x)p(x)^i \equiv 0 \pmod{p(x)^j}.$$
 Let  $r_j(x)$  be the remainder on dividing the polynomial

$(1/p(x)^j)(f(x) - \sum_{i=0}^{j-1} r_i(x)p(x)^i)$  by  $p(x)$ . Then  $\deg r_j(x) < n$  and there is  $q_j(x) \in F[x]$  with

$$(1/p(x)^j)(f(x) - \sum_{i=0}^{j-1} r_i(x)p(x)^i) = q_j(x)p(x) + r_j(x) \text{ by (4.1). Multiplying this equation by } p(x)^j$$

and rearranging gives  $f(x) - \sum_{i=0}^j r_i(x)p(x)^i = q_j(x)p(x)^{j+1} \equiv 0 \pmod{p(x)^{j+1}}$ . The inductive step is

now established. Taking  $j = m - 1$  gives  $f(x) - \sum_{i=0}^{m-1} r_i(x)p(x)^i \equiv 0 \pmod{p(x)^m}$ . As

$$\deg(f(x) - \sum_{i=0}^{m-1} r_i(x)p(x)^i) < mn = \deg p(x)^m \text{ we conclude } f(x) - \sum_{i=0}^{m-1} r_i(x)p(x)^i = 0, \text{ i.e.}$$

$f(x) = \sum_{i=0}^{m-1} r_i(x)p(x)^i$ . So  $f(x)$  is expressible in the scale of  $p(x)$  as stated. Conversely suppose

$f(x) = \sum_{i=0}^{m-1} r_i(x)p(x)^i$  where  $\deg r_i(x) < n$  for  $0 \leq i < m$ . Then  $r_0(x)$  is unique by (4.1) being the

remainder on division of  $f(x)$  by  $p(x)$ . Let the integer  $j$  satisfy  $1 \leq j < m$  and suppose inductively

that  $r_i(x)$  are unique for  $0 \leq i < j$ . Then  $r_j(x)$  is unique by (4.1) being the remainder on division of

$$(1/p(x)^j)(f(x) - \sum_{i=0}^{j-1} r_i(x)p(x)^i) \text{ by } p(x). \text{ By induction the } r_i(x) \text{ are unique for } 0 \leq i < m.$$

Each element of  $R_m$  can be expressed uniquely as  $K + f(x)$  where  $K = \langle p(x)^m \rangle$  and  $\deg f(x) < mn$ ,

$$f(x) \in \mathbb{F}_q[x]. \text{ Elements of } G_m \text{ are } K + f(x) \text{ where } f(x) = \sum_{i=0}^{m-1} r_i(x)p(x)^i \text{ with } r_0(x) \neq 0. \text{ Elements}$$

of  $H_m$  are  $K + f(x)$  where  $f(x) = \sum_{i=0}^{m-1} r_i(x)p(x)^i$  with  $r_0(x) = 1$ . Suppose  $K + f(x) \in {}^{p_0^j}H_m$  where

$$j \geq 0. \text{ Then } (K + f(x))^{p_0^j} = K + 1, \text{ i.e. } K + \sum_{i=1}^{m-1} r_i(x)^{p_0^j} p(x)^{ip_0^j} + 1 = K + 1 \text{ as raising to the power } p_0$$

is an injective ring homomorphism of  $\mathbb{F}_q[x]$ . Therefore  $\sum_{i=1}^{m-1} r_i(x)^{p_0^j} p(x)^{ip_0^j} \equiv 0 \pmod{p(x)^m}$ . Is it

possible for  $r_i(x) \neq 0$  where  $1 \leq i < m/p_0^j$ ? If so then consider the least such integer  $i$ . As  $ip_0^j < m$  we

see  $r_i(x)^{p_0^j} \equiv 0 \pmod{p(x)}$ . But  $r_i(x) \not\equiv 0 \pmod{p(x)}$  since  $\deg r_i(x) < n = \deg p(x)$  and so

$$r_i(x)^{p_0^j} \not\equiv 0 \pmod{p(x)} \text{ as } p(x) \text{ is irreducible over } \mathbb{F}_q. \text{ This contradiction shows } r_i(x) = 0 \text{ for}$$

$1 \leq i < m/p_0^j$ . Conversely  $r_i(x) = 0$  for  $1 \leq i < m/p_0^j$  implies  $K + f(x) \in {}^{p_0^j}H_m$  as each term in

$$\sum_{i=1}^{m-1} r_i(x)^{p_0^j} p(x)^{ip_0^j} \text{ is divisible by } p(x)^m. \text{ Therefore}$$

$$K + f(x) \in {}^{p_0^j}H_m \Leftrightarrow r_i(x) = 0 \text{ for } 1 \leq i < m/p_0^j \text{ where } f(x) = \sum_{i=0}^{m-1} r_i(x)p(x)^i.$$

The number of integers  $i$  with  $m/p_0^j \leq i < m$  is  $t_j = \lfloor (p_0^j - 1)m/p_0^j \rfloor$  by (c) above. The number of

choices for each  $r_i(x)$  where  $m/p_0^j \leq i < m$  so that  $K + f(x) \in {}^{p_0^j}H_m$  is  $q^n$  as  $r_i(x)$  is an arbitrary

polynomial with  $\deg r_i(x) < n$  over  $\mathbb{F}_q$ . So  $|{}^{p_0^j}H_m| = q^{nt_j} = p_0^{lnt_j}$  for  $j \geq 0$ .

As in (c) above we see

$$s_j = \ln(-t_{j-1} + 2t_j - t_{j+1}) = \ln(-\lfloor (p_0^{j-1} - 1)m/p_0^{j-1} \rfloor + 2\lfloor (p_0^j - 1)m/p_0^j \rfloor - \lfloor (p_0^{j+1} - 1)m/p_0^{j+1} \rfloor)$$

for  $j \geq 1$ .

In the case  $m=11, q=9, n=2$  we have  $p_0=3, l=2$  and so

$$t_0=0, t_1=\lfloor (3-1)11/3 \rfloor=7, t_2=\lfloor (3^2-1)11/3^2 \rfloor=9, t_3=\lfloor (3^3-1)11/3^3 \rfloor=10=t_j \text{ (} j \geq 3 \text{)}.$$

Therefore  $s_1=4(-0+2 \times 7-9)=20, s_2=4(-7+2 \times 9-10)=4, s_3=4(-9+2 \times 10-10)=4$  and

$s_j=0$  for  $j \geq 4$ . So  $H_{11}$  has 20 invariant factors 3 followed by 4 invariant factors 9 followed by 4 invariant factors 27. Finally  $G_{11}$  also has 28 invariant factors, the first 27 being the first 27 of  $H_{11}$ , the last invariant factor of  $G_{11}$  being  $27 \times (3^4 - 1) = 2160$  (see the last part of (a) above).

## Solutions 4.2 (page 199)

### Solution 1

(a)

$$\begin{pmatrix} \bar{0} & \bar{1} \\ \bar{1} & \bar{0} \end{pmatrix}, \begin{pmatrix} \bar{1} & \bar{0} \\ \bar{0} & \bar{1} \end{pmatrix}, \begin{pmatrix} \bar{1} & \bar{1} \\ \bar{0} & \bar{1} \end{pmatrix}, \begin{pmatrix} \bar{1} & \bar{0} \\ \bar{1} & \bar{1} \end{pmatrix}, \begin{pmatrix} \bar{1} & x \\ \bar{0} & \bar{1} \end{pmatrix}, \begin{pmatrix} \bar{1} & \bar{0} \\ x & \bar{1} \end{pmatrix}, \begin{pmatrix} \bar{1} & x+\bar{1} \\ \bar{0} & \bar{1} \end{pmatrix}, \begin{pmatrix} \bar{1} & \bar{0} \\ x+\bar{1} & \bar{1} \end{pmatrix}$$

the corresponding *eros* being

$$r_1 \leftrightarrow r_2, \bar{1}r_1 = \bar{1}r_2 = r_1 + \bar{0}r_2 = r_2 + \bar{0}r_1, r_1 + r_2, r_2 + r_1, r_1 + xr_2, r_2 + xr_1, r_1 + (x+\bar{1})r_2, r_2 + (x+\bar{1})r_1$$

respectively. Over  $\mathbb{Z}_p[x]$  the  $2 \times 2$  elementary matrices with entries of degree at most 1 are:

$$\begin{pmatrix} \bar{0} & \bar{1} \\ \bar{1} & \bar{0} \end{pmatrix}, \begin{pmatrix} a & \bar{0} \\ \bar{0} & \bar{1} \end{pmatrix}, \begin{pmatrix} \bar{1} & \bar{0} \\ \bar{0} & a \end{pmatrix}, \begin{pmatrix} \bar{1} & \bar{0} \\ \bar{0} & \bar{1} \end{pmatrix}, \begin{pmatrix} \bar{1} & f(x) \\ \bar{0} & \bar{1} \end{pmatrix}, \begin{pmatrix} \bar{1} & \bar{0} \\ f(x) & \bar{1} \end{pmatrix}$$

where  $a \in \mathbb{Z}_p, a \neq \bar{0}, \bar{1}$  and  $f(x) \in \mathbb{Z}_p[x], f(x) \neq 0, \deg f(x) \leq 1$ . There are  $p-2$  choices for  $a$  and  $p^2-1$  choices for  $f(x) = a_1x + a_0$ . So in all there are

$$1 + (p-2) + (p-2) + 1 + (p^2-1) + (p^2-1) = 2(p-1)(p+2) \text{ such matrices.}$$

(b) The *eros* producing the given  $2 \times 2$  elementary matrices over  $\mathbb{Q}[x]$  are:

$$r_1 \leftrightarrow r_2, r_2 + r_1, r_1 + x^2r_2, 3r_1, (1/3)r_2, r_1 - x^2r_2 \text{ respectively. The six paired } \textit{ecos} \text{ are:}$$

$$c_1 \leftrightarrow c_2, c_1 + c_2, c_2 + x^2c_1, 3c_1, (1/3)c_2, c_2 - x^2c_1 \text{ respectively. The six conjugate } \textit{ecos} \text{ are:}$$

$$c_1 \leftrightarrow c_2, c_1 - c_2, c_2 - x^2c_1, (1/3)c_1, 3c_2, c_2 + x^2c_1 \text{ respectively.}$$

(c) The *ero*  $ar_i$  and the *eco*  $ac_i$  are paired for all  $a \in F^*$ . These elementary operations are conjugate and non-identity  $\Leftrightarrow a^2 = 1, a \neq 1$ , i.e.  $a = -1, \chi(F) \neq 2$ . For  $i \neq j$  the *ero*  $r_i + f(x)r_j$  and the *eco*  $c_j + f(x)c_i$  are paired for all  $f(x) \in F[x]$ . These elementary operations are conjugate and non-identity  $\Leftrightarrow f(x) + f(x) = 0, f(x) \neq 0 \Leftrightarrow \chi(F) = 2$ . Therefore apart from  $r_i \leftrightarrow r_j, c_i \leftrightarrow c_j$  the only paired (non-identity) *eros* and *ecos* which are also conjugate are  $-r_i, -c_i$  in case  $\chi(F) \neq 2$ , and  $r_i + f(x)r_j, c_j + f(x)c_i$  where  $f(x) \neq 0$  in case  $\chi(F) = 2$ .

### Solution 2

(a) Let  $A(x), B(x), C(x) \in \mathfrak{M}_{s \times t}(F[x])$ . *Reflexive law*: taking  $P(x)$  as the  $s \times s$  identity matrix over  $F[x]$  and  $Q(x)$  as the  $t \times t$  identity matrix over  $F[x]$  we see  $P(x)A(x)Q(x)^{-1} = A(x)$  showing  $A(x) \equiv A(x)$  for all  $A(x) \in \mathfrak{M}_{s \times t}(F[x])$  by (4.11). *Symmetric law*: suppose  $A(x) \equiv B(x)$ . By (4.11) there are  $P(x) \in GL_s(F(x))$  and  $Q(x) \in GL_t(F(x))$  with  $P(x)A(x)Q(x)^{-1} = B(x)$ . Hence  $P(x)^{-1}B(x)Q(x) = P(x)^{-1}B(x)(Q(x)^{-1})^{-1} = A(x)$ . As  $P(x)^{-1} \in GL_s(F(x))$  and  $Q(x)^{-1} \in GL_t(F(x))$  we see  $B(x) \equiv A(x)$  by (4.11). *Transitive law*: suppose  $A(x) \equiv B(x)$  and  $B(x) \equiv C(x)$ . By (4.11) there are  $P_1(x), P_2(x) \in GL_s(F(x))$  and  $Q_1(x), Q_2(x) \in GL_t(F(x))$  with  $P_1(x)A(x)Q_1(x)^{-1} = B(x)$  and  $P_2(x)B(x)Q_2(x)^{-1} = C(x)$ . On substituting for  $B(x)$  we obtain  $P_2(x)P_1(x)A(x)(Q_2(x)Q_1(x))^{-1} = C(x)$  which shows  $A(x) \equiv B(x)$  by (4.11) as  $P_2(x)P_1(x) \in GL_s(F(x))$  and  $Q_2(x)Q_1(x) \in GL_t(F(x))$ . Therefore  $\equiv$  is an equivalence relation on  $\mathfrak{M}_{s \times t}(F[x])$ .

From (4.11)  $A(x) \equiv 0 \Rightarrow A(x) = 0$  and so  $\{0\}$  is the equivalence class of the zero matrix  $0$ .

Suppose  $A(x) \in \mathfrak{M}_t(F[x])$  satisfies  $A(x) \equiv I$ . Using (4.11) we see

$A(x) = P(x)^{-1}IQ(x) \in GL_t(F[x])$  as  $P(x), Q(x) \in GL_t(F[x])$ . On the other hand  $P(x) \in GL_t(F[x])$  satisfies  $P(x) \equiv I$  as  $IP(x)P(x)^{-1} = I$ . So  $GL_t(F[x])$  is the equivalence class of  $I$ .

(b) As  $d_3(x) \mid x^3$ , by the polynomial analogue of the fundamental theorem of arithmetic, the only possibilities are:

$(1, 1, 1), (1, 1, x), (1, x, x), (x, x, x), (1, 1, x^2), (1, x, x^2), (x, x, x^2), (1, x^2, x^2), (x, x^2, x^2), (x^2, x^2, x^2)$   
 $(1, 1, x^3), (1, x, x^3), (x, x, x^3), (1, x^2, x^3), (x, x^2, x^3), (x^2, x^2, x^3), (1, x^3, x^3), (x, x^3, x^3), (x^2, x^3, x^3), (x^3, x^3, x^3).$

By (4.19) each equivalence class of  $3 \times t$  matrices over  $F[x]$  corresponds to exactly one of the above triples. So there are 20 such equivalence classes.

(c) For elementary matrices  $P(x)$  over  $F[x]$  of types (i) and (ii) we see  $P(-x) = P(x)$  as  $P(x)$  is a matrix over  $F$ . Let  $P(x)$  be an elementary matrix over  $F[x]$  of type (iii). Then  $P(x)$  arises from  $I$  by applying a row operation  $r_i + f(x)r_j$  where  $f(x) \in F[x]$ . Hence  $P(-x)$  arises from  $I$  by applying  $r_i + f(-x)r_j$ , i.e.  $P(-x)$  is also an elementary matrix over  $F[x]$  as  $f(-x) \in F[x]$ . We conclude:

$P(x)$  elementary over  $F[x] \Rightarrow P(-x)$  elementary over  $F[x]$ . As  $P(-(-x)) = P(x)$  the opposite implication also holds.

By (4.11), (4.16) and (4.17) there are elementary matrices

$P_1(x), P_2(x), \dots, P_k(x), Q_1(x), Q_2(x), \dots, Q_l(x)$  over  $F[x]$  with

$$P_k(x) \cdots P_2(x)P_1(x)A(x)Q_1(x)Q_2(x) \cdots Q_l(x) = \text{diag}(d_1(x), d_2(x), \dots, d_{\min\{s,t\}}(x)).$$

Replacing  $x$  by  $-x$  gives

$$P_k(-x) \cdots P_2(-x)P_1(-x)A(-x)Q_1(-x)Q_2(-x) \cdots Q_l(-x) = \text{diag}(d_1(x), d_2(x), \dots, d_{\min\{s,t\}}(x)).$$

Using the preceding theory we see  $A(-x) \equiv \text{diag}(d_1(-x), d_2(-x), \dots, d_{\min\{s,t\}}(-x))$ . Write

$d'_i(x) = (-1)^{\deg d_i(x)} d_i(-x)$  for  $1 \leq i \leq \min\{s, t\}$ . Then  $d'_i(x)$  is monic and  $d'_i(x) \mid d'_{i+1}(x)$  for  $1 \leq i < \min\{s, t\}$ . As  $A(-x) \equiv \text{diag}(d'_1(x), d'_2(x), \dots, d'_{\min\{s,t\}}(x))$  using elementary matrices of type (ii) over  $F[x]$  we conclude  $S(A(-x)) = \text{diag}(d'_1(x), d'_2(x), \dots, d'_{\min\{s,t\}}(x))$ .

By (4.19) we see  $A(x) \equiv A(-x) \Leftrightarrow d_i(x) = d'_i(x)$  for  $1 \leq i \leq \min\{s, t\}$ . But

$d_i(x) = (-1)^{\deg d_i(x)} d_i(-x) \Leftrightarrow d_i(x)$  is either even (no odd powers of  $x$  occur in  $d_i(x)$ ) or odd (only odd powers of  $x$  occur in  $d_i(x)$ ).

Clearly  $g_1(A(x)) = 1$  as the gcd of the entries in  $A(x)$  is 1. On carrying out  $r_1 - r_2$  we see

$$\det A(x) = \begin{vmatrix} 1 & 1 \\ x^2 & x^2 + x \end{vmatrix} = x = g_2(A(x)).$$

So  $S(A(x)) = \text{diag}(1, x)$ . In this case  $d_1(x) = 1$  is even and  $d_2(x) = x$  is odd. Hence  $A(x) \equiv A(-x)$ .

As above  $g_1(B(x)) = 1$  and applying  $r_1 - r_2$  gives

$$\det B(x) = \begin{vmatrix} -1 & -1 \\ x^2 - x & x^2 + 1 \end{vmatrix} = -x - 1.$$

Hence  $g_2(B(x)) = x + 1$  and  $S(B(x)) = \text{diag}(1, x + 1)$ . In this case  $d_1(x) = 1$  is even but  $d_2(x) = x + 1$  is neither even nor odd. Hence  $B(x) \not\equiv B(-x)$ .

**Solution 3**

(a)  $A(x) = x(x+1) \begin{pmatrix} x & 0 \\ 0 & x+1 \end{pmatrix}$ . Using (4.15)

$$\begin{pmatrix} x & 0 \\ 0 & x+1 \end{pmatrix} \equiv \begin{pmatrix} x & -x \\ 0 & x+1 \end{pmatrix} \equiv \begin{pmatrix} x & 1 \\ 0 & x+1 \end{pmatrix} \equiv \begin{pmatrix} 1 & 1 \\ 1-x^2 & x+1 \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ 1-x^2 & x(x+1) \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ 0 & x(x+1) \end{pmatrix}$$

and so the sequence  $c_2 - c_1, r_1 + r_2, c_1 - (x-1)c_2, c_2 - c_1, r_2 + (x^2-1)r_1$  reduces  $A(x)$  to

$$S(A(x)) = \text{diag}(x(x+1), x^2(x+1)^2).$$

Applying  $r_1 + r_2, r_2 + (x^2-1)r_1$  (the *eros* above) to  $I$  gives  $P(x) = \begin{pmatrix} 1 & 1 \\ x^2-1 & x^2 \end{pmatrix}$ . Applying

$$r_1 + r_2, r_2 + (x-1)r_1, r_1 + r_2 \text{ (the conjugates of the above } \textit{ecos} \text{) to } I \text{ produces } Q(x) = \begin{pmatrix} x & x+1 \\ x-1 & x \end{pmatrix}.$$

Then  $P(x)A(x) = S(A(x))Q(x)$ .  $A(x)^2 = \text{diag}(x^4(x+1)^2, x^2(x+1)^4)$  has invariant factors  $x^2(x+1)^2, x^4(x+1)^4$  by (4.15).

(b) The elementary operations  $c_1 \leftrightarrow c_3, r_1 \leftrightarrow r_2, c_2 \leftrightarrow c_3, r_2 \leftrightarrow r_3$  reduce the given matrix to its Smith normal form  $\text{diag}(1, x, x^2)$ .

(c)

$$\begin{aligned} A(x) &= \begin{pmatrix} x^2 & x & x \\ x^3-2x^2 & x^2-x & x^2-2x \\ 2x^3+x-1 & x^2 & 2x^2 \end{pmatrix} \equiv \begin{pmatrix} x & 0 & x \\ x^2-x & -x^2 & x^2-2x \\ x^2 & x^3+x-1 & 2x^2 \end{pmatrix} \begin{matrix} c_1 - xc_2 \\ c_1 \leftrightarrow c_2 \end{matrix} \equiv \begin{pmatrix} x & 0 & 0 \\ x^2-x & -x^2 & -x \\ x^2 & x^3+x-1 & x^2 \end{pmatrix} \begin{matrix} c_3 - c_1 \\ c_1 \leftrightarrow c_2 \end{matrix} \\ &\equiv \begin{pmatrix} x & 0 & 0 \\ x^2-x & -x^2 & -x \\ x^2 & x^3+x-1 & x^2 \end{pmatrix} \begin{matrix} r_2 - (x-1)r_1 \\ r_3 - xr_1 \end{matrix} \equiv \begin{pmatrix} x & 0 & 0 \\ 0 & -x^2 & -x \\ 0 & x^3+x-1 & x^2 \end{pmatrix} \begin{matrix} c_2 - xc_3 \\ c_2 \leftrightarrow c_3 \end{matrix} \\ &\equiv \begin{pmatrix} x & 0 & 0 \\ 0 & x & 0 \\ 0 & 0 & x-1 \end{pmatrix} \begin{matrix} r_3 + xr_2 \\ -c_2 \end{matrix} \equiv \begin{pmatrix} x & 0 & 1 \\ 0 & x & 0 \\ 0 & 0 & x-1 \end{pmatrix} \begin{matrix} r_1 - r_3 \\ c_3 + c_1 \end{matrix} \equiv \begin{pmatrix} 1 & 0 & 0 \\ 0 & x & 0 \\ -(x-1)^2 & 0 & x(x-1) \end{pmatrix} \begin{matrix} c_1 + (1-x)c_3 \\ c_3 - c_1 \end{matrix} \\ &\equiv \begin{pmatrix} 1 & 0 & 0 \\ 0 & x & 0 \\ 0 & 0 & x(x-1) \end{pmatrix} \begin{matrix} r_3 + (x-1)^2 r_1 \end{matrix} = S(A(x)). \end{aligned}$$

Applying  $r_2 - (x-1)r_1, r_3 - xr_1, r_3 + xr_2, r_1 - r_3, r_3 + (x-1)^2 r_1$  to the  $3 \times 3$  identity matrix over  $\mathbb{Q}[x]$  gives

$$P(x) = \begin{pmatrix} 1+x^2 & -x & -1 \\ 1-x & 1 & 0 \\ 1-2x+x^2-2x^3+x^4 & x^2(2-x) & x(2-x) \end{pmatrix}.$$

Applying  $r_2 + xr_1, r_1 \leftrightarrow r_2, r_1 + r_3, r_3 + xr_2, r_2 \leftrightarrow r_3, -r_2, r_1 - r_3, r_3 + (x-1)r_1, r_1 + r_3$  to the  $3 \times 3$  identity matrix over  $\mathbb{Q}[x]$  gives

$$Q(x) = \begin{pmatrix} x^2 - x + 1 & x & x \\ -x & 0 & -1 \\ x^2 - 2x + 2 & x - 1 & x - 1 \end{pmatrix}.$$

Then  $P(x)A(x) = S(A(x))Q(x)$  and so  $P(A)A(x)Q(x)^{-1} = S(A(x))$ .

(d) As  $A(x) = AD(x)$  where  $D(x) = \text{diag}(x, x^2, \dots, x^t)$  we see  $A(x) \equiv D(x)$  and so  $D(x) = S(A(x))$ .

(e) The sequence:  $c_2 + (a^{-1} - 1)c_1, c_1 + c_2, c_2 + (a - 1)c_1, c_1 - a^{-1}c_2$  reduces  $\text{diag}(a, a^{-1})$  to  $I = \text{diag}(1, 1)$ . Use the Euclidean algorithm to find a sequence of *ecos* over  $F[x]$  of type (iii) which reduces  $A(x)$  to either  $(d(x), 0)$  or  $(0, d(x))$ . Assuming  $d(x) \neq 0(x)$  (otherwise  $A(x)$  is the zero matrix and there is nothing to do) in the latter case carry out  $c_1 + c_2, c_2 - c_1$  to get  $(d(x), 0)$ . Let  $a$  denote the leading coefficient of  $d(x)$ . Now carry out:  $c_2 + (a^{-1} - 1)c_1, c_1 + c_2, c_2 + (a - 1)c_1$  which changes  $(d(x), 0) = (\bar{d}(x), 0)\text{diag}(a, a^{-1})$  into  $S(A(x)) = (\bar{d}(x), 0)$  where  $\bar{d}(x) = a^{-1}d(x)$ . Suppose  $v$  *ecos* of type (iii) are needed to reduce  $A(x)$  to  $S(A(x))$ . Let  $Q_1(x), Q_2(x), \dots, Q_v(x)$  denote the corresponding elementary matrices over  $F[x]$ . As  $\det Q_j(x) = 1$  for  $1 \leq j \leq v$  we see  $Q(x)^{-1} = Q_1(x)Q_2(x) \cdots Q_v(x)$  satisfies  $A(x) = S(A(x))Q(x)$  and  $\det Q(x) = 1$ .

#### Solution 4

The quotient and remainder on dividing  $a_n(x)$  by  $xa_{n-1}(x)$  are  $xa_{n-2}(x)$  and  $a_{n-1}(x)$  respectively as  $a_n(x) = x^2a_{n-2}(x)a_{n-1}(x) + a_{n-1}(x)$ . Hence  $\gcd\{a_n(x), xa_{n-1}(x)\} = a_{n-1}(x)$  for  $n \geq 2$ .

As  $a_{n-1}(x) \mid a_n(x)$  for  $n \geq 1$  we obtain

$$\begin{aligned} a_n(x)/a_{n-r+1}(x) &= (a_n(x)/a_{n-1}(x))(a_{n-1}(x)/a_{n-2}(x)) \cdots (a_{n-r+2}(x)/a_{n-r+1}(x)) = \\ &= (x^2a_{n-2}(x) + 1)(x^2a_{n-3}(x) + 1) \cdots (x^2a_{n-r}(x) + 1). \end{aligned}$$

As  $a_{n-r}(x)$  is a divisor of  $a_{n-r+1}(x), a_{n-r+2}(x), \dots, a_{n-2}(x)$ , each of the above  $r-1$  factors belongs to the coset  $K_{n-r} + 1$  and so does their product, i.e.  $a_n(x)/a_{n-r+1}(x) \equiv 1 \pmod{K_{n-r}}$  for  $1 \leq r \leq n$ .

Using the definition of  $a_n(x)$  we see

$$a_2(x) = x(x^2 + 1), a_3(x) = x(x^2 + 1)(x^3 + 1), a_4(x) = x(x^2 + 1)(x^3 + 1)(x^5 + x^3 + 1).$$

Hence

$$\begin{aligned} &\begin{pmatrix} x(x^2 + 1)(x^3 + 1)(x^5 + x^3 + 1) & x^2(x^2 + 1)(x^3 + 1) \\ 0 & -1 \end{pmatrix} \equiv \begin{pmatrix} x(x^2 + 1)(x^3 + 1) & x^2(x^2 + 1)(x^3 + 1) \\ x^2(x^2 + 1) & -1 \end{pmatrix} \equiv \\ &\begin{pmatrix} x(x^2 + 1)(x^3 + 1) & 0 \\ x^2(x^2 + 1) & -(x^5 + x^3 + 1) \end{pmatrix} \equiv \begin{pmatrix} x(x^2 + 1) & x^2(x^5 + x^3 + 1) \\ x^2(x^2 + 1) & -(x^5 + x^3 + 1) \end{pmatrix} \equiv \begin{pmatrix} x(x^2 + 1) & x^2(x^5 + x^3 + 1) \\ 0 & -(x^3 + 1)(x^5 + x^3 + 1) \end{pmatrix} \equiv \\ &\begin{pmatrix} x(x^2 + 1) & x^2 \\ 0 & -(x^3 + 1)(x^5 + x^3 + 1) \end{pmatrix} \equiv \begin{pmatrix} x & x^2 \\ x(x^3 + 1)(x^5 + x^3 + 1) & -(x^3 + 1)(x^5 + x^3 + 1) \end{pmatrix} \equiv \\ &\begin{pmatrix} x & 0 \\ x(x^3 + 1)(x^5 + x^3 + 1) & -(x^2 + 1)(x^3 + 1)(x^5 + x^3 + 1) \end{pmatrix} \equiv \begin{pmatrix} x & 0 \\ 0 & -(x^2 + 1)(x^3 + 1)(x^5 + x^3 + 1) \end{pmatrix} = D(x) \end{aligned}$$

using the sequence

$$c_1 - x^2(x^2 + 1)c_2, c_2 - xc_1, r_1 - x^2r_2, r_2 - xr_1, c_2 - x^4c_1, c_1 - xc_2, c_2 - xc_1, r_2 - (x^3 + 1)(x^5 + x^3 + 1)r_1.$$

For  $1 \leq r < n$  write



$$B_r(x) = \begin{pmatrix} a_{n-r}(x) & 0 \\ xa_{n-r-1}(x)(a_n(x)/a_{n-r+1}(x)) & -(a_n(x)/a_{n-r}(x)) \end{pmatrix} \quad (r \text{ odd})$$

and

$$B_r(x) = \begin{pmatrix} a_{n-r}(x) & xa_{n-r-1}(x)(a_n(x)/a_{n-r+1}(x)) \\ 0 & -(a_n(x)/a_{n-r}(x)) \end{pmatrix} \quad (r \text{ even}),$$

the polynomial analogues of the matrices  $B_r$  in Exercises 1.2, Question 6(b).

The *ecos*  $c_1 - xa_{n-2}(x)c_2, c_2 - xc_1$  change  $A(x)$  into  $B_1(x)$ . The *eros*  $r_1 - xa_{n-3}(x)r_2, r_2 - xr_1$  change  $B_1(x)$  into  $B_2(x)$ . The *ecos*  $c_2 - xa_{n-r}(x)q_{n-r+1}(x)c_1, c_1 - xa_{n-r-1}(x)c_2, c_2 - xc_1$  change  $B_{r-1}(x)$  into  $B_r(x)$  for odd  $r$  with  $3 \leq r < n$ . The *eros*  $r_2 - xa_{n-r}(x)q_{n-r+1}(x)r_1, r_1 - xa_{n-r-1}(x)r_2, r_2 - xr_1$  change  $B_{r-1}(x)$  into  $B_r(x)$  for even  $r$  with  $4 \leq r < n$ . Finally  $B_{n-1}(x)$  is changed into

$D(x) = \text{diag}(x, -(a_n(x)/x))$  by either the *eco*  $c_2 - (a_n(x)/a_2(x))c_1$  ( $n$  odd) or the *ero*

$r_2 - (a_n(x)/a_2(x))r_1$  ( $n$  even). The number of elementary operations over  $F[x]$  used in the reduction of  $A_n(x)$  to the diagonal matrix  $D(x)$  is therefore  $2 + 2 + 3(n-3) + 1 = 3n - 4$  where  $n \geq 3$ .

There is  $q_1(x) \in F[x]$  with  $(a_n(x)/a_1(x)) = (a_n(x)/x) = 1 + x^2q_1(x)$ . The sequence

$r_1 - r_2, c_2 - xq_1(x)c_1, c_1 \leftrightarrow c_2, c_2 - xc_1, r_2 + (1 + x^2q_1(x))r_1$  of elementary operations over  $F[x]$  reduces  $D(x)$  to  $S(A_n(x)) = \text{diag}(1, a_n(x))$ .

The sequence  $r_1 \leftrightarrow r_2, c_1 \leftrightarrow c_2, r_2 + xa_{n-1}(x)r_1, -c_1$  reduces  $A_n(x)$  to its Smith normal form  $S(A(x))$ .

### Solution 5

(a) Consider  $(P_0(x), Q_0(x), (P_1(x), Q_1(x)) \in Z(D(x))$ . Then

$P_0(x)P_1(x)D(x) = P_0(x)D(x)Q_1(x) = D(x)Q_0(x)Q_1(x)$  showing

$(P_0(x), Q_0(x))(P_1(x), Q_1(x)) = (P_0(x)P_1(x), Q_0(x)Q_1(x)) \in Z(D(x))$ . Pre- and post-multiplying

$P_0(x)D(x) = D(x)Q_0(x)$  by  $P_0(x)^{-1}$  and  $Q_0(x)^{-1}$  respectively gives  $P_0(x)^{-1}D(x) = D(x)Q_0(x)^{-1}$

showing  $(P_0(x), Q_0(x))^{-1} = (P_0(x)^{-1}, Q_0(x)^{-1}) \in Z(D(x))$ . Also the identity element  $(I, I)$  of  $G$  belongs to  $Z(D(x))$  as  $ID(x) = D(x)I$ . So  $Z(D(x))$  is a subgroup of  $G$ .

(b) Suppose  $(P(x), Q(x)) \in Z(D(x))$ . Comparing  $(i, j)$ -entries in  $P(x)D(x) = D(x)Q(x)$  for  $1 \leq i, j \leq s$  gives  $p_{ij}(x)d_j(x) = d_i(x)q_{ij}(x)$ . Comparing  $(i, l)$ -entries in  $P(x)D(x) = D(x)Q(x)$  for  $1 \leq i \leq s < l \leq t$  gives  $0 = d_i(x)q_{il}(x)$  and so  $q_{il}(x) = 0$  as  $d_i(x) \neq 0$ .

Conversely suppose  $(P(x), Q(x)) \in G$  satisfies  $p_{ij}(x)d_j(x) = d_i(x)q_{ij}(x)$  for  $1 \leq i, j \leq s$  and  $q_{il}(x) = 0$  for  $1 \leq i \leq s < l \leq t$ . Then  $P(x)D(x) = D(x)Q(x)$  as corresponding entries agree.

(c) From  $P'(x)^{-1}D(x)Q'(x) = A = P''(x)^{-1}D(x)Q''(x)$  we deduce

$P'(x)P''(x)^{-1}D(x) = D(x)Q'(x)Q''(x)^{-1}$ , i.e.

$(P'(x), Q'(x))(P''(x), Q''(x))^{-1} = (P'(x)P''(x)^{-1}, Q'(x)Q''(x)^{-1}) \in Z(D(x))$ . Therefore  $(P'(x), Q'(x))$

and  $(P''(x), Q''(x))$  belong to the same left coset of  $Z(D(x))$  in  $G$ , i.e.

$Z(D(x))(P'(x), Q'(x)) = Z(D(x))(P''(x), Q''(x))$ .

(d) The sequence:  $c_1 - xc_2, c_2 - xc_1, r_2 - r_1, r_1 - r_2, c_2 - xc_1, c_1 \leftrightarrow c_2, c_2 - xc_1, r_2 + (x^2 + 1)r_1$  of elementary operations over  $F[x]$  reduces  $\begin{pmatrix} x(x^2 + 1) & x^2 \\ 0 & -1 \end{pmatrix}$  to  $\text{diag}(1, x(x^2 + 1))$  as does the sequence:

$r_1 \leftrightarrow r_2, c_1 \leftrightarrow c_2, r_2 + x^2 r_1, -c_1$ . Applying the *eros* in the first sequence to the  $2 \times 2$  identity matrix gives  $\begin{pmatrix} 2 & -1 \\ 2x^2 + 1 & -x^2 \end{pmatrix} = P'(x)$ . Applying the conjugates of the *ecos* in the first sequence to the  $2 \times 2$  identity

matrix gives  $\begin{pmatrix} 2x(x^2 + 1) & 2x^2 + 1 \\ 2x^2 + 1 & 2x \end{pmatrix} = Q'(x)$ . Using the second sequence in place of the first gives

$\begin{pmatrix} 0 & 1 \\ 1 & x^2 \end{pmatrix} = P''(x)$  and  $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = Q''(x)$ . As  $P''(x)^{-1} = \begin{pmatrix} -x^2 & 1 \\ 1 & 0 \end{pmatrix}$  and  $Q''(x)^{-1} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$  we

obtain  $(P'(x), Q'(x))(P''(x), Q''(x))^{-1} = \left( \begin{pmatrix} -2x^2 - 1 & 2 \\ -2x^2(x^2 + 1) & 2x^2 + 1 \end{pmatrix}, \begin{pmatrix} -2x^2 - 1 & 2x(x^2 + 1) \\ -2x & 2x^2 + 1 \end{pmatrix} \right)$  which

belongs to  $Z(\text{diag}(1, x(x^2 + 1)))$ .

(e) Let the  $s \times t$  matrix  $D(x) = \text{diag}(d_1(x), d_2(x), \dots, d_r(x), 0, \dots, 0)$  be in Smith normal form where

$d_r(x) \neq 0$ . Suppose  $(P(x), Q(x)) \in Z(D(x))$ . Then  $P(x)D(x) = D(x)Q(x)$ . Comparing  $(i, j)$ -entries in the above  $s \times t$  matrix equation for  $1 \leq i, j \leq r$  gives  $p_{ij}(x)d_j(x) = d_i(x)q_{ij}(x)$ .

Comparing  $(i, j)$ -entries for  $1 \leq j \leq r < i \leq s$  gives  $p_{ij}(x)d_j(x) = 0$  and so  $p_{ij}(x) = 0$ . Comparing  $(i, l)$ -entries for  $1 \leq i \leq r < l \leq t$  gives  $0 = d_i(x)q_{il}(x)$  and so  $q_{il}(x) = 0$ .

Conversely suppose  $(P(x), Q(x)) \in G$  satisfies  $p_{ij}(x)d_j(x) = d_i(x)q_{ij}(x)$  for  $1 \leq i, j \leq r$ ,  $p_{ij}(x) = 0$  for  $1 \leq j \leq r < i \leq s$  and  $q_{il}(x) = 0$  for  $1 \leq i \leq r < l \leq t$ . As all  $(i, l)$ -entries in  $P(x)D(x)$  and  $D(x)Q(x)$  are zero for  $r < i \leq s$ ,  $r < l \leq t$  we conclude  $P(x)D(x) = D(x)Q(x)$ .

### Solution 6

(a) Denote the rows of the  $s \times s$  identity matrix over  $F[x]$  by  $e_1(x), e_2(x), \dots, e_s(x)$ . So  $e_k(x)A(x)$  is row  $k$  of  $A(x)$  for  $1 \leq k \leq s$ .

Consider the *ero*  $r_i \leftrightarrow r_j$  over  $F[x]$  where  $1 \leq i < j \leq s$  and let  $P_1(x)$  be the corresponding elementary matrix over  $F[x]$ . Then

$$e_i(x)P_1(x) = e_j(x), e_j(x)P_1(x) = e_i(x), e_k(x)P_1(x) = e_k(x) \text{ for all } k \neq i, j \text{ where } 1 \leq k \leq s.$$

Post multiplying (multiplying on the right) each of these row equations by  $A(x)$  gives

$$e_i(x)P_1(x)A(x) = e_j(x)A(x), e_j(x)P_1(x)A(x) = e_i(x)A(x), e_k(x)P_1(x)A(x) = e_k(x)A(x)$$

which tell us that  $P_1(x)A(x)$  is the result of applying  $r_i \leftrightarrow r_j$  to  $A(x)$ .

Consider next the *ero*  $ar_i$  over  $F[x]$  and its corresponding elementary matrix  $P_i(x)$  over  $F[x]$  where  $a \in F^*$ . Then  $e_i(x)P_i(x) = ae_i(x)$  and  $e_j(x)P_i(x) = e_j(x)$  for all  $j \neq i$ . Postmultiplying by  $A(x)$  as above gives  $e_i(x)P_i(x)A(x) = ae_i(x)A(x)$  and  $e_j(x)P_i(x)A(x) = e_j(x)A(x)$ . These equations say that  $P_i(x)A(x)$  is the matrix which results on applying  $ar_i$  to  $A(x)$ .

Now consider the *ero*  $r_i + f(x)r_j$  over  $F[x]$  where  $i \neq j$ ,  $1 \leq i, j \leq s$  and  $f(x) \in F[x]$ . Let  $P_1(x)$  be its corresponding elementary matrix over  $F[x]$ . Then  $e_i(x)P_1(x) = e_i(x) + f(x)e_j(x)$  and  $e_k(x)P_1(x) = e_k(x)$  for all  $k \neq i$ ,  $1 \leq k \leq s$ . Postmultiplying these row equations by  $A(x)$  gives  $e_i(x)P_1(x)A(x) = (e_i(x) + f(x)e_j(x))A(x) = e_i(x)A(x) + f(x)e_j(x)A(x)$  and  $e_k(x)P_1(x)A(x) = e_k(x)A(x)$ . So  $P_1(x)A(x)$  is the result of applying  $r_i + f(x)r_j$  to  $A(x)$ .

Let  $e_1(x)^T, e_2(x)^T, \dots, e_t(x)^T$  denote the columns of the  $t \times t$  identity matrix over  $F[x]$ . So  $A(x)e_k(x)^T$  is column  $k$  of  $A(x)$  for  $1 \leq k \leq t$ . The part of (4.10) concerning *ecos* is proved by using the matrix transpose of the above theory. Let  $Q_1(x)$  be the elementary matrix corresponding to the *eco*  $c_i \leftrightarrow c_j$  over  $F[x]$ . Then  $Q_1(x)e_i(x)^T = e_j(x)^T$ ,  $Q_1(x)e_j(x)^T = e_i(x)^T$  and  $Q_1(x)e_k(x)^T = e_k(x)^T$  for  $k \neq i, j$ . Premultiplying these column equations by  $A(x)$  gives  $A(x)Q_1(x)e_i(x)^T = A(x)e_j(x)^T$ ,  $A(x)Q_1(x)e_j(x)^T = A(x)e_i(x)^T$  and  $A(x)Q_1(x)e_k(x)^T = A(x)e_k(x)^T$  for  $k \neq i, j$ . These  $t$  column equations say that  $A(x)Q_1(x)$  is the matrix which results on applying the *eco*  $c_i \leftrightarrow c_j$  to  $A(x)$ . The remaining types of *eco* can be dealt with in the same way.

(b) Let  $a_{ij}(x)$  denote the  $(i, j)$ -entry in  $A(x)$ . Suppose  $a_{1j}(x) \neq 0$  for some  $j$  with  $1 < j \leq t$ . Write  $h_j(x) = \gcd\{a_{11}(x), a_{12}(x), \dots, a_{1j}(x)\}$  for  $1 \leq j \leq t$ . Then  $h_1(x)$  is  $a_{11}(x)$  made monic (divided by its leading coefficient) in case  $a_{11}(x) \neq 0$  and  $h_1(x) = 0$  otherwise. Also  $h_{j+1}(x) = \gcd\{h_j(x), a_{1j+1}(x)\}$  for  $1 \leq j < t$ . In case  $a_{11}(x) \neq h_1(x)$  applying the *eco*  $(a_{11}(x)/h_1(x))c_1$  to  $A(x)$  produces  $A_1(x)$  with  $e_1(x)A_1(x) = (h_1(x), a_{12}(x), a_{13}(x), \dots, a_{1t}(x))$ . Suppose inductively that there is a sequence of *ecos* over  $F[x]$  which applied to  $A(x)$  produces  $A_j(x)$  with first row

$$e_1(x)A_j(x) = (h_j(x), 0, \dots, 0, a_{1j+1}(x), a_{1j+2}(x), \dots, a_{1t}(x))$$

where  $1 \leq j < t$ . Applying (4.13)(i) to the  $1 \times 2$  submatrix  $(h_j(x), a_{1j+1}(x))$  of  $A_j(x)$  gives a sequence of *ecos* over  $F[x]$  which changes  $(h_j(x), a_{1j+1}(x))$  into  $(h_{j+1}(x), 0)$ . Applying this sequence to cols 1 and  $j+1$  of  $A_j(x)$  and using the inductive hypothesis we obtain a sequence of *ecos* over  $F[x]$  which changes  $A(x)$  into  $A_{j+1}(x)$  with  $e_1(x)A_{j+1}(x) = (h_{j+1}(x), 0, \dots, 0, a_{1j+2}(x), a_{1j+3}(x), \dots, a_{1t}(x))$ .

The induction is now complete and ends with an  $s \times t$  matrix  $A_t(x)$  obtained from  $A(x)$  by means of a sequence of *ecos* over  $F[x]$  such that  $e_1(x)A_t(x) = (h_t(x), 0, 0, \dots, 0)$ . Write  $B_1(x) = A_t(x)$  and  $b_1(x) = h_t(x)$ .

In the case  $a_{1j}(x) = 0$  for all  $j$  with  $1 < j \leq t$  we let  $B_1(x) = A(x)$ ,  $b_1(x) = a_{11}(x)$ .

Should all  $(i, 1)$ -entries in  $B_1(x)$  be zero for  $2 \leq i \leq s$ , then the algorithm ends with  $B_1(x) = B(x)$ .

Otherwise let  $b_2(x)$  denote the gcd of the entries in col 1 of  $B_1(x)$ . Then  $b_2(x) \mid b_1(x)$ . Notice that  $b_2(x) = 0$  implies  $b_1(x) = 0$ , i.e. all entries in row 1 and col 1 of  $A(x)$  are zero and so  $A(x) = B(x)$ .

We may therefore assume  $b_2(x) \neq 0$ . The above technique, but matrix-transposed, is then applied to  $B_1(x)$ : there is an  $s \times t$  matrix  $B_2(x)$  which arises by applying a sequence of *eros* over  $F[x]$  to  $B_1(x)$  where  $B_2(x)e_1(x)^T = (b_2(x), 0, 0, \dots, 0)^T$ . Should  $b_2(x) = b_1(x)$  then  $e_1(x)B_2(x) = e_1(x)B_1(x)$  as these *eros* leave row 1 unchanged. So the algorithm ends with  $B_2(x) = B(x)$ .

Suppose  $B_2(x) \neq B(x)$ , i.e.  $B_2(x)$  has a non-zero  $(1, j)$ -entry for  $1 < j \leq t$ . In this case  $b_2(x) \neq b_1(x)$  and the process is restarted with  $B_2(x)$  in place of  $A(x)$ . We obtain a sequence of  $s \times t$  matrices  $B_k(x)$

over  $F[x]$  with monic  $(1,1)$ -entries  $b_k(x)$  for  $k=2,3,\dots$ . Further for  $k$  odd  $e_1(x)B_k(x)=(b_k(x),0,\dots,0)$  and  $B_k(x)$  is obtained from  $B_{k-1}(x)$  by applying *ecos* over  $F[x]$ , while for  $k$  even  $B_k(x)e_1(x)^T=(b_k(x),0,\dots,0)^T$  and  $B_k(x)$  is obtained from  $B_{k-1}(x)$  by applying *eros* over  $F[x]$ . Therefore  $b_k(x)|b_{k-1}(x)$  and the process continues provided  $B_k(x) \neq B(x)$ , i.e.  $B_k(x)$  has a non-zero off-diagonal entry in either row 1 or col 1. So  $b_k(x) \neq b_{k-1}(x)$  and  $\deg b_k(x) < \deg b_{k-1}(x)$ . Adding the  $k-2$  inequalities  $\deg b_{i-1}(x) - \deg b_i(x) \geq 1$  for  $i=3,4,\dots,k$  gives  $\deg b_2(x) - \deg b_k(x) \geq k-2$  and so  $k \leq \deg b_2(x) + 2$  as  $\deg b_k(x) \geq 0$ . Let  $l$  be the largest of the integers  $k$ . Then either  $B_l(x)=B(x)$  or  $b_l(x)=b_{l+1}(x)$  where  $l \leq \deg b_2(x) + 3$ . Suppose  $l$  is odd. Then  $e_1(x)B_l(x)=(b_l(x),0,\dots,0)$  and  $b_l(x)$  is a divisor of all entries in col 1 of  $B_l(x)$ . Hence the process terminates with  $B_{l+1}(x)=B(x)$  as none of the *eros* over  $F[x]$  which change  $B_l(x)$  into  $B_{l+1}(x)$  alter row 1. Suppose  $l$  is even. Then  $B_l(x)e_1(x)^T=(b_l(x),0,\dots,0)^T$  and  $b_l(x)$  is a divisor of all entries in row 1 of  $B_l(x)$ . Hence the process terminates with  $B_{l+1}(x)=B(x)$  as none of the *ecos* over  $F[x]$  which change  $B_l(x)$  into  $B_{l+1}(x)$  alter col 1.

(c) We use induction on  $\min\{s,t\}$ . If  $\min\{s,t\}=1$  the matrix  $B(x)$  of (4.14) is in Smith normal form. So suppose  $\min\{s,t\} > 1$ . Then

$$B = \left( \begin{array}{c|c} b_{11}(x) & 0 \\ \hline 0 & B'(x) \end{array} \right)$$

where  $B'(x)$  is an  $(s-1) \times (t-1)$  matrix over  $F[x]$ . By induction there is a finite sequence of elementary operations over  $F[x]$  which changes  $B'(x)$  into  $D'(x) = \text{diag}(d'_2(x), d'_3(x), \dots)$  in Smith normal form. Hence there is a finite sequence of elementary operations over  $F[x]$  which changes  $A(x)$  into  $D_1(x) = \text{diag}(b_{11}(x), d'_2(x), d'_3(x), \dots)$ . As in the proof of (1.11) apply (4.15) to the  $2 \times 2$  submatrix  $\text{diag}(b_{11}(x), d'_2(x))$  of  $D_1(x)$ : the elementary operations involved change  $D_1(x)$  into  $D_2(x) = \text{diag}(d_1(x), m_2(x), d'_3(x), d'_4(x), \dots)$  where  $d_1(x) = \gcd\{b_{11}(x), d'_2(x)\}$  and  $m_2(x) = \text{lcm}\{b_{11}(x), d'_2(x)\}$ . Then  $d_1(x) | m_2(x)$  and  $d_1(x) | d'_i(x)$  for  $1 < i \leq \min\{s,t\}$ . So  $d_1(x)$  is a divisor of all entries in the  $(s-1) \times (t-1)$  matrix  $D'_2(x) = \text{diag}(m_2(x), d'_3(x), d'_4(x), \dots)$ . By inductive hypothesis there is a finite sequence of elementary operations over  $F[x]$  which changes  $D'_2(x)$  into  $D'(x) = \text{diag}(d_2(x), d_3(x), \dots)$  in Smith normal form. As  $d_1(x)$  is a divisor of all the entries in  $D'(x)$  in particular we have  $d_1(x) | d_2(x)$ . So  $D(x) = \text{diag}(d_1(x), d_2(x), d_3(x), \dots)$  is in Smith normal form and there is a finite sequence of elementary operations over  $F[x]$  which changes  $D_2(x)$  into  $D(x)$  without changing row 1 and col 1. As  $A(x)$  can be changed in this way into  $D_1(x)$ ,  $D_1(x)$  into  $D_2(x)$  and  $D_2(x)$  into  $D(x)$ , we see that  $A(x)$  can be changed into  $D(x) = S(A(x))$  by a finite sequence of elementary operations over  $F[x]$ . The induction is now complete.

Denote by  $P_1(x), P_2(x), \dots, P_u(x)$  the elementary matrices corresponding to the *eros* over  $F[x]$  used in the reduction of  $A(x)$  to  $S(A(x))$ . Denote by  $Q_1(x), Q_2(x), \dots, Q_v(x)$  the elementary matrices corresponding to the *ecos* over  $F[x]$  used in the reduction of  $A(x)$  to a matrix  $S(A(x))$  in Smith normal form. Using induction on  $u+v$  and (4.10) we obtain

$$P_u(x) \cdots P_2(x) P_1(x) A(x) Q_1(x) Q_2(x) \cdots Q_v(x) = S(A(x)).$$

Write  $P(x) = P_u(x) \cdots P_2(x)P_1(x)$  and  $Q(x) = Q_v(x)^{-1} \cdots Q_2(x)^{-1}Q_1(x)^{-1}$ . Then  $P(x)$  and  $Q(x)$  are invertible over  $F[x]$  and satisfy

$$P(x)A(x)Q(x)^{-1} = S(A(x)).$$

(d) By (4.16) there are  $s \times s$  matrices  $P'(x) = P_u(x) \cdots P_2(x)P_1(x)$  and

$Q'(x) = Q_v(x)^{-1} \cdots Q_2(x)^{-1}Q_1(x)^{-1}$ , both being products of elementary matrices over  $F[x]$ , satisfying  $P'(x)P(x)Q'(x)^{-1} = S(P(x)) = \text{diag}(d_1(x), d_2(x), \dots, d_s(x))$ . As  $P'(x), P(x)$  and  $Q'(x)$  are invertible over  $F[x]$ , so also is  $P'(x)P(x)Q'(x)^{-1}$  and hence each  $d_i(x)$  is an invertible polynomial over  $F$ . So  $d_i(x) = 1$  as  $d_i(x)$  is monic. Therefore  $S(P(x)) = I$ , the  $s \times s$  identity matrix over  $F[x]$ .

The equation  $P'(x)P(x)Q'(x)^{-1} = I$  now gives

$$P(x) = P'(x)^{-1}Q'(x) = P_1(x)^{-1}P_2(x)^{-1} \cdots P_u(x)^{-1}Q_v(x)^{-1} \cdots Q_2(x)^{-1}Q_1(x)^{-1}$$

which expresses  $P(x)$  as a product of elementary matrices over  $F[x]$ . Therefore  $P^{-1}(x)P(x) = I$  can be expressed as  $Q_1(x)Q_2(x) \cdots Q_v(x)P_u(x) \cdots P_2(x)P_1(x)P(x) = I$  which shows by (4.10) that  $P(x)$  is reducible to  $I$  using only *eros* over  $F[x]$ . In the same way  $P(x)P^{-1}(x) = I$  can be written as  $P(x)Q_1(x)Q_2(x) \cdots Q_v(x)P_u(x) \cdots P_2(x)P_1(x) = I$  which shows by (4.10) that  $P(x)$  is reducible to  $I$  by means of *ecos* over  $F[x]$ .

(e) Suppose  $A(x) \equiv B(x)$ . By (4.11) there are invertible matrices  $P(x)$  and  $Q(x)$  over  $F[x]$  with  $P(x)A(x)Q(x)^{-1} = B(x)$ . Then  $g_l(A(x)) \mid g_l(P(x)A(x))$  by (4.18) and

$g_l(P(x)A(x)) \mid g_l(P(x)A(x)Q^{-1}(x))$ , i.e.  $g_l(P(x)A(x)) \mid g_l(B(x))$ , also by (4.18). Therefore  $g_l(A(x)) \mid g_l(B(x))$ . As  $B(x) \equiv A(x)$  the roles of  $A(x)$  and  $B(x)$  can be interchanged in the above to give  $g_l(B(x)) \mid g_l(A(x))$ . Hence  $g_l(A(x)) = g_l(B(x))$  as, being gcds, these polynomials are monic or zero (the polynomial analogue of non-negative integers).

By (4.16) there is an  $s \times t$  matrix  $D(x) = \text{diag}(d_1(x), d_2(x), \dots, d_{\min\{s,t\}}(x))$  in Smith normal form (4.12) with  $A(x) \equiv D(x)$ . Selecting the first  $l$  rows and first  $l$  columns of  $D(x)$  we obtain an  $l$ -minor of  $D(x)$  equal  $d_1(x)d_2(x) \cdots d_l(x)$ . Now every  $l$ -minor of  $D(x)$  is either zero or a product  $d_{j_1}(x)d_{j_2}(x) \cdots d_{j_l}(x)$  where  $1 \leq j_1 < j_2 < \cdots < j_l \leq \min\{s,t\}$ . As  $i \leq j_i$  we see  $d_i(x) \mid d_{j_i}(x)$  for  $1 \leq i \leq l$  and hence  $d_1(x)d_2(x) \cdots d_l(x) \mid d_{j_1}(x)d_{j_2}(x) \cdots d_{j_l}(x)$ . So

$$g_l(A(x)) = g_l(D(x)) = d_1(x)d_2(x) \cdots d_l(x) \text{ for } 1 \leq l \leq \min\{s,t\}.$$

These equations determine the invariant factors  $d_1(x), d_2(x), \dots, d_{\min\{s,t\}}(x)$  of  $A(x)$  in terms of the polynomials  $g_l(A(x))$  for  $1 \leq l \leq \min\{s,t\}$  using induction on  $l$ : first  $d_1(x) = g_1(A(x))$  and for  $l > 1$  we have  $d_l(x) = g_l(A(x))/g_{l-1}(A(x))$  for  $g_l(A(x)) \neq 0$ ,  $d_l(x) = 0$  for  $g_l(A(x)) = 0$ . Therefore  $A(x)$  is equivalent to a unique matrix  $D(x)$  in Smith normal form and it is legitimate to write

$$S(A(x)) = D(x) \text{ and } d_l(x) = d_l(A(x)) \text{ for } 1 \leq l \leq \min\{s,t\} \text{ as in (4.20).}$$

Suppose  $g_l(A(x)) = g_l(B(x))$  for  $1 \leq l \leq \min\{s,t\}$ . By (4.16) and the theory in the above paragraph,  $A(x)$  and  $B(x)$  are equivalent to the same matrix  $D(x)$  in Smith normal form. Hence  $A(x)$  and  $B(x)$  are equivalent to each other.

(f) By (4.16) there are invertible matrices  $P_1(x), P_2(x), Q_1(x), Q_2(x)$  over  $F[x]$  satisfying  $P_1(x)A(x)Q_1(x)^{-1} = S(A(x))$  and  $P_2(x)A(x)B(x)Q_2(x)^{-1} = S(A(x)B(x))$ . Hence  $P(x)S(A(x)B(x)) = S(A(x))C(x)$  where the  $r \times r$  matrix  $P(x) = P_1(x)P_2(x)^{-1} = (p_{ij}(x))$  is invertible over  $F[x]$  and  $C(x) = Q_1(x)B(x)Q_2(x)^{-1} = C(c_{ij}(x))$  is an  $s \times t$  matrix over  $F[x]$  equivalent to  $B(x)$ . Comparing  $(i, j)$ -entries for  $1 \leq i, j \leq r$  in the  $r \times t$  matrix equation  $P(x)S(A(x)B(x)) = S(A(x))C(x)$  gives  $p_{ij}(x)d_j(A(x)B(x)) = d_i(A(x))c_{ij}(x)$  (equation 1). Consider  $i$  and  $j$  satisfying

$1 \leq j \leq k \leq i \leq r$ . Multiplying equation 1 by the monic polynomial  $d_k(A(x)B(x))/d_j(A(x)B(x))$  gives

$p_{ij}(x)d_k(A(x)B(x)) = d_k(A(x))c'_{ij}(x)$  (equation 2) where

$c'_{ij}(x) = (d_k(A(x)B(x))/d_j(A(x)B(x)))(d_i(A(x))/d_k(A(x)))c_{ij}(x)$  is a polynomial over  $F$ .

Dividing equation 2 by  $\bar{d}_k(x) = \gcd\{d_k(A(x)), d_k(A(x)B(x))\}$  gives

$p_{ij}(x)(d_k(A(x)B(x))/\bar{d}_k(x)) = (d_k(A(x))/\bar{d}_k(x))c'_{ij}(x)$  (equation 3). As the monic polynomials

$d_k(A(x)B(x))/\bar{d}_k(x)$  and  $d_k(A(x))/\bar{d}_k(x)$  are coprime (their gcd is 1) we see that  $d_k(A(x))/\bar{d}_k(x)$

is a divisor of  $p_{ij}(x)$  for  $1 \leq j \leq k \leq i \leq r$ . On partitioning  $P(x) = \begin{pmatrix} P_{11}(x) & P_{12}(x) \\ P_{21}(x) & P_{22}(x) \end{pmatrix}$  as shown, where

$P_{11}(x)$  is the leading  $(k-1) \times (k-1)$  submatrix of  $P(x)$ , we see that all entries in  $P_{21}(x)$  are divisible by  $d_k(A(x))/\bar{d}_k(x)$  as are all entries in column 1 of  $P_{22}(x)$ . Therefore

$$\det P(x) = \begin{vmatrix} P_{11}(x) & P_{12}(x) \\ P_{21}(x) & P_{22}(x) \end{vmatrix} \equiv \begin{vmatrix} P_{11}(x) & P_{12}(x) \\ 0 & P_{22}(x) \end{vmatrix} = |P_{11}(x)| \times |P_{22}(x)| \equiv |P_{11}(x)| \times 0 = 0$$

where  $\equiv$  denotes congruence modulo  $d_k(A(x))/\bar{d}_k(x)$ . So  $d_k(A(x))/\bar{d}_k(x)$  is a monic divisor of  $\det P(x)$  which is a polynomial of degree 0 as  $P(x)$  is invertible over  $F[x]$ . Hence

$d_k(A(x))/\bar{d}_k(x) = 1$  showing  $d_k(A(x)) = \bar{d}_k(x) = \gcd\{d_k(A(x)), d_k(A(x)B(x))\}$ , i.e.  $d_k(A(x))$  is a divisor of  $d_k(A(x)B(x))$ .

There are invertible matrices  $P_2(x), P_3(x), Q_2(x), Q_3(x)$  over  $F[x]$  satisfying

$P_2(x)A(x)B(x)Q_2(x)^{-1} = S(A(x)B(x))$  and  $P_3(x)B(x)Q_3(x)^{-1} = S(B(x))$ . Hence

$S(A(x)B(x))Q(x) = E(x)S(B(x))$  where  $Q_2(x)Q_3(x)^{-1} = Q(x) = (q_{ij}(x))$  is a  $t \times t$  invertible matrix over  $F[x]$  and  $P_2(x)A(x)P_3(x)^{-1} = E(x) = (e_{ij}(x))$  is an  $r \times s$  matrix over  $F[x]$  equivalent to  $A(x)$ .

Comparing  $(i, j)$ -entries in the  $r \times t$  matrix equation  $S(A(x)B(x))Q(x) = E(x)S(B(x))$  gives

$d_i(A(x)B(x))q_{ij}(x) = e_{ij}(x)d_j(B(x))$  for  $1 \leq i \leq r, 1 \leq j \leq s$  (equation 4) and

$d_i(A(x)B(x))q_{ij}(x) = 0$  for  $1 \leq i \leq r, s < j \leq t$ . From the last equation we deduce  $q_{ij}(x) = 0$  for

$1 \leq i \leq r, s < j \leq t$  (equation 5) as  $d_i(A(x)B(x)) \neq 0$ . Consider  $q_{ij}(x)$  for  $1 \leq i \leq k \leq j \leq s$ . Multiply

equation 4 by the monic polynomial  $d_k(A(x)B(x))/d_i(A(x)B(x))$  obtaining

$d_k(A(x)B(x))q_{ij}(x) = e'_{ij}(x)d_k(B(x))$  for  $1 \leq i \leq k \leq j \leq s$  (equation 6) where

$e'_{ij}(x) = (d_k(A(x)B(x))/d_i(A(x)B(x)))(d_j(B(x))/d_k(B(x)))e_{ij}(x)$  is a polynomial over  $F$ .

Dividing equation 6 by  $\bar{d}_k(x) = \gcd\{d_k(B(x)), d_k(A(x)B(x))\}$  we see that

$q_{ij}(x)$  is divisible by  $d_k(B(x))/\bar{\bar{d}}_k(x)$  for  $1 \leq i \leq k \leq j \leq s$ . Partition the  $t \times t$  matrix  $Q = \left( \begin{array}{c|c} Q_{11} & Q_{12} \\ \hline Q_{21} & Q_{22} \end{array} \right)$

where  $Q_{11}$  is the leading  $(k-1) \times (k-1)$  submatrix of  $Q$ . By equation 5 and the foregoing divisibility condition we see that all entries in  $Q_{12}$  and all entries in row 1 of  $Q_{22}$  are divisible by  $d_k(B(x))/\bar{\bar{d}}_k(x)$ . Therefore

$$\det Q(x) = \left| \begin{array}{c|c} Q_{11}(x) & Q_{12}(x) \\ \hline Q_{21}(x) & Q_{22}(x) \end{array} \right| \equiv \left| \begin{array}{c|c} Q_{11}(x) & 0 \\ \hline Q_{21}(x) & Q_{22}(x) \end{array} \right| = |Q_{11}(x)| \times |Q_{22}(x)| \equiv |Q_{11}(x)| \times 0 = 0$$

where  $\equiv$  denotes congruence modulo  $d_k(B(x))/\bar{\bar{d}}_k(x)$ . As  $Q(x)$  is invertible over  $F[x]$  we see that the monic polynomial  $d_k(B(x))/\bar{\bar{d}}_k(x)$  is a divisor of the polynomial  $\det Q(x)$  of degree 0 over  $F$ . Hence  $d_k(B(x))/\bar{\bar{d}}_k(x) = 1$ , i.e.  $d_k(B(x)) = \bar{\bar{d}}_k(x) = \gcd\{d_k(B(x)), d_k(A(x)B(x))\}$  showing that  $d_k(B(x))$  is a divisor of  $d_k(A(x)B(x))$ .

## Solutions 5.1 (page 222)

### Solution 1

(a) As  $X^{-1} = \begin{pmatrix} 3 & -1 \\ -5 & 2 \end{pmatrix}$  we obtain  $B = \begin{pmatrix} 2 & 1 \\ 5 & 3 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} \begin{pmatrix} 3 & -1 \\ -5 & 2 \end{pmatrix} = \begin{pmatrix} -38 & 16 \\ -102 & 43 \end{pmatrix}$ .

$\det A = 1 \times 4 - 2 \times 3 = -2 = (-38) \times 43 - (-102) \times 16 = \det B$  and

$\text{trace } A = 1 + 4 = 5 = -38 + 43 = \text{trace } B$ . As  $X^2 A X^{-2} \sim A$ , by (5.5) the characteristic polynomials of  $X^2 A X^{-2}$  and  $A$  are equal namely

$$x^2 - (\text{trace } A)x + \det A = x^2 - 5x - 2.$$

(b) As  $\text{trace } A = \text{trace } B$  we have  $3 - a + a = b + 2$  giving  $b = 1$ . As  $\det A = \det B$  we see  $(3 - a)a - (-1) \times 7 = 1 \times 2 - (-1) \times 1$  giving  $a^2 - 3a - 4 = 0$ , i.e.  $(a - 4)(a + 1) = 0$ .

So either  $a = 4$  or  $a = -1$ . Incidentally

$$\begin{pmatrix} -1 & 7 \\ -1 & 4 \end{pmatrix}, \begin{pmatrix} 4 & 7 \\ -1 & -1 \end{pmatrix} \text{ and } \begin{pmatrix} 1 & 1 \\ -1 & 2 \end{pmatrix} \text{ are similar to } \begin{pmatrix} 0 & 1 \\ -3 & 3 \end{pmatrix} = C(x^2 - 3x + 3).$$

(c)

$$XAX^{-1} = \begin{pmatrix} 1 & -a/2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & a/2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & a/2 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & a/2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & a \\ 0 & -1 \end{pmatrix}.$$

All these matrices belong to the similarity class of  $A$  for all  $a \in \mathbb{R}$ , so the answer is:

Yes. As  $\chi_B(x) = x^2 - 1 = (x - 1)(x + 1)$  we see that  $B$  has distinct eigenvalues  $\pm 1$  and

so  $B$  can be diagonalised, i.e.  $B \sim \text{diag}(1, -1) = A$ . Explicitly  $(1 \pm \sqrt{2}, 1)$  are row

eigenvectors of  $B$  corresponding to  $\pm 1$ . So  $X = \begin{pmatrix} 1 + \sqrt{2} & 1 \\ 1 - \sqrt{2} & 1 \end{pmatrix}$  is invertible over  $\mathbb{R}$  and

satisfies  $XBX^{-1} = A$ .

(d) Consider  $A, A', A'' \in \mathfrak{M}_t(F)$ . Taking  $X = I$  the  $t \times t$  identity matrix over  $F$ , then  $X^{-1} = I$  and so  $XAX^{-1} = A$ . By (5.3) the *reflexive law*  $A \sim A$  for all  $A \in \mathfrak{M}_t(F)$  is obeyed.

Suppose  $A \sim A'$ . There is an invertible  $t \times t$  matrix  $X$  over  $F$  with  $XAX^{-1} = A'$ .

Hence  $X^{-1}A'X = A$ . As  $X^{-1}$  is invertible over  $F$  and  $(X^{-1})^{-1} = X$  we see  $A' \sim A$  by (5.3) showing that the *symmetric law* is obeyed.

Suppose  $A \sim A'$  and  $A' \sim A''$ . By (5.3) there are invertible  $t \times t$  matrices  $X$  and  $Y$  over  $F$  with  $XAX^{-1} = A'$  and  $YA'Y^{-1} = A''$ . Then

$(YX)A(YX)^{-1} = YXAX^{-1}Y^{-1} = YA'Y^{-1} = A''$  showing that  $A \sim A''$  by (5.3) as  $YX$  is invertible over  $F$ . So the *transitive law* is obeyed.

We conclude:  $\sim$  is an equivalence relation on  $\mathfrak{M}_t(F)$ .

(e) By hypothesis  $XAX^{-1} = A$  for all  $X \in GL_2(F)$ . So  $XA = AX$  for all

$X \in GL_2(F)$ . Write  $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$  and  $X_1 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ . Then

$X_1 \in GL_2(F)$  and so  $X_1 A = A X_1$  which gives  $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} A = A \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ , i.e.



$$\begin{pmatrix} c & d \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & a \\ 0 & c \end{pmatrix} \text{ i.e. } c=0, a=d. \text{ In the same way } X_2 A = A X_2 \text{ where } X_2 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

leads to  $b=0, a=c$ . So  $A = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} = aI$  is a scalar matrix.

### Solution 2

(a) Write  $xI - A = (b_{ij})$ . The entries in  $xI - A$  are  $x - a_{ii} = b_{ii}$  of degree 1 over  $R$  and  $-a_{ij} = b_{ij}$  for  $i \neq j$  is a constant polynomial over  $R$ . So

$(\text{sign } \pi) b_{1(1)\pi} b_{2(2)\pi} \cdots b_{t(t)\pi}$  has degree  $< t-1$  over  $R$  unless  $(i)\pi = i$  for at least  $t-1$  integers  $i \in \{1, 2, \dots, t\}$ . In this case  $(i)\pi = i$  for all  $i \in \{1, 2, \dots, t\}$  as  $\pi$  is a permutation of  $\{1, 2, \dots, t\}$ , i.e.  $\pi = t$ , the identity. Therefore the coefficient of  $x^{t-1}$  in  $\chi_A(x)$  is the coefficient of  $x^{t-1}$  in  $(\text{sign } t) b_{11} b_{22} \cdots b_{tt} = (x - a_{11})(x - a_{22}) \cdots (x - a_{tt})$  which is

$$-a_{11} - a_{22} - \dots - a_{tt} = -\text{trace } A.$$

The constant term in  $\chi_A(x)$  is  $\chi_A(0) = (\chi_A(x))\varepsilon_0$  where  $\varepsilon_0 : R[x] \rightarrow R$  is the evaluation at 0 ring homomorphism. As  $\det(xI - A) \in R[x]$  we see

$(\chi_A(x))\varepsilon_0 = (\det(xI - A))\varepsilon_0 = \det(0I - A) = \det(-A) = (-1)^t \det A$  since  $-A$  is the result of changing the sign of each of the  $t$  rows of  $A$ .

(b) Consider a permutation  $\pi$  of  $\{1, \dots, t_1, \dots, t_1 + t_2\}$  and suppose first there is  $i \in \{t_1 + 1, \dots, t_1 + t_2\}$  with  $(i)\pi \notin \{t_1 + 1, \dots, t_1 + t_2\}$ . The  $(i, (i)\pi)$ -entry in  $A$  is zero and so  $(\text{sign } \pi) a_{1(1)\pi} a_{2(2)\pi} \cdots a_{t(t)\pi} = 0$ .

Now suppose  $\pi$  satisfies  $i \in \{t_1 + 1, \dots, t_1 + t_2\} \Rightarrow (i)\pi \in \{t_1 + 1, \dots, t_1 + t_2\}$ . Then  $i \in \{1, \dots, t_1\} \Rightarrow (i)\pi \in \{1, \dots, t_1\}$  as  $\pi$  is a permutation. Let  $\pi_1$  be the restriction of  $\pi$  to  $\{1, \dots, t_1\}$ ; then  $\pi_1$  is a permutation of  $\{1, \dots, t_1\}$ . Let  $\pi_2$  be the restriction of  $\pi$  to  $\{t_1 + 1, \dots, t_1 + t_2\}$ ; then  $\pi_2$  is a permutation of  $\{t_1 + 1, \dots, t_1 + t_2\}$ . Also  $\text{sign } \pi = (\text{sign } \pi_1)(\text{sign } \pi_2)$ . Conversely each pair of permutations  $\pi_1, \pi_2$  as above arises from a unique permutation  $\pi$ . Therefore

$$\det A = \sum_{\pi} (\text{sign } \pi) a_{1(1)\pi} a_{2(2)\pi} \cdots a_{t(t)\pi}$$

where the summation is restricted to those permutations  $\pi$  satisfying

$i \in \{t_1 + 1, \dots, t_1 + t_2\} \Rightarrow (i)\pi \in \{t_1 + 1, \dots, t_1 + t_2\}$  as all other terms are zero. Hence

$$\begin{aligned} \det A &= \sum_{\pi_1, \pi_2} (\text{sign } \pi_1)(\text{sign } \pi_2) a_{1(1)\pi_2} \cdots a_{t_1(t_1)\pi_1} a_{t_1+1(t_1+1)\pi_2} \cdots a_{t_2(t_2)\pi_2} = \\ &= \left( \sum_{\pi_1} (\text{sign } \pi_1) a_{1(1)\pi_1} \cdots a_{t_1(t_1)\pi_1} \right) \left( \sum_{\pi_2} (\text{sign } \pi_2) a_{t_1+1(t_1+1)\pi_2} \cdots a_{t_2(t_2)\pi_2} \right) = (\det A_1)(\det A_2). \end{aligned}$$

Taking  $B = 0$ , the  $t_1 \times t_2$  zero matrix, we obtain  $A = A_1 \oplus A_2$ . So

$$\det(A_1 \oplus A_2) = (\det A_1)(\det A_2).$$

(c) Let  $X = \begin{pmatrix} 0 & I_{t_2} \\ I_{t_1} & 0 \end{pmatrix}$ , that is,  $X$  is a partitioned  $(t_2 + t_1) \times (t_1 + t_2)$  matrix with

entries the  $t_1 \times t_1$  identity matrix  $I_{t_1}$  and the  $t_2 \times t_2$  matrix  $I_{t_2}$  as indicated together with rectangular shaped zero matrices. Then  $\det X = (-1)^{t_1 t_2}$  as  $t_1 t_2$  interchanges of

consecutive rows changes  $X$  into  $I_t$ . So  $X$  is invertible over  $F$ . Also

$$X(A_1 \oplus A_2) = \begin{pmatrix} 0 & I_{t_2} \\ I_{t_1} & 0 \end{pmatrix} \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix} = \begin{pmatrix} 0 & A_2 \\ A_1 & 0 \end{pmatrix} = \begin{pmatrix} A_2 & 0 \\ 0 & A_1 \end{pmatrix} \begin{pmatrix} 0 & I_{t_2} \\ I_{t_1} & 0 \end{pmatrix} = (A_2 \oplus A_1)X.$$

Therefore  $X(A_1 \oplus A_2)X^{-1} = A_2 \oplus A_1$  showing  $A_1 \oplus A_2 \sim A_2 \oplus A_1$ .

Suppose  $A \oplus B = B \oplus A$ . Comparing entries in this  $5 \times 5$  matrix equation gives

$B = A \oplus C = C \oplus A$  where  $C$  is the leading  $1 \times 1$  submatrix of  $A$ . Comparing entries in the  $3 \times 3$  matrix equation  $A \oplus C = C \oplus A$  gives  $A = C \oplus C$  and so  $B = C \oplus C \oplus C$ . So  $A \oplus B = C \oplus C \oplus C \oplus C \oplus C$  is a scalar matrix.

Suppose  $A_1 \oplus A_2 = A_2 \oplus A_1$  and let  $t_1 + t_2$  be the least integer such that there is no  $t \times t$  matrix  $A$  as stated. Then  $t_1 \neq t_2$ , since  $A = A_1 = A_2$  on comparing entries in

$A_1 \oplus A_2 = A_2 \oplus A_1$  should  $t_1 = t_2$ . We may suppose  $t_1 > t_2$  by symmetry. Comparing leading  $t_1 \times t_1$  submatrices in  $A_1 \oplus A_2 = A_2 \oplus A_1$  gives  $A_1 = A_2 \oplus A_3$  where  $A_3$  is the leading  $(t_1 - t_2) \times (t_1 - t_2)$  submatrix of  $A_1$ . Comparing  $t_1 \times t_1$  submatrices in

$A_1 \oplus A_2 = A_2 \oplus A_1$  made up of the last  $t_1$  rows and last  $t_1$  columns gives

$A_3 \oplus A_2 = A_1$ . So we obtain the  $t_1 \times t_1$  matrix equation  $A_2 \oplus A_3 = A_3 \oplus A_2$ . As  $t = \gcd\{t_2, t_1 - t_2\}$  and  $t_1 < t_1 + t_2$  there is a  $t \times t$  matrix  $A$  with  $A_2 = A \oplus A \oplus \dots \oplus A$  ( $t_2/t$  terms) and  $A_3 = A \oplus A \oplus \dots \oplus A$   $((t_1 - t_2)/t)$  terms). On substituting we obtain  $A_1 = A_2 \oplus A_3 = A \oplus A \oplus \dots \oplus A$   $(t_1/t)$  terms) contrary to the above supposition. So there is no least integer  $t_1 + t_2$  as above, i.e. the matrix  $A$  exists.

(d) Suppose  $f(x) = a$ , i.e.  $f(x)$  is a constant polynomial over  $R$ . Then

$$f(A_1 \oplus A_2) = a(A_1 \oplus A_2) = (aA_1) \oplus (aA_2) = f(A_1) \oplus f(A_2).$$

Let  $\deg f(x) = t > 0$  and suppose the equation  $g(A_1 \oplus A_2) = g(A_1) \oplus g(A_2)$  holds for all  $g(x) \in R[x]$  with  $\deg g(x) < t$ . Then  $f(x) = ax^t + g(x)$  and so

$$f(A_1 \oplus A_2) = a(A_1 \oplus A_2)^t + g(A_1 \oplus A_2) = a(A_1^t \oplus A_2^t) + g(A_1) \oplus g(A_2)$$

as  $\oplus$  respects matrix multiplication. As  $\oplus$  respects matrix addition we obtain

$$\begin{aligned} f(A_1 \oplus A_2) &= (aA_1^t) \oplus (aA_2^t) + g(A_1) \oplus g(A_2) = \\ &= (aA_1^t + g(A_1)) \oplus (aA_2^t + g(A_2)) = f(A_1) \oplus f(A_2) \end{aligned}$$

which completes the induction and the proof in the case  $s = 2$ . Suppose  $s > 2$  and inductively we have

$$f(A_1 \oplus A_2 \oplus \dots \oplus A_{s-1}) = f(A_1) \oplus f(A_2) \oplus \dots \oplus f(A_{s-1}).$$

Then using the case  $s = 2$  with  $A_1$  and  $A_2$  replaced by  $A_1 \oplus A_2 \oplus \dots \oplus A_{s-1}$  and  $A_s$  respectively we obtain

$$\begin{aligned} f(A_1 \oplus A_2 \oplus \dots \oplus A_{s-1} \oplus A_s) &= \\ f(A_1 \oplus A_2 \oplus \dots \oplus A_{s-1}) \oplus f(A_s) &= f(A_1) \oplus f(A_2) \oplus \dots \oplus f(A_{s-1}) \oplus f(A_s) \end{aligned}$$

which completes the proof.

(e) Consider  $u, v \in M$  and  $r \in R$ . Then

$$\begin{aligned} (u+v)(\alpha+\beta) &= (u+v)\alpha + (u+v)\beta = (u)\alpha + (v)\alpha + (u)\beta + (v)\beta = \\ (u)\alpha + (u)\beta + (v)\alpha + (v)\beta &= (u)(\alpha+\beta) + (v)(\alpha+\beta) \end{aligned}$$

showing that  $\alpha + \beta$  is additive. Similarly

$$\begin{aligned} (u+v)(\alpha\beta) &= ((u+v)\alpha)\beta = ((u)\alpha + (v)\alpha)\beta = \\ ((u)\alpha)\beta + ((v)\alpha)\beta &= (u)(\alpha\beta) + (v)(\alpha\beta) \end{aligned}$$

showing that  $\alpha\beta$  is additive. Also

$$(rv)(\alpha + \beta) = (rv)\alpha + (rv)\beta = r((v)\alpha) + r((v)\beta) = r((v)\alpha + (v)\beta) = r((v)(\alpha + \beta))$$

and so  $\alpha + \beta$  is  $R$ -linear. Finally

$$(rv)(\alpha\beta) = ((rv)\alpha)\beta = (r((v)\alpha))\beta = r(((v)\alpha)\beta) = r((v)(\alpha\beta))$$

and so  $\alpha\beta$  is  $R$ -linear. The mapping  $a\alpha$  is additive as

$$(u+v)(a\alpha) = a(u+v)\alpha = a(u)\alpha + a(v)\alpha = (u)(a\alpha) + (v)(a\alpha).$$

Also  $(rv)(a\alpha) = a((rv)\alpha) = ar((v)\alpha) = ra((v)\alpha) = r(a(v)\alpha) = r((v)(a\alpha))$  as  $R$  is commutative. So  $a\alpha$  is  $R$ -linear.

Yes,  $\text{End } M$  is a ring. Yes,  $\text{End } M$  is an  $R$ -module.

Suppose  $\beta = \alpha^{n-1}$  is  $R$ -linear where  $n > 1$ . By the above  $\alpha^n = \alpha\beta$  is  $R$ -linear.

By induction  $\alpha^n$  is  $R$ -linear for all positive integers  $n$ .

Now the identity  $\iota: M \rightarrow M$  is  $R$ -linear. Therefore  $a_0\iota$  is  $R$ -linear by the above

theory. Suppose  $g(\alpha)$  is  $R$ -linear for all  $g(x) \in R[x]$  with  $\deg g(x) < n$ . This supposition is true for  $n=1$  as  $g(x) = a_0$  and  $g(\alpha) = a_0\iota$  in this case. Suppose  $f(x)$

has degree  $n$  over  $R$ . Then  $f(x) = a_n x^n + g(x)$  where  $\deg g(x) < n$ . By the

preceding theory  $a_n \alpha^n$  is  $R$ -linear and by the inductive hypothesis  $g(\alpha)$  is  $R$ -linear. Therefore  $f(\alpha) = a_n \alpha^n + g(\alpha)$  is  $R$ -linear, completing the induction.

Consider  $f(x) = \sum_{i \geq 0} a_i x^i$  and  $g(x) = \sum_{i \geq 0} b_i x^i$  in  $R[x]$ . Then

$$(f(x) + g(x))\epsilon_\alpha = (\sum_{i \geq 0} (a_i + b_i)x^i)\epsilon_\alpha = \sum_{i \geq 0} (a_i + b_i)\alpha^i.$$

As  $a_i, b_i, (v)\alpha^i$  belong to the commutative ring  $R$  we obtain

$$\sum_{i \geq 0} (a_i + b_i)(v)\alpha^i = \sum_{i \geq 0} a_i(v)\alpha^i + \sum_{i \geq 0} b_i(v)\alpha^i \text{ for all } v \in R.$$

So

$$\sum_{i \geq 0} (a_i + b_i)\alpha^i = \sum_{i \geq 0} a_i \alpha^i + \sum_{i \geq 0} b_i \alpha^i.$$

As  $\sum_{i \geq 0} a_i \alpha^i + \sum_{i \geq 0} b_i \alpha^i = f(\alpha) + g(\alpha) = (f(x))\epsilon_\alpha + (g(x))\epsilon_\alpha$  we obtain

$$(f(x) + g(x))\epsilon_\alpha = (f(x))\epsilon_\alpha + (g(x))\epsilon_\alpha,$$

showing that  $\epsilon_\alpha$  is additive. Similarly

$$(f(x)g(x))\epsilon_\alpha = (\sum_{i \geq 0} (a_0 b_i + a_1 b_{i-1} + \dots + a_i b_0)x^i)\epsilon_\alpha = \sum_{i \geq 0} (a_0 b_i + a_1 b_{i-1} + \dots + a_i b_0)\alpha^i$$

which is the result of collecting together terms involving  $\alpha^j \alpha^k$  where  $j+k=i$  in the product

$$(\sum_{j \geq 0} a_j \alpha^j)(\sum_{k \geq 0} b_k \alpha^k) = f(\alpha)g(\alpha) = (f(x))\epsilon_\alpha(g(x))\epsilon_\alpha.$$

So  $(f(x)g(x))\epsilon_\alpha = (f(x))\epsilon_\alpha(g(x))\epsilon_\alpha$  showing that  $\epsilon_\alpha$  is multiplicative. The equation  $(1)\epsilon_\alpha = \iota$  shows that  $\epsilon_\alpha$  maps the 1-element of  $R[x]$  to the 1-element of  $\text{End } M$ . Therefore  $\epsilon_\alpha$  is a ring homomorphism.

(f) Let  $v_1, v_2, \dots, v_t$  be the basis  $\mathcal{B}$ . Consider  $\alpha, \alpha' \in \text{End}_F V$  and let  $(\alpha)\theta = A = (a_{ij})$ ,

$(\alpha')\theta = A' = (a'_{ij})$ .  $(\alpha')\theta = A'$ . Then  $(v_i)\alpha = \sum_{j=1}^t a_{ij} v_j$  and  $(v_i)\alpha' = \sum_{j=1}^t a'_{ij} v_j$  for

$1 \leq i \leq t$ . Adding these equations together gives

$$(v_i)(\alpha + \alpha') = (v_i)\alpha + (v_i)\alpha' = \sum_{j=1}^t a_{ij}v_j + \sum_{j=1}^t a'_{ij}v_j = \sum_{j=1}^t (a_{ij} + a'_{ij})v_j$$

for  $1 \leq i \leq t$  which shows that  $\alpha + \alpha'$  has matrix  $(a_{ij} + a'_{ij}) = A + A'$  relative to  $\mathcal{B}$ .

We have shown  $(\alpha + \alpha')\theta = A + A' = (\alpha)\theta + (\alpha')\theta$  for all  $\alpha$  and  $\alpha'$  in  $\text{End}_F V$ , that is,  $\theta$  is additive.

In order to find  $(\alpha\alpha')\theta$ , first replace the dummy suffixes  $i, j$  in  $(v_i)\alpha' = \sum_{j=1}^t a'_{ij}v_j$  by

$j, k$  respectively giving  $(v_j)\alpha' = \sum_{k=1}^t a'_{jk}v_k$  for  $1 \leq j \leq t$ . Then

$$\begin{aligned} (v_i)\alpha\alpha' &= ((v_i)\alpha)\alpha' = \left(\sum_{j=1}^t a_{ij}v_j\right)\alpha' = \sum_{j=1}^t a_{ij}(v_j)\alpha' = \\ &= \sum_{j=1}^t a_{ij}\left(\sum_{k=1}^t a'_{jk}v_k\right) = \sum_{k=1}^t \left(\sum_{j=1}^t a_{ij}a'_{jk}\right)v_k \end{aligned}$$

for  $1 \leq i \leq t$  which shows that the matrix of  $\alpha\alpha'$  relative to  $\mathcal{B}$  is  $AA'$  as its

$(i, k)$ -entry is  $\sum_{j=1}^t a_{ij}a'_{jk}$ . We have shown  $(\alpha\alpha')\theta = AA' = (\alpha)\theta(\alpha')\theta$  for all  $\alpha$  and

$\alpha'$  in  $\text{End}_F V$  and so  $\theta$  is multiplicative. The 1-element of  $\text{End}_F V$  is the identity mapping  $\iota: V \rightarrow V$ . As  $(v_i)\iota = v_i$  for  $1 \leq i \leq t$  we see  $(\iota)\theta = I$ , the  $t \times t$  identity matrix over  $F$ , which is the 1-element of  $\mathfrak{M}_t(F)$ . So  $\theta$  is a mapping of rings which respects addition, multiplication and 1-elements. So  $\theta$  is a ring homomorphism.

We show that  $\theta$  is bijective. Consider  $A = (a_{ij}) \in \mathfrak{M}_t(F)$ . Suppose there is an  $\alpha$  in  $\text{End}_F V$  with  $(\alpha)\theta = A$ . Let  $v \in V$ . There are unique scalars  $\lambda_1, \lambda_2, \dots, \lambda_t$  such that  $v = \lambda_1 v_1 + \lambda_2 v_2 + \dots + \lambda_t v_t$ . Write  $(\lambda_1, \lambda_2, \dots, \lambda_t)A = (\mu_1, \mu_2, \dots, \mu_t)$ , that is,

$\sum_{i=1}^t \lambda_i a_{ij} = \mu_j$  for  $1 \leq j \leq t$ . Hence

$$(v)\alpha = \left(\sum_{i=1}^t \lambda_i v_i\right)\alpha = \sum_{i=1}^t \lambda_i (v_i)\alpha = \sum_{i=1}^t \lambda_i \left(\sum_{j=1}^t a_{ij}v_j\right) = \sum_{j=1}^t \left(\sum_{i=1}^t \lambda_i a_{ij}\right)v_j = \sum_{j=1}^t \mu_j v_j$$

showing that  $\alpha$  maps  $v = \sum_{i=1}^t \lambda_i v_i$  to the element  $\sum_{j=1}^t \mu_j v_j$  of  $V$ . So there is *at most one*

$\alpha$  in  $\text{End}_F V$  with  $(\alpha)\theta = A$ , i.e.  $\theta$  is injective.

Let  $\beta: V \rightarrow V$  be defined by  $(v)\beta = \sum_{j=1}^t \mu_j v_j$  where  $v = \sum_{i=1}^t \lambda_i v_i$  and  $\sum_{i=1}^t \lambda_i a_{ij} = \mu_j$

for  $1 \leq j \leq t$ . Is  $\beta \in \text{End}_F V$ ? Consider  $v' \in V$ . There are scalars  $\lambda'_i$  ( $1 \leq i \leq t$ ) with

$v' = \sum_{i=1}^t \lambda'_i v_i$ . Then  $(v')\beta = \sum_{j=1}^t \mu'_j v_j$  where  $\sum_{i=1}^t \lambda'_i a_{ij} = \mu'_j$  for  $1 \leq j \leq t$ . Then

$v + v' = \sum_{i=1}^t (\lambda_i + \lambda'_i)v_i$  and  $\sum_{i=1}^t (\lambda_i + \lambda'_i)a_{ij} = \mu_j + \mu'_j$  for  $1 \leq j \leq t$ . So

$(v + v')\beta = \sum_{j=1}^t (\mu_j + \mu'_j)v_j = \sum_{j=1}^t \mu_j v_j + \sum_{j=1}^t \mu'_j v_j = (v)\beta + (v')\beta$  showing that  $\beta$  is

additive. Let  $a \in F$ . Then  $av = \sum_{i=1}^t a\lambda_i v_i$  and  $\sum_{i=1}^t a\lambda_i a_{ij} = a\mu_j$  which together give

$$(av)\beta = \sum_{j=1}^t a\mu_j v_j = a\left(\sum_{j=1}^t \mu_j v_j\right) = a((v)\beta) \text{ showing } \beta \in \text{End}_F V.$$

Taking  $v = v_i$  gives  $(\lambda_1, \lambda_2, \dots, \lambda_t) = e_i$  which is row  $i$  of the  $t \times t$  identity matrix  $I$  over  $F$  for  $1 \leq i \leq t$ . Hence

$$(\lambda_1, \lambda_2, \dots, \lambda_t)A = e_i A = (a_{i1}, a_{i2}, \dots, a_{it}) = (\mu_1, \mu_2, \dots, \mu_t)$$

and so  $(v_i)\beta = \sum_{j=1}^t a_{ij} v_j$  for  $1 \leq i \leq t$ . Therefore  $\beta$  satisfies  $(\beta)\theta = A$  showing that  $\theta$  is surjective. So  $\theta$  is a ring isomorphism.

Finally  $(v_i)(a\alpha) = a((v_i)\alpha) = a\left(\sum_{j=1}^t a_{ij} v_j\right) = \sum_{j=1}^t (aa_{ij})v_j$  for  $1 \leq i \leq t$ . So  $a\alpha$  has matrix  $aA$  relative to  $\mathcal{B}$ , i.e.  $(a\alpha)\theta = a((\alpha)\theta)$  showing  $\theta$  to be  $F$ -linear.

### Solution 3

$$(a) \quad xe_1 = e_1 A = (1, 0) \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = (1, 0) = e_1 \text{ and so } x^2 e_1 = x(xe_1) = xe_1 = e_1 = (1, 0).$$

$$(x+1)e_1 = xe_1 + e_1 = e_1 + e_1 = 2e_1 = (2, 0), \quad (x-1)e_1 = xe_1 - e_1 = e_1 - e_1 = (0, 0) = 0.$$

$$xe_2 = e_2 A = (0, 1) \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = (0, 0),$$

$$x(x+1)e_2 = (x+1)xe_2 = (x+1)(0, 0) = (0, 0)(A + I) = (0, 0),$$

$$x(e_1 + e_2) = xe_1 + xe_2 = e_1 + 0 = e_1 = (1, 0),$$

$$(x-1)(e_1 + e_2) = xe_1 + xe_2 - e_1 - e_2 = e_1 + 0 - e_1 - e_2 = -e_2 = (0, -1) \text{ and so}$$

$x(x-1)(e_1 + e_2) = x(-e_2) = -xe_2 = -0 = (0, 0)$ . As  $e_1 \neq 0$ ,  $(x-1)e_1 = 0$  the order of  $e_1$  in  $M(A)$  is  $x-1$ . As  $e_2 \neq 0$ ,  $xe_2 = 0$  the order of  $e_2$  in  $M(A)$  is  $x$ . As  $e_1 + e_2 \neq 0$ ,  $x(e_1 + e_2) \neq 0$ ,  $x(x-1)(e_1 + e_2) = 0$  the order of  $e_1 + e_2$  in  $M(A)$  is  $x(x-1)$ . Yes,  $e_1 + e_2, x(e_1 + e_2)$ , i.e.  $(1, 1), (1, 0)$ , is a basis of  $\mathbb{Q}^2$ . So each element of  $M(A)$  can be expressed  $a_0(e_1 + e_2) + a_1 x(e_1 + e_2) = (a_0 + a_1 x)(e_1 + e_2)$  for some  $a_0, a_1 \in \mathbb{Q}$ . So each element of  $M(A)$  is a polynomial multiple of  $e_1 + e_2$ , i.e.  $e_1 + e_2$  is a generator of  $M(A)$ .

Let  $v \in \langle e_1 \rangle$ . Then  $v = a_1 e_1$  where  $a_1 \in \mathbb{Q}$ . So  $xv = xa_1 e_1 = a_1 xe_1 = a_1 e_1 \in \langle e_1 \rangle$ , showing that  $\langle e_1 \rangle$  is a submodule of  $M(A)$  by (5.15). Let  $v \in \langle e_2 \rangle$ . Then  $v = a_2 e_2$  where  $a_2 \in \mathbb{Q}$ . So  $xv = xa_2 e_2 = a_2 xe_2 = a_2 (0, 0) = 0 \in \langle e_2 \rangle$ , showing that  $\langle e_2 \rangle$  is a submodule of  $M(A)$  by (5.15). On the other hand  $x(e_1 + e_2) = e_1 \notin \langle e_1 + e_2 \rangle$  and so  $\langle e_1 + e_2 \rangle$  is not a submodule of  $M(A)$  by (5.14).

The four submodules of  $M(A)$  are  $0 = \{(0, 0)\}, \langle e_1 \rangle, \langle e_2 \rangle, M(A)$ ; in terms of the generator  $e_1 + e_2$  these submodules have generators

$$x(x-1)(e_1 + e_2), x(e_1 + e_2), (x-1)(e_1 + e_2), e_1 + e_2 \text{ respectively.}$$

The vectors  $v, xv$ , i.e.  $a_1 e_1 + a_2 e_2, x(a_1 e_1 + a_2 e_2) = a_1 e_1$  are linearly independent

$$\Leftrightarrow \begin{vmatrix} a_1 & a_2 \\ a_1 & 0 \end{vmatrix} \neq 0. \text{ So } v \text{ is a generator of } M(A) \Leftrightarrow -a_1 a_2 \neq 0, \text{ i.e. } a_1 \neq 0 \text{ and } a_2 \neq 0.$$

(b)  $xv_1 = v_1A = (1, -1)\begin{pmatrix} 4 & -2 \\ 5 & -3 \end{pmatrix} = (-1, 1) = -v_1$  and so  $(x+1)v_1 = 0$ . As  $v_1 \neq 0$  we see

that  $v_1$  has order  $x+1$  in  $M(A)$ .  $xv_2 = v_2A = (5, -2)\begin{pmatrix} 4 & -2 \\ 5 & -3 \end{pmatrix} = (10, -4) = 2v_2$  and so

$(x-2)v_2 = 0$ . As  $v_2 \neq 0$  we see that  $v_2$  has order  $x-2$  in  $M(A)$ . Yes,  $v_1$  and  $v_2$  are eigenvectors of  $A$  the corresponding eigenvalues being  $-1$  and  $2$  respectively.

Therefore  $\chi_A(x) = (x+1)(x-2)$  as  $\chi_A(x)$  is monic of degree 2 with zeros  $-1$  and  $2$ .

$$\chi_A(A) = (A+I)(A-2I) = \begin{pmatrix} 5 & -2 \\ 5 & -2 \end{pmatrix} \begin{pmatrix} 2 & -2 \\ 5 & -5 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Now  $xe_1 = e_1A = (1, 0)\begin{pmatrix} 4 & -2 \\ 5 & -3 \end{pmatrix} = (4, -2)$  and so

$$x^2e_1 = x(xe_1) = (4, -2)A = (4, -2)\begin{pmatrix} 4 & -2 \\ 5 & -3 \end{pmatrix} = (6, -2).$$

So  $x^2e_1 - xe_1 = (2, 0) = 2e_1$  which rearranges to give  $(x^2 - x - 2)e_1 = 0$ , i.e.

$(x+1)(x-2)e_1 = 0$ . As  $e_1, xe_1$  are linearly independent we conclude that  $e_1$  has order  $\chi_A(x) = (x+1)(x-2)$  in  $M(A)$ . Yes,  $M(A)$  is cyclic having generator  $e_1$ . The polynomials arising as orders of elements of  $M(A)$  are the monic divisors of  $\chi_A(x)$ , namely  $1, x+1, x-2, (x+1)(x-2)$  which are the orders of  $0, v_1, v_2, e_1$  respectively.

$X = \begin{pmatrix} 1 & 0 \\ 4 & -2 \end{pmatrix}$  and  $Y = \begin{pmatrix} 1 & -1 \\ 5 & -2 \end{pmatrix}$  are invertible over  $\mathbb{Q}$  as  $\det X = -2 \neq 0$  and

$\det Y = 3 \neq 0$ . Comparing the rows of  $XA = \begin{pmatrix} 1 & 0 \\ 4 & -2 \end{pmatrix} \begin{pmatrix} 4 & -2 \\ 5 & -3 \end{pmatrix} = \begin{pmatrix} 4 & -2 \\ 6 & 2 \end{pmatrix}$  with the

rows of  $X$  we see  $\begin{pmatrix} 4 & -2 \\ 6 & -2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 4 & -2 \end{pmatrix}$ , i.e.  $XA = CX$  where

$C = \begin{pmatrix} 0 & 1 \\ 2 & 1 \end{pmatrix} = C(x^2 - x - 2)$  is the companion matrix of

$$x^2 - x - 2 = (x+1)(x-2) = \chi_A(x).$$

Comparing the rows of

$$YA = \begin{pmatrix} 1 & -1 \\ 5 & -2 \end{pmatrix} \begin{pmatrix} 4 & -2 \\ 5 & -3 \end{pmatrix} = \begin{pmatrix} -1 & 1 \\ 10 & -4 \end{pmatrix}$$

with the rows of  $Y$  gives

$$\begin{pmatrix} -1 & 1 \\ 10 & -4 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 5 & -2 \end{pmatrix},$$

i.e.  $YA = DY$  where  $D = \text{diag}(-1, 2)$ . Yes,  $C$  and  $D$  are both similar to  $A$  by (5.3)

as  $XAX^{-1} = C$  and  $YAY^{-1} = D$ .

(c) Let  $x - \lambda_i$  be the order of  $e_i$  in  $F[x]$ -module  $M(A)$  for  $1 \leq i \leq t$ . Then

$(x - \lambda_i)e_i = 0$  which gives  $xe_i = e_iA = \lambda_i e_i$  for  $1 \leq i \leq t$ . So  $A = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_t)$ .

Write  $v_0 = e_1 + e_2 + \dots + e_t$  and let  $x - \lambda$  be the order of  $v_0$  in  $M(A)$ . Then

$(x - \lambda)v_0 = 0$ , i.e.  $(\lambda_1, \lambda_2, \dots, \lambda_t) = v_0A = xv_0 = \lambda v_0 = (\lambda, \lambda, \dots, \lambda)$ . Comparing entries

in these elements of  $F^t$  gives  $\lambda_i = \lambda$  for  $1 \leq i \leq t$ . So  $A = \text{diag}(\lambda, \lambda, \dots, \lambda) = \lambda I$ .

As  $B \neq \lambda I$  for all  $\lambda \in F$  one at least of  $e_1, e_2, e_1 + e_2$  has order of degree 2 in  $M(B)$ . Denote one such vector by  $v$ . Then  $v, xv$  are linearly independent and so  $v, xv$  span  $F^2$ , i.e.  $M(B)$  is cyclic with generator  $v$ .

(d) Let  $f_1(x), f_2(x) \in K$ . Then  $(v)f_1(\alpha) = 0$  and  $(v)f_2(\alpha) = 0$ . Adding gives  $(v)(f_1(\alpha) + f_2(\alpha)) = (v)f_1(\alpha) + (v)f_2(\alpha) = 0 + 0 = 0$  which shows  $f_1(x) + f_2(x) \in K$ . Also  $-f_1(x) \in K$  as  $(v)(-f_1(\alpha)) = -(v)f_1(\alpha) = -0 = 0$ . As  $0(x)v = 0$  we see that the zero polynomial  $0(x)$  belongs to  $K$ . Therefore  $K$  is a subgroup of the additive group of the ring  $F[x]$ . For  $f(x) \in F[x]$  we have  $f(x)f_1(x) \in K$  as

$$(v)f(\alpha)f_1(\alpha) = (v)f_1(\alpha)f(\alpha) = (0)f(\alpha) = 0.$$

So  $K$  is an ideal of  $F[x]$  by (4.3).

Suppose  $V$  is  $t$ -dimensional. Then  $v, (v)\alpha, (v)\alpha^2, \dots, (v)\alpha^t$  are  $t+1$  vectors of  $V$ . These vectors are linearly dependent and so there are  $a_0, a_1, \dots, a_t \in F$ , not all zero, with  $a_0v + a_1(v)\alpha + \dots + a_t(v)\alpha^t = 0$ , i.e.  $(v)f_0(\alpha) = 0$  where

$f_0(x) = a_0 + a_1x + \dots + a_tx^t \neq 0(x)$  and  $f_0(x) \in K$ . So  $K$  is a non-zero ideal of  $F[x]$ .

(e) Suppose  $\overline{f(x)} = \overline{f'(x)}$ . Then  $f(x) - f'(x) = q(x)m(x)$  for some  $q(x) \in F[x]$ . So  $f(A) - f'(A) = q(A)m(A) = q(A) \times 0 = 0$ , i.e.  $f(A) = f'(A)$ . Therefore  $\overline{f(x)}v = v f(A) = v f'(A) = \overline{f'(x)}v$  showing that the product  $\overline{f(x)}v$  is unambiguously defined. As  $M' = F'$  as sets and addition in  $M'$  is the usual vector addition, we see that module laws 1, 2, 3 and 4 (see before (2.19)) are obeyed by  $M'$ . For

$f(x), f_1(x), f_2(x) \in F[x]$  and  $v, v_1, v_2 \in M'$  we have

$$\begin{aligned} \overline{f(x)}(v_1 + v_2) &= (v_1 + v_2)f(A) = v_1f(A) + v_2f(A) = \overline{f(x)}v_1 + \overline{f(x)}v_2 \text{ and} \\ \overline{(f_1(x) + f_2(x))}v &= \overline{(f_1(x) + f_2(x))}v = v(f_1(A) + f_2(A)) = \\ &= v f_1(A) + v f_2(A) = \overline{f_1(x)}v + \overline{f_2(x)}v \end{aligned}$$

showing that  $M'$  obeys module law 5. Also

$$\begin{aligned} \overline{(f_1(x) f_2(x))}v &= \overline{f_1(x)f_2(x)}v = v f_1(A)f_2(A) = \\ &= v f_2(A)f_1(A) = \overline{(f_2(x) f_1(x))}v = \overline{f_2(x)}(\overline{f_1(x)}v) \end{aligned}$$

showing that  $M'$  obeys module law 6. The 1-element of  $F[x]/\langle m(x) \rangle$  is  $\overline{1(x)}$  and  $\overline{1(x)}v = v1(A) = vI = v$  for all  $v \in M'$  showing that  $M'$  obeys module law 7. We conclude that  $M'$  is an  $F[x]/\langle m(x) \rangle$ -module.

Suppose  $m(x)$  irreducible over  $F$  and let  $r = \deg m(x)$ . Then  $F' = F[x]/\langle m(x) \rangle$  is a field and so  $M'$  is a vector space over  $F'$ . Consider  $v \in M'$ . There are  $f_j(x) \in F[x]$

for  $1 \leq j \leq t'$  with  $v = \sum_{j=1}^{t'} \overline{f_j(x)}v_j$  as  $v_1, v_2, \dots, v_{t'}$  span  $M'$  over  $F'$ . Further we may

assume  $\deg f_j(x) < r$ . Write  $f_j(x) = \sum_{i=0}^{r-1} a_{ji}x^i$ . Substituting for each  $f_j(x)$  gives

$v = \sum_{i,j} a_{ji} \overline{x^i}v_j$  where the summation is for  $0 \leq i < r, 1 \leq j \leq t'$ . As  $a_{ji} \in F$  we see that

the  $rt'$  vectors  $\overline{x^i}v_j$  span  $M'$  over  $F$ .

Suppose  $\sum_{i,j} a_{ji} \overline{x^i}v_j = 0$ . As  $v_1, v_2, \dots, v_{t'}$  are linearly independent over  $F'$  we deduce

$\overline{f_j(x)} = \overline{0(x)}$ , i.e.  $m(x) \mid f_j(x)$ . So  $f_j(x) = 0(x)$  for  $1 \leq j \leq t'$ , i.e. the  $rt'$  vectors  $\overline{x^i} v_j$  are linearly independent. These vectors form an  $F$ -basis of  $F^t$  and so are  $t$  in number, i.e.  $t = rt'$  and so  $t/t' = \deg m(x)$ .

#### Solution 4

(a)  $xe_1 = e_1A = (0, 2, -1)$  and so  $x^2e_1 = x(xe_1) = (0, 2, -1)A = (3, 4, -2)$ . The vectors  $e_1, xe_1$  are linearly independent, but  $e_1, xe_1, x^2e_1$  are linearly dependent as  $x^2e_1 - 2xe_1 = (3, 0, 0) = 3e_1$ , i.e.  $x^2e_1 - 2xe_1 - 3e_1 = 0$ . So  $e_1$  has order  $x^2 - 2x - 3 = (x+1)(x-3)$  in  $M(A)$ . As  $(x+1)e_1 \neq 0$  and  $(x-3)((x+1)e_1) = 0$  we see that  $(x+1)e_1$  has order  $x-3$  in  $M(A)$ , i.e.  $(x+1)e_1 = (1, 2, -1)$  is a row eigenvector of  $A$  with eigenvalue 3. In the same way  $(x-3)e_1 = (-3, 2, -1) \neq 0$  satisfies  $(x+1)(-3, 2, -1) = 0$ , i.e.  $(3, 2, -1)A = -(3, 2, -1)$  which shows that  $(x-3)e_1$  is a row eigenvector of  $A$  with eigenvalue  $-1$ . As

$$(A+I)(A-3I) = \begin{pmatrix} 1 & 2 & -1 \\ 2 & 4 & -2 \\ 1 & 2 & -1 \end{pmatrix} \begin{pmatrix} -3 & 2 & -1 \\ 2 & 0 & -2 \\ 1 & 2 & -5 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

we see  $(x+1)(x-3)v = v(A+I)(A-3I) = 0$  for all  $v \in M(A)$ . Hence there is no element  $v$  having order of degree 3 in  $M(A)$ , i.e. there is no  $v \in \mathbb{Q}^3$  such that  $v, xv, x^2v$  is a basis of  $\mathbb{Q}^3$ . So  $M(A)$  is not cyclic.

As  $x(1, 0, -1) = (1, 0, -1)A = (-1, 0, 1) = -(1, 0, -1)$  we obtain  $(x+1)(1, 0, -1) = 0$ . So  $(1, 0, -1)$  has order  $x+1$  in  $M(A)$  showing that  $(1, 0, -1)$  is a row eigenvector of  $A$  with eigenvalue  $-1$ . Let

$$X = \begin{pmatrix} 1 & 2 & -1 \\ -3 & 2 & -1 \\ 1 & 0 & -1 \end{pmatrix}$$

having row eigenvectors of  $A$  as its rows. Then  $\det X = -8 \neq 0$  showing that the rows of  $X$  are linearly independent, i.e.  $X$  is invertible over  $\mathbb{Q}$ . As  $XA = DX$  where  $D = \text{diag}(3, -1, -1)$  we see  $XAX^{-1} = D$  is diagonal.

(b)  $xe_1 = e_1A = (2, 2, 1)$  and so  $x^2e_1 = x(2, 2, 1) = (2, 2, 1)A = (3, 4, 2)$ . As  $e_1, xe_1$  are linearly independent and  $x^2e_1 - 2xe_1 = (-1, 0, 0) = -e_1$ , i.e.  $(x^2 - 2x + 1)e_1 = 0$ , we see  $x^2 - 2x + 1 = (x-1)^2$  is the order of  $e_1$  in  $M(A)$ . Now

$$x(e_1 + e_2) = (1, 1, 0)A = (1, 1, 0) = e_1 + e_2$$

shows  $(x-1)(e_1 + e_2) = 0$ , i.e. the order of  $e_1 + e_2$  has order  $x-1$  in  $M(A)$ . As

$$X = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 0 \\ 2 & 2 & 1 \end{pmatrix} \text{ we see } |X| = -1 \text{ and so } X \text{ is invertible over } \mathbb{Q}. \text{ Then}$$

$$XA = \begin{pmatrix} 1 & 1 & 0 \\ 2 & 2 & 1 \\ 3 & 4 & 2 \end{pmatrix} \text{ whereas } \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 2 \end{pmatrix} X = \begin{pmatrix} 1 & 1 & 0 \\ 2 & 2 & 1 \\ 3 & 4 & 2 \end{pmatrix}, \text{ i.e. } XA = CX \text{ where}$$

$C = C(x-1) \oplus C((x-1)^2)$ . So  $XAX^{-1} = C$  on postmultiplying  $XA = CX$  by  $X^{-1}$ .



(c) Consider  $v \in N$  where  $N \subseteq \ker \alpha$ . Then  $(v)\alpha = 0 \in N$  showing that  $N$  is  $\alpha$ -invariant. Consider  $v' \in N'$  where  $\text{im } \alpha \subseteq N'$ . Then  $(v')\alpha \in \text{im } \alpha$  and so  $(v')\alpha \in N'$  showing that  $N'$  is  $\alpha$ -invariant.

Let  $N'$  be an  $\alpha$ -invariant subspace of  $V$  with  $N' \not\subseteq \ker \alpha$ . So there is  $v_0 \in N'$  with  $(v_0)\alpha \neq 0$ . By hypothesis  $\dim(\text{im } \alpha) = 1$ . As  $(v_0)\alpha \in \text{im } \alpha$  we see  $\text{im } \alpha = \langle (v_0)\alpha \rangle$ . As  $N'$  is an  $\alpha$ -invariant subspace we conclude  $\text{im } \alpha = \langle (v_0)\alpha \rangle \subseteq N'$ . So there are no further  $\alpha$ -invariant subspaces.

Suppose that  $N$  is  $\alpha$ -invariant. For all  $v \in N$  we have  $(v)\alpha \in N$  and so  $(v)(\alpha - \lambda I) = (v)\alpha - \lambda((v)I) = (v)\alpha - \lambda v \in N$  as  $N$  is closed under subtraction and scalar multiplication. Therefore  $N$  is  $(\alpha - \lambda I)$ -invariant.

Conversely suppose  $N$  to be  $(\alpha - \lambda I)$ -invariant. For all  $v \in N$  we have  $(v)(\alpha - \lambda I) \in N$  and so  $(v)\alpha = (v)\alpha - \lambda v + \lambda v = (v)(\alpha - \lambda I) + \lambda v \in N$  as  $N$  is closed under addition and scalar multiplication. Therefore  $N$  is  $\alpha$ -invariant.

(d) For  $v = (a_1, a_2, a_3) \in F^3$  we have  $(v)\alpha = (a_1, a_2, a_3)A = (a_1, 0, 0) = a_1 e_1$  showing  $\text{im } \alpha = \langle e_1 \rangle$ . So  $\text{rank } \alpha = 1$ . Also

$$\ker \alpha = \{v \in F^3 : vA = 0\} = \{(0, a_2, a_3) \in F^3\} = \langle e_2, e_3 \rangle.$$

By (c) above each  $\alpha$ -invariant subspace  $N$  of  $F^3$  satisfies either  $\langle e_1 \rangle \subseteq N$  or  $N \subseteq \langle e_2, e_3 \rangle$ .

Suppose  $F = \mathbb{Z}_2$ . There are four 1-dimensional  $\alpha$ -invariant subspaces of  $\mathbb{Z}_2^3$  namely  $\langle e_1 \rangle, \langle e_2 \rangle, \langle e_3 \rangle, \langle e_2 + e_3 \rangle$ . There are four 2-dimensional  $\alpha$ -invariant subspaces of  $\mathbb{Z}_2^3$  namely  $\langle e_1, e_2 \rangle, \langle e_1, e_3 \rangle, \langle e_1, e_2 + e_3 \rangle, \langle e_2, e_3 \rangle$ . The  $\alpha$ -invariant subspaces  $\{0\}$  and  $\mathbb{Z}_2^3$  complete the list of ten.

(e) For  $v = (a_1, a_2, a_3) \in F^3$  we have  $(v)\alpha = (a_1, a_2, a_3)A = (0, a_1, 0) = a_1 e_2$  showing  $\text{im } \alpha = \langle e_2 \rangle$ . So  $\text{rank } \alpha = 1$ . Also

$$\ker \alpha = \{v \in F^3 : vA = 0\} = \{(0, a_2, a_3) \in F^3\} = \langle e_2, e_3 \rangle.$$

By (c) above a subspace  $N$  of  $F^3$  is  $\alpha$ -invariant  $\Leftrightarrow$  either  $\langle e_2 \rangle \subseteq N$  or  $N \subseteq \langle e_2, e_3 \rangle$ .

Suppose  $F = \mathbb{Z}_2$ . There are three 1-dimensional  $\alpha$ -invariant subspaces of  $\mathbb{Z}_2^3$  namely  $\langle e_2 \rangle, \langle e_3 \rangle, \langle e_2 + e_3 \rangle$ , the 1-dimensional subspaces of  $\langle e_2, e_3 \rangle$ . There are three 2-dimensional  $\alpha$ -invariant subspaces of  $\mathbb{Z}_2^3$  namely  $\langle e_1, e_2 \rangle, \langle e_1 + e_3, e_2 \rangle, \langle e_2, e_3 \rangle$ , the 2-dimensional subspaces containing  $\langle e_2 \rangle$ . The  $\alpha$ -invariant subspaces  $\{0\}$  and  $\mathbb{Z}_2^3$  complete the list of eight.

(f) From (c) above, with  $\lambda = -1$ , the  $\alpha$ -invariant subspaces of  $\mathbb{Q}^3$  coincide with the  $(\alpha + I)$ -invariant subspaces of  $\mathbb{Q}^3$ . But  $\alpha + I$  has matrix

$$A + I = \begin{pmatrix} 1 & 2 & -1 \\ 2 & 4 & -2 \\ 1 & 2 & -1 \end{pmatrix}$$

relative to the standard basis of  $\mathbb{Q}^3$  where  $A$  is the matrix of (a) above. Now  $\text{rank}(\alpha + I) = 1$  and  $\text{im}(\alpha + I) = \langle (1, 2, -1) \rangle$ . Also  $\ker(\alpha + I) = \langle (-3, 2, -1), (1, 0, -1) \rangle$  using the eigenvectors of  $A$  with eigenvalue  $-1$ . So the  $\alpha$ -invariant subspaces of  $\mathbb{Q}^3$  are all subspaces containing  $\langle (1, 2, -1) \rangle$  and all subspaces of  $\langle (-3, 2, -1), (1, 0, -1) \rangle$ .

Now let  $\alpha: \mathbb{Q}^3 \rightarrow \mathbb{Q}^3$  be the linear mapping determined by the matrix  $A$  of (b) above. Then  $\alpha - \iota$  has matrix

$$A - I = \begin{pmatrix} 1 & 2 & 1 \\ -1 & -2 & -1 \\ 1 & 2 & 1 \end{pmatrix}$$

relative to the standard basis of  $\mathbb{Q}^3$ . From (c) above, with  $\lambda = 1$ , the  $\alpha$ -invariant subspaces of  $\mathbb{Q}^3$  coincide with the  $(\alpha - \iota)$ -invariant subspaces of  $\mathbb{Q}^3$ . But these are all subspaces containing  $\text{im}(\alpha - \iota) = \langle (1, 2, 1) \rangle$  together with all subspaces of  $\ker(\alpha - \iota) = \langle (1, 1, 0), (1, 2, 1) \rangle$ .

**Solution 5** Let  $u, u' \in N$ . Then  $(u)\beta = (a_1, a_2, \dots, a_s) \in F^s$ ,

$(u')\beta = (a'_1, a'_2, \dots, a'_s) \in F^s$  which means  $u = a_1u_1 + a_2u_2 + \dots + a_su_s$  and  $u' = a'_1u_1 + a'_2u_2 + \dots + a'_su_s$ . Adding gives

$u + u' = (a_1 + a'_1)u_1 + (a_2 + a'_2)u_2 + \dots + (a_s + a'_s)u_s$  which shows

$(u + u')\beta = (a_1 + a'_1, a_2 + a'_2, \dots, a_s + a'_s) = (u)\beta + (u')\beta$ , i.e.  $\beta$  is additive. For  $a \in F$  we have  $au = aa_1u_1 + aa_2u_2 + \dots + aa_su_s$  showing  $(au)\beta = a((u)\beta)$  and so  $\beta$  is

$F$ -linear. Also  $\beta$  is bijective as  $u_1, u_2, \dots, u_s$  is an  $F$ -basis of  $N$ . So  $\beta: N \cong F^s$  is a vector space isomorphism with  $(u_i)\beta = e_i$  for  $1 \leq i \leq s$ .

Write  $B = (b_{ij})$  for  $1 \leq i, j \leq s$ . Then

$$\begin{aligned} (xu_i)\beta &= ((u_i)\alpha)\beta = (b_{i1}u_1 + b_{i2}u_2 + \dots + b_{is}u_s)\beta = \\ &= (b_{i1}, b_{i2}, \dots, b_{is}) = e_i B = ((u_i)\beta)B = x((u_i)\beta) \end{aligned}$$

for  $1 \leq i \leq s$ . Using the  $F$ -linearity of  $\beta$  we see  $(xu)\beta = x((u)\beta)$  for all  $u \in N$ . By (5.15) with  $\theta = \beta$  we conclude that  $\beta$  is  $F[x]$ -linear, i.e.  $\beta: N \cong M(B)$  is an isomorphism of  $F[x]$ -modules.

## Solutions 5.2 (page 245)

### Solution 1

(a) The companion matrix

$$C = C(x^3 - x) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

has characteristic polynomial  $\chi_C(x) = x^3 - x = x(x-1)(x+1)$  and so  $C$  has eigenvalues  $0, 1, -1$ . The  $1 \times 3$  matrix equation  $vC = 0$  has non-zero solution  $v = (1, 0, -1)$  which is a row eigenvector of  $C$  with corresponding eigenvalue  $0$ . Also  $v(C - I) = 0$  has non-zero solution  $v = (0, 1, 1)$  which is a row eigenvector of  $C$  with corresponding eigenvalue  $1$ . Finally  $v(C + I) = 0$  has non-zero solution  $v = (0, 1, -1)$  which is a row eigenvector of  $C$  with corresponding eigenvalue  $-1$ . The matrix

$$X = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 0 & 1 & -1 \end{pmatrix}$$

having these eigenvectors as its rows is invertible over  $\mathbb{Q}$  as  $\det X = -2$  and satisfies

$$XCX^{-1} = \text{diag}(0, 1, -1).$$

(b)  $C(f(x)) = \left( \begin{array}{c|c} 0 & I \\ \hline -a_0 & c \end{array} \right)$  where  $I$  is the  $(t-1) \times (t-1)$  identity matrix over  $F$  and

$c = -(a_1, a_2, \dots, a_{t-1})$ . Expanding  $|C(f(x))|$  along col 1 gives

$$|C(f(x))| = (-1)^{t+1}(-a_0)|I| = (-1)^t a_0.$$

The first  $t-1$  rows of  $C(f(x))$  are  $e_2, e_3, \dots, e_t$  which are linearly independent. So  $t \geq \text{rank } C(f(x)) \geq t-1$ . But  $\text{rank } C(f(x)) = t \Leftrightarrow |C(f(x))| \neq 0$ . Therefore  $\text{rank } C(f(x)) = t$  or  $t-1$  according as  $a_0 \neq 0$  or  $a_0 = 0$ .

(c)  $C^2 = \begin{pmatrix} 0 & 1 \\ -b & -a \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -b & -a \end{pmatrix} = \begin{pmatrix} -b & -a \\ ab & a^2 - b \end{pmatrix} = -aC - bI$  and so  $C^2 + aC + bI = 0$ .

Also  $|xI - C| = \begin{vmatrix} x & -1 \\ b & x+a \end{vmatrix} = x^2 + ax + b$ .

$$D^2 = \begin{pmatrix} -b & -a \\ ab & a^2 - b \end{pmatrix} \begin{pmatrix} -b & -a \\ ab & a^2 - b \end{pmatrix} = \begin{pmatrix} b^2 - a^2b & a^3 - 2ab \\ a^3b - 2ab^2 & a^4 - 3a^2b + b^2 \end{pmatrix} = (a^2 - 2b)D - b^2I$$

and so  $D^2 + (2b - a^2)D + b^2I = 0$ .

(d) Suppose  $C(f(x)) \sim D$  where  $D = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_t)$ . Comparing characteristic polynomials using (5.5) and (5.26) gives  $f(x) = (x - \lambda_1)(x - \lambda_2) \cdots (x - \lambda_t)$ .

Suppose  $\lambda_i = \lambda_j$  for some pair of integers  $i, j$  with  $1 \leq i < j \leq t$ . Then  $N = \langle e_i, e_j \rangle$  is a non-cyclic submodule of  $M(D)$ : for all  $v \in N$  we have

$$xv = x(a_i e_i + a_j e_j) = (a_i e_i + a_j e_j)D = a_i \lambda_i e_i + a_j \lambda_j e_j = \lambda_i (a_i e_i + a_j e_j) = \lambda_i v$$

showing that  $v$  does not generate the 2-dimensional subspace  $N$ . But  $M(D)$  is cyclic being isomorphic to the cyclic  $F[x]$ -module  $C(f(x))$  by (5.13). So  $N$  is cyclic by (5.28). This contradiction shows  $\lambda_i \neq \lambda_j$ . Therefore  $f(x)$  has  $t$  distinct zeros in  $F$ .

Conversely suppose  $f(x)$  has  $t$  distinct zeros  $\lambda_1, \lambda_2, \dots, \lambda_t$  in  $F$ . So  $\lambda_1, \lambda_2, \dots, \lambda_t$  are  $t$  distinct eigenvalues of  $C(f(x))$  by (5.26). Denote by  $v_i$  a row eigenvector of  $C(f(x))$  corresponding to  $\lambda_i$  for  $1 \leq i \leq t$ . Let  $X$  be the  $t \times t$  matrix with  $e_i X = v_i$  for  $1 \leq i \leq t$ . Then  $X$  is invertible over  $F$  ( $\lambda_1, \lambda_2, \dots, \lambda_t$  distinct implies  $v_1, v_2, \dots, v_t$  linearly independent) and  $XC(f(x))X^{-1} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_t)$ , i.e.  $C(f(x)) \sim D$ . The polynomial  $x^4 + x = x(x-1)(x-c)(x-c-1)$  splits over  $\mathbb{F}_4$  as its zeros are  $0, 1, c, c+1$ . So  $0, 1, c, c+1$  are the eigenvalues of

$$C(x^4 + x) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix}.$$

Using four corresponding row eigenvectors of  $C(x^4 + x)$  as the rows of  $X$  gives (after some calculation)

$$X = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & c+1 & c & 1 \\ 0 & c & c+1 & 1 \end{pmatrix}$$

such that  $XC(x^4 + x)X^{-1} = \text{diag}(0, 1, c, c+1)$ . In fact there are  $3^4 \times 4! = 81 \times 24 = 1944$  invertible matrices  $X$  over  $\mathbb{F}_4$  with  $XC(x^4 + x)X^{-1}$  being diagonal, there being 3 choices for each of the 4 row eigenvectors which can appear as rows of  $X$  in  $4!$  ways.

## Solution 2

(a)  $xe_1 = e_1 A = (1, 0, 1), x^2 e_1 = x(xe_1) = (1, 0, 1)A = (2, 1, 1)$ . As

$$\begin{vmatrix} e_1 \\ xe_1 \\ x^2 e_1 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 2 & 1 & 1 \end{vmatrix} = -1 \neq 0$$

we see that  $e_1, xe_1, x^2 e_1$  are linearly independent vectors of  $\mathbb{Q}^3$  and so these three vectors span  $\mathbb{Q}^3$ . Each  $v \in \mathbb{Q}^3$  can be expressed

$$v = a_0 e_1 + a_1 x e_1 + a_2 x^2 e_1 = (a_0 + a_1 x + a_2 x^2) e_1$$

showing that  $e_1$  generates the  $\mathbb{Q}[x]$ -module  $M(A)$ . As

$$x^3 e_1 = x(x^2 e_1) = (2, 1, 1)A = (4, 2, 2) \text{ we see } x^3 e_1 = 2x^2 e_1, \text{ i.e. } (x^3 - 2x^2)e_1 = 0.$$

So  $x^3 - 2x^2 = x^2(x-2)$  is the order of  $e_1$  in  $M(A)$ . Combining (5.5), (5.26) and (5.27) gives  $\chi_A(x) = x^2(x-2)$ . Also

$$X = \begin{pmatrix} e_1 \\ xe_1 \\ x^2 e_1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 2 & 1 & 1 \end{pmatrix}$$

is invertible over  $\mathbb{Q}$  and satisfies  $XA = C(\chi_A(x))X$  as

$$\begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 2 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \\ 2 & 1 & 1 \\ 4 & 2 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 2 & 1 & 1 \end{pmatrix}.$$

By (5.23) the vector  $x(x-2)e_1 = (2,1,1) - 2(1,0,1) = (0,1,-1)$  has order  $x$  in  $M(A)$ , i.e.  $(0,1,-1)$  is a row eigenvector of  $A$  with corresponding eigenvalue  $0$ . In the same way  $x^2e_1 = (2,1,1)$  has order  $x-2$  in  $M(A)$ , i.e.  $(2,1,1)$  is a row eigenvector of  $A$  with corresponding eigenvalue  $2$ . Also  $(0,1,-1), (2,1,1)$  are linearly independent.

(b)  $xe_1 = (1,0,0)A = (1,0,1), x^2e_1 = xe_1A = (1,0,1)A = (2,0,2)$  and so  $x^2e_1 = 2xe_1$  which gives  $(x^2 - 2x)e_1 = 0$ . As  $e_1, xe_1$  are linearly independent vectors of  $\mathbb{Q}^3$  we see that  $e_1$  has order  $x^2 - 2x = x(x-2)$  in  $M(A)$ .

$$xe_2 = (0,1,0)A = (1,1,1), x^2e_2 = xe_2A = (1,1,1)A = (3,1,3)$$

and so  $x^2e_2 - 3xe_2 = (0,-2,0) = -2e_2$  which rearranges to  $(x^2 - 3x + 2)e_2 = 0$ . As  $e_2, xe_2$  are linearly independent vectors of  $\mathbb{Q}^3$  we see that  $e_2$  has order

$$x^2 - 3x + 2 = (x-1)(x-2) \text{ in } M(A).$$

As  $xe_3 = (0,0,1)A = (1,0,1), x^2e_3 = xe_3A = (1,0,1)A = (2,0,2)$  we see  $x^2e_3 = 2xe_3$  which gives  $(x^2 - 2x)e_3 = 0$ . As  $e_3, xe_3$  are linearly independent vectors of  $\mathbb{Q}^3$  we see  $e_3$  has order  $x^2 - 2x = x(x-2)$  in  $M(A)$ . As  $xe_2 = (1,1,1) = e_1 + e_2 + e_3$ , by (5.23)  $e_1 + e_2 + e_3$  has order  $(x-1)(x-2)/\gcd\{x, (x-1)(x-2)\} = (x-1)(x-2)$  in  $M(A)$ . So none of  $e_1, e_2, e_3, e_1 + e_2 + e_3$  generate  $M(A)$ .

$x(e_1 + e_2) = (1,1,0)A = (2,1,2)$  and  $x^2(e_1 + e_2) = (2,1,2)A = (5,1,5)$ . As

$$\begin{vmatrix} e_1 + e_2 \\ x(e_1 + e_2) \\ x^2(e_1 + e_2) \end{vmatrix} = \begin{vmatrix} 1 & 1 & 0 \\ 2 & 1 & 2 \\ 5 & 1 & 5 \end{vmatrix} = 3 \neq 0$$

we see  $e_1 + e_2$  generates  $M(A)$ . In fact  $x(x-1)(x-2)(e_1 + e_2) = 0$  and so  $e_1 + e_2$  has order  $x(x-1)(x-2) = \chi_A(x)$  in  $M(A)$ . Also

$$X = \begin{pmatrix} e_1 + e_2 \\ x(e_1 + e_2) \\ x^2(e_1 + e_2) \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 2 & 1 & 2 \\ 5 & 1 & 5 \end{pmatrix}$$

is invertible over  $\mathbb{Q}$  and

$$XAX^{-1} = C(\chi_A(x)) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & -2 & 3 \end{pmatrix}.$$

The eight submodules of  $M(A)$  have generators

$$\begin{aligned} e_1 + e_2 &= (1,1,0), \quad x(e_1 + e_2) = (2,1,2), \quad (x-1)(e_1 + e_2) = (1,0,2), \\ (x-2)(e_1 + e_2) &= (0,-1,2), \quad x(x-1)(e_1 + e_2) = (3,0,3), \quad x(x-2)(e_1 + e_2) = (1,-1,1), \\ (x-1)(x-2)(e_1 + e_2) &= (1,0,-1), \quad x(x-1)(x-2)(e_1 + e_2) = (0,0,0) = 0 \end{aligned}$$

corresponding to the eight monic divisors of  $\chi_A(x)$ .

(c) Does  $e_1$  generate  $M(A)$ ? Now  $xe_1 = (2,2,2), x^2e_1 = (2,2,2)A = (0,0,2)$  and

$$X = \begin{pmatrix} e_1 \\ xe_1 \\ x^2e_1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 2 & 2 \\ 0 & 0 & 2 \end{pmatrix}$$

is invertible over  $\mathbb{Q}$  since  $\det X = 4$ . So  $e_1$  does generate  $M(A)$  which is therefore cyclic. As  $x^3e_1 = (0,0,2)A = (-2,-2,-2)$  we see  $x^3e_1 = -xe_1$ . Therefore  $e_1$  has order

$x^3 + x$  in  $M(A)$  since  $(x^3 + x)e_1 = 0$  and  $e_1, xe_1, x^2e_1$  are linearly independent. So  $\chi_A(x) = x^3 + x = x(x^2 + 1)$  and  $X$  above satisfies

$$XAX^{-1} = C(x^3 + x) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}.$$

Then  $(x^2 + 1)e_1 = (1, 0, 2)$  has order  $x$  and  $xe_1/2 = (1, 1, 1)$  has order  $x^2 + 1$ . Let  $N_1 = \langle (1, 0, 2) \rangle$  and  $N_2 = \langle (1, 1, 1), x(1, 1, 1) \rangle = \langle (1, 1, 1), (0, 0, 1) \rangle$ , i.e.  $N_1$  and  $N_2$  are the cyclic submodules of  $M(A)$  generated by  $(1, 0, 2)$  and  $(1, 1, 1)$  respectively. As  $(1, 0, 2), (1, 1, 1), (0, 0, 1)$  form a basis of  $\mathbb{Q}^3$  we see  $N_1 \oplus N_2 = M(A)$  by (5.18).

(d) Suppose  $f(x)$  to be irreducible over  $F$ . Then  $f(x)$  has just two monic divisors over  $F$  namely 1 and  $f(x)$ . By (5.28)  $M(C)$  has just two submodules namely

$$M(C) = \langle 1e_1 \rangle \text{ and } \{0\} = \langle f(x)e_1 \rangle = \langle 0 \rangle.$$

Suppose  $f(x)$  is reducible over  $F$ . Then  $f(x) = (x - c)g(x)$  where  $c \in F$  and  $g(x)$  is a monic polynomial of degree 2 over  $F$ . Suppose  $g(x)$  is irreducible over  $F$ .

Then  $f(x)$  has exactly 4 monic divisors over  $F$  namely 1,  $x - c$ ,  $g(x)$ ,  $f(x)$ . By (5.28) we see  $M(C)$  has exactly 4 submodules namely the cyclic modules generated by one of  $e_1, (x - c)e_1, g(x)e_1, f(x)e_1$ .

Suppose  $g(x)$  is reducible over  $F$ . Then  $g(x) = (x - c')(x - c'')$  where  $c', c'' \in F$  and so  $f(x) = (x - c)(x - c')(x - c'')$ . Consider the case  $c = c' = c''$ . Then  $f(x)$  has exactly 4 monic divisors over  $F$  namely 1,  $x - c$ ,  $(x - c)^2$ ,  $(x - c)^3$ . By (5.28)

$M(C)$  has exactly 4 submodules namely the cyclic modules generated by one of

$$e_1, (x - c)e_1, (x - c)^2e_1, (x - c)^3e_1.$$

Suppose two of  $c, c', c''$  are equal. We may suppose  $c \neq c' = c''$ . Then  $f(x)$  has exactly 6 monic divisors over  $F$  namely

$$1, x - c, x - c', (x - c)(x - c'), (x - c')^2, (x - c)(x - c')^2.$$

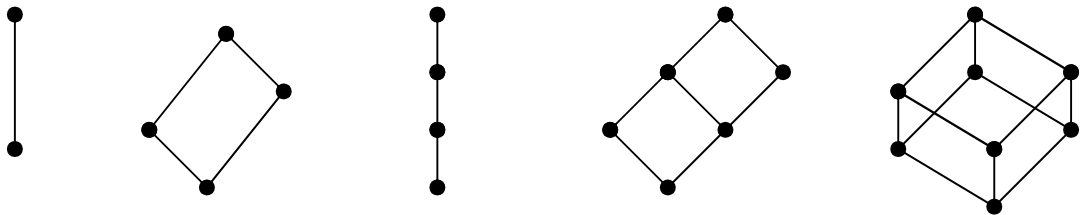
By (5.28)  $M(C)$  has exactly 6 submodules namely the cyclic modules generated by one of

$$e_1, (x - c)e_1, (x - c')e_1, (x - c)(x - c')e_1, (x - c')^2e_1, (x - c)(x - c')^2e_1.$$

Finally suppose no two of  $c, c', c''$  are equal. Then  $f(x)$  has exactly 8 monic divisors over  $F$ . By (5.28)  $M(C)$  has exactly 8 submodules namely the cyclic modules generated by one of

$$e_1, (x - c)e_1, (x - c')e_1, (x - c'')e_1, (x - c)(x - c')e_1, \\ (x - c)(x - c'')e_1, (x - c')(x - c'')e_1, (x - c)(x - c')(x - c'')e_1.$$

The five lattice diagrams are:



(i) The first lattice diagram cannot occur in the case  $F = \mathbb{R}$  as there are no irreducible polynomials of degree 3 over  $\mathbb{R}$ .

- (ii) The first two lattice diagrams cannot occur in the case  $F = \mathbb{C}$  as there are no irreducible polynomials of degree 3 or 2 over  $\mathbb{C}$ .
- (iii) The last lattice diagram cannot occur in the case  $F = \mathbb{Z}_2$  as  $\mathbb{Z}_2$  does not contain three distinct elements  $c, c', c''$ .

(e) Working in the  $F[x]$ -module  $M(A)$  we show

$$\langle e_1, e_2, \dots, e_s \rangle \subseteq \langle e_1, xe_1, \dots, x^{s-1}e_1 \rangle \text{ for } 1 \leq s \leq t \text{ by induction on } s.$$

This inclusion is true for  $s=1$ . Suppose the above inclusion is true for some  $s$  with  $1 \leq s < t$ . Row  $s$  of  $A$  gives  $xe_s = u + a_{s,s+1}e_{s+1}$  where  $u \in \langle e_1, e_2, \dots, e_s \rangle$ . But

$e_s \in \langle e_1, xe_1, \dots, x^{s-1}e_1 \rangle$  implies  $xe_s \in \langle xe_1, x^2e_1, \dots, x^se_1 \rangle \subseteq \langle e_1, xe_1, x^2e_1, \dots, x^se_1 \rangle$ . So  $xe_s, u$  and hence  $e_{s+1} = a_{s,s+1}^{-1}(xe_s - u)$  all belong to  $\langle e_1, xe_1, x^2e_1, \dots, x^se_1 \rangle$ . Therefore  $\langle e_1, e_2, \dots, e_{s+1} \rangle \subseteq \langle e_1, xe_1, \dots, x^se_1 \rangle$  completing the inductive step. We conclude

$$F^t = \langle e_1, e_2, \dots, e_t \rangle = \langle e_1, xe_1, \dots, x^{t-1}e_1 \rangle \text{ showing that } e_1 \text{ generates } M(A).$$

The given matrix  $A$  satisfies the condition  $a_{i,i+1} \neq 0$  for  $i=1, 2, 3$  and  $a_{ij} = 0$  for  $j > i+1$  (i.e.  $a_{13} = a_{14} = a_{24} = 0$ ). So  $e_1 = (1, 0, 0, 0)$  generates  $M(A)$  by the above paragraph. By (5.27)  $e_1$  has order  $\chi_A(x)$  in  $M(A)$  and

$$\chi_A(x) = |xI - A| = \begin{vmatrix} x & -1 \\ -2 & x \end{vmatrix}^2 = (x^2 - 2)^2.$$

As  $x^2 - 2$  is irreducible over  $\mathbb{Q}$  the  $\mathbb{Q}[x]$ -module  $M(A)$  has exactly three submodules:  $M(A), N, 0$  corresponding to the three monic divisors of  $(x^2 - 2)^2$  over  $\mathbb{Q}$  namely  $1, x^2 - 2, (x^2 - 2)^2$ . A generator of  $N$  is

$$(x^2 - 2)e_1 = (2, 0, 1, 0) - (2, 0, 0, 0) = (0, 0, 1, 0) = e_3.$$

As  $e_3$  has order  $x^2 - 2$  each non-zero element of  $N = \langle e_3 \rangle = \langle e_3, xe_3 \rangle = \langle e_3, e_4 \rangle$  is a generator of  $N$ . Each element  $v$  of  $M(A)$  with  $v \notin N$  is a generator of  $M(A)$ .

(f) Suppose  $e_1$  generates  $M(A)$ . Then

$$\begin{vmatrix} e_1 \\ xe_1 \\ x^2e_1 \end{vmatrix} = \begin{vmatrix} 1 & 0 & 0 \\ 0 & a & b \\ 0 & ac & ad \end{vmatrix} = a(ad - bc) \neq 0.$$

So  $a \neq 0$  and  $\begin{vmatrix} a & b \\ c & d \end{vmatrix} \neq 0$ . Conversely suppose  $a \neq 0$  and  $\begin{vmatrix} a & b \\ c & d \end{vmatrix} \neq 0$ . Then

$e_1, xe_1, x^2e_1$  are linearly independent vectors and so span  $F^3$ , i.e.  $e_1$  generates  $M(A)$ .

### Solution 3

(a) Write  $d_0(x) = \gcd\{f(x), d(x)\}$ . Then

$$(d(x)/d_0(x))f(x)v = (f(x)/d_0(x))d(x)v = (f(x)/d_0(x))0 = 0$$

which shows that the monic polynomial  $d(x)/d_0(x)$  belongs to the order ideal  $K$  of  $f(x)v$  in  $M$ . So  $K$  is non-zero and so has a unique monic generator  $g(x)$  by (4.4). By (5.11)  $g(x)$  is the order of  $f(x)v$  in  $M$  and  $g(x) \mid d(x)/d_0(x)$ . On the other hand  $g(x)f(x)v = 0$  which shows that  $g(x)f(x)$  belongs to the order ideal  $\langle d(x) \rangle$  of  $v$  in  $M$ . So there is  $h(x) \in F[x]$  with  $g(x)f(x) = h(x)d(x)$ . Dividing through by  $d_0(x)$  gives  $g(x)(f(x)/d_0(x)) = h(x)(d(x)/d_0(x))$ . As

$\gcd\{f(x)/d_0(x), d(x)/d_0(x)\} = 1$  we conclude  $d(x)/d_0(x) \mid g(x)$ . Therefore  $g(x) = d(x)/d_0(x)$ , i.e.  $f(x)v$  has order  $d(x)/\gcd\{f(x), d(x)\}$  in  $M$ .

(b) Use (5.23) throughout. As  $x^3 + x^2 + x + 1 = (x^2 + 1)(x + 1)$  we see  $\gcd\{x^2 + x, x^3 + x^2 + x + 1\} = x + 1$  and so  $(x^2 + x)v_0$  has order  $(x^3 + x^2 + x + 1)/(x + 1) = x^2 + 1$  in  $M$ . As  $x^4 + 2x^2 + 1 = (x^2 + 1)^2$  we see  $\gcd\{x^4 + 2x^2 + 1, x^3 + x^2 + x + 1\} = x^2 + 1$  and so  $(x^4 + 2x^2 + 1)v_0$  has order  $(x^3 + x^2 + x + 1)/(x^2 + 1) = x + 1$  in  $M$ . As  $2x^2 + x - 1 = (x + 1)(2x - 1)$  we see  $\gcd\{2x^2 + x - 1, x^3 + x^2 + x + 1\} = x + 1$  and so  $(2x^2 + x - 1)v_0$  has order  $\gcd\{2x^2 + x - 1, x^3 + x^2 + x + 1\} = x + 1$  in  $M$ . As

$$x^9 - x = x(x^8 - 1) = x(x^4 - 1)(x^4 + 1) = x(x - 1)(x^3 + x^2 + x + 1)(x^4 + 1)$$

we see  $(x^9 - x)v_0 = 0$  which has order 1. As

$$x^7 - x^6 + x^5 - x^4 = x^4(x^3 - x^2 + x - 1) = x^4(x - 1)(x^2 + 1)$$

we see  $\gcd\{x^7 - x^6 + x^5 - x^4, x^3 + x^2 + x + 1\} = x^2 + 1$  and so  $(x^7 - x^6 + x^5 - x^4)v_0$  has order  $(x^3 + x^2 + x + 1)/(x^2 + 1) = x + 1$  in  $M$ .

(c)  $M$  is cyclic with generator  $v_0 = K + \bar{1}$  of order  $x^4 + x = x(x + \bar{1})(x^2 + x + \bar{1})$ . By (5.23) the order of  $K + x = xv_0$  in  $M$  is

$$(x^4 + x)/\gcd\{x, x^4 + x\} = (x^4 + x)/x = x^3 + \bar{1}.$$

In the same way the order of  $K + x^2 = x^2v_0$  in  $M$  is

$$(x^4 + x)/\gcd\{x^2, x^4 + x\} = (x^4 + x)/x = x^3 + \bar{1},$$

the order of  $K + \bar{1} + x + x^2 = (x^2 + x + \bar{1})v_0$  in  $M$  is

$$(x^4 + x)/\gcd\{x^2 + x + \bar{1}, x^4 + x\} = (x^4 + x)/(x^2 + x + \bar{1}) = x^2 + x.$$

Suppose  $v = K + f(x)$  satisfies  $M = \langle v \rangle$ . We may assume  $\deg f(x) \leq 3$ . Then  $\gcd\{f(x), x(x + \bar{1})(x^2 + x + \bar{1})\} = \bar{1}$ . Now  $\deg f(x) \neq 1, 2$ ,  $f(x)$  has constant term  $\bar{1}$  and  $f(x)$  has an odd number of terms. Therefore  $f(x) = \bar{1}, x^3 + x + \bar{1}, x^3 + x^2 + \bar{1}$ , each of which satisfies  $M = \langle v \rangle$ .

(d) There is  $f(x) \in F[x]$  with  $v = f(x)v_0$ . By (5.23) the order of  $v$  in  $M$  is

$$d(x) = d_0(x)/\gcd\{f(x), d_0(x)\} \text{ and so } d(x) \mid d_0(x). \text{ Suppose } M = \langle v \rangle.$$

Interchanging the roles of  $v$  and  $v_0$  we obtain  $d_0(x) \mid d(x)$ . So  $d(x) = d_0(x)$  as  $d(x)$  and  $d_0(x)$  are both monic and each is a divisor of the other. Conversely suppose  $d(x) = d_0(x)$ . So  $\gcd\{f(x), d_0(x)\} = 1$ . By (4.6) there are  $a_1(x), a_2(x) \in F[x]$  with  $a_1(x)f(x) + a_2(x)d_0(x) = 1$ . So

$$v_0 = (a_1(x)f(x) + a_2(x)d_0(x))v_0 = a_1(x)f(x)v_0 + a_2(x)d_0(x)v_0 = a_1(x)v$$

as  $d_0(x)v_0 = 0$ . But  $v_0 = a_1(x)v$  gives  $\langle v \rangle \supseteq \langle v_0 \rangle = M$  and so  $\langle v \rangle = M$ .

By the above theory the order of  $v_0$  in  $M'$  is not monic, i.e. the order of  $v_0$  in  $M'$  is zero. So  $M' = \langle v_0 \rangle$  is free of rank 1. For example  $M' = \mathbb{Q}[x]$ ,  $v_0 = 1$ ,  $v_1 = x$ ; both  $v_0$  and  $v_1$  have order 0 in  $M'$ ,  $M' = \langle v_0 \rangle$ , but  $\langle v_1 \rangle = \{f(x) \in \mathbb{Q}[x] : f(0) = 0\} \neq M'$ .

(e) By (4.6) there are  $a_1(x), a_2(x) \in F[x]$  with  $a_1(x)f_1(x) + a_2(x)f_2(x) = 1$ . Let  $v \in M$ . There is  $f(x) \in F[x]$  with  $v = f(x)v_0$  and so

$$v = f(x)(a_1(x)f_1(x) + a_2(x)f_2(x))v_0 = v_1 + v_2 \in N_1 + N_2$$



where  $v_1 = f(x)a_1(x)f_1(x)v_0 \in N_1$  and  $v_2 = f(x)a_2(x)f_2(x)v_0 \in N_2$ . So  $M = N_1 + N_2$ . Suppose  $v'_1 + v'_2 = 0$  where  $v'_1 \in N_1, v'_2 \in N_2$ . There are  $b_1(x), b_2(x) \in F[x]$  with  $v'_1 = b_1(x)f_1(x)v_0, v'_2 = b_2(x)f_2(x)v_0$ . So  $f_1(x)v'_2 = b_2(x)f_1(x)f_2(x)v_0 = 0$  and  $f_2(x)v'_1 = b_1(x)f_1(x)f_2(x)v_0 = 0$  as  $v_0$  has order  $f_1(x)f_2(x)$  in  $M$ . Hence  $f_1(x)v'_1 = f_1(x)v'_1 + f_1(x)v'_2 = f_1(x)(v_1 + v_2) = f_1(x) \times 0 = 0$ . Therefore

$$\begin{aligned} v'_1 &= (a_1(x)f_1(x) + a_2(x)f_2(x))v'_1 = \\ &= a_1(x)f_1(x)v'_1 + a_2(x)f_2(x)v'_1 = a_1(x) \times 0 + a_2(x) \times 0 = 0 \end{aligned}$$

and so  $v'_2 = 0 + v'_2 = v'_1 + v'_2 = 0$  also. By (2.14) we see  $N_1$  and  $N_2$  are independent showing  $M = N_1 \oplus N_2$ .

Now  $g_1(x)g_2(x)v = g_1(x)g_2(x)(v_1 + v_2) = g_2(x) \times 0 + g_1(x) \times 0 = 0$  since  $g_1(x)v_1 = g_2(x)v_2 = 0$ . So  $v$  has order  $d(x)$  in  $M$  where  $d(x) \mid g_1(x)g_2(x)$ . As  $0 = d(x)v = d(x)(v_1 + v_2) = d(x)v_1 + d(x)v_2 \in N_1 + N_2$  we see  $d(x)v_1 = d(x)v_2 = 0$  using the independence of  $N_1$  and  $N_2$ . Therefore  $g_1(x) \mid d(x)$  and  $g_2(x) \mid d(x)$ . As  $g_1(x) \mid f_2(x)$  and  $g_2(x) \mid f_1(x)$  we see  $\gcd\{g_1(x), g_2(x)\} = 1$ . So  $g_1(x)g_2(x) \mid d(x)$ . As  $g_1(x)g_2(x)$  and  $d(x)$  are both monic we conclude  $d(x) = g_1(x)g_2(x)$ .

Suppose  $M = \langle v \rangle$ . Then  $v$  has the same order as  $v_0$  in  $M$ , namely  $f_1(x)f_2(x)$ . By the previous paragraph  $v_1$  has order  $f_2(x)$  in  $M$  and  $v_2$  has order  $f_1(x)$  in  $M$ . So  $N_i = \langle v_i \rangle$  for  $i=1, 2$ . Conversely suppose  $N_i = \langle v_i \rangle$  for  $i=1, 2$ . Then  $v_1$  and  $f_1(x)v_0$  have the same order in  $M$ , namely  $f_2(x)$ . Also  $v_2$  and  $f_2(x)v_0$  have the same order in  $M$ , namely  $f_1(x)$ . By the previous paragraph  $v = v_1 + v_2$  has order  $f_1(x)f_2(x)$  in  $M$  and so  $M = \langle v \rangle$ .

(f) As  $v_0$  has order  $x^4 + x^2 = x^2(x^2 + \bar{1})$  in  $M$ , by (5.23) we see  
 (i)  $x^2v_0$  has order  $(x^4 + x^2)/\gcd\{x^2, x^4 + x^2\} = (x^4 + x^2)/x^2 = x^2 + \bar{1}$  in  $M$ ,  
 (ii)  $(x^2 + \bar{1})v_0$  has order  $(x^4 + x^2)/\gcd\{x^2 + \bar{1}, x^4 + x^2\} = (x^4 + x^2)/(x^2 + \bar{1}) = x^2$  in  $M$ . Yes  $M = N_1 \oplus N_2$  by (d) above. As  $N_1 = \langle x^2v_0, x^3v_0 \rangle$  is a 2-dimensional vector space over  $\mathbb{Z}_3$  by (5.29), we see  $|N_1| = 3^2$ . As the order of the generator  $x^2v_0$  of the  $\mathbb{Z}_3$ -module  $N_1$  is irreducible over  $\mathbb{Z}_3$ , all nonzero vectors  $v_1$  of  $N_1$  satisfy  $N_1 = \langle v_1 \rangle$ ; so there are  $3^2 - 1 = 8$  such vectors  $v_1$ . As  $N_2 = \langle (x^2 + 1)v_0 \rangle$ , by (5.29)  $N_2$  has exactly 3 submodules:  $\{0\}, \langle x(x^2 + 1)v_0 \rangle$  and  $N_2$  itself, which are vector spaces over  $\mathbb{Z}_3$  of dimension 0, 1, 2 respectively. So there are  $3^2 - 3^1 = 6$  vectors  $v_2$  with  $N_2 = \langle v_2 \rangle$ , namely the vectors of  $N_2$  not in  $\langle x(x^2 + 1)v_0 \rangle$ . The  $8 \times 6 = 48$  vectors  $v = v_1 + v_2$  in  $M$  where  $N_1 = \langle v_1 \rangle$  and  $N_2 = \langle v_2 \rangle$ , and only these vectors, by (d) above, have order  $(x^2 + 1)x^2$  in  $M$  and satisfy  $M = \langle v \rangle$ .

#### Solution 4

(a) Consider  $f_1(x), f_2(x) \in K_N$ . Then  $f_1(x)v_0 \in N$  and  $f_2(x)v_0 \in N$ . As  $N$  is closed under addition we see  $(f_1(x) + f_2(x))v_0 = f_1(x)v_0 + f_2(x)v_0 \in N$ , i.e.  $f_1(x) + f_2(x) \in K_N$  showing  $K_N$  to be closed under addition. As  $0(x)v_0 = 0 \in N$  we see  $0(x) \in K_N$ , i.e.  $K_N$  contains the zero polynomial. As  $N$  is closed under negation

we have  $(-f_1(x))v_0 = -f_1(x)v_0 \in N$  showing that  $-f_1(x) \in K_N$ , i.e.  $K_N$  is closed under negation. As  $N$  is closed under polynomial multiplication we obtain  $(g(x)f_1(x))v_0 = g(x)(f_1(x)v_0) \in N$ , i.e.  $g(x)f_1(x) \in K_N$  showing  $K_N$  to be closed under polynomial multiplication. So  $K_N$  is an ideal (4.3) of  $F[x]$ . As

$$d_0(x)v_0 = 0 \in N \text{ we see } d_0(x) \in K_N \text{ and so } \langle d_0(x) \rangle \subseteq K_N.$$

Suppose  $K_N = \langle d_0(x) \rangle$  and  $v \in \langle v_0 \rangle \cap N$ . Then  $v = f(x)v_0 \in N$  for some  $f(x) \in F[x]$ . Therefore  $f(x) \in K_N$  and so  $d_0(x) \mid f(x)$ . Hence  $f(x)v_0 = 0$ , i.e.  $\langle v_0 \rangle \cap N = \{0\}$ . Conversely suppose  $\langle v_0 \rangle \cap N = \{0\}$  and consider  $f(x) \in K_N$ . So  $f(x)v_0 \in N$ . As  $f(x)v_0 \in \langle v_0 \rangle$  we see  $f(x)v_0 = 0$ , i.e.  $f(x) \in \langle d_0(x) \rangle$  and so  $K_N = \langle d_0(x) \rangle$ .

Suppose  $N_1 \subseteq N_2$  and consider  $f(x) \in K_{N_1}$ . Then  $f(x)v_0 \in N_1$  and so  $f(x)v_0 \in N_2$ . Therefore  $f(x) \in K_{N_2}$  showing  $K_{N_1} \subseteq K_{N_2}$ .

Suppose  $M = \langle v_0 \rangle$  and  $K_{N_1} \subseteq K_{N_2}$ . Let  $v \in N_1$ . There is  $f(x) \in F[x]$  with  $v = f(x)v_0$ . Therefore  $f(x) \in K_{N_1}$  and so  $f(x) \in K_{N_2}$  also. This means  $f(x)v_0 \in N_2$ , i.e.  $v \in N_2$ . So  $N_1 \subseteq N_2$ .

(b) Let  $u, v \in N_K$ . There are  $f(x), g(x) \in K$  with  $u = f(x)v_0$ ,  $v = g(x)v_0$ . As  $f(x) + g(x) \in K$  we see  $u + v = f(x)v_0 + g(x)v_0 = (f(x) + g(x))v_0 \in N_K$  showing that  $N_K$  is closed under addition. As  $0(x) \in K$  we see  $0 = 0(x)v_0 \in N_K$ . As  $-f(x) \in K$  we see  $-u = -(f(x)v_0) = (-f(x))v_0 \in N_K$ . For all  $h(x) \in F[x]$  we have  $h(x)f(x) \in K$  and so  $h(x)u = h(x)f(x)v_0 \in N_K$ . Therefore  $N_K$  is a submodule of  $M$ . By (4.4) we see  $K$  is a principal ideal with generator  $d(x)$ . Then  $N_K$  is cyclic with generator  $d(x)v_0$ .

Suppose the ideals  $K_1$  and  $K_2$  of  $F[x]$  satisfy  $K_1 \subseteq K_2$  and let  $v \in N_{K_1}$ . Then  $v = f(x)v_0$  where  $f(x) \in K_1$ . As  $f(x) \in K_2$  we see  $v \in N_{K_2}$ . So  $N_{K_1} \subseteq N_{K_2}$ .

Conversely suppose  $N_{K_1} \subseteq N_{K_2}$  where  $K_1$  and  $K_2$  are ideals of  $F[x]$  containing  $\langle d_0(x) \rangle$ . By (4.4) there are  $d_1(x)$  and  $d_2(x)$  with  $K_1 = \langle d_1(x) \rangle$ ,  $K_2 = \langle d_2(x) \rangle$ . Also  $d_1(x) \mid d_0(x)$  and  $d_2(x) \mid d_0(x)$ . Then  $N_{K_1} = \langle d_1(x)v_0 \rangle$  and  $N_{K_2} = \langle d_2(x)v_0 \rangle$ .

As  $d_1(x)v_0 \in \langle d_2(x)v_0 \rangle$  there is  $q(x) \in F[x]$  with  $d_1(x)v_0 = q(x)d_2(x)v_0$ , i.e.  $(d_1(x) - q(x)d_2(x))v_0 = 0$ . As  $v_0$  has order  $d_0(x)$  in  $M$  we see  $d_0(x) \mid (d_1(x) - q(x)d_2(x))$ , i.e. there is  $q'(x) \in F[x]$  with  $d_1(x) = q'(x)d_0(x) + q(x)d_2(x) \in \langle d_0(x) \rangle + \langle d_2(x) \rangle = \langle d_2(x) \rangle = K_2$ . Therefore  $K_1 = \langle d_1(x) \rangle \subseteq K_2$ .

(c) As  $K_N = \{f(x) \in F[x] : f(x)v_0 \in N\}$  and  $N_{K_N} = \{f(x)v_0 : f(x) \in K_N\}$  we see  $N_{K_N} \subseteq N$ . Consider  $v \in N$ . As  $M = \langle v_0 \rangle$  there is  $g(x) \in F[x]$  with  $v = g(x)v_0$ . So  $g(x) \in K_N$  and hence  $v \in N_{K_N}$ . Therefore  $N \subseteq N_{K_N}$  and so  $N = N_{K_N}$ .

As  $N_K = \{f(x)v_0 : f(x) \in K\}$  and  $K_{N_K} = \{f(x) \in F[x] : f(x)v_0 \in N_K\}$  we see

$K \subseteq K_{N_K}$ . Consider  $f(x) \in K_{N_K}$ . Then  $f(x)v_0 \in N_K$  and so there is  $g(x) \in K$  with

$f(x)v_0 = g(x)v_0$ . As above there is  $q'(x) \in F[x]$  with  $f(x) = q'(x)d_0(x) + g(x) \in K$  as  $d_0(x) \in K$ . Therefore  $K_{N_K} \subseteq K$  and so  $K = K_{N_K}$ .

$N \rightarrow K_N$  is a mapping  $\mathbb{L} \rightarrow \mathbb{L}'$  as  $d_0(x) \in K_N$  since  $d_0(x)v_0 = 0 \in N$  and this mapping is inclusion-preserving by (a) above. Using the foregoing theory we see that  $K \rightarrow N_K$  is the inverse mapping  $\mathbb{L}' \rightarrow \mathbb{L}$  and it is inclusion-preserving by (b) above.

So both these mappings are bijections.

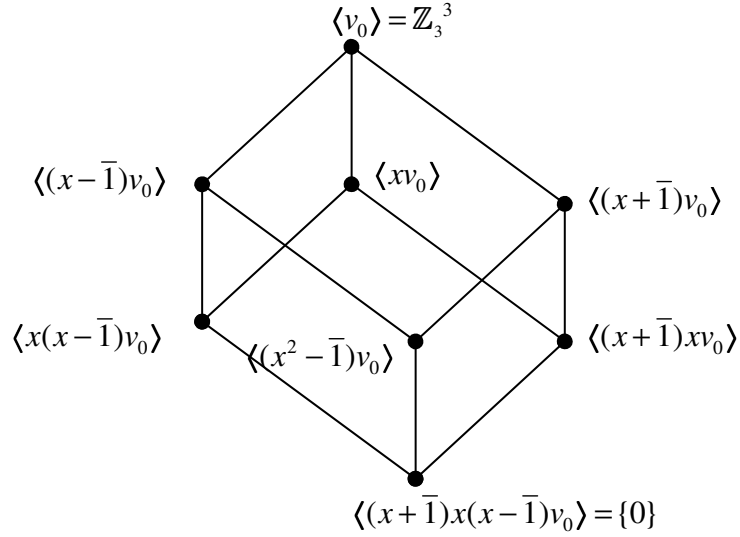
(d) The eight submodules  $N$  of  $M$  have generators

$$v_0, (x - \bar{1})v_0, xv_0, (x + \bar{1})v_0, (x^2 + x)v_0, (x^2 - \bar{1})v_0, (x^2 - x)v_0, (x^3 - x)v_0$$

corresponding to the monic generators

$$\bar{1}, x - \bar{1}, x, x + \bar{1}, x^2 + x, x^2 - \bar{1}, x^2 - x, x^3 - x$$

of the eight ideals  $K_N$  of  $\mathbb{Z}_3[x]$  with  $\langle x^3 - x \rangle \subseteq K_N$ . The lattice diagram of these submodules is:



The generators  $v$  of  $M$  are those elements of  $M$  not in any proper (i.e.  $\neq M$ ) submodule  $N$ . As a vector space of dimension  $t$  over  $\mathbb{Z}_3$  has exactly  $3^t$  vectors the sieve formula gives  $3^3 - 3^2 - 3^2 - 3^2 + 3 + 3 + 3 - 1 = 8$  for the number of  $v$  with  $M = \langle v \rangle$ .

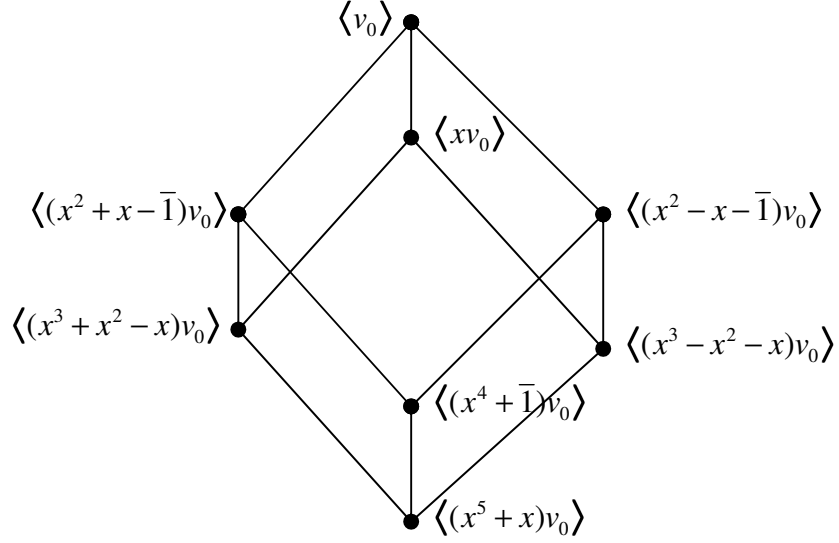
(e) Notice  $x^5 + x = x(x^2 + x - \bar{1})(x^2 - x - \bar{1})$  is the factorisation of  $x^5 - x$  into monic irreducible polynomials over  $\mathbb{Z}_3$ . The eight submodules  $N$  of  $M$  have generators

$$v_0, xv_0, (x^2 + x - \bar{1})v_0, (x^2 - x - \bar{1})v_0, (x^3 + x^2 - x)v_0, \\ (x^4 + \bar{1})v_0, (x^3 - x^2 - x)v_0, (x^5 + x)v_0$$

corresponding to the monic generators

$$1, x, (x^2 + x - \bar{1}), (x^2 - x - \bar{1}), (x^3 + x^2 - x), (x^4 + \bar{1}), (x^3 - x^2 - x), (x^5 + x)$$

of the eight ideals  $K_N$  of  $\mathbb{Z}_3[x]$  with  $\langle x^5 + x \rangle \subseteq K_N$ . The lattice diagram of these submodules is:



The generators  $v$  of  $M$  are those elements of  $M$  not in any proper (i.e.  $\neq M$ ) submodule  $N$ . As a vector space of dimension  $t$  over  $\mathbb{Z}_3$  has exactly  $3^t$  vectors, the sieve formula gives  $3^5 - 3^4 - 3^3 - 3^3 + 3^2 + 3^2 + 3 - 1 = 128$  vectors  $v$  with  $M = \langle v \rangle$ .

### Solution 5

(a) Consider  $K + g_1(x)$  and  $K + g_2(x)$  in  $N$ . Now  $f(x)g_1(x) \in K$ ,  $f(x)g_2(x) \in K$  and so  $f(x)(g_1(x) + g_2(x)) = f(x)g_1(x) + f(x)g_2(x) \in K$  as the ideal  $K$  is closed under addition. Therefore

$$(K + g_1(x)) + (K + g_2(x)) = K + (g_1(x) + g_2(x)) \in N$$

showing that  $N$  is closed under addition. Also  $-(K + g_1(x)) = K - g_1(x) \in N$  as  $f(x)(-g_1(x)) = -f(x)g_1(x) \in K$ , i.e.  $N$  is closed under negation as  $K$  is closed under negation. As  $K$  contains the zero polynomial we see  $f(x) \times 0 \in K$  showing  $K = K + 0 \in N$ , i.e.  $N$  contains the 0-element  $K$  of  $F[x]/K$ . So  $N$  is an additive subgroup of the additive group of the quotient ring  $F[x]/K$ .

Consider  $g(x) \in F[x]$ . Then  $f(x)g(x)g_1(x) = g(x)f(x)g_1(x) \in K$  by (4.3) and so  $g(x)(K + g_1(x)) = K + g(x)g_1(x) \in N$ . So  $N$  is a submodule of  $M$  by (5.14).

Write  $N' = \langle K + q(x) \rangle$  and so  $N'$  is the submodule of  $M$  generated by  $K + q(x)$ .

There is  $f'(x) \in F[x]$  with  $f(x) = f'(x)d(x)$  and so

$f(x)q(x) = f'(x)q(x)d(x) = f'(x)d_0(x) \in K$  showing  $K + q(x) \in N$ . The generator of the cyclic submodule  $N'$  belongs to the submodule  $N$  and so  $N' \subseteq N$ .

Consider a typical element  $K + g_1(x)$  of  $N$ . By (4.6) there are  $a(x), b(x) \in F[x]$  with  $a(x)f(x) + b(x)d_0(x) = d(x)$ . Multiplying through by  $g_1(x)$  gives

$d_0(x) \mid d(x)g_1(x)$  as  $d_0(x) \mid f(x)g_1(x)$ . So  $q(x)d(x) \mid d(x)g_1(x)$ . Dividing through by the monic polynomial  $d(x)$  now gives  $q(x) \mid g_1(x)$ . Therefore  $K + g_1(x)$  is a

polynomial multiple of  $K + q(x)$ , i.e.  $N \subseteq N'$ . The conclusion is  $N = N'$  and so  $N$  is cyclic with generator  $K + q(x)$ .

(b) The reader should realise that this question is a module-isomorphic version of (a) above. As  $N$  is the kernel of the linear mapping of  $F^t$  in which  $v \rightarrow v f(A)$  for all

$v \in F^t$ , we see that  $N$  is a subspace of  $F^t$ . So  $(N, +)$  is a subspace of  $(F^t, +)$ . Let  $g(x) \in F[x]$ . Then  $f(x)g(x) = g(x)f(x)$  and so  $f(A)g(A) = g(A)f(A)$ . Suppose  $v \in N$ . Working in the  $F[x]$ -module  $M(A)$  we have

$$g(x)vf(A) = vg(A)f(A) = vf(A)g(A) = 0 \times g(A) = 0$$

showing  $g(x)v \in N$ . From (5.14) we conclude that  $N$  is a submodule of  $M(A)$ .

Write  $N' = \langle q(x)v_0 \rangle$ . There is  $f'(x) \in F[x]$  with  $f(x) = f'(x)d(x)$ . So

$$(q(x)v_0)f(A) = f(x)q(x)v_0 = f'(x)d(x)q(x)v_0 = f'(x)\chi_A(x)v_0 = f'(x) \times 0 = 0$$

showing that the generator  $q(x)v_0$  of  $N'$  belongs to  $N$ . Therefore  $N' \subseteq N$ .

Conversely suppose  $v \in N$ . As  $v_0$  generates  $M(A)$  there is  $g(x) \in F[x]$  with

$$v = g(x)v_0. \text{ So } f(x)g(x)v_0 = vf(A) = 0 \text{ and so } \chi_A(x) \mid f(x)g(x), \text{ i.e.}$$

$$q(x)d(x) \mid f'(x)d(x)g(x). \text{ Cancelling the monic polynomial } d(x) \text{ gives}$$

$$q(x) \mid f'(x)g(x). \text{ As } \gcd\{f'(x), q(x)\} = 1 \text{ we conclude } q(x) \mid g(x). \text{ So } v = g(x)v_0 \text{ is}$$

a polynomial multiple of the generator  $q(x)v_0$  of  $N'$ . Therefore  $N \subseteq N'$ . As above

the conclusion is  $N = N'$  and so  $N$  is cyclic with generator  $q(x)v_0$ .

As  $q(x)v_0$  has order  $d(x)$  in  $M(A)$  we see  $\dim N = \dim N' = \deg d(x)$  by (5.29).

As  $\dim N = t - \text{rank } f(A)$  we conclude  $\text{rank } f(A) = t - \deg d(x)$ .

(c) As  $C = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -1 & 1 \end{pmatrix}$  we see  $C - I = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 1 & -1 & 0 \end{pmatrix}$  which has rank 2. As

$$C^2 = \begin{pmatrix} 0 & 0 & 1 \\ 1 & -1 & 1 \\ 1 & 0 & 0 \end{pmatrix} \text{ we obtain } C^2 + I = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix} \text{ which has rank 1. As}$$

$$\chi_C(x) = (x^2 + 1)(x - 1) \text{ and } \gcd\{x^{100}, (x^2 + 1)(x - 1)\} = 1, \text{ on taking}$$

$$A = C, f(x) = x^{100} \text{ in (b) above, we see } f(C) = C^{100} \text{ has rank } t - \deg d(x) = 3 - 0 = 3.$$

Now let  $f(x) = x^{100} - x^{50} = x^{50}(x^{50} - 1)$ . As

$$1^{50} - 1 = 0, \text{ but } i^{50} - 1 = (i^2)^{25} - 1 = (-1)^{25} - 1 = -2 \neq 0 \text{ we see that } x - 1 \text{ is a factor of}$$

$$f(x) \text{ but } x^2 + 1 \text{ is not a factor of } f(x). \text{ Therefore } \gcd\{f(x), \chi_C(x)\} = x - 1 \text{ and}$$

$$f(C) = C^{100} - C^{50} \text{ has rank } 3 - 1 = 2.$$

Now take  $f(x) = x^{100} + x^{50} = x^{50}(x^{50} + 1)$ . As

$$1^{50} + 1 \neq 0, \text{ but } i^{50} + 1 = (i^2)^{25} + 1 = (-1)^{25} + 1 = -1 + 1 = 0 \text{ we see that } x - 1 \text{ is not a}$$

$$\text{factor of } f(x) \text{ but } x^2 + 1 \text{ is a factor of } f(x). \text{ Therefore } \gcd\{f(x), \chi_C(x)\} = x^2 + 1$$

$$\text{and } f(C) = C^{100} + C^{50} \text{ has rank } 3 - 2 = 1.$$

Lastly let  $f(x) = x^{100} - x^{52} = x^{52}(x^{48} - 1)$ . As  $1^{48} - 1 = 0$  and

$$i^{48} - 1 = (-1)^{24} - 1 = 1 - 1 = 0 \text{ we see that both } x - 1 \text{ and } x^2 + 1 \text{ are factors of } f(x). \text{ So}$$

$$\gcd\{f(x), \chi_C(x)\} = (x^2 + 1)(x - 1). \text{ In this case } \chi_C(x) \mid f(x) \text{ and } f(C) = 0 \text{ has}$$

$$\text{rank } 3 - 3 = 0.$$

## Solution 6

(a) Using (5.31) the matrix

$$Y_1 = \begin{pmatrix} 1 & -2 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

is invertible over  $F$  and satisfies

$$\begin{pmatrix} 1 & -2 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 2 \end{pmatrix} \begin{pmatrix} 1 & -2 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

i.e.  $Y_1 C(x(x-1)^2) = (C(x) \oplus C((x-1)^2))Y_1$ , and so

$$Y_1 C(x(x-1)^2)Y_1^{-1} = (C(x) \oplus C((x-1)^2)).$$

Using (5.31) again, the matrix

$$Y_2 = \begin{pmatrix} 1 & 2 & 1 & 0 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

satisfies

$$\begin{pmatrix} 1 & 2 & 1 & 0 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & -2 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & -2 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & -2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 & 0 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

i.e.  $Y_2 C(x^2(x+1)^2) = (C(x^2) \oplus C((x+1)^2))Y_2$  and so

$$Y_2 C(x^2(x+1)^2)Y_2^{-1} = (C(x^2) \oplus C((x+1)^2)).$$

(b) Write  $f(x) = x^3 + ax^2 + bx + c$ . Then

$$R(x-r, x^3 + ax^2 + bx + c) = \begin{vmatrix} 1 & -r & 0 & 0 \\ 0 & 1 & -r & 0 \\ 0 & 0 & 1 & -r \\ 1 & a & b & c \end{vmatrix} = \begin{vmatrix} 1 & -r & 0 & 0 \\ 0 & 1 & -r & 0 \\ 0 & 0 & 1 & 0 \\ 1 & a & b & f(r) \end{vmatrix} = f(r)$$

on applying the *ecos*  $c_4 + r^3c_1, c_4 + r^2c_2, c_4 + rc_3$  and expanding along col 4.

$$R(x^2 + ax + b, x^2 + cx + d) = \begin{vmatrix} 1 & a & b & 0 \\ 0 & 1 & a & b \\ 1 & c & d & 0 \\ 0 & 1 & c & d \end{vmatrix} = \begin{vmatrix} 1 & a & b \\ c-a & d-b & 0 \\ 1 & c & d \end{vmatrix}$$

on performing the *ero*  $r_3 - r_1$  and expanding along col 1. Expanding along row 2 gives

$$R(x^2 + ax + b, x^2 + cx + d) = (d-b)^2 - (c-a)(ad-bc).$$

(c) Write  $S = \begin{pmatrix} 0 & I_s \\ I_t & 0 \end{pmatrix}$ , i.e.  $S$  is the partitioned  $(s+t) \times (s+t)$  matrix over  $F$  where

$I_s$  is the  $s \times s$  identity matrix over  $F$  and  $I_t$  is the  $t \times t$  identity matrix over  $F$  with zero entries elsewhere. From the matrices underlying the resultants we obtain

$R(f(x), g(x)) = |S| R(g(x), f(x))$  on taking determinants. Now  $S$  is changed into

$$\begin{pmatrix} I_t & 0 \\ 0 & I_s \end{pmatrix} = I_{s+t}$$

on applying the *st eros*  $r_{s+j} \leftrightarrow r_{s+j-1}, r_{s+j-1} \leftrightarrow r_{s+j-2}, \dots, r_{j+1} \leftrightarrow r_j$  for  $j = 1, 2, \dots, t$ . Therefore  $|S| = (-1)^{st}$  and so  $R(f(x), g(x)) = (-1)^{st} R(g(x), f(x))$ .

(d) The formula  $\det T = (-1)^{t(t-1)/2}$  holds for  $t = 1$ . Take  $t \geq 2$  and suppose inductively that the  $t \times t$  matrix

$$T = \left( \begin{array}{c|c} 0 & 1 \\ \hline \bar{T}' & 0 \end{array} \right)$$

is such that  $\det T' = (-1)^{(t-1)(t-2)/2}$ . Then

$\det T = (-1)^{t-1} \det T' = (-1)^{t-1} (-1)^{(t-1)(t-2)/2} = (-1)^{t-1+(t-1)(t-2)/2} = (-1)^{t(t-1)/2}$   
which completes the inductive step.

## Solutions 6.1 (page 267)

### Solution 1

(a) (i)

$$\begin{aligned}
 xI - A &= \begin{pmatrix} x-1 & 1 & -1 \\ 1 & x-1 & 1 \\ 2 & -2 & x+2 \end{pmatrix} \begin{matrix} \\ r_1 \leftrightarrow r_2 \\ \end{matrix} \equiv \begin{pmatrix} 1 & x-1 & 1 \\ x-1 & 1 & -1 \\ 2 & -2 & x+2 \end{pmatrix} \begin{matrix} \\ c_2 - (x-1)c_1 \\ c_3 - c_1 \end{matrix} \equiv \\
 &\begin{pmatrix} 1 & 0 & 0 \\ x-1 & 2x-x^2 & -x \\ 2 & -2x & x \end{pmatrix} \begin{matrix} \\ r_2 - (x-1)r_1 \\ r_3 - 2r_1 \end{matrix} \equiv \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2x-x^2 & -x \\ 0 & -2x & x \end{pmatrix} \begin{matrix} \\ c_2 \leftrightarrow c_3 \\ -c_2 \end{matrix} \equiv \begin{pmatrix} 1 & 0 & 0 \\ 0 & x & 2x-x^2 \\ 0 & -x & -2x \end{pmatrix} \\
 &\equiv \begin{pmatrix} 1 & 0 & 0 \\ 0 & x & 0 \\ 0 & -x & -x^2 \end{pmatrix} \begin{matrix} \\ r_3 + r_2 \\ -r_3 \end{matrix} \equiv \begin{pmatrix} 1 & 0 & 0 \\ 0 & x & 0 \\ 0 & 0 & x^2 \end{pmatrix} = S(xI - A).
 \end{aligned}$$

Therefore  $A$  has invariant factor sequence  $(x, x^2)$ , the non-constant diagonal entries in  $S(xI - A)$ . Applying the *eros* used above in sequence to  $I$  gives

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{matrix} \\ r_1 \leftrightarrow r_2 \\ r_2 - (x-1)r_1 \end{matrix} \equiv \begin{pmatrix} 0 & 1 & 0 \\ 1 & 1-x & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{matrix} \\ r_3 - 2r_1 \\ r_3 + r_2 \end{matrix} \equiv \begin{pmatrix} 0 & 1 & 0 \\ 1 & 1-x & 0 \\ 1 & -1-x & 1 \end{pmatrix} \begin{matrix} \\ \\ -r_3 \end{matrix} \equiv \begin{pmatrix} 0 & 1 & 0 \\ 1 & 1-x & 0 \\ -1 & 1+x & -1 \end{pmatrix} = P(x).$$

Applying the conjugates of the *ecos* used above in sequence to  $I$  gives

$$\begin{aligned}
 &\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{matrix} \\ r_1 + (x-1)r_2 \\ r_1 + r_3 \end{matrix} \equiv \begin{pmatrix} 1 & x-1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{matrix} \\ r_2 \leftrightarrow r_3 \\ -r_2 \end{matrix} \equiv \begin{pmatrix} 1 & x-1 & 1 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix} \\
 &\equiv \begin{pmatrix} 1 & x-1 & 1 \\ 0 & 2-x & -1 \\ 0 & 1 & 0 \end{pmatrix} \begin{matrix} \\ r_2 - (x-2)r_3 \\ \end{matrix} = Q(x) = \begin{pmatrix} \rho_1(x) \\ \rho_2(x) \\ \rho_3(x) \end{pmatrix}.
 \end{aligned}$$

Then

$$P(x)(xI - A) = \begin{pmatrix} 1 & x-1 & 1 \\ 0 & 2x-x^2 & -x \\ 0 & x^2 & 0 \end{pmatrix} = S(xI - A)Q(x).$$

Also  $\det P(x) = \det Q(x) = 1$  showing that  $P(x), Q(x)$  are invertible over  $\mathbb{Q}[x]$ . Using the evaluation homomorphism  $\theta_A : \mathbb{Q}[x]^3 \rightarrow M(A)$  we obtain

$$(\rho_1(x))\theta_A = (1, x-1, 1)\theta_A = e_1 + (x-1)e_2 + e_3 = e_1 + e_2(A - I) + e_3 = 0$$

as  $e_2(A - I) = (-1, 0, -1)$ . In the same way

$$(\rho_2(x))\theta_A = (0, 2-x, -1)\theta_A = (2-x)e_2 - e_3 = e_2(2I - A) - e_3 = (1, 1, 0) = v_1$$

has order  $x$  in  $M(A)$  by (6.5). Also  $(\rho_3(x))\theta_A = (0, 1, 0)\theta_A = (0, 1, 0) = v_2$  has order  $x^2$  in  $M(A)$  and  $M(A) = \langle v_1 \rangle \oplus \langle v_2 \rangle$  (internal direct sum). Construct

$$X = \begin{pmatrix} v_1 \\ v_2 \\ xv_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ -1 & 1 & -1 \end{pmatrix}$$



as  $xv_2 = v_2A = (-1, 1, -1)$ . Then  $X$  is invertible over  $\mathbb{Q}$  and

$$XAX^{-1} = C(x) \oplus C(x^2) = \left( \begin{array}{c|cc} 0 & 0 & 0 \\ \hline 0 & 0 & 1 \\ 0 & 0 & 0 \end{array} \right) \text{ which is the rcf of } A.$$

(ii)

$$\begin{aligned} xI - A &= \begin{pmatrix} x-1 & -1 & 1 \\ 1 & x+1 & -1 \\ -2 & -2 & x+2 \end{pmatrix} \begin{matrix} r_1 \leftrightarrow r_2 \end{matrix} \equiv \begin{pmatrix} 1 & x+1 & -1 \\ x-1 & -1 & 1 \\ -2 & -2 & x+2 \end{pmatrix} \begin{matrix} c_2 - (x+1)c_1 \\ c_3 + c_1 \end{matrix} \\ &\equiv \begin{pmatrix} 1 & 0 & 0 \\ x-1 & -x^2 & x \\ -2 & 2x & x \end{pmatrix} \begin{matrix} r_2 - (x-1)r_1 \\ r_3 + 2r_1 \end{matrix} \equiv \begin{pmatrix} 1 & 0 & 0 \\ 0 & -x^2 & x \\ 0 & 2x & x \end{pmatrix} \begin{matrix} c_2 \leftrightarrow c_3 \\ c_3 + xc_2 \end{matrix} \equiv \begin{pmatrix} 1 & 0 & 0 \\ 0 & x & 0 \\ 0 & x & x(x+2) \end{pmatrix} \\ &\equiv \begin{pmatrix} 1 & 0 & 0 \\ 0 & x & 0 \\ 0 & 0 & x(x+2) \end{pmatrix} = S(xI - A). \end{aligned}$$

The invariant factor sequence of  $A$  is therefore  $(x, x(x+2))$ . Applying the above *eros* to the identity matrix  $I$  gives

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{matrix} r_1 \leftrightarrow r_2 \\ r_2 - (x-1)r_1 \end{matrix} \equiv \begin{pmatrix} 0 & 1 & 0 \\ 1 & 1-x & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{matrix} r_3 + 2r_1 \\ r_3 - r_2 \end{matrix} \equiv \begin{pmatrix} 0 & 1 & 0 \\ 1 & 1-x & 0 \\ -1 & x+1 & 1 \end{pmatrix} = P(x).$$

Applying the conjugates of the above *ecos* to  $I$  gives

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{matrix} r_1 + (x+1)r_2 \\ r_1 - r_3 \end{matrix} \equiv \begin{pmatrix} 1 & x+1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{matrix} r_2 \leftrightarrow r_3 \\ r_2 - xr_3 \end{matrix} \equiv \begin{pmatrix} 1 & x+1 & -1 \\ 0 & -x & 1 \\ 0 & 1 & 0 \end{pmatrix} = Q(x) = \begin{pmatrix} \rho_1(x) \\ \rho_2(x) \\ \rho_3(x) \end{pmatrix}.$$

As  $\det P(x) = \det Q(x) = -1$  the matrices  $P(x), Q(x)$  are invertible over  $\mathbb{Q}[x]$ . Also

$$P(x)(xI - A) = \begin{pmatrix} 1 & x+1 & -1 \\ 0 & -x^2 & x \\ 0 & x^2 + 2x & 0 \end{pmatrix} = S(xI - A)Q(x).$$

Using the evaluation homomorphism  $\theta_A : \mathbb{Q}[x]^3 \rightarrow M(A)$  we obtain

$$(\rho_1(x))\theta_A = (1, x+1, -1)\theta_A = (1, 1, -1) + (0, 1, 0)A = (1, 1, -1) + (-1, -1, 1) = 0.$$

Also  $(\rho_2(x))\theta_A = (0, -x, 1)\theta_A = -e_2A + e_3 = (1, 1, 0) = v_1$  and

$$(\rho_3(x))\theta_A = (0, 1, 0)\theta_A = (0, 1, 0) = v_2.$$

Construct

$$X = \begin{pmatrix} \frac{v_1}{v_2} \\ xv_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ -1 & -1 & 1 \end{pmatrix}.$$

Then  $\det X = 1$  and so  $X$  is invertible over  $\mathbb{Q}$ . Also

$$XA = \begin{pmatrix} 0 & 0 & 0 \\ -1 & -1 & 1 \\ 2 & 2 & -2 \end{pmatrix} = (C(x) \oplus C(x(x+2)))X$$

giving  $XAX^{-1} = C(x) \oplus C(x(x+2))$  which is the rcf of  $A$ .

(iii)

$$\begin{aligned}
xI - A &= \begin{pmatrix} x-1 & -1 & 1 \\ -1 & x & 1 \\ -5 & -3 & x+4 \end{pmatrix} \xrightarrow{c_1 \leftrightarrow c_3} \begin{pmatrix} 1 & -1 & x-1 \\ 1 & x & -1 \\ x+4 & -3 & -5 \end{pmatrix} \xrightarrow{c_2 + c_1, c_3 - (x-1)c_1} \\
&\begin{pmatrix} 1 & 0 & 0 \\ 1 & x+1 & -x \\ x+4 & x+1 & -x^2-3x-1 \end{pmatrix} \xrightarrow{r_2 - r_1, r_3 - (x+4)r_1} \begin{pmatrix} 1 & 0 & 0 \\ 0 & x+1 & -x \\ 0 & x+1 & -x^2-3x-1 \end{pmatrix} \xrightarrow{c_2 + c_3, -c_3} \\
&\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & x \\ 0 & -x^2-2x & x^2+3x+1 \end{pmatrix} \xrightarrow{c_3 - xc_2, r_3 + x(x+2)r_2} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & (x+1)^3 \end{pmatrix} = S(xI - A).
\end{aligned}$$

So the invariant factor sequence of  $A$  is  $((x+1)^3)$ . Applying the above *eros* to  $I$  gives

$$\begin{aligned}
&\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{r_2 - r_1, r_3 - (x+4)r_1} \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -(x+4) & 0 & 1 \end{pmatrix} \\
&\xrightarrow{r_3 + x(x+2)r_2} \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -x^2-3x-4 & x(x+2) & 1 \end{pmatrix} = P(x).
\end{aligned}$$

Applying the conjugates of the above *ecos* to  $I$  gives

$$\begin{aligned}
&\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{r_1 \leftrightarrow r_3, r_1 + (x-1)r_3, r_3 - r_2} \begin{pmatrix} x-1 & -1 & 1 \\ 0 & 1 & 0 \\ 1 & -1 & 0 \end{pmatrix} \\
&\xrightarrow{-r_3, r_2 + xr_3} \begin{pmatrix} x-1 & -1 & 1 \\ -x & x+1 & 0 \\ -1 & 1 & 0 \end{pmatrix} = Q(x) = \begin{pmatrix} \rho_1(x) \\ \rho_2(x) \\ \rho_3(x) \end{pmatrix}.
\end{aligned}$$

As  $\det P(x) = \det Q(x) = 1$  we see  $P(x)$  and  $Q(x)$  are invertible over  $\mathbb{Q}[x]$  and

$$P(x)(xI - A) = \begin{pmatrix} x-1 & -1 & 1 \\ -x & x+1 & 0 \\ -(x+1)^3 & (x+1)^3 & 0 \end{pmatrix} = S(xI - A)Q(x).$$

Using the evaluation homomorphism  $\theta_A : \mathbb{Q}[x]^3 \rightarrow M(A)$  we obtain

$$(\rho_1(x))\theta_A = (x-1, -1, 1)\theta_A = e_1A + (-1, -1, 1) = 0 \text{ and}$$

$$(\rho_2(x))\theta_A = (-x, x+1, 0)\theta_A = -e_1A + e_2A + e_2 = 0. \text{ Also}$$

$$(\rho_3(x))\theta_A = (-1, 1, 0)\theta_A = (-1, 1, 0) = v_1 \text{ has order } (x+1)^3 \text{ in } M(A). \text{ So}$$

$M(A) = \langle v_1 \rangle$ , i.e.  $M(A)$  is cyclic with generator  $v_1$ . Construct

$$X = \begin{pmatrix} v_1 \\ xv_1 \\ x^2v_1 \end{pmatrix} = \begin{pmatrix} v_1 \\ v_1A \\ (v_1A)A \end{pmatrix} = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 0 \\ -1 & 0 & 1 \end{pmatrix}$$

which is invertible over  $\mathbb{Q}$  and satisfies

$$XA = \begin{pmatrix} 0 & -1 & 0 \\ -1 & 0 & 1 \\ 4 & 2 & -3 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & -3 & -3 \end{pmatrix} \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 0 \\ -1 & 0 & 1 \end{pmatrix} = C((x+1)^3)X.$$

So  $XAX^{-1} = C((x+1)^3)$  which is the rcf of  $A$ .

(b) Rank  $C(d_j(x)) \geq \deg d_j(x) - 1$  as the first  $\deg d_j(x) - 1$  rows of  $C(d_j(x))$  are linearly independent. Let  $C$  denote the rcf of  $A$ . Then

$$\text{rank } A = \text{rank } C = \sum_{j=1}^s \text{rank } C(d_j(x)) \geq \sum_{j=1}^s (\deg d_j(x) - 1) = \sum_{j=1}^s \deg d_j(x) - s = t - s.$$

Suppose  $\text{rank } A = 1$ . Then  $0 \leq t - s \leq 1$ .

Suppose  $t - s = 0$ , i.e.  $s = t$ . Then  $d_j(x) = x - a$  for some  $a \in F$  and  $1 \leq j \leq t$  as

$\deg d_j(x) = 1$  and  $d_j(x) \mid d_t(x)$ . So  $A = aI$  which has rank 0 or  $t$  according as  $a = 0$  or  $a \neq 0$ . As  $t \geq 2$  we see that  $s = t$  is impossible. So  $t - s = 1$ , i.e.  $s = t - 1$ .

Therefore  $\deg d_j(x) = 1$  for  $1 \leq j \leq t - 2$  and  $\deg d_{t-1}(x) = 2$ . As  $\text{rank } C(d_{t-1}(x)) \geq 1$  and  $\text{rank } A = 1$  we conclude  $\text{rank } C(d_{t-1}(x)) = 1$  and  $\text{rank } C(d_j(x)) = 0$  for

$1 \leq j \leq t - 2$ . So  $d_j(x) = x$  for  $1 \leq j \leq t - 2$  and  $d_{t-1}(x) = x(x - b)$ . As

$\text{trace } C(x(x - b)) = b$  we see  $\text{trace } C = b$  also. So  $b = \text{trace } A$ .

(c) Taking determinants of the equation  $P(x)(xI - A) = S(xI - A)Q(x)$  gives  $\det P(x) \chi_A(x) = \det S(xI - A) \det Q(x)$ . Now  $\det P(x)$  and  $\det Q(x)$  are non-zero constant polynomials since  $P(x)$  and  $Q(x)$  are invertible over  $F[x]$ . Also  $\chi_A(x)$  and  $\det S(xI - A) = d_1(x)d_2(x) \cdots d_s(x)$ , the product of the invariant factors of  $A$ , are monic. So equating coefficients of  $x^t$  gives  $\det P(x) = \det Q(x)$  and  $\chi_A(x) = d_1(x)d_2(x) \cdots d_s(x)$ .

### Solution 2

(a) Replacing  $x$  in  $P(x)(xI - A) = \text{diag}(1, 1, \dots, 1, d_1(x), d_2(x), \dots, d_s(x))Q(x)$  by  $x + \lambda$  gives

$$P(x + \lambda)((x + \lambda)I - A) = \text{diag}(1, 1, \dots, 1, d_1(x + \lambda), d_2(x + \lambda), \dots, d_s(x + \lambda))Q(x + \lambda).$$

Now  $\det P(x + \lambda) = \det P(x)$ , since  $\det P(x)$  is a constant polynomial ( $x$  is not involved) as  $P(x)$  is invertible over  $F[x]$  and so  $P(x + \lambda)$  is also invertible over  $F[x]$ . For the same reason  $Q(x + \lambda)$  is invertible over  $F[x]$ . Therefore the Smith normal form of  $(x + \lambda)I - A = xI - (A - \lambda I)$  is

$$\text{diag}(1, 1, \dots, 1, d_1(x + \lambda), d_2(x + \lambda), \dots, d_s(x + \lambda)),$$

i.e. the invariant factor sequence of  $A - \lambda I$  is  $(d_1(x + \lambda), d_2(x + \lambda), \dots, d_s(x + \lambda))$ .

(b)(i) The matrix  $A$  of Question 1(a)(i) above has invariant factor sequence  $(x, x^2)$ .

The matrix here is  $A + I$  which, by (a) above with  $\lambda = -1$ , has invariant factor sequence  $(x - 1, (x - 1)^2)$ .

(ii) The matrix  $A$  of Question 1(a)(ii) above has invariant factor sequence  $(x, x(x + 2))$ . The matrix here is  $A + 2I$  which, by (a) above with  $\lambda = -2$ , has invariant factor sequence  $(x - 2, (x - 2)x)$ .

(iii) The matrix  $A$  of Question 1(a)(iii) above has invariant factor sequence  $((x + 1)^3)$ . The matrix here is  $A + 5I$  which, by (a) above with  $\lambda = -5$ , has invariant factor sequence  $((x - 4)^3)$ .

(c) Write  $d(x) = a_0 + a_1x + \dots + a_{m-1}x^{m-1} + x^m$  and suppose  $vC(d(x)) = 0$  where  $v = (b_1, b_2, \dots, b_m)$ . Comparing entries gives  $v = b_m(a_1, a_2, \dots, a_{m-1}, 1)$  where  $a_0b_m = 0$ . Therefore  $a_0 \neq 0 \Rightarrow b_m = 0 \Rightarrow v = 0 \Rightarrow \text{nullity } C(d(x)) = 0$  and

$a_0 = 0 \Rightarrow \text{nullity } C(d(x)) = 1$ . As  $a_0 = d(0)$  the theory above and (4.2)(i) together show  $\text{nullity } C(d(x)) = 1 \Leftrightarrow d(0) = 0 \Leftrightarrow x \mid d(x)$ .

As  $A \sim C(d_1(x)) \oplus C(d_2(x)) \oplus \dots \oplus C(d_s(x))$  we see

$$n = \text{nullity } A = \text{nullity } (C(d_1(x)) \oplus C(d_2(x)) \oplus \dots \oplus C(d_s(x))) = \sum_{i=1}^s \text{nullity } C(d_i(x)).$$

By the preceding paragraph  $x \mid d_i(x)$  for  $n$  integers  $i$  with  $1 \leq i \leq s$ . So  $n \leq s$ . As  $d_j(x) \mid d_i(x)$  for  $j \leq i$ , the  $n$  integers  $i$  with  $x \mid d_i(x)$  are the last  $n$  integers in the above range, i.e.  $s-n+1, s-n+2, \dots, s$ , i.e.  $s-n < i \leq s$ .

(d)

$$C_0^2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

has nullity 2 and minimum polynomial  $x^2$  as  $(C_0^2)^2 = 0$ . So  $C_0^2$  has two invariant factors, the second invariant factor is  $x^2$  and their product is  $x^4$  (it must be a power of  $x$  by (6.11) and being the characteristic polynomial of a  $4 \times 4$  matrix it has degree 4). So  $C_0^2$  has invariant factor sequence  $(x^2, x^2)$ .

As

$$C_0^3 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

we see that  $C_0^3$  has  $3 = \text{nullity } C_0^3$  invariant factors, the last is  $x^2$  and the product of all three is  $x^4$ . Therefore  $(x, x, x^2)$  is the invariant factor sequence of  $C_0^3$ .

As  $C_0^4$  is the zero  $4 \times 4$  matrix its invariant factor sequence is  $(x, x, x, x)$ .

$$\text{As } C_1 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & -2 & 0 \end{pmatrix} \text{ we obtain } C_1^2 + I = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ -1 & 0 & -1 & 0 \\ 0 & -1 & 0 & -1 \end{pmatrix}.$$

As  $(e_1 + e_3)(C_1^2 + I) = (e_2 + e_4)(C_1^2 + I) = 0$  we see  $\text{nullity } C_1^2 + I = 2$ . Also  $(C_1^2 + I)^2 = 0$  by the Cayley-Hamilton theorem as  $C_1$  has characteristic polynomial  $(x^2 + 1)^2$  by (5.26). Therefore  $C_1^2 + I$  has minimum polynomial  $x^2$  and two invariant factors. So  $C_1^2 + I$  has invariant factor sequence  $(x^2, x^2)$ .

Since  $C_0^2$  and  $C_1^2 + I$  have equal invariant factor sequences these matrices are similar.

The vectors  $e_1, e_2 \in F^4$  both have order  $x^2$  in the  $F[x]$ -module  $M(C_0^2)$ . Also

$N_1 = \langle e_1, x^2 e_1 \rangle$  and  $N_2 = \langle e_2, x^2 e_2 \rangle$  are submodules of  $M(C_0^2)$  with

$M(C_0^2) = N_1 \oplus N_2$ . So in  $F[x]$ -module  $M(C_0)$ , the matrix

$$X_0 = \begin{pmatrix} e_1 \\ x^2 e_1 \\ x e_1 \\ x^3 e_1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

is invertible over  $F$  and satisfies  $X_0 C_0^2 X_0^{-1} = C(x^2) \oplus C(x^2)$  in rcf. The matrix

$$X_1 = \begin{pmatrix} e_1 \\ (x^2 + 1)e_1 \\ x e_1 \\ (x^2 + 1)x e_1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}$$

is invertible over  $F$  and satisfies  $X_1(C_1^2 + I)X_1^{-1} = C(x^2) \oplus C(x^2)$  in rcf.

For  $0 \leq i < m, 0 \leq j < n$  the  $mn$  monic polynomials  $x^i d(x)^j$  have  $mn$  different degrees  $mj + i$  accounting for all integers in the range  $0 \leq mj + i < mn$ . As  $e_1$  has order  $d(x)^n$  in the  $F[x]$ -module  $M(C)$  we see that the  $mn$  vectors  $x^i d(x)^j e_1$  are linearly independent and so are the elements of a basis of  $F^{mn}$ . In particular  $e_{i+1}, d(x)e_{i+1}, d(x)^2 e_{i+1}, \dots, d(x)^{n-1} e_{i+1}$  are linearly independent for  $0 \leq i < m$  as  $e_{i+1} = x^i e_1$ . From  $d(x)^n e_{i+1} = 0$  we deduce that  $e_{i+1}$  has order  $x^n$  in the  $F[x]$ -module  $M(d(C))$  for  $0 \leq i < m$ . Let  $N_{i+1}$  be the cyclic submodule of  $M(d(C))$  generated by  $e_{i+1}$  for  $0 \leq i < m$ . Then  $N_{i+1}$  has  $F$ -basis  $e_{i+1}, d(x)e_{i+1}, d(x)^2 e_{i+1}, \dots, d(x)^{n-1} e_{i+1}$  and

$$M(d(C)) = N_1 \oplus N_2 \oplus \dots \oplus N_m.$$

Construct the invertible  $mn \times mn$  matrix  $X$  over  $F$  having the above  $F$ -bases of  $N_1, N_2, \dots, N_m$  as its rows. Then

$$X d(C) X^{-1} = C(x^n) \oplus C(x^n) \oplus \dots \oplus C(x^n) \quad (m \text{ terms})$$

is in rcf on combining (5.2), (5.20) and (5.27). As the rcf of  $d(C)$  is independent of the particular monic polynomial  $d(x)$  of degree  $m$  over  $F$ , all such matrices  $d(C)$  are similar. Since  $d_0(x) = x^m$  is monic of degree  $m$  over  $F$  we see  $d(C) \sim d_0(C) = C^m = (C(x^{mn}))^m$ .

(e) By using the  $n$  vectors of the standard basis  $\mathcal{B}_0$  of  $F^n$ , suitably ordered, as the rows of  $X$  we obtain  $XC^m X^{-1}$  in rcf. As  $e_1$  has order  $x^n$  in the  $F[x]$ -module  $M(C)$ , the vectors of  $\mathcal{B}_0$  can be expressed  $x^{jm+i} e_1 = e_{jm+i+1}$  for  $0 \leq jm+i < n, 0 \leq i < m, 0 \leq j < q$ . Let  $N_{m-i}$  be the submodule of the  $F[x]$ -module  $M(C^m)$  generated by  $e_{i+1}$  for  $0 \leq i < m$ . Then  $N_{m-i}$  has  $F$ -basis  $e_{i+1}, x^m e_{i+1}, x^{2m} e_{i+1}, \dots, x^{qm} e_{i+1}$  for  $0 \leq i < r$  (as  $x^{qm} e_{i+1} = e_{mq+i+1} \neq 0$  but  $x^{(q+1)m} e_{i+1} = 0$  since  $mq+i+1 \leq n < (q+1)m$ , we see  $e_{i+1}$  has order  $x^{q+1}$  in  $M(C^m)$ ). In the same way  $N_{m-i}$  has  $F$ -basis  $e_{i+1}, x^m e_{i+1}, x^{2m} e_{i+1}, \dots, x^{(q-1)m} e_{i+1}$  for  $r \leq i < m$  (as  $x^{(q-1)m} e_{i+1} = e_{(q-1)m+i+1} \neq 0$  but  $x^{qm} e_{i+1} = 0$  since  $(q-1)m+i+1 \leq n \leq qm+i$ , we see  $e_{i+1}$  has order  $x^q$  in  $M(C^m)$ ). Then  $M(C^m) = N_1 \oplus N_2 \oplus \dots \oplus N_m$ . We construct the invertible  $n \times n$  matrix  $X$  over  $F$  by taking the above  $F$ -bases of

$N_1, N_2, \dots, N_m$  (in that order) as its rows. Then  $XC^mX^{-1}$  is in rcf being the direct sum of  $m-r$  companion matrices  $C(x^q)$  followed by  $r$  companion matrices  $C(x^{q+1})$ .

(f) In the  $F[x]$ -module  $M(C^2)$  the element  $e_1$  has order  $x^2 + x + 1$  since  $e_1$  and  $x^2e_1 = e_3$  are linearly independent but  $(x^4 + x^2 + 1)e_1 = 0$  in the  $F[x]$ -module  $M(C)$ . In the same way  $e_2$  and  $x^2e_2 = e_4$  are linearly independent and  $(x^4 + x^2 + 1)e_2 = 0$ , showing that  $e_2$  also has order  $x^2 + x + 1$  in  $M(C^2)$ . So

$M(C^2) = \langle e_1, e_3 \rangle \oplus \langle e_2, e_4 \rangle = \langle e_1 \rangle \oplus \langle e_2 \rangle$ . Then

$$X = \begin{pmatrix} e_1 \\ x^2e_1 \\ e_2 \\ x^2e_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

satisfies  $XC^2X^{-1} = C(x^2 + x + 1) \oplus C(x^2 + x + 1)$  by (6.5).

More generally write  $C = C(d(x^2))$ . The element  $e_1$  of the  $F[x]$ -module  $M(C^2)$  has order  $d(x)$  since  $e_1, x^2e_1, x^4e_1, \dots, x^{2(t-1)}e_1$  are linearly independent (they are  $e_1, e_3, e_5, \dots, e_{2t-1}$  in  $F^{2t}$ ) and  $d(x)e_1 = e_1d(C^2) = 0$  in  $M(C^2)$  by (5.26). In the same way  $e_2$  has order  $d(x)$  in the  $F[x]$ -module  $M(C^2)$  since  $e_2, x^2e_2, x^4e_2, \dots, x^{2(t-1)}e_2$  are linearly independent (they are  $e_2, e_4, e_6, \dots, e_{2t}$  in  $F^{2t}$ ) and  $d(x)e_2 = e_2d(C^2) = 0$  by (5.26). So

$$M(C^2) = \langle e_1, e_3, \dots, e_{2t-1} \rangle \oplus \langle e_2, e_4, \dots, e_{2t} \rangle = \langle e_1 \rangle \oplus \langle e_2 \rangle.$$

Let  $X$  be the  $2t \times 2t$  matrix having  $e_1, e_3, \dots, e_{2t-1}, e_2, e_4, \dots, e_{2t}$  in that order as its rows. Then  $XC^2X^{-1} = C(d(x)) \oplus C(d(x))$  by (6.5).

Write  $C_0 = C(d(x^4))$ . Then  $C_0^2 \sim C(d(x^2)) \oplus C(d(x^2))$  by the preceding theory with  $x^2$  in place of  $x$ . So

$$C_0^4 = (C_0^2)^2 \sim (C(d(x^2)) \oplus C(d(x^2)))^2 = C(d(x^2))^2 \oplus C(d(x^2))^2.$$

As  $C(d(x^2))^2 \sim C(d(x)) \oplus C(d(x))$  we conclude that

$C(d(x)) \oplus C(d(x)) \oplus C(d(x)) \oplus C(d(x))$  is the rcf of  $C(d(x^4))^4$ .

### Solution 3

(a) Write  $g(x) = (-1)^{\deg f(x)} f(-x)$ . Then  $g(x)$  is monic and working in the  $F[x]$ -module  $M(-A)$  we have  $g(x)v = vg(-A) = (-1)^{\deg f(x)} v f(A) = 0$  as  $v f(A) = 0$ . So  $v$  has order  $g'(x)$  in  $M(-A)$  and  $g'(x) \mid g(x)$ . As  $vg'(-A) = 0$  working in the  $F[x]$ -module  $M(A)$  we see  $g'(-x)v = vg'(-A) = 0$  giving  $g'(-x) \mid f(x)$ . Therefore  $g'(x) \mid f(-x)$  on replacing  $-x$  by  $x$ . So  $g'(x) \mid g(x)$  and hence  $g'(x) = g(x)$ , i.e.  $v$  has order  $(-1)^{\deg f(x)} f(-x)$  in  $M(-A)$ .

Replacing  $x$  by  $-x$  in  $P(x)(xI - A) = \text{diag}(1, 1, \dots, 1, d_1(x), d_2(x), \dots, d_s(x))Q(x)$  gives

$$\begin{aligned} -P(-x)(xI + A) &= P(-x)(-xI - A) = \\ \text{diag}(1, 1, \dots, 1, d_1(-x), d_2(-x), \dots, d_s(-x))Q(-x) \end{aligned}$$

showing  $xI + A \equiv \text{diag}(1, 1, \dots, 1, d_1(-x), d_2(-x), \dots, d_s(-x))$  as  $-P(-x)$  and  $Q(-x)$  are invertible over  $F[x]$ . Write  $d'_j(x) = (-1)^{\deg d_j(x)} d_j(-x)$  for  $1 \leq j \leq s$ . The polynomials  $d'_j(x)$  are monic for  $1 \leq j \leq s$  and satisfy  $d'_1(x) \neq 1, d'_j(x) \mid d'_{j+1}(x)$  for  $1 \leq j < s$ . Applying the *eros*  $-r_i$  for  $i = t - s + j$ ,  $\deg d_j(x)$  odd, to  $\text{diag}(1, 1, \dots, 1, d_1(-x), d_2(-x), \dots, d_s(-x))$  shows

$$xI - (-A) = xI + A \equiv \text{diag}(1, 1, \dots, 1, d'_1(x), d'_2(x), \dots, d'_s(x)) = S(xI + A).$$

Therefore  $(d'_1(x), d'_2(x), \dots, d'_s(x))$  is the invariant factor sequence of  $-A$  where

$$d'_j(x) = (-1)^{\deg d_j(x)} d_j(-x) \text{ for } 1 \leq j \leq s.$$

A necessary and sufficient condition for  $-A \sim A$  ( $-A$  similar to  $A$ ) is  $d'_j(x) = d_j(x)$ , i.e. each invariant factor  $d_j(x)$  of  $A$  is either an even or an odd polynomial ( $d_j(x)$  either consists exclusively of even powers of  $x$  ( $d_j(x) = d_j(-x)$ ) or exclusively of odd powers of  $x$  ( $d_j(x) = -d_j(-x)$ )).

(b)  $\chi_A(x) = x(x-3)(x+3) = x^3 - 9x$  (start the factorisation of  $|xI - A|$  by performing the row operation  $r_1 - (r_2 - r_3)$ ). Being the product of distinct irreducible factors we see  $\mu_A(x) = \chi_A(x)$  by (6.11). So  $x^3 - 9x$  is the single invariant factor of  $A$  and being odd we deduce  $-A \sim A$  from (a) above.

(c) The statement is true. Suppose  $-A \sim A$ . Then  $-A$  and  $A$  have equal traces and equal determinants. But  $\text{trace}(-A) = -\text{trace } A$  and so  $2(\text{trace } A) = 0$  giving  $\text{trace } A = 0$  as  $\chi(F) \neq 2$ . As  $A$  has an odd number of rows, changing the sign of each row gives  $|-A| = -|A|$  and so  $|-A| = |A|$  gives  $2|A| = 0$ , i.e.  $|A| = 0$  as  $\chi(F) \neq 2$ . Conversely suppose  $|A| = 0$  and  $\text{trace } A = 0$ . Then  $\chi_A(x) = x^3 + ax$  as the coefficient of  $x^2$  and the constant term in  $\chi_A(x)$  are both zero (being respectively  $-\text{trace } A$  and  $-|A|$ ). For  $a \neq 0$  the quadratic  $x^2 + a$  either has equal and opposite zeros (distinct as  $\chi(F) \neq 2$ ) in  $F$  or it is irreducible over  $F$ ; in both cases  $\chi_A(x) = \mu_A(x)$  and so  $-A \sim A$  as in (b) above. So suppose  $a = 0$  in which case  $\chi_A(x) = x^3$ . The invariant factor sequence of  $A$  is one of  $(x, x, x), (x, x^2), (x^3)$  and in each case each individual invariant factor is either an even or an odd polynomial (none have a mix of odd and even powers of  $x$ ). So  $-A \sim A$  by (a) above.

#### Solution 4

(a) Transposing  $P(x)(xI - A) = S(xI - A)Q(x)$  gives

$$(xI - A^T)P(x)^T = Q(x)^T S(xI - A) \text{ since } (xI - A)^T = xI - A^T \text{ and}$$

$$S(xI - A)^T = S(xI - A) \text{ as } S(xI - A) \text{ is symmetric being a diagonal } t \times t \text{ matrix. So}$$

$$(Q(x)^T)^{-1}(xI - A^T) = S(xI - A)(P(x)^T)^{-1}. \text{ Now } (P(x)^T)^{-1} \text{ and } (Q(x)^T)^{-1} \text{ are}$$

invertible over  $F[x]$ , their inverses being  $P(x)^T$  and  $Q(x)^T$  respectively. So we see

$$S(xI - A^T) = S(xI - A), \text{ i.e. } xI - A^T \text{ and } xI - A \text{ have the same Smith normal form.}$$

So  $A \sim A^T$  as  $A$  and  $A^T$  have the same sequence of invariant factors, namely the non-constant diagonal entries in  $S(xI - A)$ .

(b)

$$R_f C(f(x)) = \begin{pmatrix} a_1 & a_2 & a_3 & 1 \\ a_2 & a_3 & 1 & 0 \\ a_3 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -a_0 & -a_1 & -a_2 & -a_3 \end{pmatrix} = \begin{pmatrix} -a_0 & 0 & 0 & 0 \\ 0 & a_2 & a_3 & 1 \\ 0 & a_3 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

which is symmetric. As  $R_f$  is invertible and symmetric we obtain

$$R_f C(f(x)) = (R_f C(f(x)))^T = C(f(x))^T R_f^T = C(f(x))^T R_f$$

and so  $R_f C(f(x)) R_f^{-1} = C(f(x))^T$ .

Let  $f(x) = a_0 + a_1 x + a_2 x^2 + \dots + a_{t-1} x^{t-1} + a_t x^t$  where  $a_t = 1$ . Let  $R_f$  denote the  $t \times t$  matrix over  $F$  with  $(i, j)$ -entry  $a_{i+j-1}$  for  $1 \leq i+j-1 \leq t$  and  $(i, j)$ -entry 0 for  $i+j-1 > t$ . Write  $g(x) = (f(x) - a_0)/x = a_1 + a_2 x + \dots + a_{t-1} x^{t-2} + a_t x^{t-1}$ . Then

$$R_f C(f(x)) = \left( \begin{array}{c|c} a & 1 \\ \hline R_g & 0^T \end{array} \right) \left( \begin{array}{c|c} 0 & I \\ \hline -a_0 & -a \end{array} \right)$$

where  $a = (a_1, a_2, \dots, a_{t-1})$ ,  $0$  is the  $1 \times (t-1)$  zero matrix and  $I$  is the  $(t-1) \times (t-1)$  identity matrix. Therefore

$$R_f C(f(x)) = \left( \begin{array}{c|c} -a_0 & 0 \\ \hline 0^T & R_g \end{array} \right)$$

on multiplying out the indicated partitioned matrices. So  $R_f C(f(x))$  is symmetric as

$R_g$  is symmetric. As  $\det R_f = (-1)^{t(t-1)/2}$  (use induction on  $t$  here) we see that  $R_f$  is invertible and so  $R_f C(f(x)) R_f^{-1} = C(f(x))^T$  as before.

(c)

$$RC = \left( \begin{array}{cc|cc} a_1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ \hline 0 & 0 & b_1 & 1 \\ 0 & 0 & 1 & 0 \end{array} \right) \left( \begin{array}{cc|cc} 0 & 1 & 0 & 0 \\ -a_0 & -a_1 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 \\ 0 & 0 & -b_0 & -b_1 \end{array} \right) = \left( \begin{array}{cc|cc} -a_0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 0 & 0 & -b_0 & 0 \\ 0 & 0 & 0 & 1 \end{array} \right)$$

i.e.  $RC = (R_{d_1} \oplus R_{d_2})(C(d_1(x)) \oplus C(d_2(x))) = R_{d_1} C(d_1(x)) \oplus R_{d_2} C(d_2(x))$  which is symmetric. Transposing gives  $RC = (RC)^T = C^T R^T = C^T R$  as  $R$  is symmetric. As  $\det R = \det R_{d_1} \det R_{d_2} = (-1)(-1) = 1$  we see  $R$  is invertible over  $F$  and so

$$RCR^{-1} = C^T.$$

Write  $C = C(d_1(x)) \oplus C(d_2(x)) \oplus \dots \oplus C(d_s(x))$  and

$R = R_{d_1(x)} \oplus R_{d_2(x)} \oplus \dots \oplus R_{d_s(x)}$ . Then  $R$  is symmetric and invertible over  $F$  as each  $R_{d_j(x)}$  is symmetric and invertible over  $F$  for  $1 \leq j \leq s$ . As

$R_{d_j(x)} C(d_j(x)) = C(d_j(x))^T R_{d_j(x)}$  for  $1 \leq j \leq s$  we obtain

$$\begin{aligned} RC &= R_{d_1(x)} C(d_1(x)) \oplus R_{d_2(x)} C(d_2(x)) \oplus \dots \oplus R_{d_s(x)} C(d_s(x)) = \\ &C(d_1(x))^T R_{d_1(x)} \oplus C(d_2(x))^T R_{d_2(x)} \oplus \dots \oplus C(d_s(x))^T R_{d_s(x)} = C^T R \end{aligned}$$

and so  $RCR^{-1} = C^T$ .



(d) The matrix  $Y = X^T R X$  is invertible over  $F$ , being the product of invertible matrices over  $F$ , where  $X A X^{-1} = C$  is in rcf and  $R C R^{-1} = C^T$ ,  $R$  symmetric. Then  $Y^T = (X^T R X)^T = X^T R^T (X^T)^T = X^T R X = Y$ , i.e.  $Y$  is symmetric. Also

$$\begin{aligned} Y A Y^{-1} &= (X^T R X)(X^{-1} C X)(X^T R X)^{-1} = \\ X^T R C R^{-1} (X^T)^{-1} &= X^T C^T (X^{-1})^T = (X^{-1} C X)^T = A^T \end{aligned}$$

since  $(X^{-1})^T = (X^T)^{-1}$  on transposing  $XX^{-1} = X^{-1}X = I$ .

(e) From Question 1(a)(i) above

$$X = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ -1 & 1 & -1 \end{pmatrix} \text{ satisfies } X A X^{-1} = C = C(x) \oplus C(x^2) \text{ where } A = \begin{pmatrix} 1 & -1 & 1 \\ -1 & 1 & -1 \\ -2 & 2 & -2 \end{pmatrix}.$$

$$\text{In this case } R = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \text{ and } Y = X^T R X = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & -1 \\ 0 & -1 & 0 \end{pmatrix} \text{ is invertible over } \mathbb{Q} \text{ and}$$

symmetric with  $Y A Y^{-1} = A^T$ .

(f) Let  $N$  be a submodule of  $M(A)$ . Then  $N$  and  $N^o$  are subspaces of  $F^t$  and using the theory of homogeneous linear equations we know  $\dim N^o = t - \dim N$ . Consider  $v \in N^o$ . Then for  $u \in N$  we have  $uA \in N$  and so  $uAv^T = 0$ , i.e.  $u(vA^T) = 0$  which means  $vA^T = 0$ . So  $N^o$  is a submodule of  $M(A^T)$  by (5.15).

Let  $N'$  be a submodule of  $M(A^T)$ . Then  $N'$  and  $(N')\gamma$  are subspaces of  $F^t$ .

Consider  $u \in (N')\gamma$ . Then  $u = vY$  where  $v \in N'$ . So  $uA = vYA = vA^T Y \in (N')\gamma$  since  $vA^T \in N'$ . Therefore  $(N')\gamma$  is a submodule of  $M(A)$  by (5.15). Taking  $N' = N^o$ , as in the preceding paragraph, we see  $(N^o)\gamma$  is a submodule of  $M(A)$ .

(i) A typical element of  $(N)\gamma^{-1}$  is  $uY^{-1}$  where  $u \in N$ . Such an element is

'orthogonal' to all elements  $vY$  of  $(N^o)\gamma = (N)\pi$  (here  $v \in N^o$ ) since

$$uY^{-1}(vY)^T = uY^{-1}Y^T v^T = uY^{-1}Y v^T = uv^T = 0 \text{ using } Y^T = Y. \text{ So}$$

$$(N)\gamma^{-1} \subseteq ((N^o)\gamma)^o. \text{ Let } s = \dim N. \text{ Then } s = \dim (N)\gamma^{-1} \text{ and}$$

$$\dim ((N^o)\gamma)^o = t - (t - s) = s \text{ and so in fact } (N)\gamma^{-1} = ((N^o)\gamma)^o. \text{ Therefore}$$

$$N = (((N^o)\gamma)^o)\gamma, \text{ i.e. } N = (N)\pi^2.$$

(ii) Suppose  $N_1 \subseteq N_2$ . Then directly we have  $v \in N_2^o \Rightarrow v \in N_1^o$ , i.e.  $N_2^o \subseteq N_1^o$ . As  $\gamma$ , being a vector space isomorphism, is inclusion-preserving we obtain

$$(N_2^o)\gamma \subseteq (N_1^o)\gamma, \text{ i.e. } (N_2)\pi \subseteq (N_1)\pi. \text{ Suppose } (N_2)\pi \subseteq (N_1)\pi. \text{ Applying } \pi \text{ and}$$

$$\text{using the immediately foregoing theory we obtain } (N_1)\pi^2 \subseteq (N_2)\pi^2, \text{ i.e. } N_1 \subseteq N_2$$

by (i). So  $N_1 \subseteq N_2 \Leftrightarrow (N_2)\pi \subseteq (N_1)\pi$  showing that  $\pi$  is inclusion-reversing.

### Solution 5

(a) Write  $g(x) = b_0 + b_1x + b_2x^2 + \dots + b_{s-1}x^{s-1} + x^s$  where  $b_0 \neq 0$ . Then

$$g(x)^* = (x^s/b_0)g(1/x) = 1/b_0 + (b_{s-1}/b_0)x + \dots + (b_1/b_0)x^{s-1} + x^s. \text{ The polynomial}$$

$f(x)g(x)$  is monic of degree  $t + s$  with constant term  $a_0b_0$ . Also replacing  $x$  in

$f(x)g(x)$  by  $1/x$  produces  $f(1/x)g(1/x)$ . Therefore

$$(f(x)g(x))^* = (x^{t+s}/(a_0b_0))f(1/x)g(1/x) = ((x^t/a_0)f(1/x))((x^s/b_0)g(1/x)) = f(x)^*g(x)^*.$$

The polynomial  $f(x)^*$  is monic of degree  $t$  and has constant term  $1/a_0$ . So

$$f(x)^{**} = (f(x)^*)^* = (a_0x^t/a_0x^t)f(1/(1/x)) = f(x).$$

Suppose  $f(x) = f(x)^*$  and  $g(x) = g(x)^*$ , i.e. suppose that  $f(x)$  and  $g(x)$  are palindromic. Then  $(f(x)g(x))^* = f(x)^*g(x)^* = f(x)g(x)$  showing that the product  $f(x)g(x)$  of palindromic polynomials is palindromic.

Suppose now that  $g(x)$  is a monic polynomial of positive degree over  $F$  with  $g(0) \neq 0$ . Then

$$(g(x)g(x)^*)^* = g(x)^*g(x)^{**} = g(x)^*g(x) = g(x)g(x)^*$$

shows that  $g(x)g(x)^*$  is palindromic.

Let  $f(x)$  be palindromic. Comparing constant terms in  $f(x) = f(x)^*$  gives  $a_0 = a_0^{-1}$  and so  $a_0 = \pm 1$ . Suppose  $a_0 = 1$ ; comparing coefficients of  $x^i$  in  $f(x) = f(x)^*$  for  $0 \leq i \leq t$  gives  $a_i = a_{t-i}/a_0 = a_{t-i}$ . Suppose  $a_0 = -1$ ; comparing coefficients of  $x^i$  in  $f(x) = f(x)^*$  for  $0 \leq i \leq t$  gives  $a_i = a_{t-i}/a_0 = -a_{t-i}$ .

The palindromic polynomials of degrees 1, 2, 3 over  $\mathbb{Z}_3 = \{-1, 0, 1\}$  are:

$$x+1, x-1, x^2+1, x^2+x+1, x^2-x+1, x^2-1, x^3+1, x^3+x^2+x+1, \\ x^3-x^2-x+1, x^3-1, x^3+x^2-x-1, x^3-x^2+x-1.$$

(b) Comparing rows 1 and 2 in  $C(f(x))Z = I$  where  $Z = (z_{ij})$  is  $3 \times 3$  over  $F$  gives

$$Z = \begin{pmatrix} z_{11} & z_{12} & z_{13} \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \text{ and so } C(f(x))^{-1} = Z = \begin{pmatrix} -a_1/a_0 & -a_2/a_0 & -1/a_0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

Working in the  $F[x]$ -module  $M(C(f(x))^{-1})$  we see  $xe_3 = e_2$ ,  $x^2e_3 = xe_2 = e_1$  and so  $e_3$  generates this module as  $e_3, xe_3, x^2e_3$  is a basis (actually the standard basis in reverse order) of  $F^3$ . Also  $x^3e_3 = (-1/a_0)(a_1, a_2, 1) = (-1/a_0)(a_1x^2e_3 + a_2xe_3 + e_3)$  and so  $(1/a_0 + (a_2/a_0)x + (a_1/a_0)x^2 + x^3)e_3 = 0$ , i.e.  $f(x)^*e_3 = 0$ . So  $e_3$  has order  $f(x)^*$  in  $M(C(f(x))^{-1})$  as  $e_3, xe_3, x^2e_3$  are linearly independent. So by (5.27)

$$X = \begin{pmatrix} e_3 \\ xe_3 \\ x^2e_3 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

is invertible over  $F$  and satisfies  $XC(f(x))^{-1}X^{-1} = C(f(x)^*)$ .

(c) Write  $f(x) = a_0 + a_1x + a_2x^2 + \dots + a_{t-1}x^{t-1} + x^t$ . Then

$$C(f(x))^{-1} = \begin{pmatrix} -a_1/a_0 & -a_2/a_0 & \dots & -a_{t-2}/a_0 & -a_{t-1}/a_0 & -1/a_0 \\ 1 & 0 & \dots & 0 & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 & 0 \\ 0 & 0 & \dots & 0 & 1 & 0 \end{pmatrix}$$

as the product of this matrix with  $C(f(x))$  is the  $t \times t$  identity matrix  $I$ . Working in the  $F[x]$ -module  $M(C(f(x))^{-1})$  we see  $xe_t = e_{t-1}, x^2e_t = e_{t-2}, \dots, x^{t-1}e_t = e_1$  and so  $e_t$  generates this module; in fact  $e_t, xe_t, x^2e_t, \dots, x^{t-1}e_t$  is the standard basis of  $F^t$  in reverse order. Also

$x^t e_t = x e_1 = (-1/a_0)(a_1, a_2, \dots, a_{t-1}, 1) = (-1/a_0)(a_1 x^{t-1} + a_2 x^{t-2} + \dots + a_{t-1} x + 1)e_t$  and so

$$(1/a_0 + (a_{t-1}/a_0)x + (a_{t-2}/a_0)x^2 + \dots + (a_1/a_0)x^{t-1} + x^t)e_t = 0,$$

i.e.  $f(x)^* e_t = 0$ . So  $e_t$  has order  $f(x)^*$  in  $M(C(f(x))^{-1})$ . Taking

$v_0 = e_t$ ,  $A = C(f(x))^{-1}$  in (5.27) we see that  $f(x)^*$  is the characteristic polynomial of  $C(f(x))^{-1}$  and  $X$  with  $e_i X = x^{i-1} e_i = e_{i+1-i}$  for  $1 \leq i \leq t$  is invertible over  $F$  and satisfies  $X C(f(x))^{-1} X^{-1} = C(f(x)^*)$ .

(d)  $\chi_A(0) = |0I - A| = |-A| = (-1)^t |A| \neq 0$ . As  $\chi_A(x) = d_1(x)d_2(x) \cdots d_s(x)$  we see  $d_j(0) \neq 0$  for  $1 \leq j \leq s$ . There is an invertible  $t \times t$  matrix  $X$  over  $F$  with

$$XAX^{-1} = C(d_1(x)) \oplus C(d_2(x)) \oplus \dots \oplus C(d_s(x)).$$

Inverting this equation gives  $X^{-1}A^{-1}X = C(d_1(x))^{-1} \oplus C(d_2(x))^{-1} \oplus \dots \oplus C(d_s(x))^{-1}$  which shows  $A^{-1} \sim C(d_1(x))^{-1} \oplus C(d_2(x))^{-1} \oplus \dots \oplus C(d_s(x))^{-1}$ . From (c) above we know  $C(d_j(x))^{-1} \sim C(d_j(x)^*)$  for  $1 \leq j \leq s$ . Therefore

$$A^{-1} \sim C(d_1(x)^*) \oplus C(d_2(x)^*) \oplus \dots \oplus C(d_s(x)^*)$$

which is in rcf since  $d_j(x) \mid d_{j+1}(x)$ , i.e.  $d_j(x)q_j(x) = d_{j+1}(x)$  for  $q_j(x) \in F[x]$  and so  $d_j(x)^* q_j(x)^* = d_{j+1}(x)^*$ , i.e.  $d_j(x)^* \mid d_{j+1}(x)^*$  for  $1 \leq j < s$ . So  $A^{-1}$  has invariant factor sequence  $(d_1(x)^*, d_2(x)^*, \dots, d_s(x)^*)$ . By the discussion following (6.8)

$A \sim A^{-1} \Leftrightarrow d_j(x) = d_j(x)^*$  for  $1 \leq j \leq s$ , i.e.  $A \sim A^{-1} \Leftrightarrow$  each invariant factor  $d_j(x)$  of  $A$  is palindromic.

(e) The 12 sequences of invariant factors of  $3 \times 3$  matrices  $A$  over  $\mathbb{Z}_3$  such that

$A \sim A^{-1}$  are:

$$\begin{aligned} & (x-1, x-1, x-1), (x+1, x+1, x+1), (x-1, (x-1)^2), (x-1, x^2-1), \\ & (x+1, x^2-1), (x+1, (x+1)^2), (x^3-1), (x^3-x^2+x-1), (x^3+x^2-x-1), \\ & (x^3+1), (x^3+x^2+x+1), (x^3-x^2-x+1). \end{aligned}$$

(f) In case (i)  $\chi_A(x) = x^3 + x + 1$  which is not palindromic. As the product of palindromic polynomials is palindromic we see that the invariant factors of  $A$  cannot all be palindromic. So  $A \sim A^{-1}$  is false. In case (ii)  $\chi_A(x) = (x-1)^2(x+1)$  which is palindromic and further all monic non-constant divisors of  $\chi_A(x)$  are palindromic. So all the invariant factors of  $A$  are palindromic and therefore  $A \sim A^{-1}$  is true.

### Solution 6

(a) Suppose  $\mu_A(x) = \mu_B(x) = x - c$  for some  $c \in F$ . Then  $\mu_A(A) = 0$  gives  $A - cI = 0$ , i.e.  $A = cI$ . In the same way  $\mu_B(B) = 0$  gives  $B = cI$  and so  $A \sim B$  as  $A = B$  in this case. Suppose  $\mu_A(x) = \mu_B(x) \neq x - c$  for any  $c \in F$ . As  $\mu_A(x) \mid \chi_A(x)$  and  $\deg \chi_A(x) = 2$  we see  $\mu_A(x) = \chi_A(x)$ . By (6.10) there is  $v_0$  having order  $\mu_A(x)$  in  $M(A)$ . So  $M(A) = \langle v_0 \rangle$  and  $A \sim C(\mu_A(x))$  by (5.27). For the same

reason  $B \sim C(\mu_B(x))$  and so  $A \sim B$  as  $C(\mu_A(x)) = C(\mu_B(x))$ ,  $\sim$  being an equivalence relation.

Consider  $A = C(x) \oplus C(x(x+1))$  and  $B = C(x+1) \oplus C(x(x+1))$  over  $\mathbb{Q}$  (or indeed any field  $F$ ). As both the  $3 \times 3$  matrices  $A$  and  $B$  are in rcf we see  $A$  and  $B$  are not similar by (6.7). But  $\mu_A(x) = x(x+1) = \mu_B(x)$ .

Over  $\mathbb{F}_q$  there are  $q$  polynomials  $x-c$  and  $q^2$  polynomials  $x^2+ax+b$ . These are the possible minimum polynomials  $\mu_A(x)$  of  $2 \times 2$  matrices  $A$  over  $\mathbb{F}_q$ . By (6.10) and the first part of this question  $A \sim B \Leftrightarrow \mu_A(x) = \mu_B(x)$  where  $A, B \in \mathfrak{M}_2(\mathbb{F}_q)$ . So there are  $q^2 + q$  similarity classes of  $2 \times 2$  matrices over  $\mathbb{F}_q$ .

(b) Suppose  $A \sim B$ . Then  $A$  and  $B$  have the same sequence of invariant factors  $(d_1(x), d_2(x), \dots, d_s(x))$  by (6.6). Then  $\chi_A(x) = d_1(x)d_2(x) \cdots d_s(x) = \chi_B(x)$  by (6.5). Also  $\mu_A(x) = d_s(x) = \mu_B(x)$  by (6.10).

Conversely suppose  $A$  and  $B$  are  $3 \times 3$  matrices over  $F$  with  $\chi_A(x) = \chi_B(x)$  and  $\mu_A(x) = \mu_B(x)$ . Let  $\deg \mu_A(x) = 1$ ; then  $\mu_A(x) = x-c$  and  $A = B = cI$  and so  $A \sim B$ . Let  $\deg \mu_A(x) = 3$ . Then  $\mu_A(x) = \chi_A(x)$  and  $A \sim C(\chi_A(x))$  by (5.27) and (6.10). For the same reason  $B \sim C(\chi_B(x))$  and so  $A \sim B$  as  $C(\chi_A(x)) = C(\chi_B(x))$ . Suppose  $\deg \mu_A(x) = 2$  and so  $(x-c)\mu_A(x) = \chi_A(x)$  for some  $c \in F$  as  $\chi_A(x)/\mu_A(x)$  is monic of degree 1. By (6.11) we see  $(x-c) \mid \mu_A(x)$  as  $(x-c) \mid \chi_A(x)$  and the invariant factor sequence of  $A$  is

$$(x-c, \mu_A(x)) = (\chi_A(x)/\mu_A(x), \mu_A(x)).$$

So  $A$  and  $B$  have the same invariant factor sequence, i.e.  $A \sim B$ . The answer to the question is therefore: Yes!

(c) Applying the first isomorphism theorem for rings (Exercises 2.3, Question 3(b)) to the ring homomorphism  $\varepsilon_A : F[x] \rightarrow \mathfrak{M}_t(F)$  of (5.8) gives

$$\tilde{\varepsilon}_A : F[x]/\langle \mu_A(x) \rangle \cong \text{im } \varepsilon_A.$$

(d) Let  $f(x) = x^m + c_{m-1}x^{m-1} + \dots + c_1x + c_0$  be a monic polynomial of degree  $m$  over  $F$  with  $f(A) = 0$  and  $m < n$ . Then  $f(A) = A^m + c_{m-1}A^{m-1} + \dots + c_1A + c_0I = 0$  contrary to the linear independence of  $I, A, A^2, \dots, A^{n-1}$ . So there is no such polynomial  $f(x)$ . As  $I, A, A^2, \dots, A^{n-1}, A^n$  are linearly dependent there are scalars  $b_0, b_1, \dots, b_{n-1}, b_n$ , not all zero, with  $b_0I + b_1A + \dots + b_{n-1}A^{n-1} + b_nA^n = 0$ . Is it possible for  $b_n = 0$ ? If so then  $b_0 = b_1 = \dots = b_{n-1} = 0$  by the linear independence of  $I, A, A^2, \dots, A^{n-1}$ . Therefore  $b_n \neq 0$ . Write  $a_i = -b_i/b_n$  for  $0 \leq i < n$ . Then the above matrix equation gives  $A^n = a_0I + a_1A + a_2A^2 + \dots + a_{n-1}A^{n-1}$  and so  $\mu_A(x) = x^n - a_{n-1}x^{n-1} - \dots - a_1x - a_0$  is monic of degree  $n$  over  $F$  and satisfies  $\mu_A(A) = 0$ , i.e.  $\mu_A(x)$  satisfies the conditions of (6.9) and so is the minimum polynomial of  $A$ .

$$A^2 = \begin{pmatrix} 3 & 2 & 2 \\ -4 & -3 & -4 \\ 2 & 2 & 3 \end{pmatrix} = 2A - I$$

and so  $\mu_A(x) = x^2 - 2x + 1 = (x-1)^2$  as  $I, A$  are linearly independent. As  $\mu_A(x)$  and  $\chi_A(x)$  have the same irreducible factors by (6.11) we see  $\chi_A(x) = (x-1)^3$  and  $A$  has invariant factor sequence  $(x-1, (x-1)^2)$ .

(e) Write  $l(x) = \text{lcm}\{\mu_{A_1}(x), \mu_{A_2}(x)\}$ . There are  $q_i(x) \in F[x]$  with

$l(x) = q_i(x)\mu_{A_i}(x)$  for  $i=1, 2$ . Using the evaluation homomorphisms  $\mathcal{E}_{A_i}$  we see

$l(A_i) = (l(x))\mathcal{E}_{A_i} = (q_i(x)\mu_{A_i}(x))\mathcal{E}_{A_i} = q_i(A_i)\mu_{A_i}(A_i) = 0$  for  $i=1, 2$  as

$\mu_{A_i}(A_i) = 0$ . Using the theory following (5.17) we obtain

$l(A_1 \oplus A_2) = l(A_1) \oplus l(A_2) = 0 \oplus 0 = 0$  and so  $\mu_{A_1 \oplus A_2}(x) \mid l(x)$ . Conversely suppose

$f(A_1 \oplus A_2) = 0$  where  $f(x) \in F[x]$ . Then  $f(A_1) \oplus f(A_2) = 0$  which means

$f(A_i) = 0$  for  $i=1, 2$ . So  $\mu_{A_i}(x) \mid f(x)$  for  $i=1, 2$  showing that  $f(x)$  is a common multiple of  $\mu_{A_1}(x)$  and  $\mu_{A_2}(x)$ . Therefore  $l(x) \mid f(x)$ . As  $\mu_{A_1 \oplus A_2}(A_1 \oplus A_2) = 0$  we

may take  $f(x) = \mu_{A_1 \oplus A_2}(x)$  obtaining  $l(x) \mid \mu_{A_1 \oplus A_2}(x)$ . We conclude

$l(x) = \mu_{A_1 \oplus A_2}(x)$  as each of these monic polynomials is a divisor of the other.

### Solution 7

(a) There are  $q^3$  monic polynomials of degree 3 over  $\mathbb{F}_q$  as there are  $q$  choices for each of the three coefficients  $a_i$  in  $a_0 + a_1x + a_2x^2 + x^3 = f(x)$ . So there are  $q^3$  similarity classes of  $3 \times 3$  matrices  $A$  over  $\mathbb{F}_q$  having a single invariant factor  $f(x)$ . There are  $q$  similarity classes of  $3 \times 3$  scalar matrices  $A$  over  $\mathbb{F}_q$ , namely those with invariant factor sequence  $(x-a, x-a, x-a)$  for  $a \in \mathbb{F}_q$ . The remaining classes have quadratic minimum polynomial  $d_2(x)$ , which must be reducible over  $\mathbb{F}_q$  as  $d_1(x)$  is a factor of degree 1. There are  $q$  choices for  $d_1(x) = x-a$ , and having chosen  $d_1(x)$  there are  $q$  remaining choices for  $d_2(x)/d_1(x) = x-b$  where  $a, b \in \mathbb{F}_q$ . So there are  $q^2$  similarity classes with invariant factor sequence

$$(d_1(x), d_2(x)) = (x-a, (x-a)(x-b)).$$

In all there are  $q^3 + q^2 + q$  similarity classes of  $3 \times 3$  matrices over  $\mathbb{F}_q$ .

As above, with minor modifications, there are  $q^4$  similarity classes of  $4 \times 4$  matrices over  $\mathbb{F}_q$  with a single invariant factor, which must be monic of degree 4 over  $\mathbb{F}_q$ . Also there are  $q$  similarity classes of  $4 \times 4$  scalar matrices  $A$  over  $\mathbb{F}_q$ . There are  $q^3$  invariant factor sequences  $((x-a), (x-a)f(x))$  and  $q^2$  invariant factor sequences  $(f(x), f(x))$  where  $f(x)$  is monic quadratic. Also there are invariant factor sequences of length 3, namely  $q^2$  of type  $(x-a, x-a, (x-a)(x-b))$  where  $a, b \in \mathbb{F}_q$  and  $a=b$  is allowed. So there are  $q^4 + q^3 + 2q^2 + q$  similarity classes of  $4 \times 4$  matrices over  $\mathbb{F}_q$ . More generally consider an  $n \times n$  matrix  $A$  over  $\mathbb{F}_q$  having minimum polynomial of degree  $m$ . Let  $(d_1(x), d_2(x), \dots, d_s(x))$  be the sequence of invariant factors of  $A$  and write  $t_j = \deg d_j(x)$  for  $1 \leq j \leq s$ . By (6.5) and (6.10) the sequence  $(t_1, t_2, \dots, t_s)$  is a partition of  $n$  with largest part  $t_s = m$ . How many invariant factor sequences

$(d_1(x), d_2(x), \dots, d_s(x))$  give rise to the same partition  $(t_1, t_2, \dots, t_s)$  ? There are  $q^{t_1}$  choices for  $d_1(x)$ , namely any monic polynomial over  $\mathbb{F}_q$  of degree  $t_1$ . Assuming  $s \geq 2$ , having chosen  $d_1(x)$  there remain  $q^{t_2-t_1}$  choices for  $d_2(x)/d_1(x)$ , namely any monic polynomial of degree  $t_2-t_1$  over  $\mathbb{F}_q$ . So in all there are  $q^{t_1} \times q^{t_2-t_1} = q^{t_2}$  choices for the pair  $(d_1(x), d_2(x))$  as  $d_2(x) = d_1(x) \times (d_2(x)/d_1(x))$ . Assuming inductively that there are  $q^{t_j}$  choices for  $(d_1(x), d_2(x), \dots, d_j(x))$  where  $1 \leq j < s$  we see that there are  $q^{t_{j+1}-t_j}$  choices for  $d_{j+1}(x)/d_j(x)$ , namely any monic polynomial of degree  $t_{j+1}-t_j$  over  $\mathbb{F}_q$ . Hence there are  $q^{t_j} \times q^{t_{j+1}-t_j} = q^{t_{j+1}}$  choices for  $(d_1(x), d_2(x), \dots, d_{j+1}(x))$  thereby completing the induction. So  $q^m$  invariant factor sequences  $(d_1(x), d_2(x), \dots, d_s(x))$  give rise to each of the  $P(n, m)$  partitions  $(t_1, t_2, \dots, t_s)$  of  $n$  with largest part  $m$ . By the discussion following (6.8) we see that there are  $P(n, m)q^m$  similarity classes of  $n \times n$  matrices over  $\mathbb{F}_q$  with minimum polynomial of degree  $m$ .

Remember  $A \in GL_n(F) \Leftrightarrow A$  is invertible over the field  $F \Leftrightarrow \mu_A(0) = d_s(0) \neq 0$ . The number of monic polynomials  $f(x)$  over  $\mathbb{F}_q$  with  $f(0) \neq 0$  is

$N(\deg f(x)) = q^{\deg f(x)-1}(q-1)$  as there are  $q$  choices for the coefficient of  $x^i$  in such a polynomial for  $0 < i < \deg f(x)$  and  $q-1$  choices for the constant term (any non-zero element). Consider a partition  $(t'_1, \dots, t'_l)$  of  $n$  in which there are  $m_i$  parts  $t'_i$  for  $1 \leq i \leq l$  and  $t'_1 < t'_2 < \dots < t'_l$ ,  $t'_l = m$ ; we have adopted a change of notation here as the number of parts in  $(t'_1, \dots, t'_l)$  is  $s = m_1 + m_2 + \dots + m_l$ . There are  $P(n, m, l)$  such partitions. Taking repetitions into account we write

$$(d_1(x), d_2(x), \dots, d_s(x)) = (d'_1(x), \dots, d'_l(x))$$

for the sequence of invariant factors of a matrix in  $GL_n(\mathbb{F}_q)$ . How many satisfy  $(\deg d'_1(x), \dots, \deg d'_s(x)) = (t'_1, \dots, t'_l)$ ? Such a sequence has  $m_i$  factors equal  $d'_i(x)$  for  $1 \leq i \leq l$  where  $1 \leq \deg d'_1(x) < \deg d'_2(x) < \dots < \deg d'_l(x)$  and is specified by the  $l$  non-constant monic polynomials  $d'_1(x), d'_2(x)/d'_1(x), d'_3(x)/d'_2(x), \dots, d'_l(x)/d'_{l-1}(x)$  none of which vanish (take the value 0) at 0. The number of such sequences is therefore

$$N(t_1) \times N(t_2 - t_1) \times N(t_3 - t_2) \times \dots \times N(t_l - t_{l-1}) = q^{t_1-1}(q-1)q^{t_2-t_1-1}(q-1)q^{t_3-t_2-1}(q-1) \dots q^{t_l-t_{l-1}-1}(q-1) = q^{m-l}(q-1)^l$$

as  $t_l = m$ . So the number of conjugacy (similarity) classes in  $GL_n(\mathbb{F}_q)$  is

$$\sum_{m=1}^n \left( \sum_{l=1}^m P(n, m, l) q^{m-l} (q-1)^l \right).$$

The partitions contributing to  $P(8, 3)$  are

$(1, 1, 1, 1, 3), (1, 1, 1, 2, 3), (1, 2, 2, 3), (1, 1, 3, 3), (2, 3, 3)$ . So  $P(8, 3) = 5$ ,  $P(8, 3, 1) = 0$ ,  $P(8, 3, 2) = 3$  and  $P(8, 3, 3) = 2$ . Therefore the number of conjugacy classes of elements of  $GL_8(\mathbb{Z}_5)$  having cubic minimum polynomial is

$$0 \times 5^2 \times (5-1)^1 + 3 \times 5^1 \times (5-1)^2 + 2 \times 5^0 \times (5-1)^3 = 368.$$

(b) As  $A^2 = 0$  the last invariant factor  $\mu_A(x)$  of  $A$ , being also the minimum polynomial of  $A$ , is a (monic, non-constant) divisor of  $x^2$ . So  $\mu_A(x) = x$  or  $\mu_A(x) = x^2$ . The characteristic polynomial  $\chi_A(x)$  of  $A$  is a power of  $x$  and of degree  $t$ ; so  $\chi_A(x) = x^t$  is the product of the invariant factors of  $A$ . Now  $\mu_A(x) = x \Leftrightarrow A = 0$  and the zero  $t \times t$  matrix has invariant factor sequence  $(x, x, \dots, x) \in F[x]^t$ . In the case  $\mu_A(x) = x^2$  the invariant factor sequence of  $A$  is  $(x, x, \dots, x, x^2, x^2, \dots, x^2)$  where there are  $r$  invariant factors  $x^2$  for  $1 \leq r \leq \lfloor t/2 \rfloor$ ; in fact  $r = \text{rank } A$  on comparing  $A$  with its rcf. So there are  $1 + \lfloor t/2 \rfloor$  similarity classes of  $t \times t$  matrices  $A$  over  $F$  satisfying  $A^2 = 0$  corresponding to the  $1 + \lfloor t/2 \rfloor$  possible values of  $\text{rank } A$ .

Let  $A$  be a  $t \times t$  matrix over  $\mathbb{Z}_2$ . Then  $A^2 - I = (A - I)^2$  and so  $A^2 = I \Leftrightarrow (A - I)^2 = 0$ . Also  $A \sim B \Leftrightarrow A - I \sim B - I$ . By the above paragraph there are  $\lfloor t/2 \rfloor$  similarity classes of  $t \times t$  involutions  $A$  over  $\mathbb{Z}_2$  corresponding to the  $\lfloor t/2 \rfloor$  values of  $r = \text{rank}(A - I)$ , i.e.  $1 \leq r \leq \lfloor t/2 \rfloor$ .

Suppose  $\mathbb{Z}_2$  is replaced by a field  $F$  of characteristic 2. The answer is unchanged as the theory of the above paragraph applies equally well to  $F$ : the group  $GL_t(F)$  has  $\lfloor t/2 \rfloor$  conjugacy classes of involutions.

(c) Split the set of  $N(t)$  sequences  $(d_1(x), d_2(x), \dots, d_s(x))$  of invariant factors of the  $t \times t$  matrix  $A$  satisfying  $A^3 = 0$  into two subsets: the first with  $d_s(x) \neq x^3$  and the second with  $d_s(x) = x^3$ . By (b) above there are  $\lfloor t/2 \rfloor + 1 = \lfloor (t+2)/2 \rfloor$  sequences in the first subset as  $d_s(x) \mid x^2$  and so  $A^2 = 0$ . There are  $N(t-3)$  sequences  $(d_1(x), d_2(x), \dots, d_{s-1}(x), x^3)$  in the second subset as they are precisely those with  $(d_1(x), d_2(x), \dots, d_{s-1}(x))$  being the invariant factor sequence of a  $(t-3) \times (t-3)$  matrix  $B$  over  $F$  with  $B^3 = 0$ . Counting sequences gives  $N(t) = \lfloor (t+2)/2 \rfloor + N(t-3)$ .

The sequences for  $t = 1, 2, 3$  are  $(x)$ ;  $(x, x), (x^2)$ ;  $(x, x, x), (x, x^2), (x^3)$  and so  $N(1) = 1, N(2) = 2, N(3) = 3$ . Using the above formula:  $N(4) = 3 + 1 = 4$ ,  $N(5) = 4 + 1 = 5$ ,  $N(6) = 5 + 1 = 6$ ,  $N(7) = 6 + 1 = 7$ ,  $N(8) = 7 + 1 = 8$ ,  $N(9) = 8 + 1 = 9$ ,  $N(10) = 9 + 1 = 10$ .

Suppose  $t$  to be odd. Then  $t-3$  is even. Write  $t = 2u + 1$  and so  $\lfloor t/2 \rfloor = u$ ,  $(t-3)/2 = u - 1$ . Therefore

$$N'(t) - N'(t-3) = (1/2)(u+1)(u+2) - (1/2)(u-1)u = u + 1$$

as the second terms in  $N'(t)$  and  $N'(t-3)$  are equal (and so cancel out). As

$$\lfloor (t+2)/2 \rfloor = \lfloor (2u+3)/2 \rfloor = \lfloor u + 1 + 1/2 \rfloor = u + 1 \text{ we see}$$

$$N'(t) = \lfloor (t+2)/2 \rfloor + N'(t-3) \text{ for } t \text{ odd.}$$

Suppose  $t$  to be even. Then  $t-3$  is odd. Write  $t = 2u$  and so

$$\lfloor (t-3)/2 \rfloor = \lfloor u - 3/2 \rfloor = u - 2, \lfloor t/6 \rfloor = \lfloor u/3 \rfloor, \lfloor (t-6)/6 \rfloor = \lfloor u/3 - 1 \rfloor = \lfloor u/3 \rfloor - 1.$$

Therefore  $N'(t) = (u+1)(u+2)/2 - (\lfloor u/3 \rfloor + 1)(u - (3/2)\lfloor u/3 \rfloor)$  and

$$N'(t-3) = (u-1)u/2 - \lfloor u/3 \rfloor(u - 3 - (3/2)(\lfloor u/3 \rfloor - 1)). \text{ Subtracting gives}$$

$N'(t) - N'(t-3) = 2u + 1 - u = u + 1 = (t/2) + 1 = (t+2)/2 = \lfloor (t+2)/2 \rfloor$  as the terms involving  $\lfloor u/3 \rfloor$  cancel. So  $N'(t) = \lfloor (t+2)/2 \rfloor + N'(t-3)$  for  $t$  even.

As  $\lfloor (1-3)/6 \rfloor = -1$  we obtain  $N'(1) = 1 - 0 = 1$ . Also  $N'(2) = 3 - 1 = 2$  and  $N'(3) = 3 - 0 = 3$ . So  $N'(t) = N(t)$  for  $t = 1, 2, 3$ . Let  $t > 3$  and assume inductively that  $N'(t-3) = N(t-3)$ . Then

$$N'(t) = \lfloor (t+2)/2 \rfloor + N'(t-3) = \lfloor (t+2)/2 \rfloor + N(t-3) = N(t)$$

completing the induction, i.e.  $N'(t) = N(t)$  for  $t \geq 1$ .

Let  $B$  be a  $t \times t$  matrix over  $\mathbb{Z}_3$  (or any field of characteristic 3). Write  $B - I = A$ .

Then  $B \neq I \Leftrightarrow A \neq 0$  and as  $A^3 = (B - I)^3 = B^3 - 3B^2 + 3B - I = B^3 - I$  we see

$B^3 = I \Leftrightarrow A^3 = 0$ . Also  $B \sim B' \Leftrightarrow A \sim A'$  where  $B'$  be a  $t \times t$  matrix over  $\mathbb{Z}_3$  and

$B' - I = A'$ . So the number of conjugacy classes of elements of order 3 in  $GL_t(\mathbb{Z}_3)$  is

$N(t) - 1$ , the number of similarity classes of  $t \times t$  matrices  $A$  over  $\mathbb{Z}_3$  satisfying

$A \neq 0, A^3 = 0$ . So there are  $N(100) - 1 = 883$  conjugacy classes of elements of order 3 in  $GL_{100}(\mathbb{Z}_3)$ .

### Solution 8

(a) The proof is by induction on  $t$ . First consider  $t = 1$ . Submodules  $K$  of  $F[x]^1 = F[x]$  are precisely ideals  $K$  of  $F[x]$ . By (4.4) there is  $d(x)$  in  $K$  with  $K = \langle d(x) \rangle$ . By convention the empty set  $\emptyset$  is regarded as being an  $F[x]$ -basis of  $K = \langle 0(x) \rangle$ , that is,  $K = \{0(x)\}$  is free of rank  $s = 0$ . For  $K \neq \{0(x)\}$  the single non-zero polynomial  $d(x)$  is an  $F[x]$ -basis of  $K$  and so  $s = 1$ .

Now suppose  $t > 1$ . It is convenient to regard  $F[x]^{t-1}$  as being the submodule of

$F[x]^t$  consisting of  $t$ -tuples of polynomials having last entry zero, that is,

$F[x]^{t-1} = \{(f_1(x), f_2(x), \dots, f_{t-1}(x), 0(x)) \in F[x]^t\}$ . Let  $K$  be a submodule of  $F[x]^t$

and consider  $K' = \{f_t(x) : (f_1(x), f_2(x), \dots, f_t(x)) \in K\}$ , that is,  $K'$  consists of those polynomials  $f_t(x)$  which occur in the last place of  $t$ -tuples in  $K$ . Then  $K'$  is an ideal of  $F[x]^t$ . By (4.4) there is  $d(x)$  in  $F[x]$  with  $K' = \langle d(x) \rangle$ . The intersection

$K \cap F[x]^{t-1}$  is a submodule of  $F[x]^{t-1}$  and so by inductive hypothesis  $K \cap F[x]^{t-1}$

has an  $F[x]$ -basis  $z_1(x), z_2(x), \dots, z_{s-1}(x)$  where  $s \leq t$ . If  $d(x) = 0(x)$  then

$K \subseteq F[x]^{t-1}$  and  $K = K \cap F[x]^{t-1}$  has  $F[x]$ -basis as above; so  $K$  is free of rank

$s - 1 < t$ . If  $d(x) \neq 0(x)$  there is a  $t$ -tuple  $z_s(x)$  in  $K$  having last entry  $d(x)$ . We finish the proof by showing that  $z_1(x), z_2(x), \dots, z_{s-1}(x), z_s(x)$  is an  $F[x]$ -basis of  $K$ .

Let  $k \in K$ . The last entry in the  $t$ -tuple  $k$  belongs to  $K'$  and so is  $q(x)d(x)$  for some  $q(x)$  in  $F[x]$ . Hence  $k - q(x)z_s(x)$  has last entry zero. As  $z_s(x)$  belongs to  $K$  so also does  $k - q(x)z_s(x)$ . Therefore

$k - q(x)z_s(x) \in K \cap F[x]^{t-1} = \langle z_1(x), z_2(x), \dots, z_{s-1}(x) \rangle$ . Hence there are

$f_1(x), f_2(x), \dots, f_{s-1}(x)$  in  $F[x]$  such that

$k = f_1(x)z_1(x) + f_2(x)z_2(x) + \dots + f_{s-1}(x)z_{s-1}(x) + q(x)z_s(x)$ . So

$K = \langle z_1(x), z_2(x), \dots, z_{s-1}(x), z_s(x) \rangle$ , that is,  $z_1(x), z_2(x), \dots, z_{s-1}(x), z_s(x)$  generate  $K$  according to (2.19)(i).



We now show that  $z_1(x), z_2(x), \dots, z_{s-1}(x), z_s(x)$  are  $F[x]$ -independent. So suppose there are  $f_1(x), f_2(x), \dots, f_{s-1}(x), f_s(x)$  in  $F[x]$  with  $f_1(x)z_1(x) + f_2(x)z_2(x) + \dots + f_{s-1}(x)z_{s-1}(x) + f_s(x)z_s(x) = 0(x)$ . Comparing last entries gives  $f_s(x)d(x) = 0(x)$  as the last entry in each of  $z_1(x), z_2(x), \dots, z_{s-1}(x)$  is the zero polynomial and  $z_s(x)$  has last entry  $d(x)$ . Since  $d(x) \neq 0(x)$  we deduce  $f_s(x) = 0(x)$ . This leaves  $f_1(x)z_1(x) + f_2(x)z_2(x) + \dots + f_{s-1}(x)z_{s-1}(x) = 0(x)$ . As  $z_1(x), z_2(x), \dots, z_{s-1}(x)$  form a  $F[x]$ -basis of  $K \cap F[x]^{t-1}$  they are  $F[x]$ -independent (2.19) and so  $f_1(x) = f_2(x) = \dots = f_{s-1}(x) = 0(x)$ . Hence  $z_1(x), z_2(x), \dots, z_{s-1}(x), z_s(x)$  are indeed  $F[x]$ -independent and so form an  $F[x]$ -basis of  $K$ . The induction is now complete as  $\text{rank } K = s \leq t$ .

(b) Consider the  $t$ -tuples  $v(x) = (f_1(x), f_2(x), \dots, f_t(x)) \in F[x]^t$  and

$v'(x) = (f'_1(x), f'_2(x), \dots, f'_t(x)) \in F[x]^t$  and let  $f(x) \in F[x]$ . Then  $\theta_A$  is additive as

$$(v(x) + v'(x))\theta_A = \sum_{i=1}^t (f_i(x) + f'_i(x))e_i = \sum_{i=1}^t f_i(x)e_i + \sum_{i=1}^t f'_i(x)e_i = (v(x))\theta_A + (v'(x))\theta_A$$

using the module laws which hold in  $M(A)$  by (5.7) and (5.8). For the same reason

$$(f(x)v(x))\theta_A = \sum_{i=1}^t (f(x)f_i(x))e_i = f(x)\left(\sum_{i=1}^t f_i(x)e_i\right) = f(x)((v(x))\theta_A)$$

and so  $\theta_A$  is  $F[x]$ -linear.

(c) Take  $d(x) = d_1(x)$  and consider the isomorphism  $\alpha|: M_{(d_1)} \cong M'_{(d_1)}$ . Then

$$(F[x]/\langle d_i(x) \rangle)_{(d_1(x))} \cong F[x]/\langle d_1(x) \rangle \text{ since } \gcd\{d_1(x), d_i(x)\} = d_1(x) \text{ for } 1 \leq i \leq s.$$

From  $M \cong F[x]/\langle d_1(x) \rangle \oplus F[x]/\langle d_2(x) \rangle \oplus \dots \oplus F[x]/\langle d_s(x) \rangle$  we deduce

$$\begin{aligned} M_{(d_1(x))} &\cong (F[x]/\langle d_1(x) \rangle)_{(d_1(x))} \oplus (F[x]/\langle d_2(x) \rangle)_{(d_1(x))} \oplus \dots \oplus (F[x]/\langle d_s(x) \rangle)_{(d_1(x))} \\ &\cong F[x]/\langle d_1(x) \rangle \oplus F[x]/\langle d_1(x) \rangle \oplus \dots \oplus F[x]/\langle d_1(x) \rangle = (F[x]/\langle d_1(x) \rangle)^s. \end{aligned}$$

The  $F[x]$ -module  $M_{(d_1(x))}$  is therefore isomorphic to the free  $F[x]/\langle d_1(x) \rangle$ -module

$$(F[x]/\langle d_1(x) \rangle)^s \text{ of rank } s. \text{ By (2.25) both } M_{(d_1(x))} \text{ and } M'_{(d_1(x))} \text{ are free}$$

$F[x]$ -modules of rank  $s$ . Combining

$$(F[x]/\langle d'_i(x) \rangle)_{(d_1(x))} \cong F[x]/\langle \gcd\{d_1(x), d'_i(x)\} \rangle \text{ for } 1 \leq i \leq s' \text{ and}$$

$$M' \cong F[x]/\langle d'_1(x) \rangle \oplus F[x]/\langle d'_2(x) \rangle \oplus \dots \oplus F[x]/\langle d'_{s'}(x) \rangle \text{ gives}$$

$$M'_{(d_1(x))} \cong \bigoplus_{i=1}^{s'} F[x]/\langle \gcd\{d_1(x), d'_i(x)\} \rangle \text{ showing that } M'_{(d_1(x))} \text{ is the direct sum of}$$

$s'$  cyclic submodules and so is generated by  $s'$  of its elements. From (2.20) we

deduce  $s' \geq s$ . As  $\alpha^{-1}: M' \cong M$  the preceding theory 'works' with  $M$  and

$$M' \text{ interchanged. Using } \alpha^{-1}|: M'_{(d'_1(x))} \cong M_{(d'_1(x))} \text{ the } F[x]/\langle d'_1(x) \rangle\text{-module } M'_{(d'_1(x))}$$

is seen to be isomorphic to the free  $F[x]/\langle d'_1(x) \rangle$ -module  $(F[x]/\langle d'_1(x) \rangle)^{s'}$  of rank  $s'$ . Hence

$$M_{(d'_1(x))} \cong \bigoplus_{i=1}^s F[x]/\langle \gcd\{d'_1(x), d_i(x)\} \rangle$$

is a free  $F[x]/\langle d'_1(x) \rangle$ -module of rank  $s'$  and is generated by  $s$  of its elements.

Therefore  $s \geq s'$  by (2.20) and so  $s = s'$ . From (2.18) and (2.21) the  $s$  generators

of  $M'_{(d_1(x))}$  form an  $F[x]/\langle d_1(x) \rangle$  – basis of the  $F[x]/\langle d_1(x) \rangle$  – module  $M'_{(d_1(x))}$  (Exercises 2.3, Question 7(b)). Therefore each of these  $s$  generators has order ideal  $\{\bar{0}\} = \{\langle d_1(x) \rangle\}$  in the  $F[x]/\langle d_1(x) \rangle$  – module  $M'_{(d_1(x))}$  and order  $d_1(x)$  in the

$F[x]$  – module  $M'_{(d_1(x))}$ . From the first of these generators we deduce

$\gcd\{d_1(x), d'_1(x)\} = d_1(x)$  showing  $d_1(x) \mid d'_1(x)$ . Interchanging  $M$  and  $M'$  gives  $d'_1(x) \mid d_1(x)$  and so  $d_1(x) = d'_1(x)$ .

Let  $m_1$  denote the number of  $i$  with  $d_i(x) = d_1(x)$  and let  $m'_1$  denote the number of  $i$  with  $d'_i(x) = d_1(x)$ . As

$$d_1(x)(F[x]/\langle d_i(x) \rangle) \cong F[x]/\langle d_i(x)/\gcd\{d_1(x), d_i(x)\} \rangle = F[x]/\langle d_i(x)/d_1(x) \rangle$$

we obtain

$$d_1(x)M \cong \sum_{m_1 < i \leq s} \oplus F[x]/\langle d_i(x)/d_1(x) \rangle.$$

So  $d_1(x)M$  is the direct sum of  $s - m_1$  non-trivial cyclic submodules. As

$d_i(x)/d_1(x) \mid d_j(x)/d_1(x)$  for  $m_1 < i \leq j \leq s$  this decomposition of  $d_1(x)M$  is again as in (6.6). In the same way

$$d_1(x)M' \cong \sum_{m'_1 < i \leq s} \oplus F[x]/\langle d'_i(x)/d_1(x) \rangle$$

which is a decomposition of  $d_1(x)M'$  into  $s' - m'_1$  non-trivial cyclic submodules as in (6.6). As  $\alpha \mid d_1(x)M \cong d_1(x)M'$  the proof can be completed by induction on the

number,  $r$  say, of different polynomials among  $d_1(x), d_2(x), \dots, d_s(x)$ . Take  $r = 1$ .

Then  $m_1 = s$  and  $d_1(x)M$  is trivial. So  $d_1(x)M'$  is also trivial and  $m'_1 = s$ . Therefore

$d_i(x) = d_1(x) = d'_i(x)$  for  $1 \leq i \leq s$ . Now take  $r > 1$ . There are  $r - 1$  different

polynomials among  $d_{m_1+1}(x)/d_1(x), d_{m_1+2}(x)/d_1(x), \dots, d_s(x)/d_1(x)$  and so the

conclusion of (6.6) holds on replacing  $\alpha: M \cong M'$  by  $\alpha \mid d_1(x)M \cong d_1(x)M'$ , that is,

$s - m_1 = s - m'_1$  (showing  $m_1 = m'_1$ ) and also  $d_i(x)/d_1(x) = d'_i(x)/d_1(x)$  for

$m_1 < i \leq s$ . Therefore  $s = s'$ ,  $d_i(x) = d_1(x) = d'_i(x)$  for  $1 \leq i \leq m_1 = m'_1$  and

$d_i(x) = d'_i(x)$  for  $m_1 < i \leq s'$  on multiplying by  $d_1(x)$ . The induction is now complete.

## Solutions 6.2 (page 299)

### Solution 1

(a) (i) One way of reducing  $xI - A$  to its Smith normal form  $S(xI - A)$  is set out below:

$$\begin{aligned}
 xI - A &= \begin{pmatrix} x-1 & 1 & -2 \\ -1 & x+1 & -2 \\ 1 & -1 & x+2 \end{pmatrix} \begin{matrix} \equiv \\ r_1 \leftrightarrow r_3 \end{matrix} \begin{pmatrix} 1 & -1 & x+2 \\ -1 & x+1 & -2 \\ x-1 & 1 & -2 \end{pmatrix} \begin{matrix} \equiv \\ c_2 + c_1 \\ c_3 - (x+2)c_1 \end{matrix} \\
 &\begin{pmatrix} 1 & 0 & 0 \\ -1 & x & x \\ x-1 & x & -x(x+1) \end{pmatrix} \begin{matrix} \equiv \\ r_2 + r_1 \\ r_3 - (x-1)r_1 \end{matrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & x & x \\ 0 & x & -x(x+1) \end{pmatrix} \begin{matrix} \equiv \\ c_3 - c_2 \\ r_3 - r_2 \end{matrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & x & 0 \\ 0 & 0 & -x(x+2) \end{pmatrix} \\
 &\equiv \begin{pmatrix} 1 & 0 & 0 \\ 0 & x & 0 \\ 0 & 0 & x(x+2) \end{pmatrix} \begin{matrix} \equiv \\ -r_3 \end{matrix} = S(xI - A).
 \end{aligned}$$

Applying the above *eros* to  $I$  we obtain  $P(x)$ :

$$I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{matrix} \equiv \\ r_1 \leftrightarrow r_3 \end{matrix} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{matrix} \equiv \\ r_2 + r_1 \\ r_3 - (x-1)r_1 \end{matrix} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & -x+1 \end{pmatrix} \begin{matrix} \equiv \\ r_3 - r_2 \\ -r_3 \end{matrix} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ -1 & 1 & x \end{pmatrix} = P(x).$$

Applying the conjugates of the above *ecos* to  $I$  produces  $Q(x)$ :

$$I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{matrix} \equiv \\ r_1 - r_2 \\ r_1 + (x+2)r_3 \end{matrix} \begin{pmatrix} 1 & -1 & x+2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{matrix} \equiv \\ r_2 + r_3 \end{matrix} \begin{pmatrix} 1 & -1 & x+2 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} = Q(x).$$

Then  $P(x)$  and  $Q(x)$  are invertible over  $\mathbb{Q}[x]$  and satisfy

$P(x)(xI - A) = S(xI - A)Q(x)$ , i.e.  $P(x)(xI - A)Q(x)^{-1} = \text{diag}(1, x, x(x+2))$  is in Smith normal form. As usual write  $\rho_i(x) = e_i Q(x)$  for  $i = 1, 2, 3$ . Using the linear mapping  $\theta_A : \mathbb{Q}[x]^3 \rightarrow M(A)$  and (6.5),  $(\rho_1(x))\theta_A = 0$ , the vector  $v_1 = (\rho_2(x))\theta_A = (0, 1, 1)$  has order  $x$  in  $M(A)$  and generates the submodule  $N_1$ , and  $v_2 = (\rho_3(x))\theta_A = (0, 0, 1)$  has order  $x(x+2)$  and generates the submodule  $N_2$  where  $M(A) = N_1 \oplus N_2$ . So

$$X = \begin{pmatrix} v_1 \\ v_2 \\ xv_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ -1 & 1 & -2 \end{pmatrix}$$

is invertible over  $\mathbb{Q}$  as  $\det X = -1$  and satisfies

$$XAX^{-1} = C(x) \oplus C(x(x+2)) = \left( \begin{array}{c|cc} 0 & 0 & 0 \\ \hline 0 & 0 & 1 \\ \hline 0 & 0 & -2 \end{array} \right)$$

which is the rcf of  $A$ . Now  $M(A)_x$  has  $\mathbb{Q}$ -basis  $v_1, (x+2)v_2$ , i.e.  $(0, 1, 1), (-1, 1, 0)$ . Also  $M(A)_{x+2}$  has  $\mathbb{Q}$ -basis  $xv_2 = (-1, 1, -2)$ . Using these bases as the rows of  $Y$  gives

$$Y = \begin{pmatrix} v_1 \\ (x+2)v_2 \\ -xv_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 1 \\ -1 & 1 & 0 \\ -1 & 1 & -2 \end{pmatrix}$$

which is invertible over  $\mathbb{Q}$  as  $\det Y = -2$  and

$YAY^{-1} = C(x) \oplus C(x) \oplus C(x+2) = \text{diag}(0, 0, -2)$  is in pcf.

(ii) Reducing  $xI - A$  to its Smith normal form  $S(xI - A)$ :

$$\begin{aligned} xI - A &= \begin{pmatrix} x+1 & -1 & -1 \\ -1 & x+2 & 1 \\ 1 & -1 & x-1 \end{pmatrix} \begin{matrix} \\ r_1 \leftrightarrow r_3 \\ \end{matrix} \equiv \begin{pmatrix} 1 & -1 & x-1 \\ -1 & x+2 & 1 \\ x+1 & -1 & -1 \end{pmatrix} \begin{matrix} \\ c_2 + c_1 \\ c_3 - (x-1)c_1 \end{matrix} \\ &\equiv \begin{pmatrix} 1 & 0 & 0 \\ -1 & x+1 & x \\ x+1 & x & -x^2 \end{pmatrix} \begin{matrix} \\ r_2 + r_1 \\ r_3 - (x+1)r_1 \end{matrix} \equiv \begin{pmatrix} 1 & 0 & 0 \\ 0 & x+1 & x \\ 0 & x & -x^2 \end{pmatrix} \begin{matrix} \\ c_2 - c_3 \\ c_3 - xc_2 \end{matrix} \equiv \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & x^2 + x & -(x+2)x^2 \end{pmatrix} \\ &\equiv \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & (x+2)x^2 \end{pmatrix} \begin{matrix} \\ r_3 - (x^2 + x)r_2 \\ -r_3 \end{matrix} = S(xI - A). \end{aligned}$$

Calculating  $P(x)$ :

$$\begin{aligned} I &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{matrix} \\ r_1 \leftrightarrow r_3 \\ \end{matrix} \equiv \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{matrix} \\ r_2 + r_1 \\ r_3 - (x+1)r_1 \end{matrix} \\ &\equiv \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & -x-1 \end{pmatrix} \begin{matrix} \\ r_3 - x(x+1)r_2 \\ -r_3 \end{matrix} \equiv \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ -1 & x(x+1) & (x+1)^2 \end{pmatrix} = P(x). \end{aligned}$$

Calculating  $Q(x)$ :

$$I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{matrix} \\ r_1 - r_2 \\ r_1 + (x-1)r_3 \end{matrix} \equiv \begin{pmatrix} 1 & -1 & x-1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{matrix} \\ r_3 + r_2 \\ r_2 + xr_3 \end{matrix} \equiv \begin{pmatrix} 1 & -1 & x-1 \\ 0 & x+1 & x \\ 0 & 1 & 1 \end{pmatrix} = Q(x).$$

Then  $|P(x)| = |Q(x)| = 1$  and so  $P(x)$  and  $Q(x)$  are invertible over  $\mathbb{Q}[x]$ . As

$P(x)(xI - A) = S(xI - A)Q(x)$  we deduce  $P(x)(xI - A)Q(x)^{-1} = \text{diag}(1, 1, (x+2)x^2)$  is in Smith normal form. In this case  $M(A)$  is cyclic with generator

$v_1 = (\rho_3(x))\theta_A = (0, 1, 1)$  of order  $(x+2)x^2$ . The matrix

$$X = \begin{pmatrix} v_1 \\ xv_1 \\ x^2v_1 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 1 \\ 0 & -1 & 0 \\ -1 & 2 & 1 \end{pmatrix}$$

is invertible over  $\mathbb{Q}$  and satisfies  $XA = C((x+2)x^2)X$ . So  $XAX^{-1} = C((x+2)x^2)$

which is in rcf. The vector  $(x+2)v_1 = (0, -1, 0) + (0, 2, 2) = (0, 1, 2)$  has order  $x^2$  and generates the submodule  $M(A)_x$ . So the vectors  $(x+2)v_1, x(x+2)v_1$ , that is,

$(0, 1, 2), (-1, 0, 1)$  form a  $\mathbb{Q}$ -basis of  $M(A)_x$ . The vector  $x^2v_1 = (-1, 2, 1)$  has order  $x+2$  and so is a  $\mathbb{Q}$ -basis of  $M(A)_{x+2}$ . As  $\mathbb{Q}^3 = M(A)_x \oplus M(A)_{x+2}$  we see that

$$Y = \begin{pmatrix} (x+2)v_1 \\ \frac{x(x+2)v_1}{x^2v_1} \end{pmatrix} = \begin{pmatrix} 0 & 1 & 2 \\ -1 & 0 & 1 \\ -1 & 2 & 1 \end{pmatrix}$$

is invertible over  $\mathbb{Q}$  (its determinant is  $-4$ ) and satisfies  $YA = (C(x^2) \oplus C(x+2))Y$ .

So  $YAY^{-1} = C(x^2) \oplus C(x+2)$  is in pcf.

(iii) Reducing  $xI - A$  to its Smith normal form  $S(xI - A)$ :

$$\begin{aligned} xI - A &= \begin{pmatrix} x+3 & -1 & -2 \\ 1 & x+1 & -2 \\ 1 & -1 & x \end{pmatrix} \xrightarrow{r_1 \leftrightarrow r_3} \begin{pmatrix} 1 & -1 & x \\ 1 & x+1 & -2 \\ x+3 & -1 & -2 \end{pmatrix} \\ &\equiv \begin{pmatrix} 1 & 0 & 0 \\ 1 & x+2 & -x-2 \\ x+3 & x+2 & -x^2-3x-2 \end{pmatrix} \xrightarrow{r_2 - r_1, r_3 - (x+3)r_1} \begin{pmatrix} 1 & 0 & 0 \\ 0 & x+2 & -x-2 \\ 0 & x+2 & -(x+1)(x+2) \end{pmatrix} \\ &\equiv \begin{pmatrix} 1 & 0 & 0 \\ 0 & x+2 & 0 \\ 0 & x+2 & -x(x+2) \end{pmatrix} \xrightarrow{r_3 - r_2, -r_3} \begin{pmatrix} 1 & 0 & 0 \\ 0 & x+2 & 0 \\ 0 & 0 & x(x+2) \end{pmatrix} = S(xI - A). \end{aligned}$$

Calculating  $P(x)$ :

$$\begin{aligned} I &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{r_1 \leftrightarrow r_3} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \xrightarrow{r_2 - r_1, r_3 - (x+3)r_1} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & -1 \\ 1 & 0 & -x-3 \end{pmatrix} \\ &\equiv \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & -1 \\ -1 & 1 & x+2 \end{pmatrix} \xrightarrow{r_3 - r_2, -r_3} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & -1 \\ -1 & 1 & x+2 \end{pmatrix} = P(x). \end{aligned}$$

Calculating  $Q(x)$ :

$$I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{r_1 - r_2, r_1 + xr_3} \begin{pmatrix} 1 & -1 & x \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{r_2 - r_3} \begin{pmatrix} 1 & -1 & x \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} = Q(x).$$

Then  $|P(x)| = |Q(x)| = 1$  and so  $P(x)$  and  $Q(x)$  are invertible over  $\mathbb{Q}[x]$ . As

$P(x)(xI - A) = S(xI - A)Q(x)$  we deduce

$$P(x)(xI - A)Q(x)^{-1} = \text{diag}(1, x+2, x(x+2))$$

is in Smith normal form. Using row 2 of  $Q$ , the vector  $v_1 = (0, 1, -1)\theta_A = (0, 1, -1)$  has

order  $x+2$  in the  $\mathbb{Q}[x]$ -module  $M(A)$ , i.e.  $v_1$  is a row eigenvector of  $A$

corresponding to the eigenvalue  $-2$ . Using row 3 of  $Q$  the vector

$v_2 = (0, 0, 1)\theta_A = (0, 0, 1)$  has order  $x(x+2)$  in  $M(A)$ . So

$$X = \begin{pmatrix} \frac{v_1}{v_2} \\ xv_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 & -1 \\ 0 & 0 & 1 \\ -1 & 1 & 0 \end{pmatrix}$$

is invertible over  $\mathbb{Q}$  and satisfies

$$XAX^{-1} = C(x+2) \oplus C(x(x+2)) = \left( \begin{array}{c|cc} -2 & 0 & 0 \\ \hline 0 & 0 & 1 \\ \hline 0 & 0 & -2 \end{array} \right)$$

in rcf. So  $M(A)_x$  has basis  $(x+2)v_2 = (-1, 1, 0) + 2(0, 0, 1) = (-1, 1, 2)$ . Also  $M(A)_{x+2}$  has basis  $v_1, xv_2$ , i.e.  $(0, 1, -1), (-1, 1, 0)$ . Using these bases we construct

$$Y = \begin{pmatrix} (x+2)v_2 \\ v_1 \\ xv_2 \end{pmatrix} = \begin{pmatrix} -1 & 1 & 2 \\ 0 & 1 & -1 \\ -1 & 1 & 0 \end{pmatrix}$$

which is invertible over  $\mathbb{Q}$  as  $\det Y = 2$  and satisfies  $YAY^{-1} = \text{diag}(0, -2, -2)$  in pcf.

(iv) Reducing  $xI - A$  to its Smith normal form  $S(xI - A)$ :

$$\begin{aligned} xI - A &= \begin{pmatrix} x+1 & -1 & -1 \\ -1 & x & -1 \\ 1 & 1 & x+3 \end{pmatrix} \xrightarrow{r_1 \leftrightarrow r_3} \begin{pmatrix} 1 & 1 & x+3 \\ -1 & x & -1 \\ x+1 & -1 & -1 \end{pmatrix} \xrightarrow{\substack{c_2 - c_1 \\ c_3 - (x+3)c_1}} \begin{pmatrix} 1 & 0 & 0 \\ -1 & x+1 & x+2 \\ x+1 & -x-2 & -(x+2)^2 \end{pmatrix} \\ &\xrightarrow{\substack{r_2 + r_1 \\ r_3 - (x+1)r_1}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & x+1 & x+2 \\ 0 & -(x+2) & -(x+2)^2 \end{pmatrix} \xrightarrow{\substack{c_2 - c_3 \\ -c_2}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & x+2 \\ 0 & -(x+1)(x+2) & -(x+2)^2 \end{pmatrix} \\ &\xrightarrow{\substack{c_3 - (x+2)c_2 \\ r_3 + (x+1)(x+2)r_2}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & x(x+2)^2 \end{pmatrix} = S(xI - A). \end{aligned}$$

Calculating  $P(x)$  using the *eros* in the above reduction:

$$I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{r_1 \leftrightarrow r_3} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix} \xrightarrow{\substack{r_3 - (x+1)r_1 \\ r_3 + (x+1)(x+2)r_1}} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & (x+1)(x+2) & (x+1)^2 \end{pmatrix} = P(x).$$

Calculating  $Q(x)$  using the conjugates of the *ecos* in the above reduction:

$$\begin{aligned} I &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{\substack{r_1 + r_2 \\ r_1 + (x+3)r_3}} \begin{pmatrix} 1 & 1 & x+3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{\substack{r_3 + r_2 \\ -r_2}} \begin{pmatrix} 1 & 1 & x+3 \\ 0 & -1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \\ &\xrightarrow{\substack{r_2 + (x+2)r_3}} \begin{pmatrix} 1 & 1 & x+3 \\ 0 & x+1 & x+2 \\ 0 & 1 & 1 \end{pmatrix} = Q(x). \end{aligned}$$

Then  $|P(x)| = |Q(x)| = -1$  and so  $P(x)$  and  $Q(x)$  are invertible over  $\mathbb{Q}[x]$ . As

$P(x)(xI - A) = S(xI - A)Q(x)$  we deduce  $P(x)(xI - A)Q(x)^{-1} = \text{diag}(1, 1, x(x+2)^2)$

is in Smith normal form. Using row 3 of  $Q(x)$  we see that  $M(A)$  is cyclic with

generator  $v_1 = (0, 1, 1)$  of order  $x(x+2)^2$ . Therefore

$$X = \begin{pmatrix} v_1 \\ xv_1 \\ x^2v_1 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 1 \\ 0 & -1 & -2 \\ 1 & 2 & 5 \end{pmatrix}$$

is invertible over  $\mathbb{Q}$  and satisfies

$$XAX^{-1} = C(x(x+2)^2) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & -4 & -4 \end{pmatrix}$$

in rcf. A basis of  $M(A)_x$  is the single vector

$$(x+2)^2 v_1 = (x^2 + 4x + 4)v_1 = (1, 2, 5) + 4(0, -1, -2) + 4(0, 1, 1) = (1, 2, 1)$$

and so  $(1, 2, 1)$  is a row eigenvector of  $A$  corresponding to the eigenvalue 0. A basis of  $M(A)_{x+2}$  is  $xv_1, x^2v_1$ , i.e.  $(0, -1, -2), (1, 2, 5)$ . Using the vectors in these bases we construct

$$Y = \begin{pmatrix} \frac{(x+2)^2 v_1}{xv_1} \\ xv_1 \\ x^2v_1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 1 \\ 0 & -1 & -2 \\ 1 & 2 & 5 \end{pmatrix}$$

which is invertible over  $\mathbb{Q}$  as  $\det Y = -4$  and satisfies

$$YAY^{-1} = C(x) \oplus C((x+2)^2) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -4 & -4 \end{pmatrix}$$

The invariant factor sequences of the above four matrices  $A$  are

$$(x, x(x+2)), (x^2(x+2)), (x+2, x(x+2)), (x(x+2)^2).$$

No two of these sequences are equal. By the theory following (6.8) no two of the matrices  $A$  are similar.

(b) Let  $p(x)$  be an irreducible factor of  $\chi_A(x)$ . Then  $p(x)$  is an irreducible factor of  $\mu_A(x)$  by (6.11). So  $\mu_A(x) = p(x)q(x)$  for  $q(x) \in F[x]$  and evaluating this polynomial equality at  $A$  gives  $0 = \mu_A(A) = p(A)q(A)$ . Is  $\det p(A) \neq 0$  possible? If so then the matrix  $p(A)$  is invertible over  $F$  and hence  $q(A) = p(A)^{-1} \times 0 = 0$ . This is contrary to  $\mu_A(x)$  being the minimum polynomial (6.9) of  $A$  as  $q(x)$  is monic and  $\deg q(x) < \deg \mu_A(x)$ . So  $\det p(A) = 0$ .

Conversely suppose  $\det p(A) = 0$  where  $p(x)$  is irreducible over  $F$ . We suppose  $p(x)$  is not a divisor of  $\chi_A(x)$  and (as above) aim for a contradiction. As  $\gcd\{p(x), \chi_A(x)\} = 1$ , by (4.6) there are  $a(x), b(x) \in F[x]$  with  $a(x)p(x) + b(x)\chi_A(x) = 1$ . Evaluation at  $A$  gives  $a(A)p(A) + b(A)\chi_A(A) = I$ , i.e.  $a(A)p(A) = I$  as  $\chi_A(A) = 0$  by (6.11), showing  $p(A)$  to be an invertible matrix over  $F$  as  $p(A)^{-1} = a(A)$ . Therefore  $\det p(A) \neq 0$  contrary to hypothesis. So  $p(x)$  is a divisor of  $\chi_A(x)$ .

(c)

$$A^2 = \begin{pmatrix} 2 & 1 & -1 \\ 1 & 0 & -1 \\ 2 & 2 & -1 \end{pmatrix} \begin{pmatrix} 2 & 1 & -1 \\ 1 & 0 & -1 \\ 2 & 2 & -1 \end{pmatrix} = \begin{pmatrix} 3 & 0 & -2 \\ 0 & -1 & 0 \\ 4 & 0 & -3 \end{pmatrix} \text{ and so } A^2 + I = \begin{pmatrix} 4 & 0 & -2 \\ 0 & 0 & 0 \\ 4 & 0 & -2 \end{pmatrix}.$$

So  $\det(A^2 + I) = 0$  and  $(x^2 + 1) \mid \chi_A(x)$  by (b) above. The coefficient of  $x^2$  in  $\chi_A(x)$  is  $-\text{trace } A = -1$  which is enough to find the factorisation  $\chi_A(x) = (x-1)(x^2 + 1)$ . By

inspection  $e_2 = (0, 1, 0) \in M(A)_{x^2+1}$  and  $xe_2 = (1, 0, -1)$  is not proportional to  $e_2$ . So  $e_2, xe_2$  is a basis of  $M(A)_{x^2+1}$ . By the Cayley-Hamilton theorem  $(A^2 + I)(A - I) = 0$  and so the non-zero rows of  $A^2 + I$  are eigenvectors of  $A$  associated with the eigenvalue 1. So  $v_0 = (2, 0, -1)$  is such an eigenvector and is a basis of  $M(A)_{x-1}$ .

Therefore

$$Y = \begin{pmatrix} v_0 \\ e_2 \\ xe_2 \end{pmatrix} = \begin{pmatrix} 2 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & -1 \end{pmatrix}$$

is invertible over  $\mathbb{Q}$  and

$$YAY^{-1} = C(x-1) \oplus C(x^2+1) = \left( \begin{array}{c|cc} 1 & 0 & 0 \\ \hline 0 & 0 & 1 \\ 0 & -1 & 0 \end{array} \right) \text{ is in pcf.}$$

### Solution 2

(a) Write  $N_j = \{v \in M(A) : p_j(x)^{t_{sj}} v = 0\}$ . By (6.10) and (6.11) we see

$1 \leq t_{sj} \leq n_j$ . Therefore  $p_j(x)^{t_{sj}} v = 0 \Rightarrow p_j(x)^{n_j} v = 0$  showing  $N_j \subseteq M(A)_{p_j(x)}$ .

Consider now  $v \in M(A)_{p_j(x)}$ . Then  $p_j(x)^{n_j} v = 0$  and  $\mu_A(x)v = v\mu_A(A) = v \times 0 = 0$ .

The order (5.11) of  $v$  in  $M(A)$  is therefore a divisor of both  $p_j(x)^{n_j}$  and  $\mu_A(x)$ , and so is a divisor of  $\gcd\{p_j(x)^{n_j}, \mu_A(x)\} = p_j(x)^{t_{sj}}$ . So  $p_j(x)^{t_{sj}} v = 0$  and hence  $v \in N_j$  showing  $M(A)_{p_j(x)} \subseteq N_j$ . The conclusion is:

$$M(A)_{p_j(x)} = N_j = \{v \in M(A) : p_j(x)^{t_{sj}} v = 0\}.$$

(b) The  $t \times t$  matrix  $A$  has  $t$  different eigenvalues  $\lambda_1, \lambda_2, \dots, \lambda_t \in F$ . So

$(x - \lambda_j) \mid \chi_A(x)$  for  $1 \leq j \leq t$  and hence  $\chi_A(x) = (x - \lambda_1)(x - \lambda_2) \cdots (x - \lambda_t)$  is the factorisation of  $\chi_A(x)$  into irreducible polynomials  $p_j(x) = x - \lambda_j$  over  $F$ . The primary component  $M(A)_{p_j(x)}$  (the row eigenspace of  $A$  associated with the

eigenvalue  $\lambda_j$ ) has dimension 1 by (6.13) and so there is  $v_j \in F^t$  with

$M(A)_{p_j(x)} = \langle v_j \rangle$  for  $1 \leq j \leq t$ . By (6.12) the eigenvectors  $v_1, v_2, \dots, v_t$  form a basis of  $F^t$  and so, using (6.13) with  $k = t$  and  $1 \times 1$  matrices  $A_j = (\lambda_j)$ , are the rows of an invertible  $t \times t$  matrix  $X$  over  $F$  with  $XAX^{-1} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_t)$ , which is in pcf.

As  $\mu_A(x)$  and  $\chi_A(x)$  have the same irreducible factors by (6.11) we see

$\mu_A(x) = \chi_A(x)$ . So  $M(A)$  is cyclic (in fact  $v_1 + v_2 + \dots + v_t$  generates  $M(A)$ ).

Suppose  $A \sim D$  where  $D = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_t)$ . We suppose (as we may) that the first  $n_1$  diagonal entries in  $D$  are  $\lambda_1$ , the next  $n_2$  diagonal entries in  $D$  are  $\lambda_2, \dots$ , the last  $n_k$  diagonal entries in  $D$  are  $\lambda_k$  where  $\lambda_1, \lambda_2, \dots, \lambda_k$  are distinct and

$n_1 + n_2 + \dots + n_k = t$ . As similar matrices have equal characteristic polynomials

$\chi_A(x) = \chi_D(x) = \det(xI - D) = (x - \lambda_1)^{n_1} (x - \lambda_2)^{n_2} \cdots (x - \lambda_k)^{n_k}$  showing that



$\lambda_1, \lambda_2, \dots, \lambda_k$  are the eigenvalues of  $A$ . Also  $\mu_A(x) = \mu_D(x)$  by (6.12). The irreducible factors over  $F$  of  $\chi_A(x)$  are  $x - \lambda_1, x - \lambda_2, \dots, x - \lambda_k$  and these  $k$  polynomials are divisors of  $\mu_A(x)$  by (6.11). So their product  $f(x) = (x - \lambda_1)(x - \lambda_2) \cdots (x - \lambda_k)$  is also a divisor of  $\mu_A(x)$ , that is,  $f(x) \mid \mu_A(x)$ . The matrix  $f(D) = (D - \lambda_1 I)(D - \lambda_2 I) \cdots (D - \lambda_k I)$  is itself diagonal, being a product of diagonal matrices. For  $1 \leq i \leq t$  the  $(i, i)$ -entry in  $D$  is  $\lambda_j$  for some  $j$  with  $1 \leq j \leq k$ . Then the  $(i, i)$ -entry in  $f(D)$  is  $(\lambda_j - \lambda_1)(\lambda_j - \lambda_2) \cdots (\lambda_j - \lambda_k)$  which is zero for  $1 \leq i \leq t$  since one of the factors in this product is zero. So  $f(D) = 0$  and therefore  $f(x) \in K_D = \langle \mu_D(x) \rangle$  the annihilator ideal (6.10) of  $D$ . This means  $\mu_A(x) \mid f(x)$  as  $\mu_A(x) = \mu_D(x)$ . So  $\mu_A(x) = f(x)$  as  $\mu_A(x)$  and  $f(x)$  are monic polynomials, each being a divisor of the other. Therefore

$$\mu_A(x) = (x - \lambda_1)(x - \lambda_2) \cdots (x - \lambda_k)$$

as we wanted to show.

Conversely suppose  $\mu_A(x) = (x - \lambda_1)(x - \lambda_2) \cdots (x - \lambda_k)$  where  $\lambda_1, \lambda_2, \dots, \lambda_k$  are  $k$  different elements of  $F$ . Write  $p_j(x) = x - \lambda_j$  for  $1 \leq j \leq k$ . Then

$$\chi_A(x) = (x - \lambda_1)^{n_1} (x - \lambda_2)^{n_2} \cdots (x - \lambda_k)^{n_k}.$$

By (a) above and (6.13)

$$M(A)_{x-\lambda_j} = \{v \in M(A) : (x - \lambda_j)v = 0\} = \{v \in F^t : Av = \lambda_j v\}$$

which is the (row) eigenspace associated with the eigenvalue  $\lambda_j$  of  $A$  and  $\dim M(A)_{x-\lambda_j} = n_j$ . The primary components of  $M(A)$  are the eigenspaces of  $A$ .

As in (6.13) let  $\mathcal{B} = \mathcal{B}_1 \cup \mathcal{B}_2 \cup \dots \cup \mathcal{B}_k$  be a basis of  $F^t$  where  $\mathcal{B}_j$  is a basis of  $M(A)_{p_j(x)}$  for  $1 \leq j \leq k$ . The vectors in  $\mathcal{B}$  are linearly independent eigenvectors

of  $A$ . Let  $X$  be the invertible  $t \times t$  matrix over  $F$  having the vectors in  $\mathcal{B}$  as its rows. Then  $XAX^{-1}$  is diagonal, there being  $n_1$  entries  $\lambda_1$ ,  $n_2$  entries  $\lambda_2$ , ...,  $n_k$  entries  $\lambda_k$  on the diagonal. So  $A$  is similar to a diagonal matrix over  $F$ .

Statement (i) is true. Suppose  $A$  is similar to a diagonal matrix over  $F$ . By (a) above each elementary divisor  $(x - \lambda_j)^{t_{ij}}$  of  $A$  is a divisor of

$\mu_A(x) = (x - \lambda_1)(x - \lambda_2) \cdots (x - \lambda_k)$  and so  $t_{ij} = 1$ . Conversely suppose the elementary divisors are  $x - \lambda_j$  and  $n_j$  in number for  $1 \leq j \leq k$ . Then each

$M(A)_{x-\lambda_j} = \{v \in F^t : (x - \lambda_j)v = 0\}$  is the  $n_j$ -dimensional eigenspace of  $A$

associated with the eigenvalue  $\lambda_j$ . So  $F^t$  has a basis consisting of eigenvectors of  $A$  and so  $A$  is similar to a diagonal matrix over  $F$ .

Statement (ii) is also true. Suppose  $A$  is similar to a diagonal matrix over  $F$ . Each invariant factor  $d_i(x)$  is the product of  $k_i$  elementary divisors at most one from each primary component. By (i) above  $d_i(x)$  is the product of  $k_i$  distinct factors  $x - \lambda_j$ . Conversely suppose each invariant factor of  $A$  splits into distinct factors of degree 1.

The minimum polynomial  $\mu_A(x)$  is an invariant factor of  $A$  by (6.10). So  $\mu_A(x)$  splits into distinct factors of degree 1. By the earlier part of the question,  $A$  is similar to a diagonal matrix over  $F$ .

(c)  $(0,0,3)\alpha = (0,0,1,1,1)$ ,  $(0,1,2)\alpha = (0,1,0,1,1)$ ,  $(0,2,1)\alpha = (0,1,1,0,1)$ ,  
 $(0,3,0)\alpha = (0,1,1,1,0)$ ,  $(1,0,2)\alpha = (1,0,0,1,1)$ ,  $(1,1,1)\alpha = (1,0,1,0,1)$ ,  
 $(1,2,0)\alpha = (1,0,1,1,0)$ ,  $(2,0,1)\alpha = (1,1,0,0,1)$ ,  $(2,1,0)\alpha = (1,1,0,1,0)$ ,  
 $(3,0,0)\alpha = (1,1,1,0,0)$ .

As  $(l_1, l_2, \dots, l_s)\alpha = (m_1, m_2, \dots, m_{s+t-1})$  is a sequence of  $s+t-1$  zeros and ones with  $s-1$  zeros, we see  $m_1 + m_2 + \dots + m_{s+t-1} = t$  showing  $(l_1, l_2, \dots, l_s)\alpha \in M(s, t)$ . So  $\alpha: L(s, t) \rightarrow M(s, t)$ , that is,  $\alpha$  maps  $L(s, t)$  to  $M(s, t)$ .

Consider  $(m_1, m_2, \dots, m_{s+t-1}) \in M(s, t)$ . This sequence of  $s+t-1$  zeros and ones satisfies  $m_1 + m_2 + \dots + m_{s+t-1} = t$  and so there are  $s-1$  zeros in the sequence. Let the zeros occur in the  $j_i$ -position for  $1 \leq i < s$  where  $0 < j_1 < j_2 < \dots < j_{s-1}$ . Bearing in mind that we are trying to construct  $\alpha^{-1}: M(s, t) \rightarrow L(s, t)$ , we write  $l_1 = j_1 - 1$ ,  $l_2 = j_2 - j_1 - 1, \dots, l_{s-1} = j_{s-1} - j_{s-2} - 1$ . Then  $l_1, l_2, \dots, l_{s-1}$  are non-negative integers satisfying  $l_1 + l_2 + \dots + l_i = j_i - i$  for  $1 \leq i < s$ . So

$0 \leq l_1 + l_2 + \dots + l_{s-1} = j_{s-1} - (s-1) \leq t$ . Finally write  $l_s = s+t-1 - j_{s-1}$ . Then  $l_s$  is non-negative and  $l_1 + l_2 + \dots + l_s = t$ . Therefore  $(l_1, l_2, \dots, l_s) \in L(s, t)$ . Put  $(l_1, l_2, \dots, l_s) = (m_1, m_2, \dots, m_{s+t-1})\beta$ . Then  $\beta: M(s, t) \rightarrow L(s, t)$ . As the zeros in  $(m_1, m_2, \dots, m_{s+t-1})$  occur in positions  $l_1 + l_2 + \dots + l_i + i = j_i$  for  $1 \leq i < s$ , we see  $\beta\alpha$  is the identity mapping of  $M(s, t)$ . As the zeros in  $(l_1, l_2, \dots, l_s)\alpha$  occur in positions  $l_1 + l_2 + \dots + l_i + i = j_i$  for  $1 \leq i < s$ , we see  $\alpha\beta$  is the identity mapping of  $L(s, t)$ . The conclusion is:  $\beta = \alpha^{-1}$ . So  $\alpha: L(s, t) \rightarrow M(s, t)$  is a bijection and  $|L(s, t)| = |M(s, t)|$ .

The number of ways of choosing  $t$  things from  $s+t-1$  things is the binomial coefficient  $\binom{s+t-1}{t}$ . So  $|M(s, t)| = \binom{s+t-1}{t}$ , the  $t$  things being the positions in

$(m_1, m_2, \dots, m_{s+t-1}) \in M(s, t)$  with entry 1. Therefore  $|L(s, t)| = \binom{s+t-1}{t}$ .

(d) As  $f(A) = 0$  we see  $f(x) \in K_A$  the annihilator ideal (6.10) of  $A$ . But  $K_A = \langle \mu_A(x) \rangle$  and so  $\mu_A(x) \mid f(x)$ . As  $f(x)$  is a product of distinct factors of degree 1 over  $F$  the same is true of  $\mu_A(x)$ . By (b) above  $A$  is similar to a diagonal matrix over  $F$ .

Write  $l_j = \dim U_j$  where  $U_j = \{v \in F^t : vA = c_j v\}$ . Suppose  $(x - c_j) \mid \mu_A(x)$ . Then  $M(A)_{x-c_j} = U_j$  by (a) above with  $t_{s_j} = 1$ , i.e. the zero  $c_j$  of  $f(x)$  is an eigenvalue of  $A$  and the primary component  $M(A)_{x-c_j}$  is the corresponding (row) eigenspace of  $A$ . So  $l_j > 0$ . Conversely suppose  $l_j > 0$ . Then  $c_j$  is an eigenvalue of  $A$  and so  $(x - c_j) \mid \chi_A(x)$ . By (6.11) we see  $(x - c_j) \mid \mu_A(x)$ . Therefore the polynomial  $x - c_j$  appears  $l_j$  times in the list of elementary divisors of  $A$  for  $1 \leq j \leq s$ . By (6.12) we see  $F^t = U_1 \oplus U_2 \oplus \dots \oplus U_s$  where the non-zero terms in this decomposition are the

primary components of  $M(A)$ . Comparing dimensions gives  $t = l_1 + l_2 + \dots + l_s$ . From the theory following (6.13) the similarity classes of  $t \times t$  matrices  $A$  over  $F$  with  $f(A) = 0$  correspond to the elements  $(l_1, l_2, \dots, l_s) \in L(s, t)$  of (a) above. So there are

$$|L(s, t)| = \binom{s+t-1}{t} \text{ similarity classes of such matrices.}$$

(e) (i) Take  $s = 2$  and  $f(x) = (x-1)(x+1)$  in (d) above. Then  $A^2 = I \Leftrightarrow f(A) = 0$ .

So there are  $\binom{2+t-1}{t} = t+1$  similarity classes of  $t \times t$  matrices  $A$  over a given field

$F$  of characteristic  $\neq 2$  satisfying  $A^2 = I$ .

(ii) Take  $s = 3$  and  $f(x) = x(x-1)(x+1)$  in (d) above. Then  $A^3 = A \Leftrightarrow f(A) = 0$ .

So there are  $\binom{3+t-1}{t} = (t+2)(t+1)/2$  similarity classes of  $t \times t$  matrices  $A$  over a given field  $F$  of characteristic  $\neq 2$  satisfying  $A^3 = A$ .

(f) Suppose that  $x^2 + x + 1$  is irreducible over  $\mathbb{Z}_p$ . As  $x^2 + x + 1 = (x-1)^2$  over  $\mathbb{Z}_3$  we see  $p \neq 3$  and so  $p \not\equiv 0 \pmod{3}$  as  $p$  is prime. So  $p \equiv \pm 1 \pmod{3}$ . In the case  $p \equiv 1 \pmod{3}$  the multiplicative group  $\mathbb{Z}_p^*$  of non-zero elements of  $\mathbb{Z}_p$  is cyclic with generator  $c$  by (3.17) and so  $d = c^{(p-1)/3}$  is an element of  $\mathbb{Z}_p$  satisfying  $d^3 = 1$ ,  $d \neq 1$ . So  $0 = d^3 - 1 = (d-1)(d^2 + d + 1)$  from which we deduce  $d^2 + d + 1 = 0$ . Therefore  $x^2 + x + 1$  is not irreducible over  $\mathbb{Z}_p$  by (4.8)(i) as  $x^2 + x + 1$  has zero  $d$  in  $\mathbb{Z}_p$ . So  $p \equiv -1 \pmod{3}$  as the other possibilities have been ruled out.

Conversely assume  $p \equiv -1 \pmod{3}$ . We aim for a contradiction by supposing that  $x^2 + x + 1$  is reducible over  $\mathbb{Z}_p$ . By (4.8)(ii) there is a zero  $d$  of  $x^2 + x + 1$  in  $\mathbb{Z}_p$ . Then  $d \neq 1$  as  $1^2 + 1 + 1 \neq 0$  since  $p \not\equiv 0 \pmod{3}$ . But  $d^3 = 1$  as  $d^3 - 1 = (d-1)(d^2 + d + 1) = 0$  showing that  $H = \{1, d, d^2\}$  is a subgroup of  $\mathbb{Z}_p^*$ . The multiplicative group  $\mathbb{Z}_p^*$  of order  $p-1$  has a subgroup  $H$  of order 3. By Lagrange's theorem  $3 \mid (p-1)$ , that is  $p \equiv 1 \pmod{3}$  contrary to our hypothesis. Therefore  $x^2 + x + 1$  is irreducible over  $\mathbb{Z}_p$ .

Suppose  $p \equiv 1 \pmod{3}$ . Then  $x^2 + x + 1$  has zero  $d$  in  $\mathbb{Z}_p$ . Another zero is  $1/d$  and  $d \neq 1/d$ . Also the zeros  $1, d, 1/d$  of  $x^3 - 1$  are distinct. By (d) above, with  $s = 3$ , the number of similarity classes of  $t \times t$  matrices  $A$  over  $\mathbb{Z}_p$  with  $A^3 = I$  is

$$\binom{3+t-1}{t} = (t+2)(t+1)/2.$$

Suppose  $p \equiv -1 \pmod{3}$ . The possible elementary divisors of  $t \times t$  matrices  $A$  over  $\mathbb{Z}_p$  satisfying  $A^3 = I$  are  $x-1$  and  $x^2 + x + 1$  as  $\mu_A(x) \mid (x^3 - 1)$ . So

$\chi_A(x) = (x-1)^{n_1} (x^2 + x + 1)^{n_2}$  where  $n_1$  and  $n_2$  are non-negative integers with

$n_1 + 2n_2 = t$ . There are  $\lfloor t/2 \rfloor + 1$  such pairs  $(n_1, n_2)$  as  $0 \leq n_2 \leq \lfloor t/2 \rfloor$  and

$n_1 = t - 2n_2 \geq 0$  for  $n_2$  in the above interval. So there are  $\lfloor t/2 \rfloor + 1$  similarity classes

of  $t \times t$  matrices  $A$  over  $\mathbb{Z}_p$  with  $A^3 = I$ , namely those with  $n_1$  elementary divisors  $x-1$  and  $n_2$  elementary divisors  $x^2+x+1$  for  $n_2 = 0, 1, 2, \dots, \lfloor t/2 \rfloor$ .

Notice that the case  $p=3$  is covered in Exercises 6.1, Question 7(c). There are  $N(t)$  similarity classes of  $t \times t$  matrices  $A$  over  $\mathbb{Z}_3$  with  $A^3 = I$ .

(g) By the multiplicative version of the  $|G|$ -lemma, all  $q-1$  elements  $\lambda \in \mathbb{F}_q^*$  satisfy  $\lambda^{q-1} = 1$  the 1-element of  $\mathbb{F}_q$ . So all  $q$  elements  $\lambda$  of  $\mathbb{F}_q$  satisfy  $\lambda^q = \lambda$ . Hence the diagonal matrix  $D = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_t)$  over  $\mathbb{F}_q$  satisfies  $D^q = D$  since  $D^q = \text{diag}(\lambda_1^q, \lambda_2^q, \dots, \lambda_t^q) = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_t) = D$ .

Suppose  $A$  to be similar to  $D$ . There is an invertible  $t \times t$  matrix  $X$  over  $\mathbb{F}_q$  with  $XAX^{-1} = D$ . Taking  $f(x) = x^q - x$  in (5.12) we obtain  $Xf(A)X^{-1} = f(D) = D^q - D = 0$  the  $t \times t$  zero matrix over  $\mathbb{F}_q$ . Therefore  $f(A) = X^{-1}0X = 0$ , that is,  $A^q - A = 0$  and so  $A^q = A$ .

Conversely suppose the  $t \times t$  matrix  $A$  over  $\mathbb{F}_q$  satisfies  $A^q = A$ . Then

$\mu_A(x) \mid (x^q - x)$ . As the  $q$  elements  $\lambda$  of  $\mathbb{F}_q$  are zeros of the polynomial  $x^q - x$  of degree  $q$  over  $\mathbb{F}_q$  we see  $x^q - x = \prod_{\lambda \in \mathbb{F}_q} (x - \lambda)$  and so we may apply (d) above: taking

$s = q$  there are  $\binom{q+t-1}{t}$  similarity classes of  $t \times t$  matrices  $A$  over  $\mathbb{F}_q$  with  $A^q = A$ .

(i) Representatives of the  $\binom{2+4-1}{4} = 5$  similarity classes of  $4 \times 4$  matrices  $A$  over  $\mathbb{Z}_2$  with  $A^2 = A$  are:

$$\text{diag}(0, 0, 0, 0), \text{diag}(1, 0, 0, 0), \text{diag}(1, 1, 0, 0), \text{diag}(1, 1, 1, 0), \text{diag}(1, 1, 1, 1).$$

(ii) Representatives of the  $\binom{4+2-1}{2} = 10$  similarity classes of  $2 \times 2$  matrices  $A$  over  $\mathbb{F}_4 = \{0, 1, c, c^2 : 1+1=0, c^2+c+1=0\}$  with  $A^4 = A$  are:

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} c & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} c^2 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & c \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & c^2 \end{pmatrix}, \begin{pmatrix} c & 0 \\ 0 & c \end{pmatrix}, \begin{pmatrix} c & 0 \\ 0 & c^2 \end{pmatrix}, \begin{pmatrix} c^2 & 0 \\ 0 & c^2 \end{pmatrix}.$$

### Solution 3

(a) Write  $N = M(A)_{p(x)}$ . Consider  $u, v \in N$ . Then  $p(x)^n u = 0$  and  $p(x)^n v = 0$ .

Using the distributive law (which holds in the  $F[x]$ -module  $M(A)$ ) we have

$$p(x)^n(u+v) = p(x)^n u + p(x)^n v = 0 + 0 = 0 \text{ showing } u+v \in N. \text{ Also } -v \in N \text{ as}$$

$$p(x)^n(-v) = -p(x)^n v = -0 = 0. \text{ The zero row vector } 0 \text{ in } F^t \text{ belongs to } N \text{ as}$$

$$p(x)^n 0 = 0 \text{ and so } N \text{ is a subgroup of the additive group } (F^t, +). \text{ For } f(x) \in F[x]$$

$$p(x)^n(f(x)v) = (p(x)^n f(x))v = (f(x)p(x)^n)v = f(x)(p(x)^n v) = f(x)0 = 0$$

showing that  $f(x)v \in N$ . By (2.26) we conclude that  $N$  is a submodule of  $M(A)$ .

Suppose  $N = \{0\}$  and look for a contradiction. By the Cayley-Hamilton theorem

$$p(x)^n(q(x)v) = (p(x)^n q(x))v = (\chi_A(x)v = v\chi_A(A) = v0 = 0 \text{ for all } v \in M(A). \text{ So}$$

$$q(x)v \in N, \text{ i.e. } q(x)v = 0, \text{ for all } v \in M(A). \text{ Therefore } e_i q(A) = q(x)e_i = 0 \text{ for}$$

$1 \leq i \leq t$  showing that all the rows of the  $t \times t$  matrix  $q(A)$  are zero. So  $q(A) = 0$  which means  $\mu_A(x) \mid q(x)$  by (6.10). But  $p(x) \mid \mu_A(x)$  as  $\chi_A(x)$  and  $\mu_A(x)$  have the same irreducible factors by (6.11). So  $p(x) \mid q(x)$  which is contrary to  $\gcd\{p(x), q(x)\} = 1$ . We conclude:  $N \neq \{0\}$ , i.e. the primary components of  $M(A)$  are non-trivial.

(b) Suppose  $M = M(C(p(x)^n))$  has submodules  $N_1$  and  $N_2$  such that  $M = N_1 \oplus N_2$ . Then  $M$  is cyclic with generator  $e_1$ , the first element of the standard basis of  $F^t$  where  $t = \deg p(x)^n$ , by (5.26). The monic divisors of  $p(x)^n$  are the  $n+1$  polynomials  $p(x)^m$  for  $0 \leq m \leq n$ . By (5.28) the submodules  $N_1$  and  $N_2$  are themselves cyclic with generators  $p(x)^{m_1} e_1$  and  $p(x)^{m_2} e_1$  respectively, where we may assume  $0 \leq m_1 \leq m_2 \leq n$ . So  $N_2 \subseteq N_1$  as the generator  $p(x)^{m_2} e_1$  of  $N_2$  is a polynomial multiple (namely  $p(x)^{m_2 - m_1}$ ) of the generator  $p(x)^{m_1} e_1$  of  $N_1$ . As  $M = N_1 \oplus N_2$  we know  $N_1$  and  $N_2$  are independent, i.e.  $(N_1, +)$  and  $(N_2, +)$  are independent subgroups (2.14) of  $(M, +)$ . So  $N_1 \cap N_2 = \{0\}$ . But  $N_1 \cap N_2 = N_2$  as  $N_2 \subseteq N_1$  and so  $N_2 = \{0\}$ . As  $M$  is non-zero we see that  $M$  is indecomposable.

Let  $N$  be an indecomposable submodule of  $M(A)$ . As  $N \neq \{0\}$  by

Exercises 5.1, Question 5 there is an  $s \times s$  matrix  $B$  over  $F$  with  $N \cong M(B)$ . So  $M(B)$  is indecomposable.

Should  $\chi_B(x)$  be divisible by two or more irreducible polynomials over  $F$  the primary decomposition (6.12) of  $M(B)$  would contradict its indecomposability by (a) above. So  $\chi_B(x) = p(x)^n$  where  $p(x)$  is a monic irreducible polynomial over  $F$  and  $n$  is a positive integer. Also  $M(B)$  has only one invariant factor, since otherwise the invariant factor decomposition (6.5) of  $M(B)$  would contradict its indecomposability. So  $M(B)$  is cyclic. By (5.27) we conclude  $N \cong M(C(p(x)^n))$ .

(c) As  $\dim N_j \geq 1$  for  $1 \leq j \leq r$ , on comparing dimensions of subspaces we obtain

$t = \dim F^t = \dim M(A) = \dim(N_1 \oplus N_2 \oplus \dots \oplus N_r) = \dim N_1 + \dim N_2 + \dots + \dim N_r \geq r$  showing  $r \leq t$ .

Consider a decomposition  $M(A) = N_1 \oplus N_2 \oplus \dots \oplus N_r$  into a direct sum of  $r$  non-zero submodules  $N_j$  with  $r$  as large as possible (as  $r$  is bounded above by  $t$  there is such an  $r$ ). Suppose  $N_1$  is decomposable and so  $N_1 = N'_1 \oplus N''_1$  where  $N'_1$  and  $N''_1$  are non-zero. Then

$$M(A) = N'_1 \oplus N''_1 \oplus N_2 \oplus \dots \oplus N_r$$

is a decomposition with  $r+1$  summands (terms) contrary to the choice of  $r$ . So  $N_1$  is indecomposable and in the same way we see that each  $N_j$  is indecomposable for

$1 \leq j \leq r$ . Therefore  $N_j \cong M(C(p_j(x)^{n_j}))$  where  $p_j(x)$  is monic and irreducible over  $F$  and  $n_j$  is a positive integer by (b) above. From (5.20) we conclude

$$A \sim C(p_1(x)^{n_1}) \oplus C(p_2(x)^{n_2}) \oplus \dots \oplus C(p_r(x)^{n_r})$$

showing that the polynomials  $p_j(x)^{n_j}$  for  $1 \leq j \leq r$  are the elementary divisors of  $A$ . So  $r$  is the number of elementary divisors of  $A$ .

(d) Write  $m_j(x) = \chi_A(x) / p_j(x)^{n_j} = \prod_{i=1, i \neq j}^k p_i(x)^{n_i}$  which is a monic polynomial of degree  $t - n_j \deg p_j(x)$  over  $F$ . Then  $\text{lcm}\{m_1(x), m_2(x), \dots, m_k(x)\} = \chi_A(x)$  and (more to the point)  $\gcd\{m_1(x), m_2(x), \dots, m_k(x)\} = 1$  as  $p_j(x)$  is not a divisor of  $m_j(x)$  for  $1 \leq j \leq k$ , i.e. the polynomials  $m_1(x), m_2(x), \dots, m_k(x)$  have no common irreducible divisor. By (4.6) there are  $a_j(x) \in F[x]$  for  $1 \leq j \leq k$  such that  $a_1(x)m_1(x) + a_2(x)m_2(x) + \dots + a_k(x)m_k(x) = 1$ . We are now ready to prove (6.12).

Consider  $v \in M(A)$ . Then  $v = 1v = (\sum_{j=1}^k a_j(x)m_j(x))v = \sum_{j=1}^k a_j(x)m_j(x)v = \sum_{j=1}^k v_j$

where  $v_j = a_j(x)m_j(x)v$  for  $1 \leq j \leq k$ . Now  $v_j \in M(A)_{p_j(x)}$  as

$p_j(x)^{n_j} v_j = p_j(x)^{n_j} a_j(x)m_j(x)v = a_j(x)\chi_A(x)v = a_j(x)(v\chi_A(A)) = a_j(x)0 = 0$  for  $1 \leq j \leq k$ . Therefore each vector  $v$  in  $M(A)$  is expressible as a sum of vectors  $v_j$  each belonging to the primary component  $M(A)_{p_j(x)}$ , i.e.

$$M(A) = M(A)_{p_1(x)} + M(A)_{p_2(x)} + \dots + M(A)_{p_k(x)}.$$

Finally we show that the above sum of primary components is direct using (2.15).

Suppose  $v_1 + v_2 + \dots + v_k = 0$  where  $v_j \in M(A)_{p_j(x)}$  for  $1 \leq j \leq k$ . We concentrate on

one particular term  $v_j$ . For  $i \neq j$ ,  $1 \leq i \leq k$  we have  $m_j(x)v_i = 0$  as  $p_i(x)^{n_i} \mid m_j(x)$

and  $p_i(x)^{n_i} v_i = 0$ . Inserting  $k-1$  zero terms  $m_j(x)v_i = 0$  produces

$$m_j(x)v_j = m_j(x)v_j + \sum_{i=1, i \neq j}^k m_j(x)v_i = \sum_{i=1}^k m_j(x)v_i = m_j(x)(\sum_{i=1}^k v_i) = m_j(x)0 = 0.$$

The polynomial

$$1 - a_j(x)m_j(x) = \sum_{i=1, i \neq j}^k a_i(x)m_i(x)$$

is divisible by  $p_j(x)^{n_j}$  since  $p_j(x)^{n_j} \mid m_i(x)$  for  $i \neq j$ ,  $1 \leq i \leq k$ . So

$$a'_j(x) = (1 - a_j(x)m_j(x)) / p_j(x)^{n_j}$$

is a polynomial over  $F$ . Then

$$\begin{aligned} v_j = 1v_j &= (a_j(x)m_j(x) + (1 - a_j(x)m_j(x)))v_j = a_j(x)m_j(x)v_j + (1 - a_j(x)m_j(x))v_j = \\ &= a_j(x)m_j(x)v_j + a'_j(x)p_j(x)^{n_j}v_j = a_j(x)0 + a'_j(x)0 = 0 \quad \text{for } 1 \leq j \leq k. \end{aligned}$$

So  $v_1 + v_2 + \dots + v_k = 0$  implies  $v_1 = v_2 = \dots = v_k = 0$ . Therefore the primary components of  $M(A)$  are independent (2.14). By (2.15)

$$M(A) = M(A)_{p_1(x)} \oplus M(A)_{p_2(x)} \oplus \dots \oplus M(A)_{p_k(x)},$$

i.e.  $M(A)$  is the internal direct sum of its primary components.

#### Solution 4

(a) Consider  $X, Y \in \text{im } \mathcal{E}_A$ . As  $XY = YX$  by the binomial theorem

$$(X + Y)\varphi = (X + Y)^p = \sum_{i=0}^p \binom{p}{i} X^{p-i} Y^i = X^p + Y^p = (X)\varphi + (Y)\varphi$$

since  $p$  is a divisor of the binomial coefficient  $\binom{p}{i}$  for  $0 < i < p$  and  $pZ = 0$  for all

$Z \in \text{im } \mathcal{E}_A$ . So  $\varphi$  is additive. Also  $(XY)\varphi = (XY)^p = X^p Y^p = (X)\varphi(Y)\varphi$  and  $(I)\varphi = I^p = I$ . Therefore  $\varphi$  is a **ring endomorphism** of  $\text{im } \mathcal{E}_A$ , i.e.  $\varphi$  is a ring homomorphism between  $\text{im } \mathcal{E}_A$  and itself. As  $a^p = a$  for all  $a \in \mathbb{Z}_p$  we obtain

$$(aX)\varphi = ((aI)X)\varphi = (aI)\varphi(X)\varphi = (aI)(X)\varphi = a((X)\varphi)$$

since  $(aI)\varphi = (aI)^p = a^p I = aI$ . So  $\varphi$  is  $\mathbb{Z}_p$ -linear.

Write  $\mu_A(x) = a_0 + a_1x + \dots + a_{r-1}x^{r-1} + x^r$  for the minimum polynomial of  $A$ .

Applying  $\varphi$  to the matrix equation  $\mu_A(A) = 0$ , i.e.  $a_0I + a_1A + \dots + a_{r-1}A^{r-1} + A^r = 0$ , gives

$$(a_0I + a_1A + \dots + a_{r-1}A^{r-1} + A^r)\varphi = a_0I + a_1A^p + \dots + a_{r-1}A^{p(r-1)} + A^{pr} = (0)\varphi = 0$$

showing  $\mu_A(A^p) = 0$ . We have shown that  $\mu_A(x)$  belongs to the annihilator ideal

$$K_{A^p} \quad (6.10) \text{ of } A^p. \text{ As } K_{A^p} = \langle \mu_{A^p}(x) \rangle \text{ we conclude } \mu_{A^p}(x) \mid \mu_A(x).$$

Suppose now that  $\mu_A(x)$  is irreducible over  $\mathbb{Z}_p$ . By (6.10) the  $s$  invariant factors of  $A$  are non-constant monic divisors of  $\mu_A(x)$ . So the invariant factors of  $A$  are all equal  $\mu_A(x)$ . Incidentally we see  $\chi_A(x) = (\mu_A(x))^s$  and  $t = rs$  on comparing degrees.

But  $\mu_{A^p}(x) \mid \mu_A(x)$  gives  $\mu_{A^p}(x) = \mu_A(x)$  and so  $A^p$  also has  $s$  invariant factors

$\mu_A(x)$ . Therefore  $A^p \sim A$  as  $A^p$  and  $A$  have the same sequence of invariant factors.

As  $\ker \mathcal{E}_A = \langle \mu_A(x) \rangle$  we obtain  $\tilde{\mathcal{E}}_A : \mathbb{Z}_p[x] / \langle \mu_A(x) \rangle \cong \text{im } \mathcal{E}_A$  by the first isomorphism theorem for rings Exercises 2.3, Question 3(b). In this case  $\text{im } \mathcal{E}_A$  is a finite field of order  $p^r$ . Write  $K = \langle \mu_A(x) \rangle$ . Then  $(K + f(x))\theta = K + f(x)^p$  for all

$f(x) \in \mathbb{Z}_p[x]$ . So  $(K + f(x))\theta\tilde{\mathcal{E}}_A = f(A)^p$  whereas

$$(K + f(x))\tilde{\mathcal{E}}_A\varphi = (f(A))\varphi = f(A)^p.$$

Therefore  $\theta\tilde{\mathcal{E}}_A = \tilde{\mathcal{E}}_A\varphi$ . Also  $\varphi = (\tilde{\mathcal{E}}_A)^{-1}\theta\tilde{\mathcal{E}}_A$  is an automorphism of  $\text{im } \mathcal{E}_A$ .

In the case of  $A = C(x^2 + x + 1) \oplus C(x^2 + x + 1)$  over  $\mathbb{Z}_2$  we see  $\mu_A(x) = x^2 + x + 1$  which is irreducible over  $\mathbb{Z}_2$ . Therefore  $A^2 \sim A$  by the preceding theory.

In the case of  $B = C((x^2 + x + 1)^2) = C(x^4 + x^2 + 1)$  we see  $B^4 + B^2 + I = 0$  by (5.26). So  $(B^2)^2 + B^2 + I = 0$  showing  $x^2 + x + 1 \in K_{B^2}$ , i.e.  $x^2 + x + 1$  belongs to

the annihilator ideal of  $B^2$ . Therefore  $\mu_{B^2}(x) \mid (x^2 + x + 1)$  and so

$\mu_{B^2}(x) = x^2 + x + 1$ . As  $B$  and  $B^2$  have different minimum polynomials we conclude  $B^2 \not\sim B$ .

(b) Suppose  $1 < m \leq p$ . Then  $\mu_A(A) = 0$  gives  $p_0(A)^m = 0$ . Multiplication by  $p_0(A)^{p-m}$  produces  $p_0(A)^p = 0$ . So

$$p_0(A^p) = p_0((A)\varphi) = (p_0(A))\varphi = (p_0(A))^p = 0$$

as  $\varphi$  is an **algebra homomorphism** ( $\varphi$  is a homomorphism of the ring  $\text{im } \mathcal{E}_A$  and a linear mapping of the vector space  $\text{im } \mathcal{E}_A$  over  $\mathbb{Z}_p$ ) by (a) above. Therefore

$\mu_{A^p}(x) \mid p_0(x)$  on applying (6.10) to  $A^p$ . As  $p_0(x)$  is irreducible we see

$\mu_{A^p}(x) = p_0(x)$ . As  $\mu_{A^p}(x) \neq \mu_A(x)$  we conclude  $A^p \not\sim A$ .

More generally let  $k$  be the positive integer with  $p^{k-1} < m \leq p^k$ . As above  $\mu_A(A) = 0$  gives  $p_0(A)^m = 0$ . Multiplication by  $p_0(A)^{p^{k-m}}$  gives  $p_0(A)^{p^k} = 0$ . Using  $\varphi$  as in (a) above gives

$$p_0(A^p)^{p^{k-1}} = p_0((A)\varphi)^{p^{k-1}} = (p_0(A)^{p^{k-1}})\varphi = (p_0(A)^{p^{k-1}})^p = p_0(A)^{p^k} = 0$$

which shows  $\mu_{A^p}(x) \mid (p_0(x))^{p^{k-1}}$ . As  $\deg p_0(x)^{p^{k-1}} = np^{k-1} < mn$  we see

$\mu_{A^p}(x) \neq \mu_A(x)$  as  $\deg \mu_{A^p}(x) < mn = \deg \mu_A(x)$ . Therefore  $A^p \nmid A$ .

$A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$  and so  $A^2 = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}$ . As  $(A^2)^2 + A^2 + I = A + A^2 + I = 0$  we see

$\mu_{A^2}(x)$  is a non-constant divisor of the irreducible polynomial  $x^2 + x + 1$  over  $\mathbb{Z}_2$  by

(6.10). Therefore  $\mu_{A^2}(x) = x^2 + x + 1$  and so  $A^2$  has the single invariant factor

$x^2 + x + 1$ . As  $A$  also has the single invariant factor  $x^2 + x + 1$  we deduce  $A^2 \sim A$ .

Alternatively  $A^2 = XAX^{-1}$  where  $X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ .

$$B = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \text{ and so } B^2 = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}. \text{ As } (B^2)^3 + B^2 + I = (B^3 + B + I)^2 = 0,$$

since  $B^3 + B + I = 0$  by (5.26), we see  $\mu_{B^2}(x)$  is a non-constant divisor of the

irreducible polynomial  $x^3 + x + 1$  over  $\mathbb{Z}_2$  by (6.10). Therefore  $\mu_{B^2}(x) = x^3 + x + 1$

and so  $B^2$  has the single invariant factor  $x^3 + x + 1$ . As  $B$  also has the single invariant factor  $x^3 + x + 1$  we deduce  $B^2 \sim B$ .

More generally by (5.26) with  $d_0(x) = f(x)$  we obtain  $f(A) = 0$ . Write

$\varphi = (\tilde{\mathcal{E}}_A)^{-1} \theta \tilde{\mathcal{E}}_A$ . Being a composition of isomorphisms  $\varphi$  is an automorphism of the field  $\text{im } \mathcal{E}_A$ . For  $X \in \text{im } \mathcal{E}_A$  there is  $y \in F[x]/\langle f(x) \rangle$  with  $(y)\tilde{\mathcal{E}}_A = X$ . As

$(y)\theta = y^p$  and  $\tilde{\mathcal{E}}_A$  respects multiplication we obtain

$$(X)\varphi = (y)\tilde{\mathcal{E}}_A(\tilde{\mathcal{E}}_A)^{-1} \theta \tilde{\mathcal{E}}_A = (y)\theta \tilde{\mathcal{E}}_A = (y^p)\tilde{\mathcal{E}}_A = ((y)\tilde{\mathcal{E}}_A)^p = X^p.$$

So  $(X)\varphi = X^p$  for all  $X \in \text{im } \mathcal{E}_A$  and in particular  $(aI)\varphi = (aI)^p = a^p I = aI$  for all  $a \in \mathbb{Z}_p$ . So

$$(aX)\varphi = (aIX)\varphi = (aI)\varphi(X)\varphi = aIX^p = aX^p.$$

Let  $f(x) = a_0 + a_1x + \dots + a_{t-1}x^{t-1} + x^t$  where  $a_i \in \mathbb{Z}_p$  for  $0 \leq i < t$ . Then

$$\begin{aligned} 0 &= (0)\varphi = (f(A))\varphi = (a_0I + a_1A + \dots + a_{t-1}A^{t-1} + A^t)\varphi = \\ &= (a_0I)\varphi + (a_1A)\varphi + \dots + (a_{t-1}A^{t-1})\varphi + (A^t)\varphi = \\ &= a_0I + a_1A^p + \dots + a_{t-1}A^{(t-1)p} + A^{tp} = f(A^p), \end{aligned}$$

i.e.  $f(A^p) = 0$ . By (6.10) the minimum polynomial  $\mu_{A^p}(x)$  of  $A^p$  is a monic and non-constant divisor of  $f(x)$ . As  $f(x)$  is monic and irreducible over  $\mathbb{Z}_p$  we see

$\mu_{A^p}(x) = f(x)$ . Both  $A$  and  $A^p$  have single invariant factor  $f(x)$  and so  $A \sim A^p$ .

By (5.26) we see  $(A^2 + A + I)^2 = 0$ , i.e.  $(A^2)^2 + A^2 + I = 0$ . By (6.10) the

minimum polynomial  $\mu_{A^2}(x)$  of  $A^2$  is a monic and non-constant divisor of the

irreducible polynomial  $x^2 + x + 1$  over  $\mathbb{Z}_2$ . Therefore  $\mu_{A^2}(x) = x^2 + x + 1$ . As



$\mu_A(x) = (x^2 + x + 1)^2$  we see  $A \not\sim A^2$  since similar matrices have equal minimum polynomials.

By (5.26) we see  $(B^2 + B + I)^3 = 0$  and so  $(B^2 + B + I)^4 = 0$ . The last equation can be expressed  $(B^2)^4 + (B^2)^2 + I = 0$  showing  $\mu_{A^2}(x) \mid (x^4 + x^2 + 1)$ . Certainly  $\mu_{A^2}(x) \neq \mu_A(x) = (x^2 + x + 1)^3$  and so  $A \not\sim A^2$ .

(c) Factorise  $\chi_A(x) = p_1(x)^{n_1} p_2(x)^{n_2} \cdots p_k(x)^{n_k}$  into powers of distinct monic irreducible polynomials  $p_j(x)$  over  $\mathbb{Z}_p$  for  $1 \leq j \leq k$ . As in the proof of (6.13) there is an invertible  $t \times t$  matrix  $X$  with  $XAX^{-1} = A_1 \oplus A_2 \oplus \cdots \oplus A_k$  where

$$\chi_{A_j}(x) = p_j(x)^{n_j} \text{ for } 1 \leq j \leq k. \text{ Then } XA^p X^{-1} = (XAX^{-1})^p = A_1^p \oplus A_2^p \oplus \cdots \oplus A_k^p$$

using properties of the direct sum (5.17) of square matrices. As  $\mu_{A_j^p}(x) \mid \mu_{A_j}(x)$  by

(a) above, we see  $\mu_{A_j^p}(x)$  is a power of  $p_j(x)$  and so  $M(A^p)_{p_j(x)} \cong M(A_j^p)$  for

$1 \leq j \leq k$ . Therefore  $\chi_{A_j^p}(x) = p_j(x)^{n_j}$  as  $A_j^p$  is an  $n_j \times n_j$  matrix for  $1 \leq j \leq k$  and

$$\text{so } \chi_{A^p}(x) = p_1(x)^{n_1} p_2(x)^{n_2} \cdots p_k(x)^{n_k} = \chi_A(x).$$

Suppose  $A \sim A^p$ . Then  $M(A) \cong M(A^p)$  by (5.13). Taking  $B = A^p$  in the theory following (6.12) we obtain  $M(A)_{p_j(x)} \cong M(A^p)_{p_j(x)}$  and so  $M(A_j) \cong M(A_j^p)$ , i.e.

$A_j \sim A_j^p$  for  $1 \leq j \leq k$ . By (b) above  $\mu_{A_j}(x) = p_j(x)$  as otherwise  $A_j \not\sim A_j^p$  for  $1 \leq j \leq k$ . Therefore  $\mu_A(x) = p_1(x)p_2(x)\cdots p_k(x)$  is a product of distinct monic irreducible polynomials over  $\mathbb{Z}_p$  by Exercise 6.1, Question 6(e).

Conversely suppose  $\mu_A(x) = p_1(x)p_2(x)\cdots p_k(x)$ . Then  $\mu_{A_j}(x) = p_j(x)$  and so

$A_j \sim A_j^p$  by (a) above for  $1 \leq j \leq k$ . Therefore

$$A_1 \oplus A_2 \oplus \cdots \oplus A_k = A_1^p \oplus A_2^p \oplus \cdots \oplus A_k^p \text{ and so } A \sim A^p.$$

### Solution 5

(a) Using the partition function  $p(n)$  of (3.13), there are

$p(3)p(4)p(5) = 4 \times 5 \times 7 = 140$  similarity classes of  $12 \times 12$  matrices  $A$  over  $\mathbb{Q}$  with  $\chi_A(x) = (x-1)^3 x^4 (x+1)^5$ . As  $\mu_A(x) = (x-1)^i x^j (x+1)^k$  where  $1 \leq i \leq 3$ ,  $1 \leq j \leq 4$ ,  $1 \leq k \leq 5$  by (6.11) there are  $3 \times 4 \times 5 = 60$  different minimum polynomials  $\mu_A(x)$  among these classes.

(i)  $A$  is cyclic  $\Leftrightarrow$  the partitions of 3, 4, 5 are (3), (4), (5), i.e. each of these partitions has one part only and  $\mu_A(x) = \chi_A(x)$ .

(ii)  $A$  is diagonalizable  $\Leftrightarrow$  the partitions of 3, 4, 5 are (1, 1, 1), (1, 1, 1, 1), (1, 1, 1, 1, 1), i.e. all the parts in these partitions are 1 and  $\mu_A(x) = (x-1)x(x+1)$ .

The table of the isomorphism class of  $M(A)$  is

$\cong M(A)$	$x-1$	$x$	$x+1$
$d_1(x)$	0	0	1
$d_2(x)$	0	0	1
$d_3(x)$	1	1	1
$d_4(x)$	2	3	2

and so the invariant factor sequence of  $A$  is

$$(d_1(x), d_2(x), d_3(x), d_4(x)) = (x+1, x+1, (x-1)x(x+1), (x-1)^2 x^3 (x+1)^2).$$

Therefore  $\mu_A(x) = (x-1)^2 x^3 (x+1)^2$  and the rcf and pcf of  $A$  are

$$C(x+1) \oplus C(x+1) \oplus C((x-1)x(x+1)) \oplus C((x-1)^2 x^3 (x+1)^2) \text{ and}$$

$$C(x-1) \oplus C((x-1)^2) \oplus C(x) \oplus C(x^3) \oplus C(x+1) \oplus C(x+1) \oplus C(x+1) \oplus C((x+1)^2)$$

respectively.

(b) Since  $x-1, x^2+1, x+1$  are irreducible over  $\mathbb{Q}$ , using the partition function  $p(n)$  of (3.13), there are  $p(5)p(4)p(5) = 7 \times 5 \times 7 = 245$  similarity classes of  $18 \times 18$  matrices  $A$  over  $\mathbb{Q}$  with  $\chi_A(x) = (x-1)^5 (x^2+1)^4 (x+1)^5$ . The isomorphism class of  $M(A_0)$  has table

$\cong M(A_0)$	$x-1$	$x^2+1$	$x+1$
$d_1(x)$	1	1	1
$d_2(x)$	2	1	1
$d_3(x)$	2	2	3

and so the invariant factor sequence of  $A_0$  is

$$(d_1(x), d_2(x), d_3(x)) = ((x-1)(x^2+1)(x+1), (x-1)^2 (x^2+1)(x+1), (x-1)^2 (x^2+1)^2 (x+1)^3).$$

The invariant factors of  $-A_0$  are

$$d_1(-x) = (-x+1)((-x)^2+1)(x+1) = d_1(x),$$

$$-d_2(-x) = -(-x-1)^2((-x)^2+1)(-x+1) \neq d_2(x),$$

$$-d_3(-x) = -(-x-1)^2((-x)^2+1)^2(-x+1)^3 \neq -d_3(x)$$

by Exercises 6.1, Question 2(a). The table of  $\cong M(-A_0)$  is therefore

$\cong M(-A_0)$	$x-1$	$x^2+1$	$x+1$
$d_1(-x)$	1	1	1
$-d_2(-x)$	1	1	2
$-d_3(-x)$	3	2	2

and  $-A_0 \not\sim A_0$  as the invariant factors of  $A_0$  and  $-A_0$  are different. By

Exercises 6.1, Question 2(a) we see  $-A \sim A$  if and only if the  $(x-1)$ -column equals the  $(x+1)$ -column in the table of  $\cong M(A)$ . So there are  $p(5) \times p(4) = 7 \times 5 = 35$  similarity classes of matrices  $A$  as above with  $-A \sim A$ , as the exponent of  $x+1$  must partition in the same way as the exponent of  $x-1$  for such  $A$ . As  $x-1, x^2+1, x+1$  are all palindromic, we see  $A \sim A^{-1}$  for all 245 similarity classes as above by Exercises 6.1, Question 5(d).

(c) There are  $p(5)p(5)p(6)p(6) = 7 \times 7 \times 11 \times 11 = 5929$  similarity classes of  $22 \times 22$  matrices  $A$  over the rational field  $\mathbb{Q}$  with  $\chi_A(x) = (x-1)^5 (x+1)^5 (x-2)^6 (x+1/2)^6$ .

As  $\chi_{-A}(x) = \chi_A(-x) = (x-1)^5 (x+1)^5 (x+2)^6 (x+1/2)^6 \neq \chi_A(x)$  none of these

similarity classes satisfy  $-A \sim A$  by (5.5). Using the terminology and notation of Exercises 6.1, Question 5(d) the polynomials  $x-1$  and  $x+1$  are palindromic whereas  $(x-2)^* = x-1/2$  and  $(x-1/2)^* = x-2$ . So  $A \sim A^{-1}$  if and only if the partition of the exponent 6 of  $x-2$  in  $\chi_A(x)$  is equal to the partition of the exponent 6 of  $x-1/2$  in  $\chi_A(x)$ . Therefore there are  $p(5)p(5)p(6) = 7 \times 7 \times 11 = 539$  similarity classes of matrices  $A$  over  $\mathbb{Q}$  with  $\chi_A(x)$  as above and  $A \sim A^{-1}$ .

### Solution 6

(a) Working in the  $F[x]$ -module  $M(A)$  we have  $xe_1 = e_1A = a_{11}e_1 + a_{12}e_2$  and  $x^2e_1 = (a_{11}e_1 + a_{12}e_2)A = v_2 + a_{12}a_{23}e_3$  where  $v_2 \in \langle e_1, e_2 \rangle$ . Using induction on  $i$  we obtain  $x^ie_1 = v_i + a_{12}a_{23} \cdots a_{i+1}e_{i+1}$  where  $v_i \in \langle e_1, e_2, \dots, e_i \rangle$  for  $1 \leq i < t$ . Let  $X$  be the  $t \times t$  matrix with  $e_iX = x^{i-1}e_1$  for  $1 \leq i \leq t$ . Then  $X$  is a lower triangular matrix, i.e. the  $(i, j)$ -entries in  $X$  are all zero for  $i < j$ . Also the  $(1, 1)$ -entry in  $X$  is 1 and the  $(i, i)$ -entry in  $X$  is  $a_{12}a_{23} \cdots a_{i-1,i}$  for  $1 < i \leq t$ . As  $\det X$  is the product of the entries on the main diagonal of  $X$  we obtain  $\det X = a_{12}^{t-1}a_{23}^{t-2} \cdots a_{t-2,t-1}^2a_{t-1,t}$ . Now  $e_1$  generates  $M(A)$  if and only if the rows  $e_1, xe_1, x^2e_1, \dots, x^{t-1}e_1$  of  $X$  are linearly independent, i.e. if and only if  $\det X \neq 0 \Leftrightarrow a_{i+1,i} \neq 0$  for  $1 \leq i < t$ . Incidentally when this is the case  $XAX^{-1} = C(\chi_A(x))$  by (5.27).

(b) Taking  $f(x) = x+1, n=1, l=3, F = \mathbb{Q}$  in (6.16) we obtain the  $3 \times 3$  matrix

$$Z_1 = \begin{pmatrix} e_1 \\ (x+1)e_1 \\ (x+1)^2e_1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 2 & 1 \end{pmatrix}$$

which is invertible over  $\mathbb{Q}$ . Now

$$J(x+1, 3) = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{pmatrix}$$

and  $Z_1$  satisfies  $Z_1C((x+1)^3) = J(x+1, 3)Z_1$  by (6.16), i.e.

$$Z_1C((x+1)^3)Z_1^{-1} = J(x+1, 3).$$

Taking  $f(x) = x^2+1, n=2, l=2, F = \mathbb{Q}$  in (6.16) we obtain the  $4 \times 4$  matrix

$$Z_2 = \begin{pmatrix} e_1 \\ xe_1 \\ (x^2+1)e_1 \\ x(x^2+1)e_1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}$$

which is invertible over  $\mathbb{Q}$  as  $\det Z_2 = 1$ . Now

$$J(x^2+1, 2) = \left( \begin{array}{cc|cc} 0 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{array} \right)$$

and  $Z_2$  satisfies  $Z_2C((x^2+1)^2)Z_2^{-1} = J(x^2+1, 2)$  by (6.16).

The  $7 \times 7$  matrix  $Z = Z_1 \oplus Z_2$  is invertible over  $\mathbb{Q}$  and as  $Z^{-1} = Z_1^{-1} \oplus Z_2^{-1}$  we see

$$Z(C((x+1)^3) \oplus C((x^2+1)^2))Z^{-1} = Z_1 C((x+1)^3)Z_1^{-1} \oplus Z_2 C((x^2+1)^2)Z_2^{-1} = J(x+1, 3) \oplus J(x^2+1, 2)$$

which is in Jnf by (6.15) as  $x+1$  and  $x^2+1$  are monic and irreducible over  $\mathbb{Q}$ .

(c) Taking  $f(x) = (x+1)^2$  and  $l=3$  in (6.16) gives

$$C((x+1)^6) = C(((x+1)^2)^3) \sim J((x+1)^2, 3).$$

All the matrices in the list are similar as each is similar as above to  $C((x+1)^6)$ . The matrix  $J(x+1, 6)$  is in Jnf as  $x+1$  is irreducible over  $F$ .

(d) Using the method of (5.31) we construct

$$Y = \begin{pmatrix} (x+1)^2 e_1 \\ x(x+1)^2 e_1 \\ x^2(x+1)^2 e_1 \\ \hline x^3 e_1 \\ x^4 e_1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 1 & 0 & 0 \\ 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & 1 & 2 & 1 \\ \hline 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

which is invertible over  $F$  as  $\det Y = 1$  and satisfies

$$YAY^{-1} = YC(x^3(x+1)^2)Y^{-1} = C(x^3) \oplus C((x+1)^2)$$

in pcf. As  $C(x^3) = J(x, 3)$  the first three rows of  $Z$  can be taken to be the first three

rows of  $Y$ . As  $\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & -2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix} = J(x+1, 2)$  by (6.16) we see

$$Z = Y \left( \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \oplus \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \right) = \begin{pmatrix} 1 & 2 & 1 & 0 & 0 \\ 0 & 1 & 2 & 1 & 0 \\ 0 & 0 & 1 & 2 & 1 \\ \hline 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix}$$

satisfies  $ZAZ^{-1} = ZC(x^3(x+1)^2)Z^{-1} = J(x, 3) \oplus J(x+1, 2)$  in Jnf.

Now  $ZA^2Z^{-1} = (ZAZ^{-1})^2 = J(x, 3)^2 \oplus J(x+1, 2)^2$ . Also

$$J(x, 3)^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \sim J(x, 1) \oplus J(x, 2) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \text{ and } J(x+1, 2)^2 = \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix}.$$

Therefore in the case  $\chi(F) = 2$  we see  $J(x, 1) \oplus J(x, 2) \oplus J(x-1) \oplus J(x-1)$  is the Jnf of  $A^2$ . In the case  $\chi(F) \neq 2$ , as

$$\begin{pmatrix} -1/2 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} -2 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = J(x-1, 2),$$

the Jnf of  $A^2$  is  $J(x, 1) \oplus J(x, 2) \oplus J(x-1, 2)$ .

Now  $Z(A + A^2)Z^{-1} = (J(x, 3) + J(x, 3)^2) \oplus (J(x+1, 2) + J(x+1, 2)^2)$ . But

$$X_1(J(x, 3) + J(x, 3)^2)X_1^{-1} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} = J(x, 3) \text{ where } X_1 = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \text{ Also}$$

$$X_2(J(x+1, 2) + J(x+1, 2)^2)X_2^{-1} = X_2 \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix} X_2^{-1} = J(x, 2) \text{ where } X_2 = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Therefore  $(X_1 \oplus X_2)Z(A + A^2)Z^{-1}(X_1 \oplus X_2)^{-1} = J(x, 3) \oplus J(x, 2)$  which is in Jnf.

(e) Suppose that  $A$  has  $s$  invariant factors  $d_1(x), d_2(x), \dots, d_s(x)$ . From  $d_j(x) \mid \chi_A(x)$  we deduce  $d_j(x) = (x - \lambda)^{t_j}$  where  $1 \leq t_1 \leq t_2 \leq \dots \leq t_s$  as  $d_j(x) \mid d_{j+1}(x)$  for  $1 \leq j < s$ . As  $d_1(x)d_2(x) \cdots d_s(x) = \chi_A(x) = (x - \lambda)^t$  we see  $t_1 + t_2 + \dots + t_s = t$  and so  $(t_1, t_2, \dots, t_s)$  is a partition (3.13) of  $t$  with  $s$  parts. All partitions of  $t$  arise in this way and so each partition  $(t_1, t_2, \dots, t_s)$  of  $t$  corresponds to a similarity class of  $t \times t$  matrices  $A$ , namely those having invariant factor sequence  $((x - \lambda)^{t_1}, (x - \lambda)^{t_2}, \dots, (x - \lambda)^{t_s})$ . So the number  $p(t)$  of partitions of  $t$  is the number of similarity classes of  $t \times t$  matrices  $A$  over  $F$  with  $\chi_A(x) = (x - \lambda)^t$ .

(i)

$$\begin{aligned} & C(x - \lambda) \oplus C(x - \lambda) \oplus C(x - \lambda) \oplus C(x - \lambda) \oplus C(x - \lambda) = \lambda I, \\ & C(x - \lambda) \oplus C(x - \lambda) \oplus C(x - \lambda) \oplus C((x - \lambda)^2), \\ & C(x - \lambda) \oplus C(x - \lambda) \oplus C((x - \lambda)^3), \\ & C(x - \lambda) \oplus C((x - \lambda)^2) \oplus C((x - \lambda)^2), \\ & C(x - \lambda) \oplus C((x - \lambda)^4), \\ & C((x - \lambda)^2) \oplus C((x - \lambda)^3), \\ & C((x - \lambda)^5). \end{aligned}$$

(ii)

$$\begin{aligned} & J(x - \lambda, 1) \oplus J(x - \lambda, 1) \oplus J(x - \lambda, 1) \oplus J(x - \lambda, 1) \oplus J(x - \lambda, 1), \\ & J(x - \lambda, 1) \oplus J(x - \lambda, 1) \oplus J(x - \lambda, 1) \oplus J(x - \lambda, 2), \\ & J(x - \lambda, 1) \oplus J(x - \lambda, 1) \oplus J(x - \lambda, 3), \\ & J(x - \lambda, 1) \oplus J(x - \lambda, 2) \oplus J(x - \lambda, 2), \\ & J(x - \lambda, 1) \oplus J(x - \lambda, 4), \\ & J(x - \lambda, 2) \oplus J(x - \lambda, 3), \\ & J(x - \lambda, 5). \end{aligned}$$

As  $J(x - \lambda, t_j) - \lambda I = C(x^{t_j})$  we see  $\text{rank}(J(x - \lambda, t_j) - \lambda I) = t_j - 1$ . Therefore

$$\begin{aligned} \text{rank}(J - \lambda I) &= \text{rank}\left(\sum_{j=1}^s \oplus (J(x - \lambda, t_j) - \lambda I)\right) = \\ & \sum_{j=1}^s \text{rank}(J(x - \lambda, t_j) - \lambda I) = \sum_{j=1}^s (t_j - 1) = \sum_{j=1}^s t_j - s = t - s. \end{aligned}$$

As  $B^2 = (I + C)^2 = I + 2C + C^2$  where  $C = C(x^t)$  we see

$$\chi_{B^2}(x) = |xI - I - 2C - C^2| = (x - 1)^t$$

since  $2C - C^2$  is strictly upper triangular (its  $(i, j)$ -entries are zero for  $i \geq j$ ).

Therefore the Jnf of  $B^2$  is  $\sum_{j=1}^s \oplus J(x - 1, t_j)$  for some partition  $(t_1, t_2, \dots, t_s)$  of  $t$ . By

the preceding part of the question  $\text{rank}(B^2 - I) = t - s$ . But directly  $B^2 - I = 2C + C^2$  has rank  $t - 1$  in the case  $\chi(F) \neq 2$  and rank  $t - 2$  in the case  $\chi(F) = 2$ .

Suppose  $B \sim B^2$ . Then  $B - I \sim B^2 - I$  and so

$t - 1 = \text{rank}(B - I) = \text{rank}(B^2 - I)$ . Therefore  $\chi(F) \neq 2$ .

Conversely suppose  $\chi(F) \neq 2$ . Then  $\text{rank}(B^2 - I) = t - 1 = t - s$  showing  $s = 1$  and so  $B^2 \sim J(x-1, t) = B$ .

Suppose  $\chi(F) = 2$ . Then  $B^2 = I + C^2$ . By Exercises 6.1, Question 2(e) the rcf of  $C^2$  is  $C(x^{t/2}) \oplus C(x^{t/2})$  for  $t$  even and  $C(x^{(t-1)/2}) \oplus C(x^{(t+1)/2})$  for  $t$  odd. Hence the Inf of  $B^2$  is  $J(x-1, t/2) \oplus J(x-1, t/2)$  for  $t$  even and  $J(x-1, (t-1)/2) \oplus J(x-1, (t+1)/2)$  for  $t$  odd.

### Solution 7

(a)(i) The coefficient of  $x^i$  in  $f(x) + g(x)$  is  $a_i + b_i$  where  $a_i = b_i = 0$  for  $i > m, j > n$ . The coefficient of  $x^{i-1}$  in  $(f(x) + g(x))'$  is therefore  $i(a_i + b_i)$  for  $i > 0$ . As the coefficient of  $x^{i-1}$  in  $f'(x) + g'(x)$  is  $ia_i + ib_i$  for  $i > 0$  by (6.20) we conclude  $(f(x) + g(x))' = f'(x) + g'(x)$ . (We write  $(f(x) + g(x))'$  and  $(f(x)g(x))'$  rather than  $(f + g)'(x)$  and  $(fg)'(x)$ .)

The coefficient of  $x^{i-1}$  in  $cf'(x)$  equals the coefficient of  $x^{i-1}$  in  $(cf(x))'$  as  $cia_i = ica_i$  for  $i > 0$ . Therefore  $(cf(x))' = cf'(x)$ .

(ii) The coefficient of  $x^{i-1}$  in  $f'(x)g(x)$  is

$$a_1b_{i-1} + 2a_2b_{i-2} + \dots + (i-1)a_{i-1}b_1 + ia_ib_0$$

and the coefficient of  $x^{i-1}$  in  $f(x)g'(x)$  is

$$a_0ib_i + a_1(i-1)b_{i-1} + \dots + a_{i-1}b_1.$$

The sum of these coefficients is

$$i(a_0b_i + a_1b_{i-1} + \dots + a_{i-1}b_1 + a_ib_0)$$

which is the coefficient of  $x^{i-1}$  in  $(f(x)g(x))'$  for  $i > 0$ . Therefore

$$(f(x)g(x))' = f'(x)g(x) = f(x)g'(x).$$

Putting  $i = 1$  in  $(g(x)^i)' = ig(x)^{i-1}g'(x)$  gives  $g'(x) = g'(x)$  as  $g(x)^0 = 1$ . Suppose inductively  $i > 1$  and  $(g(x)^{i-1})' = (i-1)g(x)^{i-2}g'(x)$ . From (ii) above

$$(g(x)^i)' = (g(x)^{i-1}g(x))' = (g(x)^{i-1})'g(x) + g(x)^{i-1}g'(x).$$

So

$$\begin{aligned} (g(x)^i)' &= (i-1)g(x)^{i-2}g'(x)g(x) + g(x)^{i-1}g'(x) = \\ &= (i-1)g(x)^{i-1}g'(x) + g(x)^{i-1}g'(x) = ig(x)^{i-1}g'(x) \end{aligned}$$

completing the inductive step. Therefore  $(g(x)^i)' = ig(x)^{i-1}g'(x)$  for  $i \geq 1$ .

Now  $f(g(x)) = \sum_{i=0}^m a_i g(x)^i$ . Using (i) and (ii) above we obtain

$$f(g(x))' = \sum_{i=0}^m a_i (g(x)^i)' = \sum_{i=1}^m ia_i g(x)^{i-1}g'(x) = f'(g(x))g'(x) \text{ as } (g(x)^0)' = 0.$$

Therefore  $f(g(x))' = f'(g(x))g'(x)$ .

(b) The binomial coefficient  $\binom{p}{i}$  is divisible by the prime  $p$  for  $0 < i < p$

(see Exercises 2.3, Question 5(a)). By the binomial theorem

$$(x-c)^p = x^p + (-c)^p = x^p + (-1)^p c^p = x^p - y$$

as  $(-1)^p \equiv -1 \pmod{p}$  for all primes  $p$ . Therefore  $x^p - y$  has the factorisation

$(x-c)^p$  into irreducible polynomials over  $E$ . Let  $p(x)$  and  $q(x)$  be monic

irreducible factors of  $x^p - y$  over  $F$ . Then  $p(x) = (x - c)^l$  and  $q(x) = (x - c)^m$  for positive integers  $l, m$ . So either  $p(x) \mid q(x)$  or  $q(x) \mid p(x)$ . As  $p(x)$  and  $q(x)$  are irreducible over  $F$  we see  $p(x) = q(x)$  and so  $x^p - y = p(x)^k$  where  $p = k \deg p(x)$ . As  $p$  is prime either  $\deg p(x) = 1$  or (as we wish to show)  $k = 1$  and  $x^p - y = p(x)$  is irreducible over  $F$ . So suppose  $\deg p(x) = 1$  and so  $p(x) = x - c'$  where  $c' \in F$ . Then  $x^p - y = (x - c')^p$  and also  $x^p - y = (x - c)^p$ . As the resolution of  $x^p - y$  into monic irreducible factors over  $E$  is unique we see  $c' = c$ . So  $c = f(y)/g(y)$  is a zero of  $x^p - y$  where  $f(y), g(y) \in \mathbb{Z}_p[y]$ ,  $g(y) \neq 0$ . So  $c^p = y$  which gives

$f(y)^p = yg(y)^p$ . This last equation is impossible as  $\deg f(y)^p \equiv 0 \pmod{p}$  and  $\deg yg(y)^p \equiv 1 \pmod{p}$ . So  $x^p - y$  is irreducible over  $F$ .

Yes,  $x^p - y$  is inseparable over  $F$ . The above reasoning uses the fact that  $\mathbb{Z}_p$  is a field of characteristic  $p$  but nothing further. On replacing  $\mathbb{Z}_p$  by any field  $F_0$  of characteristic  $p$  the properties of  $x^p - y$  are unchanged: it is irreducible and inseparable over the field of fractions  $F = F_0(y)$  of  $F_0[x]$ .

(c) Let  $g(x) = b_n x^n + b_{n-1} x^{n-1} + b_1 x + b_0$  be a polynomial over  $F$  and let  $c \in F$ . For  $i \leq j \leq n$  the coefficients of  $x^{j-i}$  in  $H_i(f(x) + g(x))$  and in  $H_i(f(x)) + H_i(g(x))$

are equal as  $(a_j + b_j) \binom{j}{i} = a_j \binom{j}{i} + b_j \binom{j}{i}$ . Therefore

$H_i(f(x) + g(x)) = H_i(f(x)) + H_i(g(x))$ . Also  $cH_i(f(x)) = H_i(cf(x))$  as the coefficient of  $x^{j-i}$  in both these polynomials is  $ca_j \binom{j}{i}$  for  $i \leq j \leq n$ . So

$f(x) \rightarrow H_i(f(x))$  is a linear mapping of the (infinite dimensional) vector space  $F[x]$  over  $F$ .

With  $f(x) = x^j$  we have  $f''(x) = 0$  for  $0 \leq j < 2$  and  $f''(x) = j(j-1)x^{j-2}$  for  $j \geq 2$ . As  $j(j-1)$  is even we see  $f''(x) = 0$  for all  $j \geq 0$  in the case  $F = \mathbb{Z}_2$ . Applying part (a)(i) above twice in succession, we conclude:  $f''(x) = 0$  for all  $f(x) \in \mathbb{Z}_2[x]$ .

Now  $H_2(x^j) = 0$  for  $0 \leq j < 2$  and  $H_2(x^j) = \binom{j}{2} x^{j-2}$  for  $j \geq 2$ . Also

$\binom{j}{2} = j(j-1)/2$  is even if and only if either  $j \equiv 0 \pmod{4}$  or  $j \equiv 1 \pmod{4}$ . So

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + a_1 x + a_0 \in \mathbb{Z}_2[x]$$

satisfies  $H_2(f(x)) = 0$  if and only if  $a_j = 0$  for  $j \equiv 2 \pmod{4}$  and  $a_j = 0$  for

$j \equiv 3 \pmod{4}$ . In other words  $\ker H_2 = \langle 1, x, x^4, x^5, x^8, x^9, \dots \rangle$  and

$\text{im } H_2 = \langle 1, x, x^4, x^5, x^8, x^9, \dots \rangle$  and so  $\ker H_2 = \text{im } H_2$ . Hence  $H_2^2 = 0$ , i.e.

$H_2(H_2(f(x)))$  is the zero polynomial for all  $f(x) \in \mathbb{Z}_2[x]$ .

As  $i! \binom{j}{i} = j!/(j-i)!$  we see  $i!H_i(x^j) = j(j-1)(j-2)\cdots(j-i+1)x^{j-i}$  which is the

result of formally differentiating  $i$  times the polynomial  $x^j$  for  $j \geq i$ . As both  $i!H_i$  and formally differentiating  $i$  times are linear mappings of  $F[x]$  with  $x^j$  in their

kernels for  $0 \leq j < i$ , we conclude  $i!H_i(f(x)) = f^{(i)}(x)$  for all  $f(x) \in F[x]$ . As

$$(j-i) \binom{j}{i} = (i+1) \binom{j}{i+1} \text{ we deduce}$$

$$H'_i(x^j) = (j-i) \binom{j}{i} x^{j-i-1} = (i+1) \binom{j}{i+1} x^{j-i-1} = (i+1)H_{i+1}(x^j)$$

for all  $j$ . As above we conclude  $H'_i(f(x)) = (i+1)H_{i+1}(f(x))$  for all  $f(x) \in F[x]$ .

Suppose  $\chi(F)$  is not a divisor of  $i!$ , i.e.  $i! \neq 0$  in  $F$ . Dividing the equation

$$i!H_i(f(x)) = f^{(i)}(x) \text{ through by } i! \text{ gives } H_i(f(x)) = f^{(i)}(x)/i!.$$

Suppose  $\chi(F)$  is not a divisor of  $i+1$ . Then  $i+1 \neq 0$  in  $F$ . Dividing the equation

$$H'_i(f(x)) = (i+1)H_{i+1}(f(x)) \text{ through by } i+1 \text{ gives } H_{i+1}(f(x)) = H_i(f(x))/(i+1)$$

for all  $f(x) \in F[x]$ .

(d)

$$\text{As } J_S(B, 4) = \begin{pmatrix} B & I & 0 & 0 \\ 0 & B & I & 0 \\ 0 & 0 & B & I \\ 0 & 0 & 0 & B \end{pmatrix} \text{ we see } J_S(B, 4)^2 = \begin{pmatrix} B^2 & 2B & I & 0 \\ 0 & B^2 & 2B & I \\ 0 & 0 & B^2 & 2B \\ 0 & 0 & 0 & B^2 \end{pmatrix},$$

$$J_S(B, 4)^3 = \begin{pmatrix} B^3 & 3B^2 & 3B & I \\ 0 & B^3 & 3B^2 & 3B \\ 0 & 0 & B^3 & 3B^2 \\ 0 & 0 & 0 & B^3 \end{pmatrix} \text{ and } J_S(B, 4)^4 = \begin{pmatrix} B^4 & 4B^3 & 6B^2 & 4B \\ 0 & B^4 & 4B^3 & 6B^2 \\ 0 & 0 & B^4 & 4B^3 \\ 0 & 0 & 0 & B^4 \end{pmatrix}.$$

Let  $f(x) = a_n x^n + a_{n-1} x^{n-1} + a_1 x + a_0$ . As  $J_S(B, l)^j = \sum_{i=0}^{l-1} \binom{j}{i} \tilde{B}^{j-i} \tilde{N}^i$ , on multiplying

this equation by  $a_j$  and summing over  $j$  we obtain

$$f(J_S(B, l)) = \sum_{j=0}^n a_j J_S(B, l)^j = \sum_{j=0}^n a_j \left( \sum_{i=0}^{l-1} \binom{j}{i} \tilde{B}^{j-i} \tilde{N}^i \right) =$$

$$\sum_{i=0}^{l-1} \left( \sum_{j=0}^n a_j \binom{j}{i} \tilde{B}^{j-i} \right) \tilde{N}^i = \sum_{i=0}^{l-1} H_i(f(\tilde{B})) \tilde{N}^i$$

where

$H_i(f(\tilde{B})) = H_i(f(B \oplus B \oplus \dots \oplus B)) = H_i(f(B)) \oplus H_i(f(B)) \oplus \dots \oplus H_i(f(B))$ ,  
the direct sum of  $l$  matrices  $H_i(f(B)) = (H_i(f(x)))\mathcal{E}_B$ . So the  $(k, i+k)$ -entries in  
the partitioned matrices  $f(J_S(B, l))$  and  $H_i(f(\tilde{B}))\tilde{N}^i$  are both equal to the  $n \times n$   
matrix  $H_i(f(B))$  for  $1 \leq k \leq l-i, 0 \leq i < l$ .

Let  $1 \leq k \leq k' \leq l$ . The  $(k, k')$ -entry in the  $l \times l$  partitioned matrix  $\chi_B(J_S(B, l))$  is  
therefore  $(H_{k'-k}(\chi_B(x)))\mathcal{E}_B$  on taking  $i = k' - k$ . In particular the  $(i, i)$ -entries are  
 $(H_0(\chi_B(x)))\mathcal{E}_B = (\chi_B(x))\mathcal{E}_B = \chi_B(B) = 0$  by (6.11) and the  $(i, i+1)$ -entries are

$$(H_1(\chi_B(x)))\mathcal{E}_B = (\chi'_B(x))\mathcal{E}_B = \chi'_B(B).$$

### Solution 8

(a) Note  $\tilde{C}^2 = 0$  and  $\tilde{N}^l = 0$ . Write  $C = C(x^2)$ . Using the binomial expansion we see



$$A^{l-1} = (\tilde{C} + \tilde{N})^{l-1} = \binom{l-1}{l-2} \tilde{C} \tilde{N}^{l-2} + \binom{l-1}{l-1} \tilde{N}^{l-1} = \left( \begin{array}{c|c|c|c|c} 0 & \cdots & 0 & (l-1)C & I \\ 0 & \cdots & \cdots & 0 & (l-1)C \\ 0 & \cdots & \cdots & \cdots & 0 \\ \vdots & & & & \vdots \\ 0 & \cdots & \cdots & \cdots & 0 \end{array} \right) \neq 0,$$

$$A^l = (\tilde{C} + \tilde{N})^l = \binom{l}{l-1} \tilde{C} \tilde{N}^{l-1} = \left( \begin{array}{c|c|c|c|c} 0 & \cdots & \cdots & 0 & lC \\ 0 & \cdots & \cdots & \cdots & 0 \\ \vdots & & & & \vdots \\ \vdots & & & & \vdots \\ 0 & \cdots & \cdots & \cdots & 0 \end{array} \right)$$

and  $A^{l+1} = 0$ . As  $C \neq 0$  we see  $A^l = 0 \Leftrightarrow \chi(F) \mid l$ . As  $\chi_A(x) = x^{2l}$  from (6.11) we deduce  $\mu_A(x) = x^m$  where  $m$  is the least integer with  $A^m = 0$ . Therefore  $m = l$  or  $m = l+1$  according as  $\chi(F)$  is or is not a divisor of  $l$ .

By inspection the first  $2l-2$  rows of  $A$  are linearly independent. As  $e_{2l-1}A = e_{2l-2}A$  and  $e_{2l}A = 0$  we see  $\text{rank } A = 2l-2$ . So nullity  $A = 2$  which is the number of invariant factors of  $A$ . We conclude that the rcf of  $A$  is either  $C(x^l) \oplus C(x^l)$  or  $C(x^{l-1}) \oplus C(x^{l+1})$  according as  $\chi(F)$  is or is not a divisor of  $l$ .

(b) Consider first the  $1 \times 1$  matrices  $A = (a)$  and  $A' = (a')$  over  $F$ . Then

$$J_S(A, 2) \oplus J_S(A', 2) = \left( \begin{array}{c|c|c|c} a & 1 & 0 & 0 \\ 0 & a & 0 & 0 \\ 0 & 0 & a' & 1 \\ 0 & 0 & 0 & a' \end{array} \right) \text{ and } J_S(A \oplus A', 2) = \left( \begin{array}{c|c|c|c} a & 0 & 1 & 0 \\ 0 & a' & 0 & 1 \\ 0 & 0 & a & 0 \\ 0 & 0 & 0 & a' \end{array} \right).$$

$$\text{Let } X = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \text{ Then}$$

$$X(J_S(A, 2) \oplus J_S(A', 2)) = \begin{pmatrix} a & 1 & 0 & 0 \\ 0 & 0 & a' & 1 \\ 0 & a & 0 & 0 \\ 0 & 0 & 0 & a' \end{pmatrix} = J_S(A \oplus A', l)X.$$

Now consider the case  $n, n'$  arbitrary,  $l = 3$ . The  $3(n+n') \times 3(n+n')$  matrices

$$J_S(A, 3) \oplus J_S(A', 3) = \left( \begin{array}{c|c|c|c|c|c} A & I_n & 0 & 0 & 0 & 0 \\ 0 & A & I_n & 0 & 0 & 0 \\ 0 & 0 & A & 0 & 0 & 0 \\ 0 & 0 & 0 & A' & I_{n'} & 0 \\ 0 & 0 & 0 & 0 & A' & I_{n'} \\ 0 & 0 & 0 & 0 & 0 & A' \end{array} \right)$$

and

$$J_S(A \oplus A', 3) = \left( \begin{array}{c|c|c|c|c|c} A & 0 & I_n & 0 & 0 & 0 \\ \hline 0 & A' & 0 & I_{n'} & 0 & 0 \\ \hline 0 & 0 & A & 0 & I_n & 0 \\ \hline 0 & 0 & 0 & A' & 0 & I_{n'} \\ \hline 0 & 0 & 0 & 0 & A & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & A' \end{array} \right)$$

can be conveniently viewed as  $6 \times 6$  partitioned matrices, their entries being as indicated ( $I_n$  denotes the  $n \times n$  identity matrix,  $I_{n'}$  denotes the  $n' \times n'$  identity matrix and 0 denotes zero matrices of the appropriate size:  $n \times n, n' \times n', n \times n', n' \times n$ ).

Write

$$X = \left( \begin{array}{c|c|c|c|c|c} I_n & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & I_{n'} & 0 & 0 \\ \hline 0 & I_n & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & I_{n'} & 0 \\ \hline 0 & 0 & I_n & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & I_{n'} \end{array} \right).$$

As  $X$  can be changed into the  $3(n+n') \times 3(n+n')$  identity matrix by applying  $2nn' + nn'$  elementary column operations of the type  $c_{i-1} \leftrightarrow c_i$  we see  $\det X = (-1)^{3nn'}$ . So  $X$  is invertible over  $F$ . Also

$$X(J_S(A, 3) \oplus J_S(A', 3)) = \left( \begin{array}{c|c|c|c|c|c} A & I_n & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & A' & I_{n'} & 0 \\ \hline 0 & A & I_n & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & A' & I_{n'} \\ \hline 0 & 0 & A & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & A' \end{array} \right) = J_S(A \oplus A', 3)X.$$

Now consider the general case. The  $l(n+n') \times l(n+n')$  matrices  $J_S(A, l) \oplus J_S(A', l)$  and  $J_S(A \oplus A', l)$  over  $F$  are  $2l \times 2l$  partitioned matrices with entries  $A, I_n, A', I_{n'}$  and  $n \times n, n' \times n', n \times n', n' \times n$  zero matrices 0. In fact the  $2l \times 2l$  partitioned matrix  $J_S(A, l) \oplus J_S(A', l)$  has  $(k, k)$ -entry  $A$  for  $1 \leq k \leq l$ ,  $(k, k+1)$ -entry  $I_n$  for  $1 \leq k < l$ ,  $(k+l, k+l)$ -entry  $A'$  for  $1 \leq k \leq l$ ,  $(k+l, k+l+1)$ -entry  $I_{n'}$  for  $1 \leq k < l$ , all other entries being zero matrices. The  $2l \times 2l$  partitioned matrix  $J_S(A \oplus A', l)$  has  $(2k-1, 2k-1)$ -entry  $A$  and  $(2k, 2k)$ -entry  $A'$  for  $1 \leq k \leq l$ ,  $(2k-1, 2k+1)$ -entry  $I_n$  and  $(2k, 2k+2)$ -entry  $I_{n'}$  for  $1 \leq k < l$ , all other entries being zero matrices.

Denote by  $X$  the  $2l \times 2l$  partitioned matrix with  $(2k-1, k)$ -entry  $I_n$  and  $(2k, k+l)$ -entry  $I_{n'}$  for  $1 \leq k \leq l$ , all other entries being zero matrices. For  $1 \leq j \leq n', 1 \leq k \leq l$  column  $ln + (k-1)n' + j$  of  $X$  can be moved to column  $kn + (k-1)n' + j$  using  $(l-k)n$  interchanges of consecutive columns. Carrying out this operation, starting with column  $ln+1$  and ending with column  $l(n+n')$ , changes  $X$  into the  $l(n+n') \times l(n+n')$  identity matrix  $I$  and uses

$(l-1+l-2+\dots+2+1+0)nn' = l(l-1)nn'/2$  column interchanges. Hence

$\det X = (-1)^{l(l-1)nn'/2}$  showing that  $X$  is invertible over  $F$ .

The  $2l \times 2l$  partitioned matrices  $X$  and  $J_S(A, l) \oplus J_S(A', l)$  are compatible, i.e.

$X(J_S(A, l) \oplus J_S(A', l))$  is a  $2l \times 2l$  partitioned matrix and has  $(2k-1, k)$ -entry  $A$  and  $(2k, k+l)$ -entry  $A'$  for  $1 \leq k \leq l$ ,  $(2k-1, k+1)$ -entry  $I_n$  and  $(2k, k+l+1)$ -entry  $I_{n'}$  for  $1 \leq k < l$ , all other entries being zero matrices. In the same way the  $2l \times 2l$  partitioned matrices  $J_S(A \oplus A', l)$  and  $X$  are compatible, i.e.

$J_S(A \oplus A', l)X$  is a  $2l \times 2l$  partitioned matrix. On comparing matrix entries we obtain

$$X(J_S(A, l) \oplus J_S(A', l)) = J_S(A \oplus A', l)X.$$

The  $ln \times ln$  matrices  $J_S(A, l)$ ,  $J_S(B, l)$  and  $\tilde{X}$  over  $F$  partition into  $l \times l$  matrices over the ring  $\mathfrak{M}_n(F)$  of  $n \times n$  matrices over  $F$ . The  $l \times l$  partitioned matrix

$J_S(A, l)\tilde{X}$  has  $(k, k)$ -entry  $AX$  for  $1 \leq k \leq l$ ,  $(k, k+1)$ -entry  $X$  for  $1 \leq k < l$ , all other entries being zero matrices. In the same way the  $l \times l$  partitioned matrix  $\tilde{X}J_S(B, l)$  has  $(k, k)$ -entry  $XB$  for  $1 \leq k \leq l$ ,  $(k, k+1)$ -entry  $X$  for  $1 \leq k < l$ , all other entries being zero matrices. As  $AX = XB$  we deduce  $J_S(A, l)\tilde{X} = \tilde{X}J_S(B, l)$ .

As  $\det \tilde{X} = (\det X)^l \neq 0$  we conclude  $J_S(A, l) \sim J_S(B, l)$  as  $\tilde{X}$  is invertible over  $F$ .

As  $\gcd\{f(x), g(x)\} = 1$  we know  $C(f(x)g(x)) \sim C(f(x)) \oplus C(g(x))$  by (5.31). So  $J_S(C(f(x)g(x)), l) \sim J_S(C(f(x)) \oplus C(g(x)), l)$  using the preceding paragraph. By the first part of (b) above

$$J_S(C(f(x)) \oplus C(g(x)), l) \sim J_S(C(f(x)), l) \oplus J_S(C(g(x)), l).$$

Therefore

$$J_S(C(f(x)g(x)), l) \sim J_S(C(f(x)), l) \oplus J_S(C(g(x)), l) \text{ for } \gcd\{f(x), g(x)\} = 1.$$

### Solution 9

(a) The composition  $\alpha\beta$  of the additive mappings  $\alpha$  and  $\beta$  is itself additive by Exercises 2.1, Question 4(d). As  $\alpha$  and  $\beta$  are semi-linear there are  $\theta, \varphi \in \text{Aut } R$  with  $(av)\alpha = (a)\theta(v)\alpha$  and  $(bw)\beta = (b)\varphi(w)\beta$  for all  $a, b \in R$  and  $v, w \in M$ . Setting  $b = (a)\theta$ ,  $w = (v)\alpha$  gives

$$(av)\alpha\beta = ((a)\theta(v)\alpha)\beta = (bw)\beta = (b)\varphi(w)\beta = (a)\theta\varphi(v)\alpha\beta$$

for all  $a \in R, v \in M$  showing that  $\alpha\beta$  is semilinear as  $\theta\varphi \in \text{Aut } R$  by

Exercises 2.3, Question 3(d).

Let  $\alpha$  be bijective. Then  $\alpha^{-1}$  is additive by Exercises 2.1, Question 4(d). For  $a \in R, v \in M$  we see  $av \in M$  and so there is  $w \in M$  with  $(w)\alpha = av$ . By

Exercises 2.3, Question 3(d) we know  $\theta^{-1} \in \text{Aut } R$ . Also

$$((a)\theta^{-1}(v)\alpha^{-1})\alpha = ((a)\theta^{-1}\theta)((v)\alpha^{-1}\alpha) = av = (w)\alpha.$$

Therefore

$$(av)\alpha^{-1} = w = (a)\theta^{-1}(v)\alpha^{-1} \text{ showing that } \alpha^{-1} \text{ is semi-linear.}$$

Denote two elements of  $R^t$  by  $v = (a_1, a_2, \dots, a_t)$  and  $w = (b_1, b_2, \dots, b_t)$ . As  $\theta$  is additive we have  $(a_i + b_i)\theta = (a_i)\theta + (b_i)\theta$  for  $1 \leq i \leq t$  showing that the  $i$ th entries in  $(v+w)\hat{\theta}$  and  $(v)\hat{\theta} + (w)\hat{\theta}$  are equal. So  $(v+w)\hat{\theta} = (v)\hat{\theta} + (w)\hat{\theta}$  for all  $v, w \in R^t$  showing  $\hat{\theta}$  to be additive. For  $a \in R$  we have  $(aa_i)\theta = (a)\theta(a_i)\theta$  showing that the

$i$ th entries in  $(av)\hat{\theta}$  and  $(a)\theta(v)\hat{\theta}$  are equal for  $1 \leq i \leq t$ . Therefore

$(av)\hat{\theta} = (a)\theta(v)\hat{\theta}$  for all  $a \in R, v \in R^t$  showing  $\hat{\theta}$  semi-linear.

(b) Applying the eros  $r_i - r_{l+i}$  for  $1 \leq i \leq l$  over  $E$  to  $Y$  followed by  $(c - c')^{-1}r_i$  for

$1 \leq i \leq l$  and finally  $r_{l+i} - c'r_i$  for  $1 \leq i \leq l$  produces the matrix  $Z'_S$  with rows

$e_i Z'_S = v_{1i}, e_{l+i} Z'_S = v_{0i}$  for  $1 \leq i \leq l$ . Comparing determinants gives

$\det Y = (c - c')^l \det Z'_S$ . For  $i = 1, 2, \dots, l$  in turn we apply the  $l - i + 1$  eros  $r_j \leftrightarrow r_{j-1}$  to  $Z'_S$  for  $j = l + i, l + i - 1, \dots, 2i$  which produces  $Z_S$ . Therefore

$\det Z'_S = (-1)^{l(l+1)/2} \det Z_S$  as each of these  $l + l - 1 + l - 2 + \dots + 2 + 1 = l(l+1)/2$  eros gives a sign change in the determinant. So  $\det Y = (-1)^{l(l+1)/2} (c - c')^l \det Z_S$ .

(c) The quadratic polynomial  $p_0(x) = x^2 - 2x + 2$  is irreducible over  $\mathbb{R}$  as it has no real zeros. As  $\chi(\mathbb{R}) = 0$  and  $p_0(x)$  is irreducible over  $\mathbb{R}$  we see that  $p_0(x)$  is separable over  $\mathbb{R}$ . The zeros in  $\mathbb{C}$  of  $p_0(x)$  are  $c = 1 + i$  and  $c' = 1 - i$ . The element

$e_1$  of the  $\mathbb{R}[x]$ -module  $M = M(C(p_0(x)^3))$  has order  $p_0(x)^3$  and so

$w = (x - c')^3 e_1 = (x + c - 2)^3 e_1$  has order  $(x - c)^3$  in  $M$ . Also

$$w = ((c - 2)^3, 3(c - 2)^2, 3(c - 2), 1, 0, 0) = (2c, -6c + 6, 3c - 6, 1, 0, 0)$$

using  $c^2 = 2c - 2$ . So

$$\begin{pmatrix} w \\ xw \\ x^2w \end{pmatrix} = \begin{pmatrix} 2c & -6c + 6 & 3c - 6 & 1 & 0 & 0 \\ 0 & 2c & -6c + 6 & 3c - 6 & 1 & 0 \\ 0 & 0 & 2c & -6c + 6 & 3c - 6 & 1 \end{pmatrix}$$

and

$$\begin{pmatrix} w \\ (x - c)w \\ (x - c)^2w \end{pmatrix} = \begin{pmatrix} 2c & -6c + 6 & 3c - 6 & 1 & 0 & 0 \\ -4c + 4 & 8c - 12 & -6c + 12 & 2c - 6 & 1 & 0 \\ 4c - 8 & -8c + 20 & 8c - 24 & -4c + 16 & c - 6 & 1 \end{pmatrix} = \begin{pmatrix} v_{10} + cv_{11} \\ v_{20} + cv_{21} \\ v_{30} + cv_{31} \end{pmatrix}$$

on using  $c^2 = 2c - 2$  again. Therefore

$$Z_S = \begin{pmatrix} v_{10} \\ v_{11} \\ v_{20} \\ v_{21} \\ v_{30} \\ v_{31} \end{pmatrix} = \begin{pmatrix} 0 & 6 & -6 & 1 & 0 & 0 \\ 2 & -6 & 3 & 0 & 0 & 0 \\ \hline 4 & -12 & 12 & -6 & 1 & 0 \\ -4 & 8 & -6 & 2 & 0 & 0 \\ \hline -8 & 20 & -24 & 16 & -6 & 1 \\ 4 & -8 & 8 & -4 & 1 & 0 \end{pmatrix}.$$

Now

$$C(p_0(x)^3) = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ -8 & 24 & -36 & 32 & -18 & 6 \end{pmatrix},$$

$$J_S(C(p_0(x))^T, 3) = \left( \begin{array}{cc|cc|cc} 0 & -2 & 1 & 0 & 0 & 0 \\ 1 & 2 & 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & -2 & 1 & 0 \\ 0 & 0 & 1 & 2 & 0 & 1 \\ \hline 0 & 0 & 0 & 0 & 0 & -2 \\ 0 & 0 & 0 & 0 & 1 & 2 \end{array} \right)$$

and so

$$Z_S C(p_0(x)^3) = \left( \begin{array}{cccccc} 0 & 0 & 6 & -6 & 1 & 0 \\ 0 & 2 & -6 & 3 & 0 & 0 \\ 0 & 4 & -12 & 12 & -6 & 3 \\ 0 & -4 & 8 & -6 & 2 & 0 \\ -8 & 16 & -16 & 8 & -2 & 0 \\ 0 & 4 & -8 & 8 & -4 & 1 \end{array} \right) = J_S(C(p_0(x))^T, 3) Z_S.$$

A calculation shows  $\det Z_S = -64 \neq 0$  and so  $Z_S$  is invertible over  $\mathbb{Q}$ . Therefore

$Z_S C(p_0(x)^3) Z_S^{-1} = J_S(C(p_0(x))^T, 3)$ . The rotation matrix  $R_{\pi/4} = (1/\sqrt{2}) \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$  is

such that  $C(x^2 - 2x + 2)^T$  is similar to  $\sqrt{2} R_{\pi/4}$  as  $Z_0 = \begin{pmatrix} 1 & 2 \\ 1 & 0 \end{pmatrix}$  is invertible over  $\mathbb{Q}$

and satisfies  $Z_0 \begin{pmatrix} 0 & -2 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} Z_0$ . Also

$$J_S(\sqrt{2} R_{\pi/4}, 3) = \left( \begin{array}{cc|cc|cc} 1 & 1 & 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & -1 & 1 & 0 & 1 \\ \hline 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & -1 & 1 \end{array} \right)$$

is in real Jordan form (6.24). Let

$$Z = (Z_0 \oplus Z_0 \oplus Z_0) Z_S = \left( \begin{array}{cccccc} 4 & -6 & 0 & 1 & 0 & 0 \\ 0 & 6 & -6 & 1 & 0 & 0 \\ \hline -4 & 4 & 0 & -2 & 1 & 0 \\ 4 & -12 & 12 & -6 & 1 & 0 \\ \hline 0 & 4 & -8 & 8 & -4 & 1 \\ -8 & 20 & -24 & 16 & -6 & 1 \end{array} \right).$$

Then  $\det Z = 512$  and so  $Z$  is invertible over  $\mathbb{Q}$  and over  $\mathbb{R}$ . Further

$$ZC((x^2 - 2x + 2)^3) = \left( \begin{array}{cccccc} 0 & 4 & -6 & 0 & 1 & 0 \\ 0 & 0 & 6 & -6 & 1 & 0 \\ \hline 0 & -4 & 4 & 0 & -2 & 1 \\ 0 & 4 & -12 & 12 & -6 & 1 \\ \hline -8 & 24 & -32 & 24 & -10 & 2 \\ -8 & 16 & -16 & 8 & -2 & 0 \end{array} \right) = J_S(\sqrt{2} R_{\pi/4}, 3) Z$$

and so  $ZC((x^2 - 2x + 2)^3) Z^{-1} = J_S(\sqrt{2} R_{\pi/4}, 3)$ .

### Solutions 6.3 (page 332)

#### Solution 1

(a)

$$A^2 = \begin{pmatrix} 1 & 1 & -2 & 2 \\ -5 & -2 & 6 & -4 \\ 2 & 1 & -4 & 3 \\ 3 & 1 & -5 & 3 \end{pmatrix} \begin{pmatrix} 1 & 1 & -2 & 2 \\ -5 & -2 & 6 & -4 \\ 2 & 1 & -4 & 3 \\ 3 & 1 & -5 & 3 \end{pmatrix} = \begin{pmatrix} -2 & -1 & 2 & -2 \\ 5 & 1 & -6 & 4 \\ -2 & -1 & 3 & -3 \\ -3 & -1 & 5 & -4 \end{pmatrix} = -A - I.$$

So  $\mu_A(x)$  is a monic and non-constant divisor of  $x^2 + x + 1$  by (6.11) since

$A^2 + A + I = 0$ . As  $x^2 + x + 1$  is irreducible over  $\mathbb{Q}$  we deduce  $\mu_A(x) = x^2 + x + 1$ .

Also  $\chi_A(x)$  is of degree 4 and is a power of  $x^2 + x + 1$  by (6.11). Therefore

$\chi_A(x) = (x^2 + x + 1)^2$ . The invariant factors of  $A$  are  $x^2 + x + 1, x^2 + x + 1$ . Construct

$X$  with  $XAX^{-1}$  in rcf by taking  $e_1 = (1, 0, 0, 0)$  as its first row. Then its second row must be  $e_1 A = (1, 1, -2, 2)$ . Any vector in  $\mathbb{Q}^4$  but not in the 2-dimensional subspace  $\langle e_1, e_1 A \rangle$  can be taken as the third row of  $X$ , for example  $e_2 = (0, 1, 0, 0)$ . The fourth row of  $X$  is then  $e_2 A = (-5, -2, 6, -4)$ . So

$$X = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & -2 & 2 \\ 0 & 1 & 0 & 0 \\ -5 & -2 & 6 & -4 \end{pmatrix}$$

is invertible over  $\mathbb{Q}$  as

$$\det X = - \begin{vmatrix} -2 & 2 \\ 6 & -4 \end{vmatrix} = 4.$$

Also  $X$  satisfies  $XAX^{-1} = C(x^2 + x + 1) \oplus C(x^2 + x + 1)$ .

(b)

Consider  $B' = \begin{pmatrix} a_0 + ia_1 & b_0 + ib_1 \\ c_0 + ic_1 & d_0 + id_1 \end{pmatrix}$  and  $B'' = \begin{pmatrix} a'_0 + ia'_1 & b'_0 + ib'_1 \\ c'_0 + ic'_1 & d'_0 + id'_1 \end{pmatrix}$  in  $\mathfrak{M}_2(\mathbb{Q}(i))$ .

Then

$$\begin{aligned} (B' + B'')\varphi &= \begin{pmatrix} a_0 + a'_0 + i(a_1 + a'_1) & b_0 + b'_0 + i(b_1 + b'_1) \\ c_0 + c'_0 + i(c_1 + c'_1) & d_0 + d'_0 + i(d_1 + d'_1) \end{pmatrix} \varphi = \\ &= \begin{pmatrix} a_0 + a'_0 & a_1 + a'_1 & b_0 + b'_0 & b_1 + b'_1 \\ -a_1 - a'_1 & a_0 + a'_0 & -b_1 - b'_1 & b_0 + b'_0 \\ c_0 + c'_0 & c_1 + c'_1 & d_0 + d'_0 & d_1 + d'_1 \\ -c_1 - c'_1 & c_0 + c'_0 & -d_1 - d'_1 & d_0 + d'_0 \end{pmatrix} \begin{pmatrix} a_0 & a_1 & b_0 & b_1 \\ -a_1 & a_0 & -b_1 & b_0 \\ c_0 & c_1 & d_0 & d_1 \\ -c_1 & c_0 & -d_1 & d_0 \end{pmatrix} + \begin{pmatrix} a'_0 & a'_1 & b'_0 & b'_1 \\ -a'_1 & a'_0 & -b'_1 & b'_0 \\ c'_0 & c'_1 & d'_0 & d'_1 \\ -c'_1 & c'_0 & -d'_1 & d'_0 \end{pmatrix} \begin{pmatrix} a'_0 & a'_1 & b'_0 & b'_1 \\ -a'_1 & a'_0 & -b'_1 & b'_0 \\ c'_0 & c'_1 & d'_0 & d'_1 \\ -c'_1 & c'_0 & -d'_1 & d'_0 \end{pmatrix} = \\ &= (B')\varphi + (B'')\varphi \end{aligned}$$

showing  $\varphi$  to be additive. Also

$$\begin{aligned}
(B'B'')\varphi &= \begin{pmatrix} a_0a'_0 - a_1a'_1 + b_0c'_0 - b_1c'_1 + & a_0b'_0 - a_1b'_1 + b_0d'_0 - b_1d'_1 + \\ i(a_0a'_1 + a_1a'_0 + b_0c'_1 + b_1c'_0) & i(a_0b'_1 + a_1b'_0 + b_0d'_1 + b_1d'_0) \\ c_0a'_0 - c_1a'_1 + d_0c'_0 - d_1c'_1 + & c_0b'_0 - c_1b'_1 + d_0d'_0 - d_1d'_1 + \\ i(c_0a'_1 + c_1a'_0 + d_0c'_1 + d_1c'_0) & i(c_0b'_1 + c_1b'_0 + d_0d'_1 + d_1d'_0) \end{pmatrix} \varphi = \\
&= \begin{pmatrix} a_0a'_0 - a_1a'_1 + & a_0a'_1 + a_1a'_0 + & a_0b'_0 - a_1b'_1 + & a_0b'_1 + a_1b'_0 + \\ b_0c'_0 - b_1c'_1 & b_0c'_1 + b_1c'_0 & b_0d'_0 - b_1d'_1 & b_0d'_1 + b_1d'_0 \\ -a_0a'_1 - a_1a'_0 & a_0a'_0 - a_1a'_1 + & -a_0b'_1 - a_1b'_0 & a_0b'_0 - a_1b'_1 + \\ -b_0c'_1 - b_1c'_0 & b_0c'_0 - b_1c'_1 & -b_0d'_1 - b_1d'_0 & b_0d'_0 - b_1d'_1 \\ c_0a'_0 - c_1a'_1 + & c_0a'_1 + c_1a'_0 + & c_0b'_0 - c_1b'_1 + & c_0b'_1 + c_1b'_0 + \\ d_0c'_0 - d_1c'_1 & d_0c'_1 + d_1c'_0 & d_0d'_0 - d_1d'_1 & d_0d'_1 + d_1d'_0 \\ -c_0a'_1 - c_1a'_0 & c_0a'_0 - c_1a'_1 + & -c_0b'_1 - c_1b'_0 & c_0b'_0 - c_1b'_1 + \\ -d_0c'_1 - d_1c'_0 & d_0c'_0 - d_1c'_1 & -d_0d'_1 - d_1d'_0 & d_0d'_0 - d_1d'_1 \end{pmatrix} = \\
&= \begin{pmatrix} a_0 & a_1 & b_0 & b_1 \\ -a_1 & a_0 & -b_1 & b_0 \\ c_0 & c_1 & d_0 & d_1 \\ -c_1 & c_0 & -d_1 & d_0 \end{pmatrix} \begin{pmatrix} a'_0 & a'_1 & b'_0 & b'_1 \\ -a'_1 & a'_0 & -b'_1 & b'_0 \\ c'_0 & c'_1 & d'_0 & d'_1 \\ -c'_1 & c'_0 & -d'_1 & d'_0 \end{pmatrix} = (B')\varphi(B'')\varphi
\end{aligned}$$

showing  $\varphi$  to be multiplicative. As  $(I_2)\varphi = I_4$ , i.e. the image of the  $2 \times 2$  identity matrix over  $\mathbb{Q}$  by  $\varphi$  is the  $4 \times 4$  identity matrix over  $\mathbb{Q}$ , and so  $\varphi$  respects

1-elements. So  $\varphi: \mathfrak{M}_2(\mathbb{Q}(i)) \rightarrow \mathfrak{M}_4(\mathbb{Q})$  is a ring homomorphism.

Suppose  $B' \in \ker \varphi$ . Then  $(B')\varphi = 0$ , the  $4 \times 4$  zero matrix over  $\mathbb{Q}$ . On comparing entries we obtain  $B' = 0$ , the  $2 \times 2$  zero matrix over  $\mathbb{Q}$ , i.e.  $\ker \varphi = 0$ .

Consider an arbitrary matrix

$$B = \begin{pmatrix} a_0 & a_1 & b_0 & b_1 \\ a_2 & a_3 & b_2 & b_3 \\ c_0 & c_1 & d_0 & d_1 \\ c_2 & c_3 & d_2 & d_3 \end{pmatrix}$$

in  $\mathfrak{M}_4(\mathbb{Q})$ . Suppose  $B \in Z(A)$  where  $A = C(x^2 + 1) \oplus C(x^2 + 1)$ . So  $BA = AB$ , i.e.

$$\begin{pmatrix} -a_1 & a_0 & -b_1 & b_0 \\ -a_3 & a_2 & -b_3 & b_2 \\ -c_1 & c_0 & -d_1 & d_0 \\ -c_3 & c_2 & -d_3 & d_2 \end{pmatrix} = \begin{pmatrix} a_2 & a_3 & b_2 & b_3 \\ -a_0 & -a_1 & -b_0 & -b_1 \\ c_2 & c_3 & d_2 & d_3 \\ -c_0 & -c_1 & -d_0 & -d_1 \end{pmatrix}$$

giving  $a_2 = -a_1, a_3 = a_0, b_2 = -b_1, b_3 = b_0, c_2 = -c_1, c_3 = c_0, d_2 = -d_1, d_3 = d_0$  and so

$$B = \begin{pmatrix} a_0 & a_1 & b_0 & b_1 \\ -a_1 & a_0 & -b_1 & b_0 \\ c_0 & c_1 & d_0 & d_1 \\ -c_1 & c_0 & -d_1 & d_0 \end{pmatrix}$$

showing  $B \in \text{im } \varphi$ . Therefore  $Z(A) \subseteq \text{im } \varphi$ . Conversely suppose  $B \in \text{im } \varphi$ . The preceding steps can be reversed to give  $B \in Z(A)$ . So  $\text{im } \varphi \subseteq Z(A)$  and therefore  $Z(A) = \text{im } \varphi$ .

For  $B' = \begin{pmatrix} a_0 + ia_1 & b_0 + ib_1 \\ c_0 + ic_1 & d_0 + id_1 \end{pmatrix}$  we see

$$\det B' = a_0 d_0 - a_1 d_1 - b_0 c_0 + b_1 c_1 + i(a_0 d_1 + a_1 d_0 - b_0 c_1 - b_1 c_0)$$

and so  $|\det B'|^2 = (a_0 d_0 - a_1 d_1 - b_0 c_0 + b_1 c_1)^2 + (a_0 d_1 + a_1 d_0 - b_0 c_1 - b_1 c_0)^2$ . On multiplying out the squares  $16 + 16 = 32$  terms are obtained, but 8 of these cancel ( $-2a_0 d_0 a_1 d_1 + 2a_0 d_1 a_1 d_0 = 0$  and  $-2b_0 c_0 b_1 c_1 + 2b_0 c_1 b_1 c_0 = 0$ ) to leave the 24 terms

$$(a_0^2 + a_1^2)(d_0^2 + d_1^2) + (b_0^2 + b_1^2)(c_0^2 + c_1^2) +$$

$$2a_0(b_1 c_1 d_0 - b_0 c_0 d_0 - b_0 c_1 d_1 - b_1 c_0 d_1) + 2a_1(b_0 c_0 d_1 - b_1 c_1 d_1 - b_1 c_0 d_0 - b_0 c_1 d_0)$$

of  $\det(B')\varphi$ . Therefore  $\det(B')\varphi = |\det B'|^2$ . There must be a better way!

(c) As  $x^2 + 1$  is irreducible over  $\mathbb{R}$  we see that  $t$  is even and  $A$  has  $t/2$  invariant factors  $x^2 + 1$  by (6.26). Write  $B = r(A) = aA + bI$ . Then

$$B^2 - 2bB + (a^2 + b^2)I = (aA + bI)^2 - 2b(aA + bI) + (a^2 + b^2)I =$$

$$a^2 A^2 + 2abA + b^2 I - 2abA - 2b^2 I + a^2 I + b^2 I = a^2(A^2 + I) = 0.$$

The polynomial  $x^2 - 2bx + a^2 + b^2$  is irreducible over  $\mathbb{R}$  as  $(-2b)^2 < 4(a^2 + b^2)$  since  $a \neq 0$  and so the above calculation shows  $\mu_B(x) = x^2 - 2bx + a^2 + b^2$ . Applying (6.26) with  $B$  in place of  $A$ , we see that  $B$  has  $t/2$  invariant factors

$\mu_B(x) = x^2 - 2bx + a^2 + b^2$ . In the case  $a = 0$  the scalar matrix  $r(A) = bI$  has  $t$  invariant factors  $x - b$ .

By (4.1) there are polynomials  $q(x)$  and  $r(x) = ax + b$  over  $\mathbb{R}$  with

$$f(x) = q(x)(x^2 + 1) + r(x).$$

Then  $f(A) = q(A)(A^2 + I) + r(A) = q(A) \times 0 + r(A) = r(A)$ . So the invariant factors of  $f(A) = r(A)$  are as described in the preceding paragraph.

As  $A^{10} - A^9 = I - A = r_1(A)$  where  $r_1(x) = -x + 1$  we see that  $A^{10} - A^9$  has  $t/2$  invariant factors  $x^2 - 2x + 2$ . As  $A^{11} + A^{10} = A + I = r_2(A)$  where  $r_2(x) = x + 1$  we see that  $A^{11} + A^{10}$  also has  $t/2$  invariant factors  $x^2 - 2x + 2$ . So  $A^{10} - A^9$  and  $A^{11} + A^{10}$  are similar.

(d) The invertible  $t \times t$  matrices  $X$  over  $F$  with  $XAX^{-1}$  in rcf correspond to bases  $v_1, v_2, \dots, v_s$  of the  $s$ -dimensional vector space  $F^t$  over the field  $F(c)$  where  $\mu_A(c) = 0$  by (6.26). In fact  $e_{(j-1)m+1}X = v_j$  for  $1 \leq j \leq s$ , i.e.  $v_j$  is row  $(j-1)m + 1$  of  $X$ . As  $|F(s)| = q^m$  the number of these bases is the product of the  $s$  factors  $(q^{ms} - 1)(q^{ms} - q^m) \cdots (q^{ms} - q^{m(s-1)})$  as explained after (2.18). The size of the similarity class of  $A$  is  $|GL_t(F)|$  divided by the above number by (6.29), i.e.

$$(q^t - 1)(q^t - q) \cdots (q^t - q^{t-1}) / (q^t - 1)(q^t - q^m) \cdots (q^t - q^{t-m}) \text{ as } t = ms.$$

Notice first that  $x^2 + 1$  and  $x^3 - x + 1$  are irreducible over  $\mathbb{Z}_3$  by (4.8)(ii). Taking  $q = 3, t = 2, m = 2, s = 1$  we see that there are  $(3^2 - 1)(3^2 - 3)/(3^2 - 1) = 6$  matrices over  $\mathbb{Z}_3$  similar to  $C(x^2 + 1)$ . Taking  $q = 3, t = 3, m = 3, s = 1$  we see that there are  $(3^3 - 1)(3^3 - 3)(3^3 - 3^3)/(3^3 - 1) = 24 \times 18 = 432$  matrices over  $\mathbb{Z}_3$  similar to  $C(x^3 - x + 1)$ . Taking  $q = 3, t = 4, m = 2, s = 2$  there are

$$(3^4 - 1)(3^4 - 3)(3^4 - 3^2)(3^4 - 3^3)/(3^4 - 1)(3^4 - 3^2) = 78 \times 54 = 4212$$



matrices over  $\mathbb{Z}_3$  similar to  $A = C(x^2 + 1) \oplus C(x^2 + 1)$  as  $\mu_A(x) = x^2 + 1$ .

(e) As  $\mu_A(x)$  is irreducible over  $F$  we see  $\chi_A(x) = \mu_A(x)^s$  by (6.13). Comparing degrees gives  $s = t/m$  where  $m = \deg \mu_A(x)$ . A typical element of the field  $F(c)$ , where  $\mu_A(c) = 0$ , is  $f(c)$  where  $f(x) \in F[x]$ . Also  $f(c) = g(c) \Leftrightarrow f(x) \equiv g(x) \pmod{\mu_A(x)}$  where  $f(x), g(x) \in F[x]$  (see the discussion after (4.9)). So the product of  $f(c)$  and the element  $v$  of the  $F[x]$ -module  $M(A)$  is unambiguously defined by  $f(c)v = f(x)v = vf(A)$ . The seven module laws listed before (2.19) hold in  $M(A)$  and immediately give rise to the seven laws of a vector space over  $F(c)$ , i.e.  $F^t$  has the structure of a vector space over  $F(c)$  which we denote by  $M(A)'$ . Then  $\dim M(A)' = s'$  where  $s' \leq t$  as  $F(c)$  is an extension field of  $F$ . Let  $v_1, v_2, \dots, v_{s'}$  be a basis of  $M(A)'$ . The  $ms'$  vectors  $x^{i-1}v_j$  for  $1 \leq i \leq m, 1 \leq j \leq s'$  form a basis of  $F^t$  and so  $ms' = t = ms$ . Therefore

$$\dim M(A)' = s = t/m.$$

Let  $N$  be a submodule of  $M(A)$ . Then  $f(x)v \in N$  where  $f(x) \in F[x], v \in N$ . So  $f(c)v \in N$  where  $f(c) \in F(c), v \in N$ . Therefore  $N$  is a subspace of  $M(A)'$ .

Conversely  $N$  a subspace of  $M(A)'$  implies  $N$  a submodule of  $M(A)$ .

Let  $\beta \in \text{End } M(A)$ . Then  $(f(c)v)\beta = (f(x)v)\beta = f(x)((v)\beta) = f(c)((v)\beta)$  showing that  $\beta$  is a linear mapping of  $M(A)'$ . Conversely each linear mapping of  $M(A)'$  belongs to  $\text{End } M(A)$ . Let  $\mathcal{B}$  denote a basis of  $M(A)'$ . For each  $\beta \in \text{End } M(A)$  write  $(\beta)\theta$  for the matrix of the linear mapping  $\beta$  of  $M(A)'$  relative to  $\mathcal{B}$ . Then  $(\beta)\theta \in \mathfrak{M}_s(F(c))$  and, mimicking the proof of (3.15) with  $\mathbb{Z}, t$ , replaced by  $F(c), s$ , we see that  $\theta: \text{End } M(A) \cong \mathfrak{M}_s(F(c))$  is a ring isomorphism. Restricting  $\theta$  to  $\text{Aut } M(A)$ , the group of invertible elements of  $\text{End } M(A)$ , produces the group isomorphism  $\theta|: \text{Aut } M(A) \cong GL_s(F(c))$ .

## Solution 2

(a) Let  $B \in Z(A)$  where  $A = C(x^3)$  over  $F$ . By (6.27) there is

$f(x) = a_0 + a_1x + a_2x^2 \in F[x]$  with

$$B = f(A) = a_0I + a_1A + a_2A^2 = \begin{pmatrix} a_0 & a_1 & a_2 \\ 0 & a_0 & a_1 \\ 0 & 0 & a_0 \end{pmatrix}.$$

Yes  $Z(A)$  is a ring. In fact  $Z(A)$  is a commutative subring of  $\mathfrak{M}_3(F)$ . Yes  $Z(A)$  is a vector space over  $F$  with basis  $I, A, A^2$  and so  $\dim Z(A) = 3$ . Also  $B \in U(Z(A)) \Leftrightarrow a_0 \neq 0$  where  $B$  is as above. Yes  $\text{End } M(A)$  and  $Z(A)$  are isomorphic rings (see (b) below). Yes, on comparing the sets of invertible elements in these rings,  $\text{Aut } M(A)$  and  $U(Z(A))$  are isomorphic groups.

In the case  $F = \mathbb{Z}_2$  we have  $|\text{End } M(A)| = 2^3 = 8$  and  $|\text{Aut } M(A)| = 2^2 = 4$ . Now  $(I + A)^2 = I + A^2, (I + A)^3 = I + A + A^2$  showing that  $U(Z(A))$  is cyclic of order 4 with generator  $I + A$ . As  $\text{Aut } M(A) \cong U(Z(A))$  we conclude that  $\text{Aut } M(A)$  is a cyclic group.

In the case  $F = \mathbb{F}_q$  we have  $|\text{End } M(A)| = q^3$  as there are  $q$  choices for each of  $a_0, a_1, a_2$ . Also  $|\text{Aut } M(A)| = (q-1)q^2$  as  $a_0 \neq 0$  and so there are  $q-1$  choices for  $a_0$ ,  $q$  choices for  $a_1, a_2$  as before.

(b) Consider  $\beta, \beta' \in \text{End } M(A)$  with matrices  $B = (b_{ij}), B' = (b'_{ij})$  respectively relative to the standard basis  $\mathcal{B}_0$  of  $F^t$ . So  $(e_i)\beta = \sum_{j=1}^t b_{ij}e_j$  and  $(e_i)\beta' = \sum_{j=1}^t b'_{ij}e_j$  for  $1 \leq i \leq t$ . Adding these equations gives

$$(e_i)(\beta + \beta') = (e_i)\beta + (e_i)\beta' = \sum_{j=1}^t b_{ij}e_j + \sum_{j=1}^t b'_{ij}e_j = \sum_{j=1}^t (b_{ij} + b'_{ij})e_j \text{ for } 1 \leq i \leq t$$

which shows that  $\beta + \beta'$  has matrix  $B + B' = (b_{ij} + b'_{ij})$  relative to  $\mathcal{B}_0$ , i.e.

$$(\beta + \beta')\theta = B + B' = (\beta)\theta + (\beta')\theta. \text{ So } \theta \text{ respects addition. Now } (e_j)\beta' = \sum_{k=1}^t b'_{jk}e_k$$

for  $1 \leq j \leq t$ . So applying  $\beta'$  to  $(e_i)\beta = \sum_{j=1}^t b_{ij}e_j$  gives

$$\begin{aligned} (e_i)(\beta\beta') &= ((e_i)\beta)\beta' = \left(\sum_{j=1}^t b_{ij}e_j\right)\beta' = \\ &= \sum_{j=1}^t b_{ij}(e_j)\beta' = \sum_{j=1}^t b_{ij}\left(\sum_{k=1}^t b'_{jk}e_k\right) = \sum_{k=1}^t \left(\sum_{j=1}^t b_{ij}b'_{jk}\right)e_k \end{aligned}$$

for  $1 \leq j \leq t$  which shows that  $\beta\beta'$  has matrix  $BB' = \left(\sum_{j=1}^t b_{ij}b'_{jk}\right)$  relative to  $\mathcal{B}_0$ , i.e.

$(\beta\beta')\theta = BB' = (\beta)\theta(\beta')\theta$ . So  $\theta$  respects multiplication. The identity mapping  $\iota$  of  $F^t$  is the 1-element of the ring  $\text{End } M(A)$  and  $(\iota)\theta = I$ , i.e. the matrix of  $\iota$  relative to  $\mathcal{B}_0$  is the  $t \times t$  identity matrix  $I$  over  $F$ . As  $I$  is the 1-element of  $\mathfrak{M}_t(F)$  we conclude that  $\theta': \text{End } M(A) \rightarrow \mathfrak{M}_t(F)$ , defined by  $(\beta)\theta' = (\beta)\theta$  for all

$\beta \in \text{End } M(A)$ , is a ring homomorphism.

Let  $\beta \in \ker \theta'$ . Then  $(\beta)\theta = B = 0$ , the zero  $t \times t$  matrix over  $F$ , i.e.  $B = (b_{ij})$

where  $b_{ij} = 0$  for  $1 \leq i, j \leq t$ . This gives  $(e_i)\beta = \sum_{j=1}^t b_{ij}e_j = \sum_{j=1}^t 0e_j = 0$  for  $1 \leq i \leq t$ .

For  $v \in F^t$  we have  $v = \sum_{i=1}^t a_i e_i$  and so  $(v)\beta = \sum_{i=1}^t a_i (e_i)\beta = \sum_{i=1}^t a_i 0 = 0$  showing that

$\beta = 0$ . So  $\ker \theta' = 0$  showing  $\theta'$  is injective by Exercises 2.3, Question 1(a)(i).

By (6.27) and (6.28) we see  $\text{im } \theta' = Z(A)$ . So  $Z(A)$  is a subring of  $\mathfrak{M}_t(F)$  by

Exercises 2.3, Question 3(b). Write  $\theta = \theta'\iota'$  where  $\iota': Z(A) \rightarrow \mathfrak{M}_t(F)$  is the

inclusion. As  $\iota'$  is an injective ring homomorphism we see that  $\theta: \text{End } M(A) \rightarrow Z(A)$

is a ring isomorphism, i.e.  $\theta: \text{End } M(A) \cong Z(A)$ . Now  $\text{Aut } M(A) = U(\text{End } M(A))$ ,

i.e. the automorphisms of the  $F[x]$ -module  $M(A)$  are exactly the invertible elements

of the ring  $\text{End } M(A)$ . Therefore  $\theta|: \text{Aut } M(A) \cong U(Z(A))$ , i.e.  $\theta|$  is a group

isomorphism (Exercises 2.3, Question 4(c)) between the corresponding groups of

invertible elements of these rings.

(c) Each element of the ring  $F[x]/\langle g(x) \rangle$  is uniquely expressible  $\langle g(x) \rangle + f(x)$  where  $\deg f(x) < \deg g(x)$ . Also  $\langle g(x) \rangle + f(x)$  is an invertible element of  $F[x]/\langle g(x) \rangle$  if

and only if  $\gcd\{f(x), g(x)\} = 1$ . Therefore in the case of a finite field  $F$  of order  $q$  we have  $\Phi_q(g(x)) = |U(F[x]/\langle g(x) \rangle)|$ , i.e. the number of polynomials  $f(x)$  over  $F$  with  $\gcd\{f(x), g(x)\} = 1$  and  $\deg f(x) < \deg g(x)$  is the order of the multiplicative group  $U(F[x]/\langle g(x) \rangle)$ .

For non-zero  $g(x), h(x) \in F[x]$  with  $\gcd\{g(x), h(x)\} = 1$  we have

$$\begin{aligned} \Phi_q(g(x)h(x)) &= |U(F[x]/\langle g(x)h(x) \rangle)| = |U(F[x]/\langle g(x) \rangle) \times U(F[x]/\langle h(x) \rangle)| = \\ &= |U(F[x]/\langle g(x) \rangle)| \times |U(F[x]/\langle h(x) \rangle)| = \Phi_q(g(x)) \Phi_q(h(x)) \end{aligned}$$

showing that  $\Phi_q$  has the multiplicative property.

As  $\deg p(x)^n = mn$  there are  $q^{mn}$  polynomials  $f(x)$  over  $F$  with  $\deg f(x) < \deg p(x)^n$  as there are  $q$  choices for each of the  $mn$  coefficients of  $x^i$  in  $f(x)$  for  $0 \leq i < mn$ . Suppose  $\gcd\{f(x), p(x)^n\} \neq 1$ . Then  $p(x) \mid \gcd\{f(x), p(x)^n\}$  as  $p(x)$  is irreducible over  $F$ . So  $p(x) \mid f(x)$  and so  $f(x)/p(x)$  is a polynomial of degree less than  $mn - m = m(n-1)$  over  $F$ . There are  $q^{m(n-1)}$  such polynomials over  $F$  and hence there are exactly  $q^{m(n-1)}$  polynomials  $f(x)$  as above with  $\gcd\{f(x), p(x)^n\} \neq 1$ . Therefore  $\Phi_q(p(x)^n) = q^{mn} - q^{m(n-1)} = q^{m(n-1)}(q^m - 1)$ .

(d) In each case  $|\text{End } M(A)| = 2^6 = 64$  as  $I, A, A^2, A^3, A^4, A^5$  is a basis of the 6-dimensional vector space  $Z(A)$  over  $\mathbb{Z}_2$  and  $\text{End } M(A) \cong Z(A)$  by (b) above.

(i) Taking  $p(x) = x$ ,  $m = 1$ ,  $n = 6$ ,  $q = 2$  we have  $\chi_A(x) = x^6$  and so  $\Phi_2(x^6) = 2^6 - 2^5 = 32$ , i.e.  $|\text{Aut } M(A)| = 2^5 = 32$ . By (6.29) there are exactly

$$\begin{aligned} |GL_6(\mathbb{Z}_2)| / |\text{Aut } M(A)| &= \\ (2^6 - 1)(2^6 - 2)(2^6 - 2^2)(2^6 - 2^3)(2^6 - 2^4)(2^6 - 2^5) / 2^5 &= 629959680 \end{aligned}$$

matrices over  $\mathbb{Z}_2$  similar to  $A = C(x^6)$ .

(ii) Taking  $p(x) = x^2 + x + 1$  which is irreducible over  $\mathbb{Z}_2$ ,  $m = 2$ ,  $n = 3$ ,  $q = 2$  we have  $\Phi_2(\chi_A(x)) = \Phi_2((x^2 + x + 1)^3) = 2^{2 \times 3} - 2^{2 \times 2} = 48$ . So  $|\text{Aut } M(A)| = 48$ . By (6.29) there are exactly

$$\begin{aligned} |GL_6(\mathbb{Z}_2)| / |\text{Aut } M(A)| &= \\ (2^6 - 1)(2^6 - 2)(2^6 - 2^2)(2^6 - 2^3)(2^6 - 2^4)(2^6 - 2^5) / (2^6 - 2^4) &= 419973120 \end{aligned}$$

matrices over  $\mathbb{Z}_2$  similar to  $A = C((x^2 + x + 1)^3)$ .

(iii) As  $\Phi_2(x^2) = 2^2 - 2 = 2$  and  $\Phi_2((x^2 + x + 1)^2) = 2^4 - 2^2 = 12$  we see  $\Phi_2(x^2(x^2 + x + 1)^2) = 2 \times 12 = 24$ . So  $|\text{Aut } M(A)| = 24$  and the number of  $6 \times 6$  matrices over  $\mathbb{Z}_2$  which are similar to  $A = C(x^2(x^2 + x + 1)^2)$  is

$$\begin{aligned} |GL_6(\mathbb{Z}_2)| / |\text{Aut } M(A)| &= \\ (2^6 - 1)(2^6 - 2)(2^6 - 2^2)(2^6 - 2^3)(2^6 - 2^4)(2^6 - 2^5) / (2^5 - 2^3) &= 839946240. \end{aligned}$$

(iv) As  $x^6 + x^2 = x^2(x^4 + 1) = x^2(x + 1)^4$  over  $\mathbb{Z}_2$  we see

$$|\text{Aut } M(A)| = \Phi_2(x^6 + x^2) = \Phi_2(x^2) \Phi_2((x + 1)^4) = (2^2 - 2)(2^4 - 2^3) = 2^4 = 16.$$

So the number of  $6 \times 6$  matrices over  $\mathbb{Z}_2$  which are similar to  $A = C(x^6 + x^2)$  is

$$|GL_6(\mathbb{Z}_2)|/|\text{Aut } M(A)| = \\ (2^6 - 1)(2^6 - 2)(2^6 - 2^2)(2^6 - 2^3)(2^6 - 2^4)(2^6 - 2^5)/2^4 = 1259919360.$$

(e) Each of the  $q$  scalar matrices  $aI$  for  $a \in \mathbb{F}_q$  belongs to a singleton similarity class  $\{aI\}$ . Let  $A$  be a non-scalar  $2 \times 2$  matrix over  $\mathbb{F}_q$ . Then  $M(A)$  is cyclic with quadratic minimum polynomial  $\mu_A(x)$  by Exercises 6.1, Question 6(a). There are  $q$  polynomials  $\mu_A(x) = (x-a)^2$  giving  $\Phi_q((x-a)^2) = q^2 - q$ , and so the similarity class of  $A$  has size  $(q^2 - 1)(q^2 - q)/(q^2 - q) = q^2 - 1$  by (6.29). There are  $q(q-1)/2$  polynomials  $\mu_A(x) = (x-a)(x-b)$ ,  $a \neq b$ , giving  $\Phi_q((x-a)(x-b)) = (q-1)^2$  and so the size of the similarity class of  $A$  is  $(q^2 - 1)(q^2 - q)/(q-1)^2 = (q+1)q$ . There remain  $q^2 - q - q(q-1)/2 = q(q-1)/2$  monic quadratics which are the irreducible  $\mu_A(x)$  and so  $\Phi_q(\mu_A(x)) = q^2 - 1$  and the size of the similarity class of  $A$  is  $(q^2 - 1)(q^2 - q)/(q^2 - 1) = q(q-1)$ . Adding up the number of matrices in  $\mathfrak{M}_2(\mathbb{F}_q)$  according to their similarity classes gives

$$q + q(q^2 - 1) + (q(q-1)/2)(q+1)q + (q(q-1)/2)q(q-1) = q + q^3 - q + q^4 - q^3 = q^4$$

as expected since  $|\mathfrak{M}_2(\mathbb{F}_q)| = q^4$ . There are  $q-1$  scalar matrices  $aI$  in  $GL_2(\mathbb{F}_q)$  namely those with  $a \neq 0$ . The  $q-1$  similarity classes of matrices  $A$  with  $\mu_A(x) = (x-a)^2$ ,  $a \neq 0$  are contained in  $GL_2(\mathbb{F}_q)$ . The  $(q-1)(q-2)/2$  similarity classes of matrices  $A$  with  $\mu_A(x) = (x-a)(x-b)$ ,  $a \neq 0, b \neq 0, a \neq b$  are contained in  $GL_2(\mathbb{F}_q)$ . All the  $q(q-1)/2$  similarity classes of matrices  $A$  with  $\mu_A(x)$  irreducible are contained in  $GL_2(\mathbb{F}_q)$  as  $\mu_A(0) \neq 0$ . Adding up the number of matrices in  $GL_2(\mathbb{F}_q)$  according to their  $q^2 - 1$  conjugacy classes gives

$$q - 1 + (q-1)(q^2 - 1) + ((q-1)(q-2)/2)(q+1)q + (q(q-1)/2)q(q-1) = \\ (q^2 - 1)(q^2 - q) = |GL_2(\mathbb{F}_q)|.$$

### Solution 3

(a) Let  $e$  denote the identity element of  $G$ . Then  $(e)\theta$  is the identity mapping  $\iota$  of  $\Omega$ . So  $x^{(e)\theta} = x^\iota = x$  showing  $x \sim x$  for all  $x \in \Omega$ . Suppose  $x \sim y$  for  $x, y \in \Omega$ . There is  $g \in G$  with  $x^{(g)\theta} = y$ . As  $\theta: G \rightarrow S(\Omega)$  is a group homomorphism we see  $((g)\theta)^{-1} = (g^{-1})\theta$  and so  $y^{(g^{-1})\theta} = x$  showing  $y \sim x$  as  $g^{-1} \in G$ . Suppose  $x \sim y$  and  $y \sim z$  where  $x, y, z \in \Omega$ . There are  $g, h \in G$  with  $x^{(g)\theta} = y$  and  $y^{(h)\theta} = z$ . As  $(g)\theta(h)\theta = (gh)\theta$  we see  $x^{(gh)\theta} = x^{(g)\theta(h)\theta} = (x^{(g)\theta})^{(h)\theta} = y^{(h)\theta} = z$  showing  $x \sim z$  as  $gh \in G$ . We conclude that  $\sim$  is an equivalence relation on  $\Omega$ . The equivalence class of  $x$  is

$$O_x = \{y \in \Omega : y \sim x\} = \{y \in \Omega : y = x^{(g)\theta} \text{ for } g \in G\} = \{x^{(g)\theta} : g \in G\}.$$

We verify that  $G_x$  is a subgroup of  $G$  by showing that  $G_x$  contains the identity  $e$  of  $G$  and  $G_x$  is closed under multiplication and inversion. As  $(e)\theta = \iota$  we see

$x^{(e)\theta} = x^\iota = x$  showing  $e \in G_x$ . Consider  $g, h \in G_x$ . Then  $x^{(g)\theta} = x$  and  $x^{(h)\theta} = x$ . As above we see  $x^{(gh)\theta} = x^{(g)\theta(h)\theta} = (x^{(g)\theta})^{(h)\theta} = x^{(h)\theta} = x$  showing  $gh \in G_x$ .

Also  $x^{(g)\theta} = x$  gives  $x^{((g)\theta)^{-1}} = x$  and so  $x^{(g^{-1})\theta} = x$  showing  $g^{-1} \in G_x$ . Therefore  $G_x$  is a subgroup of  $G$ .

To show that the correspondence  $G_x g \rightarrow x^{(g)\theta}$  is unambiguously defined suppose

$G_x g = G_x h$  for  $g, h \in G$ . Then  $gh^{-1} \in G_x$  which means  $x^{(gh^{-1})\theta} = x$ . As  $x^{(g)\theta(h^{-1})\theta} = x^{(g)\theta((h)\theta)^{-1}}$  on applying  $(h)\theta$  we obtain  $x^{(g)\theta} = x^{(h)\theta}$  showing that

the above correspondence from the set of left cosets of  $G_x$  to  $O_x$  is indeed unambiguously defined. This correspondence is surjective directly from the definition of

$O_x$ . This correspondence is injective because  $x^{(g)\theta} = x^{(h)\theta}$  implies  $x^{(gh^{-1})\theta} = x$  which implies  $gh^{-1} \in G_x$  and hence  $G_x g = G_x h$ , on reversing the above steps. So this correspondence is bijective, completing the proof of the orbit-stabilizer theorem.

In the case of  $G_x$  having finite index  $n$  in  $G$  we see  $|O_x| = n$  as there are exactly  $n$  distinct left cosets of  $G_x$  in  $G$ . Should  $G$  be finite we have  $n = |G|/|G_x|$  as the  $n$  left cosets each consist of  $|G_x|$  elements and they partition  $G$ . So  $|G|/|G_x| = |O_x|$ .

Suppose  $\theta$  to be trivial. Then  $O_x = \{x^{(g)\theta} : g \in G\} = \{x^I : g \in G\} = \{x : g \in G\} = \{x\}$  and  $G_x = \{g \in G : x^{(g)\theta} = x\} = \{g \in G : x^I = x\} = \{g \in G : x = x\} = G$ .

Suppose  $G = S(\Omega)$  and let  $\theta : S(\Omega) \rightarrow S(\Omega)$  be the identity mapping. Then

$O_x = \{x^{(g)\theta} : g \in S(\Omega)\} = \{x^g : g \in S(\Omega)\}$ . Taking  $g = I$ , the identity mapping of  $\Omega$ , gives  $x \in O_x$ . For  $y \in \Omega$  with  $y \neq x$  let  $g$  denote the bijection of  $\Omega$  which

interchanges  $x$  and  $y$  and fixes all other elements, i.e.  $x^g = y$ ,  $y^g = x$ ,  $z^g = z$  for all  $z \in \Omega$  with  $z \neq x, z \neq y$ . As  $x^g = y$  we see  $y \in O_x$ . Therefore  $O_x = \Omega$ . Suppose

$\Omega = \{x\}$ . Then  $G_x = G$  is trivial. So suppose  $\Omega \neq \{x\}$  and write

$\Omega_x = \{y \in \Omega : y \neq x\}$ . Then  $G_x = \{g \in S(\Omega) : x^g = x\} \cong S(\Omega_x)$ , i.e. the bijections of  $\Omega$  which fix  $x$  correspond (by restriction) to the bijections of  $\Omega_x$ .

(b) Is  $(X^{-1})\theta : \Omega \rightarrow \Omega$  the inverse of  $(X)\theta : \Omega \rightarrow \Omega$ ? If so then  $(X)\theta$  is a bijection as only bijections have inverses. As  $(X^{-1})^{-1} = X$  we have

$$A^{(X^{-1})\theta(X)\theta} = (A^{(X^{-1})\theta})^{(X)\theta} = (XAX^{-1})^{(X)\theta} = X^{-1}(XAX^{-1})X = (X^{-1}X)A(X^{-1}X) = A$$

for all  $A \in \Omega$ . So  $(X^{-1})\theta(X)\theta = I$  the identity mapping of  $\Omega$ . Interchanging  $X$  and  $X^{-1}$  gives  $(X)\theta(X^{-1})\theta = I$ . So  $(X^{-1})\theta$  is the (two-sided) inverse of  $(X)\theta$  which is therefore bijective. So  $(X)\theta \in S(\Omega)$  for  $X \in G$ . For  $A \in \Omega$  and  $X, Y \in G$  we have

$$\begin{aligned} A^{(XY)\theta} &= (XY)^{-1}A(XY) = Y^{-1}X^{-1}AXY = Y^{-1}(X^{-1}AX)Y = \\ &= (X^{-1}AX)^{(Y)\theta} = (A^{(X)\theta})^{(Y)\theta} = (A)^{(X)\theta(Y)\theta} \end{aligned}$$

showing  $(XY)\theta = (X)\theta(Y)\theta$ , i.e.  $\theta : G \rightarrow S(\Omega)$  is a permutation representation of  $G$  on  $\Omega$ . By (5.4) the similarity class of  $A$  is

$$\{XAX^{-1} : X \in G\} = \{X^{-1}AX : X^{-1} \in G\} = \{X^{-1}AX : X \in G\} = \{A^{(X)\theta} : X \in G\} = O_A.$$

So  $O_A$  is equal to the similarity class of  $A$ . Now

$$\begin{aligned} G_A &= \{X \in G : A^{(X)\theta} = A\} = \{X \in G : X^{-1}AX = A\} = \\ &= \{X \in G : AX = XA\} = \{X \in G : X \in Z(A)\} = G \cap Z(A) \end{aligned}$$

by (6.28). The invertible elements of the ring  $Z(A)$  are precisely the matrices in  $G \cap Z(A)$ , i.e.  $G \cap Z(A) = U(Z(A))$ . So yes  $G_A = U(Z(A))$ .

**Solution 4**

(a) Let  $u, v \in M$  and  $r' \in R$ . As  $\alpha$  and  $\beta$  are additive we have

$(u+v)\alpha = (u)\alpha + (v)\alpha$  and  $(u+v)\beta = (u)\beta + (v)\beta$ . As  $(M', +)$  is commutative

$$\begin{aligned} (u+v)(\alpha + \beta) &= (u+v)\alpha + (u+v)\beta = (u)\alpha + (v)\alpha + (u)\beta + (v)\beta = \\ &= (u)\alpha + (u)\beta + (v)\alpha + (v)\beta = (u)(\alpha + \beta) + (v)(\alpha + \beta) \end{aligned}$$

and so  $\alpha + \beta$  is additive. As  $\alpha$  and  $\beta$  are  $R$ -linear and  $M'$  is an  $R$ -module we obtain

$$(r'v)(\alpha + \beta) = (r'v)\alpha + (r'v)\beta = r'((v)\alpha) + r'((v)\beta) = r'((v)\alpha + (v)\beta) = r'((v)(\alpha + \beta))$$

and so  $\alpha + \beta$  is  $R$ -linear. Treating  $r\alpha: M \rightarrow M'$  in the same way:

$$\begin{aligned} (u+v)(r\alpha) &= r((u+v)\alpha) = r((u)\alpha + (v)\alpha) = r((u)\alpha) + r((v)\alpha) = (u)(r\alpha) + (v)(r\alpha), \\ (r'v)(r\alpha) &= r((r'v)\alpha) = rr'((v)\alpha) = r'r((v)\alpha) = r'((v)(r\alpha)) \end{aligned}$$

using the commutativity of  $R$  and the fact that  $M'$  is an  $R$ -module. Therefore  $r\alpha$  is  $R$ -linear.

Consider  $\alpha \in \text{Hom}(M, M')$ . We show  $-\alpha \in \text{Hom}(M, M')$  where  $(v)(-\alpha) = -(v)\alpha$

for all  $v \in M$ . For  $u, v \in M$  and  $r' \in R$  we have

$$(u+v)(-\alpha) = -(u+v)\alpha = -((u)\alpha + (v)\alpha) = -(u)\alpha - (v)\alpha = (u)(-\alpha) + (v)(-\alpha) \text{ and}$$

$$(r'v)(-\alpha) = -(r'v)\alpha = -r'(v)\alpha = r'(-(v)\alpha) = r'((v)(-\alpha)) \text{ showing that } -\alpha \text{ is}$$

$R$ -linear, i.e.  $-\alpha \in \text{Hom}(M, M')$ . As the zero mapping  $0: M \rightarrow M'$  (which maps all elements of  $M$  to the zero element of  $M'$ ) is in  $\text{Hom}(M, M')$  we see that

$\text{Hom}(M, M')$  is an additive group, being a subgroup of the additive group of

$\text{Hom}((M, +), (M', +))$  by Exercises 3.3, Question 6(a).

We show that laws 5, 6, 7 of an  $R$ -module, stated before (2.19), hold in

$\text{Hom}(M, M')$ . Consider  $\alpha, \beta \in \text{Hom}(M, M')$  and  $r, r_1, r_2 \in R$ . Then

$r(\alpha + \beta) = r\alpha + r\beta$  as law 5 holds in  $M'$  and so

$$\begin{aligned} (v)(r(\alpha + \beta)) &= r((v)(\alpha + \beta)) = r((v)\alpha + (v)\beta) = \\ &= r((v)\alpha) + r((v)\beta) = (v)(r\alpha) + (v)(r\beta) = (v)(r\alpha + r\beta) \end{aligned}$$

for all  $v \in M$ .

Also  $(r_1 + r_2)\alpha = r_1\alpha + r_2\alpha$  as

$$(v)((r_1 + r_2)\alpha) = (r_1 + r_2)((v)\alpha) = r_1((v)\alpha) + r_2((v)\alpha) = (v)(r_1\alpha) + (v)(r_2\alpha) = (v)(r_1\alpha + r_2\alpha)$$

for all  $v \in M$ . Therefore law 5 holds in  $\text{Hom}(M, M')$ .

Now  $(r_1 r_2)\alpha = r_1(r_2\alpha)$  as for all  $v \in M$  we have

$$(v)((r_1 r_2)\alpha) = (r_1 r_2)((v)\alpha) = r_1(r_2((v)\alpha)) = r_1((v)(r_2\alpha)) = (v)(r_1(r_2\alpha))$$

on using law 6 in  $M'$ . So law 6 holds in  $\text{Hom}(M, M')$ . Finally law 7 holds in

$\text{Hom}(M, M')$  as  $1\alpha = \alpha$  since  $(v)(1\alpha) = 1(v)\alpha = (v)\alpha$  on using law 7 in  $M'$ . We conclude that  $\text{Hom}(M, M')$  is an  $R$ -module.

(b) Let  $\iota_1: M_1 \rightarrow M_1 \oplus M_2$  be defined by  $(v_1)\iota_1 = (v_1, 0)$  for all  $v_1 \in M_1$ . Also

let  $\iota_2: M_2 \rightarrow M_1 \oplus M_2$  be defined by  $(v_2)\iota_2 = (0, v_2)$  for all  $v_2 \in M_2$ . Then  $\iota_1$  and  $\iota_2$

are injective and  $R$ -linear. For  $\alpha \in \text{Hom}(M_1 \oplus M_2, M')$  write  $(\alpha)\theta = (\iota_1\alpha, \iota_2\alpha)$ .

As  $\iota_j\alpha \in \text{Hom}(M_j, M')$  for  $j=1, 2$ , we see

$$\theta: \text{Hom}(M_1 \oplus M_2, M') \rightarrow \text{Hom}(M_1, M') \oplus \text{Hom}(M_2, M').$$

From Exercises 3.3, Question 6(b) we see that  $\theta$  is an isomorphism of additive groups. We show  $\theta$  is  $R$ -linear. Consider  $\alpha \in \text{Hom}(M_1 \oplus M_2, M')$  and  $r \in R$ . By (a) above

$$r\alpha \in \text{Hom}(M_1 \oplus M_2, M') \text{ and so } (r\alpha)\theta = (\iota_1(r\alpha), \iota_2(r\alpha)).$$

As  $(v_1)(\iota_1(r\alpha)) = (v_1, 0)(r\alpha) = r((v_1, 0)\alpha) = r((v_1)\iota_1\alpha) = (v_1)(r(\iota_1\alpha))$  for all  $v_1 \in M_1$  we see  $\iota_1(r\alpha) = r(\iota_1\alpha)$ . Also  $\iota_2(r\alpha) = r(\iota_2\alpha)$  and so

$$(r\alpha)\theta = (r(\iota_1\alpha), r(\iota_2\alpha)) = r(\iota_1\alpha, \iota_2\alpha) = r((\alpha)\theta)$$

showing that  $\theta$  is  $R$ -linear. Therefore

$$\theta: \text{Hom}(M_1 \oplus M_2, M') \cong \text{Hom}(M_1, M') \oplus \text{Hom}(M_2, M').$$

Let  $\pi_1: M'_1 \oplus M'_2 \rightarrow M'_1$  and  $\pi_2: M'_1 \oplus M'_2 \rightarrow M'_2$  be the projections defined by  $(v'_1, v'_2)\pi_1 = v'_1$  and  $(v'_1, v'_2)\pi_2 = v'_2$  for all  $(v'_1, v'_2) \in M'_1 \oplus M'_2$ . Then  $\pi_1$  and  $\pi_2$  are surjective  $R$ -linear mappings. For  $\alpha \in \text{Hom}(M, M'_1 \oplus M'_2)$  write

$$(\alpha)\varphi = (\alpha\pi_1, \alpha\pi_2). \text{ As } \alpha\pi_i \in \text{Hom}(M, M'_i) \text{ for } i=1, 2 \text{ we see}$$

$$\varphi: \text{Hom}(M, M'_1 \oplus M'_2) \rightarrow \text{Hom}(M, M'_1) \oplus \text{Hom}(M, M'_2).$$

From Exercises 3.3, Question 6(b) we see that  $\varphi$  is an isomorphism of additive groups.

We show that  $\varphi$  is  $R$ -linear. Consider  $\alpha \in \text{Hom}(M, M'_1 \oplus M'_2)$  and  $r \in R$ . By (a) above  $r\alpha \in \text{Hom}(M, M'_1 \oplus M'_2)$  and so  $(r\alpha)\varphi = ((r\alpha)\pi_1, (r\alpha)\pi_2)$ . As

$$(v)((r\alpha)\pi_1) = ((v)(r\alpha))\pi_1 = (r((v)\alpha))\pi_1 = r((v)\alpha\pi_1) = (v)(r(\alpha\pi_1))$$

for all  $v \in M$  we have  $(r\alpha)\pi_1 = r(\alpha\pi_1)$  and in the same way  $(r\alpha)\pi_2 = r(\alpha\pi_2)$ . So

$$(r\alpha)\varphi = ((r\alpha)\pi_1, (r\alpha)\pi_2) = (r(\alpha\pi_1), r(\alpha\pi_2)) = r(\alpha\pi_1, \alpha\pi_2) = r((\alpha)\varphi)$$

showing that  $\varphi$  is  $R$ -linear. Therefore

$$\varphi: \text{Hom}(M, M'_1 \oplus M'_2) \cong \text{Hom}(M, M'_1) \oplus \text{Hom}(M, M'_2).$$

(c) Consider  $\beta \in \text{Hom}(M, M')$ . There is  $f(x) \in F[x]$  with  $(v_0)\beta = f(x)v'_0$ . As  $0 = (0)\beta = (d_0(x)v_0)\beta = d_0(x)((v_0)\beta) = d_0(x)f(x)v'_0$  we obtain  $d'_0(x) \mid d_0(x)f(x)$  by (5.11). Conversely suppose  $f(x) \in F[x]$  satisfies  $d'_0(x) \mid d_0(x)f(x)$ . We show

that there is  $\beta \in \text{Hom}(M, M')$  with  $(v_0)\beta = f(x)v'_0$ . For  $v \in M$  write

$$(v)\beta = f(x)g(x)v'_0 \text{ where } v = g(x)v_0. \text{ This definition of } (v)\beta \text{ is unambiguous as } v = h(x)v_0 \text{ gives } g(x) \equiv h(x) \pmod{d_0(x)} \text{ and so } f(x)g(x) \equiv f(x)h(x) \pmod{d'_0(x)},$$

i.e.  $f(x)g(x)v'_0 = f(x)h(x)v'_0$ . With  $v_1 = g_1(x)v_0$ ,  $v_2 = g_2(x)v_0$  we have

$$(v_1 + v_2)\beta = ((g_1(x) + g_2(x))v_0)\beta = f(x)(g_1(x) + g_2(x))v'_0 = f(x)g_1(x)v'_0 + f(x)g_2(x)v'_0 = (v_1)\beta + (v_2)\beta$$

and  $(k(x)v)\beta = (k(x)g(x)v_0)\beta = f(x)k(x)g(x)v'_0 = k(x)f(x)g(x)v'_0 = k(x)((v)\beta)$

for  $k(x) \in F[x]$ . Therefore  $\beta \in \text{Hom}(M, M')$  and  $(v_0)\beta = f(x)v'_0$  on taking

$$g(x) = 1(x).$$

Suppose  $d_0(x) = d'_0(x) = 0(x)$ . The condition  $d'_0(x) \mid d_0(x)f(x)$  is satisfied by all

$f(x) \in F[x]$ . The mapping  $\beta \rightarrow f(x)$  is an isomorphism  $\text{Hom}(M, M') \cong F[x]$  and  $\beta_0$ , where  $\beta_0 \rightarrow 1(x)$ , generates  $\text{Hom}(M, M')$  as  $\beta = f(x)\beta_0$ . Further  $\beta_0$  has order  $0(x) = \gcd\{0(x), 0(x)\}$  in  $\text{Hom}(M, M')$ .

Suppose  $d_0(x) = 0(x)$  but  $d'_0(x) \neq 0(x)$ . Then again the condition  $d'_0(x) \mid d_0(x)f(x)$  is satisfied by all  $f(x) \in F[x]$ . The mapping  $\beta \rightarrow \langle d'_0(x) \rangle + f(x)$  is an isomorphism

$\text{Hom}(M, M') \cong F[x]/\langle d'_0(x) \rangle$ . Then  $\beta_0$  specified by  $(v_0)\beta_0 = v'_0$  generates

$\text{Hom}(M, M')$  and has order  $d'_0(x) = \gcd\{0(x), d'_0(x)\}$  in  $\text{Hom}(M, M')$ .

Suppose  $d_0(x) \neq 0(x)$  but  $d'_0(x) = 0(x)$ . In this case the condition  $d'_0(x) \mid d_0(x)f(x)$  is satisfied by  $f(x) = 0(x)$  only. In this case  $\text{Hom}(M, M')$  is trivial.

Suppose  $d_0(x) \neq 0(x)$ ,  $d'_0(x) \neq 0(x)$ . In this case

$$d'_0(x) \mid d_0(x)f(x) \Leftrightarrow (d'_0(x)/\gcd\{d_0(x), d'_0(x)\}) \mid f(x).$$

The mapping  $\beta_0$ , where  $(v_0)\beta_0 = (d'_0(x)/\gcd\{d_0(x), d'_0(x)\})v'_0$ , generates

$\text{Hom}(M, M')$  and has the same order in  $\text{Hom}(M, M')$  as

$d'_0(x)/\gcd\{d_0(x), d'_0(x)\}v'_0$  has in  $M'$ , namely  $\gcd\{d_0(x), d'_0(x)\}$ .

(d)  $M = N_1 \oplus N_2 \oplus \dots \oplus N_s$  where  $N_i$  is a cyclic submodule of  $M$  being generated by  $v_i$  of order  $d_i(x)$  in  $M$  for  $1 \leq i \leq s$ . In the same way  $M' = N'_1 \oplus N'_2 \oplus \dots \oplus N'_t$  where  $N'_j$  is a cyclic submodule of  $M'$  being generated by  $v'_j$  of order  $d'_j(x)$  in  $M'$  for  $1 \leq j \leq t$ . The condition  $d_s(x) \neq 0(x)$  means that  $M$  is a **torsion module**, i.e.

$M = T(M)$  using the notation of Exercises 3.1, Question 8(a). In the same way

$d'_t(x) \neq 0(x)$  gives  $M' = T(M')$ . Using induction and (b) above we obtain the decomposition

$$\text{Hom}(M, M') \cong \sum_{i=1}^s \sum_{j=1}^t \oplus \text{Hom}(N_i, N'_j).$$

From (c) above  $\text{Hom}(N_i, N'_j)$  is a cyclic  $F[x]$ -module with generator of order

$\gcd\{d_i(x), d'_j(x)\}$  and so  $\text{Hom}(N_i, N'_j)$  is a vector space of dimension

$\deg(\gcd\{d_i(x), d'_j(x)\})$  over  $F$  by (5.24) for  $1 \leq i \leq s, 1 \leq j \leq t$ . Comparing dimensions gives

$$\dim(\text{Hom}(M, M')) = \sum_{i=1}^s \sum_{j=1}^t \dim(\text{Hom}(N_i, N'_j)) = \sum_{i=1}^s \sum_{j=1}^t \deg(\gcd\{d_i(x), d'_j(x)\}).$$

The necessary and sufficient condition is:  $\gcd\{d_i(x), d'_j(x)\} \neq 1(x)$  for at most one pair  $i, j$ . Suppose this condition is satisfied. Then at most one of the summands

$\text{Hom}(N_i, N'_j)$  is non-trivial by (c) above. As  $\gcd\{d_i(x), d'_j(x)\} \mid \gcd\{d_s(x), d'_t(x)\}$

for  $1 \leq i \leq s, 1 \leq j \leq t$  we see  $\text{Hom}(M, M') \cong \text{Hom}(N_s, N'_t)$  which is cyclic.

Conversely suppose  $\text{Hom}(M, M')$  to be cyclic. Then  $\text{Hom}(M, M')$  has at most one invariant factor. Either  $\text{Hom}(M, M')$  is trivial, and so  $\gcd\{d_i(x), d'_j(x)\} = 1(x)$  for all

pairs  $i, j$ , or  $\text{Hom}(M, M')$  is non-trivial with single invariant factor

$\gcd\{d_s(x), d'_t(x)\}$ . So  $\text{Hom}(M, M')$  has dimension  $\deg(\gcd\{d_s(x), d'_t(x)\})$  and

$\deg(\gcd\{d_i(x), d'_j(x)\}) = 0$  for  $(i, j) \neq (s, t)$  on comparing dimensions. Therefore

$\text{Hom}(M, M')$  cyclic implies  $\gcd\{d_i(x), d'_j(x)\} \neq 1(x)$  for at most one pair  $i, j$ .

### Solution 5

(a) Suppose  $\beta \in \text{Hom}((M(A), M(A')))$ . Then  $(xe_i)\beta = x((e_i)\beta)$  for  $1 \leq i \leq m$ . As  $e_i \in M(A)$  and  $e_i B \in M(A')$  we obtain

$$e_i AB = (e_i A)\beta = (xe_i)\beta = x((e_i)\beta) = x(e_i B) = e_i BA' \text{ for } 1 \leq i \leq m.$$

So  $AB = BA'$  as the rows of these  $m \times n$  matrices over  $F$  are equal showing that  $B$  intertwines  $A$  and  $A'$ .



Conversely suppose  $B$  intertwines  $A$  and  $A'$ , i.e.  $AB = BA'$ . As  $\beta: F^m \rightarrow F^n$  is  $F$ -linear, we use the second part of (5.15) to show that  $\beta: M(A) \rightarrow M(A')$  is  $F[x]$ -linear. Consider  $v \in M(A)$ . As  $(v)\beta = vB$  is in  $M(A')$  we obtain  $(xv)\beta = (vA)B = v(AB) = vBA' = (vB)A' = x(vB) = x((v)\beta)$  for all  $v \in M(A)$ . By (5.15) with  $\beta$  in place of  $\theta$  we conclude that  $\beta: M(A) \rightarrow M(A')$  is  $F[x]$ -linear.

(b)

Let  $B = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{pmatrix}$ . Then  $C(x^2)B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} B = \begin{pmatrix} b_{21} & b_{22} & b_{23} \\ 0 & 0 & 0 \end{pmatrix}$  and

$$BC(x^3) = B \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & b_{11} & b_{12} \\ 0 & b_{21} & b_{22} \end{pmatrix}.$$

Comparing entries we see  $C(x^2)B = BC(x^3)$  if and only if  $B = \begin{pmatrix} 0 & b_{12} & b_{13} \\ 0 & 0 & b_{12} \end{pmatrix}$ .

Consider  $B_0 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ . Then  $x\beta_0$  is determined by  $xB_0 = C(x^2)B_0 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ .

So  $B = b_{12}B_0 + b_{13}xB_0$  showing that each  $\beta \in \text{Hom}(M(C(x^2)), M(C(x^3)))$  can be expressed  $\beta = (b_{12} + b_{13}x)\beta_0$ , i.e. the  $F[x]$ -module  $\text{Hom}(M(C(x^2)), M(C(x^3)))$  is generated by  $\beta_0$ . As  $x^2B_0 = 0$  we see  $x^2\beta_0 = 0$  and so  $\beta_0$  has order  $x^2$  as  $\beta_0, x\beta_0$  are  $F$ -independent. Both  $M(C(x^2))$  and  $\text{Hom}(M(C(x^2)), M(C(x^3)))$  are cyclic with generator of order  $x^2$  and so these  $F[x]$ -modules are isomorphic.

Now let  $B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{pmatrix}$ . Then  $C(x^3)B = \begin{pmatrix} b_{21} & b_{22} \\ b_{31} & b_{32} \\ 0 & 0 \end{pmatrix}$  and  $BC(x^2) = \begin{pmatrix} 0 & b_{11} \\ 0 & b_{21} \\ 0 & b_{31} \end{pmatrix}$ .

Comparing entries gives  $C(x^3)B = BC(x^2)$  if and only if  $B = \begin{pmatrix} b_{11} & b_{12} \\ 0 & b_{11} \\ 0 & 0 \end{pmatrix}$ . Consider

$B_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$ . Then  $xB_0 = B_0C(x^2) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$  and so  $B = (b_{11} + b_{12}x)B_0$ . Therefore

the linear mapping  $\beta_0$  determined by  $B_0$  generates the  $F[x]$ -module

$\text{Hom}(M(C(x^3)), M(C(x^2)))$ . As  $\beta_0, x\beta_0$  are  $F$ -independent and  $x^2\beta_0 = 0$  we see that  $\beta_0$  has order  $x^2$ . As the  $F[x]$ -module  $M(C(x^3))$  is cyclic with generator  $e_1$  of order  $x^3$  it is not isomorphic to the  $F[x]$ -module  $\text{Hom}(M(C(x^3)), M(C(x^2)))$ .

(c) The  $1 \times n$  vector  $e'_1$  has order  $d'(x)$  in  $M(A')$  by (5.26) as  $A' = C(d'(x))$ . Write  $f(x) = d'(x)/\gcd\{d(x), d'(x)\}$ . Then  $\gcd\{f(x), d'(x)\} = f(x)$  as  $f(x) \mid d'(x)$  and so  $u_0 = f(x)e'_1$  has order

$$d'(x)/\gcd\{f(x), d'(x)\} = d'(x)/f(x) = \gcd\{d(x), d'(x)\}$$

by (5.23).

For  $1 \leq i < m$  we have  $e_i AB_0 = e_{i+1} B_0 = x^i u_0 = x^{i-1} u_0 A' = e_i B_0 A'$ . Let  $d(x) = c_0 + c_1 x + \dots + c_{m-1} x^{m-1} + x^m$ . By (5.26) we know  $e_1$  has order  $d(x)$  in  $M(A)$  as  $A = C(d(x))$ . So

$$\begin{aligned} e_m AB_0 &= -(c_0 e_1 + c_1 e_2 + \dots + c_{m-1} e_m) B_0 = \\ &= -(c_0 + c_1 x + \dots + c_{m-1} x^{m-1}) u_0 = x^m u_0 = x^{m-1} u_0 A' = e_m B_0 A' \end{aligned}$$

since  $d(x) u_0 = 0$  as  $\gcd\{d(x), d'(x)\} \mid d(x)$ . Therefore  $AB_0 = B_0 A'$  as these  $m \times n$  matrices have equal rows, i.e.  $B_0$  intertwines  $A$  and  $A'$ .

For  $1 \leq i \leq r$  we have

$$\begin{aligned} e_i B_0 &= x^{i-1} (b_0 + b_1 x + \dots + b_{n-r-1} x^{n-r-1} + x^{n-r}) e'_1 = \\ &= b_0 e'_i + b_1 e'_{i+1} + \dots + b_{n-r-1} e'_{i+n-r-1} + b_{n-r} e'_{i+n-r} \end{aligned}$$

as  $x^{j-1} e'_1 = e'_j$  for  $1 \leq j \leq n$  since  $A'$  is a companion matrix.

For  $r < i \leq m$  we show that row  $i$  of  $B_0$  is a certain linear combination of the preceding  $r$  rows of  $B_0$ . In fact  $(a_0 + a_1 x + \dots + a_{r-1} x^{r-1} + x^r) u_0 = \gcd\{d(x), d'(x)\} u_0 = 0$ .

Multiplying this equation by  $x^{i-r-1}$  and rearranging gives

$$\begin{aligned} e_i B_0 &= x^{i-1} u_0 = -(a_0 x^{i-r-1} + a_1 x^{i-r} + \dots + a_{r-1} x^{i-2}) u_0 = \\ &= -(a_0 e_{i-r} B_0 + a_1 e_{i-r+1} B_0 + \dots + a_{r-1} e_{i-1} B_0). \end{aligned}$$

The first  $r$  rows of  $B_0$  are the vectors of the basis  $\mathcal{B}_{u_0}$  as in (5.24) and so are linearly independent over  $F$ . By the above equation the remaining rows of  $B_0$  are linear combinations of the first  $r$  rows of  $B_0$ . Therefore  $\text{rank } B_0 = r$ .

As  $B_0$  intertwines  $A$  and  $A'$ , by (a) above  $\beta_0 \in \text{Hom}(M(A), M(A'))$ . As

$(e_1) \beta_0 = e_1 B_0 = u_0 = (d'_0(x) / \gcd\{d_0(x), d'_0(x)\}) e'_1$ , taking  $v_0 = e_1$ ,  $v'_0 = e'_1$  in Qu. 4 (c) above we see that  $\beta_0$  generates  $\text{Hom}(M(A), M(A'))$ .

(i) In the case  $d'(x) \mid d(x)$  we see  $\gcd\{d(x), d'(x)\} = d'(x)$  and so

$d'(x) / \gcd\{d(x), d'(x)\} = 1$ . So  $u_0 = 1 \times e'_1 = e'_1$  and  $r = n = \deg d'(x) \leq m$ . Therefore  $e_i B_0 = e'_i$  for  $1 \leq i \leq n$  and  $e_i B_0 = e'_n (A')^{i-n}$  for  $n < i \leq m$ .

(ii) In the case  $d(x) \mid d'(x)$  we see  $\gcd\{d(x), d'(x)\} = d(x)$  and so

$r = m = \deg d(x) \leq n$ . Therefore all the rows of  $B_0$  are given by the above formula

$e_i B_0 = b_0 e'_i + b_1 e'_{i+1} + \dots + b_{n-r-1} e'_{i+n-r-1} + b_{n-r} e'_{i+n-r}$  for  $1 \leq i \leq m$ .

(iii) In the case  $\gcd\{d(x), d'(x)\} = 1$  we have  $r = 0$  and  $u_0 = d'(x) e'_1 = 0$ . So  $B_0 = 0$  showing that the only matrix which intertwines  $C(d(x))$  and  $C(d'(x))$  is the zero  $m \times n$  matrix.

(d) The  $t \times t$  matrix  $CB$  is partitioned, in the same way as  $B$ , into

$\deg d_i(x) \times \deg d_j(x)$  submatrices  $C(d_i(x)) B_{ij}$  for  $1 \leq i, j \leq s$ . Also the  $t \times t$  matrix

$BC$  is partitioned, in the same way as  $B$ , into  $\deg d_i(x) \times \deg d_j(x)$  submatrices

$B_{ij} C(d_j(x))$  for  $1 \leq i, j \leq s$ . Therefore

$$B \in Z(C) \Leftrightarrow CB = BC \Leftrightarrow C(d_i(x)) B_{ij} = B_{ij} C(d_j(x)) \text{ for all } 1 \leq i, j \leq s.$$

So  $B \in Z(C)$  if and only if  $B_{ij}$  intertwines  $C(d_i(x))$  and  $C(d_j(x))$  for all  $1 \leq i, j \leq s$ .

By (c) above there is a  $\deg d_i(x) \times \deg d_j(x)$  matrix  $(B_{ij})_0$  such that

$B_{ij} = f_{ij}(x) (B_{ij})_0$  where  $f_{ij}(x)$  is a polynomial of degree less than

$\deg \gcd\{d_i(x), d_j(x)\}$  over  $F$  for all  $1 \leq i, j \leq s$ . Therefore the  $F[x]$ -module  $\text{End } M(C)$  is the direct sum of  $s^2$  cyclic submodules  $M_{ij}$  generated by  $(\beta_{ij})_0$  determined by  $(B_{ij})_0$  for  $1 \leq i, j \leq s$ .

(e) Let both  $C = C(d_1(x)) \oplus C(d_2(x)) \oplus \dots \oplus C(d_s(x))$  and

$C' = C(d'_1(x)) \oplus C(d'_2(x)) \oplus \dots \oplus C(d'_{s'}(x))$  be in rcf over a field  $F$ . Write

$t = \sum_{i=1}^s \deg d_i(x)$  and  $t' = \sum_{j=1}^{s'} \deg d'_j(x)$ . Partition an arbitrary  $t \times t'$  matrix  $B$  over  $F$  into  $ss'$  submatrices  $B_{ij}$ , i.e.

$$B = \begin{pmatrix} B_{11} & B_{12} & \cdots & B_{1s'} \\ B_{21} & B_{22} & \cdots & B_{2s'} \\ \vdots & \vdots & \ddots & \vdots \\ B_{s1} & B_{s2} & \cdots & B_{ss'} \end{pmatrix}$$

where each  $B_{ij}$  is a  $\deg d_i(x) \times \deg d_j(x)$  submatrix of  $B$  for  $1 \leq i \leq s, 1 \leq j \leq s'$ .

Then  $CB$  is partitioned in the same way by the  $\deg d_i(x) \times \deg d_j(x)$  submatrices

$C(d_i(x))B_{ij}$  and  $BC'$  is partitioned in the same way by the  $\deg d_i(x) \times \deg d_j(x)$

submatrices  $B_{ij}C(d'_j(x))$  for  $1 \leq i \leq s, 1 \leq j \leq s'$ . Comparing these submatrices we see

that  $B$  intertwines  $C$  and  $C'$ , i.e.  $CB = BC'$ , if and only if

$$C(d_i(x))B_{ij} = B_{ij}C(d'_j(x)) \text{ for } 1 \leq i \leq s, 1 \leq j \leq s'.$$

Suppose  $B_{ij}$  intertwines  $C(d_i(x))$  and  $C(d'_j(x))$ . By (c) above there is a

$\deg d_i(x) \times \deg d'_j(x)$  matrix  $(B_{ij})_0$  such that  $B_{ij} = f_{ij}(x)(B_{ij})_0$  where  $f_{ij}(x)$  is a polynomial of degree less than  $\deg \gcd\{d_i(x), d'_j(x)\}$  over  $F$ . The matrices  $B$  which

intertwine  $C$  and  $C'$  are the elements of an  $F[x]$ -module isomorphic to

$\text{Hom}(M(C), M(C'))$  which is the direct sum of  $ss'$  cyclic modules  $M_{ij}$  for

$1 \leq i \leq s, 1 \leq j \leq s'$  where  $M_{ij}$  has generator of order  $\gcd\{d_i(x), d'_j(x)\}$  and

$F$ -dimension  $\deg \gcd\{d_i(x), d'_j(x)\}$ .

Suppose  $CB = BC'$ . Substituting  $C = XAX^{-1}$  and  $C' = X'A'(X')^{-1}$  gives

$A(X^{-1}BX') = (X^{-1}BX')A'$  which says that  $X^{-1}BX'$  intertwines  $A$  and  $A'$ . The steps in the foregoing theory can be reversed to show that every matrix which intertwines  $A$  and  $A'$  is of the form  $X^{-1}BX'$  where  $B$  intertwines  $C$  and  $C'$ .

There is an invertible  $B$  over  $F$  with  $CB = BC'$ , i.e.  $C \sim C'$ , if and only if  $C = C'$  as the rcf is unique by (6.7).

(f)(i) Let  $d(x) = (x+1)(x^2+1)$ ,  $d'(x) = (x+1)^3$  and so  $\gcd\{d(x), d'(x)\} = x+1$ ,

$d'(x)/\gcd\{d(x), d'(x)\} = (x+1)^2$  with  $F = \mathbb{Q}$ . Using (c) above

$u_0 = (x+1)^2 e'_1 = (1, 2, 1)$  has order  $x+1$  in  $M(A')$  and

$$B_0 = \begin{pmatrix} u_0 \\ u_0 A' \\ u_0 (A')^2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 1 \\ -1 & -2 & -1 \\ 1 & 2 & -1 \end{pmatrix}$$

intertwines  $A$  and  $A'$ . In this case  $B_0$  is a  $\mathbb{Q}$ -basis of the 1-dimensional space of

matrices  $B$  with  $AB = BA'$ . On replacing  $\mathbb{Q}$  by  $\mathbb{Z}_2$  we see  $d(x) = d'(x)$  and so

$d'(x)/\gcd\{d(x), d'(x)\} = 1$ . So  $u_0 = e'_1$  has order  $(x+1)^3$  in  $M(A')$  and

$$B_0 = \begin{pmatrix} u_0 \\ u_0 A' \\ u_0 (A')^2 \end{pmatrix} = \begin{pmatrix} e'_1 \\ e'_2 \\ e'_3 \end{pmatrix} = I.$$

In this case  $B_0, B_0 A', B_0 (A')^2$ , i.e.  $I, A', (A')^2$  is a  $\mathbb{Z}_2$  - basis of the 3 - dimensional space of matrices  $B$  with  $AB = BA'$ .

(ii) Let  $d(x) = (x+1)(x^2+1)$ ,  $d'(x) = (x^2+1)^2$  and so  $\gcd\{d(x), d'(x)\} = x^2+1$ ,  $d'(x)/\gcd\{d(x), d'(x)\} = x^2+1$  with  $F = \mathbb{Q}$ . Using (c) above

$u_0 = (x^2+1)e'_1 = (1, 0, 1, 0)$  has order  $x^2+1$  in  $M(A')$  and

$$B_0 = \begin{pmatrix} u_0 \\ u_0 A' \\ u_0 (A')^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ -1 & 0 & -1 & 0 \end{pmatrix}. \text{ Then } B_0 A' = \begin{pmatrix} u_0 A' \\ u_0 (A')^2 \\ u_0 (A')^3 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 1 \\ -1 & 0 & -1 & 0 \\ 0 & -1 & 0 & -1 \end{pmatrix}$$

and  $B_0, B_0 A'$  is a  $\mathbb{Q}$  - basis of the 2 - dimensional space of matrices  $B$  with  $AB = BA'$ . On replacing  $\mathbb{Q}$  by  $\mathbb{Z}_2$  we obtain

$$d(x) = (x+1)^3, d'(x) = (x+1)^4 = x^4+1$$

and  $d'(x)/\gcd\{d(x), d'(x)\} = d'(x)/d(x) = x+1$ . So  $u_0 = (x+1)e'_1 = (1, 1, 0, 0)$  has

$$\text{order } (x+1)^3 \text{ in } M(A') \text{ and } B_0 = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}. \text{ Then } B_0 A' = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \end{pmatrix} \text{ and}$$

$$B_0 (A')^2 = \begin{pmatrix} 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \end{pmatrix}. \text{ In this case } B_0, B_0 A', B_0 (A')^2 \text{ is a } \mathbb{Z}_2 \text{ - basis of the}$$

3 - dimensional space of matrices  $B$  with  $AB = BA'$ .

(iii) Write  $A = C(d_1(x)) \oplus C(d_2(x))$  and  $A' = C(d'_1(x)) \oplus C(d'_2(x))$ . Partition a

typical  $3 \times 4$  matrix  $B$  intertwining  $A$  and  $A'$  as  $B = \left( \begin{array}{c|c} B_{11} & B_{12} \\ \hline B_{21} & B_{22} \end{array} \right)$  where

$B_{11}, B_{12}, B_{21}, B_{22}$  are  $1 \times 1, 1 \times 3, 2 \times 1, 2 \times 3$  submatrices respectively. Using (c) and (d) above

$$(B_{11})_0 = (1), (B_{12})_0 = ((x^2+1)e'_1) = \begin{pmatrix} 1 & 0 & 1 \end{pmatrix}, (B_{21})_0 = \begin{pmatrix} e'_1 \\ e'_1 C(x+1) \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \text{ and}$$

$$(B_{22})_0 = \begin{pmatrix} (x^2+1)e'_1 \\ x(x^2+1)e'_1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \\ -1 & 0 & -1 \end{pmatrix}. \text{ The four matrices}$$

$$\left( \begin{array}{c|c} (B_{11})_0 & 0 \\ \hline 0 & 0 \end{array} \right), \left( \begin{array}{c|c} 0 & (B_{12})_0 \\ \hline 0 & 0 \end{array} \right), \left( \begin{array}{c|c} 0 & 0 \\ \hline (B_{21})_0 & 0 \end{array} \right), \left( \begin{array}{c|c} 0 & 0 \\ \hline 0 & (B_{22})_0 \end{array} \right),$$

i.e.

$$\left( \begin{array}{c|ccc} 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right), \left( \begin{array}{c|ccc} 0 & 1 & 0 & 1 \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right), \left( \begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \hline 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{array} \right), \left( \begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \hline 0 & 1 & 0 & 1 \\ 0 & -1 & 0 & -1 \end{array} \right)$$

form a  $\mathbb{Q}$  - basis of the 4 - dimensional space of matrices  $B$  with  $AB = BA'$ .

On replacing  $\mathbb{Q}$  by  $\mathbb{Z}_2$  we obtain  $d'_2(x) = (x+1)^3$ . The matrices  $(B_{11})_0 = (1)$ ,

$(B_{12})_0 = ((x+1)^2 e'_1) = (1, 0, 1)$  and  $(B_{21})_0 = \begin{pmatrix} e'_1 \\ e'_1 C(x+1) \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$  are essentially

unchanged. However  $\gcd\{d_2(x), d'_2(x)\} = (x+1)^2$  and so

$$(B_{22})_0 = \begin{pmatrix} (x+1)e'_1 \\ x(x+1)e'_1 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \text{ and } (B_{22})_0 C((x+1)^3) = \begin{pmatrix} x(x+1)e'_1 \\ x^2(x+1)e'_1 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

form a  $\mathbb{Z}_2$ -basis of the 2-dimensional vector space of all matrices intertwining

$C((x+1)^2)$  and  $C((x+1)^3)$ . So the five matrices

$$\left( \begin{array}{c|ccc} 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \end{array} \right), \left( \begin{array}{c|ccc} 0 & 1 & 0 & 1 \\ \hline 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \end{array} \right), \left( \begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \hline 1 & 0 & 0 & 0 \\ \hline 1 & 0 & 0 & 0 \end{array} \right), \left( \begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \hline 0 & 1 & 1 & 0 \\ \hline 0 & 0 & 1 & 1 \end{array} \right), \left( \begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 1 & 1 \\ \hline 0 & 1 & 1 & 0 \end{array} \right)$$

form a  $\mathbb{Z}_2$ -basis of the 5-dimensional vector space of all matrices intertwining  $A$  and  $A'$ .

### Solution 6

(a) Here  $d_1(x) = x$ ,  $d_2(x) = x^2$ ,  $d_3(x) = x^2$ . So  $B(x)$  satisfies the endomorphism condition (6.31) if and only if

$d_2(x)/d_1(x) \mid b_{12}(x)$ ,  $d_3(x)/d_1(x) \mid b_{13}(x)$ ,  $d_3(x)/d_2(x) \mid b_{23}(x)$ . As  $d_3(x)/d_2(x) = 1$  and  $1 \mid b_{23}(x)$  for all  $b_{23}(x) \in F[x]$ , the last division is automatically satisfied leaving  $x \mid b_{12}(x), x \mid b_{13}(x)$ .

Let  $B(x) = (b_{ij}(x))$  and  $B'(x) = (b'_{ij}(x))$  be  $3 \times 3$  matrices over  $F$  satisfying

e.c.rel.  $M(A)$ . Then  $x \mid b_{12}(x), x \mid b_{13}(x), x \mid b'_{12}(x), x \mid b'_{13}(x)$ . So  $x \mid (b_{12}(x) + b'_{12}(x))$

and  $x \mid (b_{13}(x) + b'_{13}(x))$  showing that  $B(x) + B'(x)$  satisfies e.c.rel.  $M(A)$ .

The  $(1, 2)$ -entry in  $B(x)B'(x)$  is  $b_{11}(x)b'_{12}(x) + b_{12}(x)b'_{22}(x) + b_{13}(x)b'_{32}(x)$  which is divisible by  $x$  as each term is divisible by  $x$ . In the same way the  $(1, 3)$ -entry in  $B(x)B'(x)$  is  $b_{11}(x)b'_{13}(x) + b_{12}(x)b'_{23}(x) + b_{13}(x)b'_{33}(x)$  is divisible by  $x$ . Therefore  $B(x)B'(x)$  satisfies e.c.rel.  $M(A)$ . As  $-B(x)$  and the  $3 \times 3$  zero matrix over  $F[x]$  satisfy e.c.rel.  $M(A)$  we see that  $R_A$  is a subgroup of the additive group of  $\mathfrak{M}_3(F)$ .

As the  $3 \times 3$  identity matrix over  $F[x]$  satisfies e.c.rel.  $M(A)$  we conclude that  $R_A$  is a subring of  $\mathfrak{M}_3(F[x])$  by Exercises 2.3, Question 3(b).

Yes,  $K \subseteq R_A$  as  $x^2 \mid b_{1j}(x)$  for  $j = 2, 3$  and so  $x \mid b_{1j}(x)$  for  $j = 2, 3$ , i.e. all matrices in  $K$  satisfy e.c.rel.  $M(A)$ .

It is straightforward to verify that  $K$  is an additive subgroup of both the additive group of  $R_A$  and the additive group of  $\mathfrak{M}_3(F[x])$ . Consider  $A(x) = (a_{ij}(x)) \in \mathfrak{M}_3(F[x])$

and  $B(x) = (b_{ij}(x)) \in K$ . As  $d_k(x) \mid b_{jk}(x)$  for  $1 \leq j, k \leq 3$  we see

$$d_k(x) \mid (a_{i1}(x)b_{1k}(x) + a_{i2}(x)b_{2k}(x) + a_{i3}(x)b_{3k}(x)) \text{ for } 1 \leq i, k \leq 3$$

which shows  $A(x)B(x) \in K$ . However  $B(x) = \text{diag}(x, 0, 0) \in K$  and

$$A(x) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \in \mathfrak{M}_3(F[x]). \text{ But } B(x)A(x) = \begin{pmatrix} 0 & x & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \notin K \text{ and so } K \text{ is not an}$$

ideal of  $\mathfrak{M}_3(F[x])$ , i.e.  $K$  is not a 2-sided ideal of  $\mathfrak{M}_3(F[x])$  as defined in

Exercises 2.3, Question 3(a). Suppose  $A(x) = (a_{ij}(x)) \in R_A$ . The  $(i, k)$ -entry in

$B(x)A(x)$  is  $c_{ik}(x) = b_{i1}(x)a_{1k}(x) + b_{i2}(x)a_{2k}(x) + b_{i3}(x)a_{3k}(x)$  for  $1 \leq i, k \leq 3$ . As

$x \mid b_{ij}(x)$  for  $1 \leq j \leq 3$  we obtain  $x \mid c_{i1}(x)$  for  $1 \leq i \leq 3$ . As  $x \mid b_{i1}(x)$ , and  $x \mid a_{1j}(x)$ ,

$x^2 \mid b_{ij}(x)$  for  $2 \leq j \leq 3$  we obtain  $x^2 \mid c_{ij}(x)$  for  $1 \leq i \leq 3$ ,  $2 \leq j \leq 3$ . Therefore

$B(x)A(x) \in K$  and so  $K$  is an ideal of  $R_A$ .

The reduced matrices in  $R_A$  are of the type

$$A(x) = \begin{pmatrix} a_{11} & a_{12}x & a_{13}x \\ a_{21} & a_{22}x + a'_{22} & a_{23}x + a'_{23} \\ a_{31} & a_{32}x + a'_{32} & a_{33}x + a'_{33} \end{pmatrix}$$

where  $a_{ij} \in \mathbb{Z}_2$  for  $1 \leq i, j \leq 3$  and  $a'_{ij} \in \mathbb{Z}_2$  for  $2 \leq i, j \leq 3$ . The number of reduced

matrices is therefore  $2^{9+4} = 2^{13}$ . Each element of  $R_A/K$  contains a unique reduced

matrix by (6.32) and (6.33) and so  $|R_A/K| = 2^{13}$ . Each element of  $U(R_A/K)$

contains a unique reduced and invertible matrix. As  $\chi_A(x) = d_1(x)d_2(x)d_3(x) = x^5$  we can use (6.35):

$$\det A(x) = \begin{vmatrix} a_{11} & a_{12}x & a_{13}x \\ a_{21} & a_{22}x + a'_{22} & a_{23}x + a'_{23} \\ a_{31} & a_{32}x + a'_{32} & a_{33}x + a'_{33} \end{vmatrix} \equiv \begin{vmatrix} a_{11} & 0 & 0 \\ a_{21} & a'_{22} & a'_{23} \\ a_{31} & a'_{32} & a'_{33} \end{vmatrix} \pmod{\langle x \rangle}$$

and so  $\det A(x)$  and  $\chi_A(x)$  are coprime if and only if  $a_{11} \neq 0$  and  $\begin{vmatrix} a'_{22} & a'_{23} \\ a'_{32} & a'_{33} \end{vmatrix} \neq 0$ . For

$A(x)$  to be invertible there is only one 'choice' for  $a_{11}$ , two choices for each of the 8

remaining  $a_{ij}$  and  $|GL_2(\mathbb{Z}_2)| = 6$  choices for  $\begin{pmatrix} a'_{22} & a'_{23} \\ a'_{32} & a'_{33} \end{pmatrix}$ . Therefore

$|U(R_A/K)| = 1 \times 2^8 \times 6 = 1536$ . Restricting the ring isomorphism

$\tilde{\varphi}: R_A/K \cong \text{End } M(A)$  of (6.33) to  $U(R_A/K)$  gives a group isomorphism

$U(R_A/K) \cong \text{Aut } M(A)$  and so  $|\text{Aut } M(A)| = 1536$ . By (6.29) the size of the similarity class of  $A$  is

$$|GL_5(\mathbb{Z}_2)| / |\text{Aut } M(A)| = 31 \times 30 \times 28 \times 24 \times 16 / 1536 = 31 \times 30 \times 7 = 6510.$$

(b)(i) Let  $B(x) = (b_{ij}(x))$  belong to the ring  $R_A$ . Multiplying

$d_i(x)b_{ij}(x) \equiv 0 \pmod{d_j(x)}$  by  $-1$  gives  $d_i(x)(-b_{ij}(x)) \equiv 0 \pmod{d_j(x)}$  for

$1 \leq i, j \leq s$ . Therefore  $-B(x) \in R_A$ , i.e.  $R_A$  is closed under negation.

(ii) As  $d_i(x)0(x) \equiv 0 \pmod{d_j(x)}$  for  $1 \leq i, j \leq s$  the zero matrix of  $\mathfrak{M}_s(F[x])$  is in  $R_A$ . As  $d_i(x)0(x) \equiv 0 \pmod{d_j(x)}$  for  $1 \leq i, j \leq s, i \neq j$  and

$d_i(x)1(x) \equiv 0 \pmod{d_i(x)}$  for  $1 \leq i \leq s$  we see that the identity matrix of  $\mathfrak{M}_s(F[x])$  is in  $R_A$ . So  $R_A$  is a subring of  $\mathfrak{M}_s(F[x])$  which completes the proof of (6.32).

(c) Let  $B(x) = (b_{ij}(x))$  belong to the ring  $R_A$  and suppose  $v, w \in M(A)$ . For  $1 \leq i \leq s$

there are  $f_i(x), g_i(x) \in F[x]$  with  $v = \sum_{i=1}^s f_i(x)v_i$ ,  $w = \sum_{i=1}^s g_i(x)v_i$  and so

$v + w = \sum_{i=1}^s (f_i(x) + g_i(x))v_i$  as module law 5 holds in  $M(A)$ . As

$(v)\beta = \sum_{i,j=1}^s f_i(x)b_{ij}(x)v_j$  and  $(w)\beta = \sum_{i,j=1}^s g_i(x)b_{ij}(x)v_j$ , on adding and using the

module laws  $(v)\beta + (w)\beta = \sum_{i,j=1}^s (f_i(x) + g_i(x))b_{ij}(x)v_j = (v+w)\beta$  which shows  $\beta$  to

be additive. For  $f(x) \in F[x]$  we have  $f(x)v = \sum_{i=1}^s f(x)f_i(x)v_i$  and so

$(f(x)v)\beta = \sum_{i,j=1}^s f(x)f_i(x)b_{ij}(x)v_j = f(x)((v)\beta)$  as the module laws are obeyed by

$M(A)$ . So  $\beta$  is  $F[x]$ -linear, i.e.  $\beta \in \text{End } M(A)$ . For  $1 \leq i \leq s$  on taking

$f_i(x) = 1(x), f_j(x) = 0(x)$  for  $j \neq i$  we obtain  $v_i = \sum_{i=1}^s f_i(x)v_i$  and

$(v_i)\beta = \sum_{i,j=1}^s f_i(x)b_{ij}(x)v_j = \sum_{j=1}^s b_{ij}(x)v_j$  showing  $B(x)$  represents  $\beta$  as in (6.30).

### Solution 7

(a) The invariant factor sequences of non-invertible  $3 \times 3$  matrices  $A$  over  $\mathbb{Z}_2$  are  $(x, x, x), (x, x^2), (x, x(x+1)), (x+1, x(x+1)), (x^3), (x^3 + x^2), (x^3 + x), (x^3 + x^2 + x)$  as  $x \mid \chi_A(x)$ . The corresponding orders  $|\text{Aut}(M(A))|$  are 168, 8, 6, 6, 4, 2, 2, 3 using (6.36), (6.37), (6.38) in the first 4 cases and (6.27) in the last 4 cases with

$$\Phi_2(x^3) = 4, \Phi_2(x^3 + x^2) = \Phi_2(x^2(x+1)) = \Phi_2(x^2)\Phi_2(x+1) = 2 \times 1 = 2,$$

$$\Phi_2(x^3 + x) = \Phi_2(x(x+1)^2) = \Phi_2(x)\Phi_2((x+1)^2) = 1 \times 2 = 2,$$

$$\Phi_2(x^3 + x^2 + x) = \Phi_2(x(x^2 + x + 1)) = \Phi_2(x)\Phi_2(x^2 + x + 1) = 1 \times 3 = 3.$$

By (6.29) the numbers of matrices in the 8 similarity classes are

$$168/168 = 1, 168/8 = 21, 168/6 = 28, 168/6 = 28,$$

$$168/4 = 42, 168/2 = 84, 168/2 = 84, 168/3 = 56.$$

As  $1 + 21 + 28 + 28 + 42 + 84 + 84 + 56 = 344$  all non-invertible  $3 \times 3$  matrices over  $\mathbb{Z}_2$  are accounted for.

(b) Using (4.8)(ii) the 8 monic irreducible polynomials of degree 3 over  $\mathbb{Z}_3$  are

$$x^3 - x + 1, x^3 - x - 1, x^3 + x^2 - 1, x^3 - x^2 + 1, x^3 + x^2 + x - 1,$$

$$x^3 + x^2 - x + 1, x^3 - x^2 + x + 1, x^3 - x^2 - x - 1$$

as none of these have a zero in  $\mathbb{Z}_3$ .

There are 12 invariant factor sequences of non-cyclic similarity classes (those with more than one invariant factor) namely

$$(x, x, x), (x-1, x-1, x-1), (x+1, x+1, x+1), (x, x^2), (x-1, (x-1)^2), \\ (x+1, (x+1)^2), (x, x(x-1)), (x, x(x+1)), (x-1, x(x-1)), \\ (x-1, (x+1)(x-1)), (x+1, x(x+1)), (x+1, (x+1)(x-1)).$$

The corresponding values of  $|\text{Aut } M(A)|$  are

$$|GL_3(\mathbb{Z}_3)| = 26 \times 24 \times 18 = 11232 \quad (3 \text{ times}), 2 \times 2 \times 3^3 = 108 \quad (3 \text{ times}), \\ |GL_2(\mathbb{Z}_3)| \times |GL_1(\mathbb{Z}_3)| = 48 \times 2 = 96 \quad (6 \text{ times})$$

and the corresponding sizes of the similarity classes are

$$1 \quad (3 \text{ times}), 11232/108 = 104 \quad (3 \text{ times}), 11232/96 = 117 \quad (6 \text{ times}).$$

There are  $3^3 = 27$  cyclic similarity classes corresponding to the 27 monic cubics over  $\mathbb{Z}_3$ . These 27 cubics  $\chi_A(x)$  factorise

$$(x-a)^3 \quad (3 \text{ of these}), (x-a)^2(x-b), a \neq b \quad (6 \text{ of these}), x(x-1)(x+1), \\ (x-a) \times (\text{monic irreducible quadratic over } \mathbb{Z}_3) \quad (9 \text{ of these}), \\ \text{monic irreducible cubic over } \mathbb{Z}_3 \quad (8 \text{ of these}).$$

The corresponding values of  $\Phi_3(\chi_A(x)) = |\text{Aut } M(A)|$  are

$$27-9=18 \quad (3 \text{ times}), (9-3) \times 2 = 12 \quad (6 \text{ times}), 2 \times 2 \times 2 = 8 \quad (\text{once}), \\ 2 \times |GL_1(\mathbb{F}_9)| = 2 \times 8 = 16 \quad (9 \text{ times}), |GL_1(\mathbb{F}_{27})| = 26 \quad (8 \text{ times})$$

and the corresponding sizes of the similarity classes are

$$11232/18 = 624 \quad (3 \text{ times}), 11232/12 = 936 \quad (6 \text{ times}), 11232/8 = 1404 \quad (\text{once}), \\ 11232/16 = 702 \quad (9 \text{ times}), 11232/26 = 432 \quad (8 \text{ times}).$$

The number of matrices belonging to the union of these similarity classes is

$$3 \times 1 + 3 \times 104 + 6 \times 117 + 3 \times 624 + 6 \times 936 + 1 \times 1404 + 9 \times 702 + 8 \times 432 = 19683 = 3^9$$

and so all matrices in  $\mathfrak{M}_3(\mathbb{Z}_3)$  are accounted for.

Selecting similarity classes with  $\gcd\{x, \chi_A(x)\} = 1$ , i.e. the conjugacy classes of  $GL_3(\mathbb{Z}_3)$ , and counting up gives

$$2 \times 1 + 2 \times 104 + 2 \times 117 + 2 \times 624 + 2 \times 936 + 6 \times 702 + 8 \times 432 = 11232 = |GL_3(\mathbb{Z}_3)|.$$

So  $GL_3(\mathbb{Z}_3)$  has 2 conjugacy classes with one matrix, 2 conjugacy classes with 104 matrices, ..., 8 conjugacy classes with 432 matrices.

The matrix  $C = C(x^3 - x + 1)$  has  $\det C = -1$ . From Exercises 4.1, Question 3(c) we deduce that  $x^3 - x + 1$ , being irreducible over  $\mathbb{Z}_3$ , is a divisor of  $x^{27} - x$ . So

$C^{27} - C = 0$  as  $C^3 - C + I = 0$  by (5.26) giving  $C^{26} = I$ . Now  $26 = 2 \times 13$ . But  $C^2 \neq I$  as  $x^3 - x + 1$  is the minimum polynomial of  $C$  and  $C^{13} \neq I$  as (comparing determinants)  $(-1)^{13} \neq 1$ . Therefore  $C$  has multiplicative order 26. As  $x^3 - x + 1$  is not a divisor of  $x$  or  $x^{13} - 1$  (if so then  $C^{13} = I$ ) but  $(x^3 - x + 1) \mid x(x^{13} - 1)(x^{13} + 1)$  we conclude  $(x^3 - x + 1) \mid (x^{13} + 1)$ . For the same reason all 4 monic irreducible

polynomials over  $\mathbb{Z}_3$  with constant term 1 are divisors of  $x^{13} + 1$ , i.e.

$$x^{13} + 1 = (x+1)(x^3 - x + 1)(x^3 - x^2 + 1)(x^3 + x^2 - x + 1)(x^3 - x^2 + x + 1)$$

is the irreducible factorisation of  $x^{13} + 1$  over  $\mathbb{Z}_3$ . Replacing  $x$  by  $-x$  and changing the sign of each factor gives

$$x^{13} - 1 = (x-1)(x^3 + x - 1)(x^3 + x^2 - 1)(x^3 + x^2 + x - 1)(x^3 - x^2 - x - 1).$$



So all 4 monic irreducible polynomials over  $\mathbb{Z}_3$  with constant term  $-1$  are divisors of  $x^{13}-1$ . So the companion matrices of these 4 polynomials have determinant 1 and multiplicative order 13. The matrices

$$I, C(x-1) \oplus C((x-1)^2), C(x+1) \oplus C((x-1)(x+1)), C((x-1)^3), C((x-1)(x+1)^2), \\ C((x-1)(x^2+1)), C((x+1)(x^2+x-1)), C((x+1)(x^2-x-1))$$

in ref all have determinant 1 and multiplicative orders 1, 3, 2, 3, 6, 4, 8, 8 respectively.

These 8 matrices together with the above companion matrices, i.e.

$$C(x^3+x-1), C(x^3+x^2-1), C(x^3+x^2+x-1), C(x^3-x^2-x-1),$$

represent the 12 similarity classes of matrices contained in  $SL_3(\mathbb{Z}_3)$ . Therefore

$SL_3(\mathbb{Z}_3)$  does not contain an element of multiplicative order 26.

(c) We refine the study of similarity classes of matrices  $A$  in  $\mathfrak{M}_3(\mathbb{F}_q)$  carried out in

Exercises 6.1, Question 7(a). There are  $q$  classes with invariant factor sequence

$(x-a, x-a, x-a)$ ,  $q$  classes with invariant factor sequence  $(x-a, (x-a)^2)$ , and

$q(q-1)$  classes with invariant factor sequence  $(x-a, (x-a)(x-b))$  where  $a, b \in \mathbb{F}_q$ ,

$a \neq b$ . The remaining  $q^3$  cyclic classes have a single cubic invariant factor:

$(x-a)^3$  ( $q$  of these),  $(x-a)^2(x-b)$  ( $q(q-1)$  of these),  $(x-a)(x-b)(x-c)$

( $q(q-1)(q-2)/6$  of these),  $(x-a)p_2(x)$  where  $p_2(x)$  is a monic irreducible

quadratic over  $\mathbb{F}_q$  ( $q \times q(q-1)/2 = q^2(q-1)/2$  of these),  $p_3(x)$  where  $p_3(x)$  is a

monic irreducible cubic over  $\mathbb{F}_q$  ( $(q+1)q(q-1)/3$  of these), where  $a, b, c$  are distinct elements of  $\mathbb{F}_q$ .

The corresponding values of  $|\text{Aut } M(A)|$  are

$$|GL_3(\mathbb{F}_q)| = (q^3-1)(q^3-q)(q^3-q^2) \quad (q \text{ of these}) \text{ by (6.36),}$$

$$(q-1)^2 q^3 \quad (q \text{ of these}) \text{ by (6.37),}$$

$$|GL_2(\mathbb{F}_q)| \times |GL_1(\mathbb{F}_q)| = (q^2-1)(q^2-q)(q-1) \quad (q(q-1) \text{ of these}) \text{ by (6.36) and (6.38),}$$

$$\Phi_q((x-a)^3) = q^3 - q^2 \quad (q \text{ of these}),$$

$$\Phi_q((x-a)^2(x-b)) = (q^2-q)(q-1) \quad (q(q-1) \text{ of these}),$$

$$\Phi_q((x-a)(x-b)(x-c)) = (q-1)^3 \quad (q(q-1)(q-2)/6 \text{ of these}),$$

$$\Phi_q((x-a)p_2(x)) = (q-1)(q^2-1) \quad (q^2(q-1)/2 \text{ of these}),$$

$$\Phi_q(p_3(x)) = (q^3-1) \quad ((q+1)q(q-1)/3 \text{ of these}).$$

By (6.29) the corresponding sizes of the similarity classes partitioning  $\mathfrak{M}_3(\mathbb{F}_q)$  are

$$1 \quad (q \text{ of these}), \quad |GL_3(\mathbb{F}_q)| / ((q-1)^2 q^3) = (q^3-1)(q+1) \quad (q \text{ of these}),$$

$$|GL_3(\mathbb{F}_q)| / (|GL_2(\mathbb{F}_q)| \times |GL_1(\mathbb{F}_q)|) = (q^2+q+1)q^2 \quad (q(q-1) \text{ of these}),$$

$$(q^3-1)(q^3-q) \quad (q \text{ of these}), \quad (q^3-1)(q+1)q^2 \quad (q(q-1) \text{ of these}),$$

$$(q^2+q+1)(q+1)q^3 \quad (q(q-1)(q-2)/6 \text{ of these}),$$

$$(q^2+q+1)(q-1)q^3 \quad (q^2(q-1)/2 \text{ of these}),$$

$$(q^3-q)(q^3-q^2) \quad ((q+1)q(q-1)/3 \text{ of these}).$$

Counting the number of matrices by similarity class gives

$$\begin{aligned} & q + (q^3 - 1)(q + 1)q + (q^2 + q + 1)q^3(q - 1) + (q^3 - 1)(q^2 - 1)q^2 + (q^3 - 1)(q^2 - 1)q^3 + \\ & (q^3 - 1)(q + 1)q^4(q - 2)/6 + (q^3 - 1)(q - 1)q^5/2 + (q + 1)^2(q - 1)^3q^4/3 = \\ & q + q^5 + q^4 - q^2 - q + q^6 - q^3 + q^7 - q^5 - q^4 + q^2 + q^8 - q^6 - q^5 + q^3 + \\ & (q^4/6)\{(q^3 - 1)(q + 1)(q - 2) + 3(q^3 - 1)(q - 1)q + 2(q + 1)^2(q - 1)^3\} = \\ & -q^5 + q^7 + q^8 + (q^4/6)\{6q^5 - 6q^4 - 6q^3 + 6q\} = q^9 \end{aligned}$$

verifying that all  $q^9$  matrices in  $\mathfrak{M}_3(\mathbb{F}_q)$  are accounted for.

The similarity class of  $A \in \mathfrak{M}_3(\mathbb{F}_q)$  is not in  $GL_3(\mathbb{F}_q) \Leftrightarrow \chi_A(0) = 0$ . So the similarity classes not in  $GL_3(\mathbb{F}_q)$  are those with invariant factor sequences  $(x, x, x)$ ,  $(x, x^2)$ ,  $(x, x(x - a))$ ,  $a \neq 0$  ( $q - 1$  of these),  $(x - a, x(x - a))$ ,  $a \neq 0$  ( $q - 1$  of these) and  $q^2$  cyclic classes namely those with  $\chi_A(x) = x^3 + cx^2 + dx$ . The number of similarity classes not in  $GL_3(\mathbb{F}_q)$  is therefore  $1 + 1 + q - 1 + q - 1 + q^2 = q^2 + 2q$ . Of the  $q^2$  cyclic classes not in  $GL_3(\mathbb{F}_q)$ , there is 1 class with  $\chi_A(x) = x^3$ , there are  $q - 1$  classes with  $\chi_A(x) = x^2(x - a)$ ,  $a \neq 0$ , there are  $q - 1$  classes with  $\chi_A(x) = x(x - a)^2$ ,  $a \neq 0$ , there are  $(q - 1)(q - 2)/2$  classes with  $\chi_A(x) = x(x - a)(x - b)$ ,  $a \neq 0, b \neq 0, a \neq b$  and  $q(q - 1)/2$  classes with  $\chi_A(x) = xp_2(x)$  where  $p_2(x)$  is a monic irreducible quadratic polynomial over  $\mathbb{F}_q$ .

Counting the number of matrices not in  $GL_3(\mathbb{F}_q)$  by similarity class gives

$$\begin{aligned} & 1 \times 1 + (q^3 - 1)(q + 1) \times 1 + (q^2 + q + 1)q^2 \times (q - 1) + (q^2 + q + 1)q^2 \times (q - 1) + \\ & (q^3 - 1)(q^3 - q) \times 1 + (q^3 - 1)(q^3 + q^2) \times (2q - 2) + \\ & (q^2 + q + 1)(q + 1)q^3 \times (q - 1)(q - 2)/2 + (q^3 - 1)q^3 \times q(q - 1)/2. \end{aligned}$$

Using  $(q^2 + q + 1)(q - 1) = q^3 - 1$  the above expression simplifies to

$q^8 + q^7 - q^5 - q^4 + q^3$ , showing that all singular ( $\det A = 0$ ) matrices  $A \in \mathfrak{M}_3(\mathbb{F}_q)$  are accounted for.

(d) Using Exercises 4.1, Question 3(c) and the field  $\mathbb{F}_{81}$ , the polynomial  $x^{81} - x$  splits into distinct factors of degree 1 over  $\mathbb{F}_{81}$ . Also  $x^{81} - x$  factorizes into monic irreducible factors of degrees 1, 2, 4 over  $\mathbb{Z}_3$  and further all such polynomials occur once only in this factorization. As  $x^9 - x = x(x - 1)(x + 1)(x^2 + 1)(x^2 + x - 1)(x^2 - x + 1)$  accounts for all monic irreducible polynomials of degrees 1 and 2 over  $\mathbb{Z}_3$ , the polynomial  $(x^{81} - x)/(x^9 - x) = (x^{80} - 1)/(x^8 - 1) = x^{72} + x^{64} + x^{56} + x^{48} + x^{40} + x^{32} + x^{24} + x^{16} + x^8 + 1$  is the product of the monic irreducible polynomials of degree 4 over  $\mathbb{Z}_3$ , each appearing exactly once. Therefore the number of monic irreducible polynomials of degree 4 over  $\mathbb{Z}_3$  is  $72/4 = 18$ .

Let  $a$  and  $b$  be distinct elements of  $\mathbb{Z}_3$  and write  $p_2(x), p_3(x), p_4(x)$  for monic irreducible polynomials of degree 2, 3, 4 over  $\mathbb{Z}_3$ . The invariant factor sequences of  $4 \times 4$  matrices  $A$  over  $\mathbb{Z}_3$ , listed according to the irreducible factorisation over  $\mathbb{Z}_3$  of the invariant factors together with the order of  $\text{Aut } M(A)$ , are:

3 classes  $(x - a, x - a, x - a, x - a)$  with  $|\text{Aut } M(A)| = |GL_4(\mathbb{Z}_3)| = 24261120$  by (6.36),

3 classes  $(x-a, x-a, (x-a)^2)$  with  $|\text{Aut } M(A)| = 23328$  by (6.37) with  $q=3, m=1, r=2, l_1=1, l_2=2, s_1=2, s_2=1$ ,  
 6 classes  $(x-a, x-a, (x-a)(x-b))$  with  $|\text{Aut } M(A)| = |GL_3(\mathbb{Z}_3)| \times |GL_1(\mathbb{Z}_3)| = 22464$  by (6.36) and (6.38),  
 3 classes  $(x-a, (x-a)^3)$  with  $|\text{Aut } M(A)| = 324$  by (6.37) with  $q=3, m=1, r=2, l_1=1, l_2=3, s_1=1, s_2=1$ ,  
 6 classes  $((x-a), (x-a)(x-b)^2)$  with  $|\text{Aut } M(A)| = 48 \times 6 = 288$  by (6.36) and (6.38),  
 6 classes  $((x-a), (x-a)^2(x-b))$  with  $|\text{Aut } M(A)| = 108 \times 2 = 216$ ,  
 3 classes  $(x-a, x(x-1)(x+1))$  with  $|\text{Aut } M(A)| = 48 \times 2 \times 2 = 192$ ,  
 9 classes  $(x-a, (x-a)p_2(x))$  with  $|\text{Aut } M(A)| = 48 \times 8 = 384$ ,  
 3 classes  $((x-a)^2, (x-a)^2)$  with  $|\text{Aut } M(A)| = 3888$ ,  
 3 classes  $((x-a)(x-b), (x-a)(x-b))$  with  $|\text{Aut } M(A)| = 48 \times 48 = 2304$ ,  
 3 classes  $(p_2(x), p_2(x))$  with  $|\text{Aut } M(A)| = 5760$ , which are the non-cyclic classes.

Continuing, the cyclic classes with the single indicated invariant factor  $\chi_A(x)$  are:

3 classes  $(x-a)^4$  having  $|\text{Aut } M(A)| = 54$ ,  
 6 classes  $(x-a)^3(x-b)$  having  $|\text{Aut } M(A)| = 18 \times 2 = 36$ ,  
 3 classes  $(x-a)^2(x-b)^2$  having  $|\text{Aut } M(A)| = 6 \times 6 = 36$ ,  
 3 classes  $(x-a)x(x-1)(x+1)$  having  $|\text{Aut } M(A)| = 6 \times 2 \times 2 = 24$ ,  
 9 classes  $(x-a)^2 p_2(x)$  having  $|\text{Aut } M(A)| = 6 \times 8 = 48$ ,  
 9 classes  $(x-a)(x-b)p_2(x)$  having  $|\text{Aut } M(A)| = 2 \times 2 \times 8 = 32$ ,  
 3 classes  $p_2(x)p_2'(x)$  where  $p_2(x)$  and  $p_2'(x)$  are different monic irreducible quadratics over  $\mathbb{Z}_3$  having  $|\text{Aut } M(A)| = 8 \times 8 = 64$ ,  
 3 classes  $p_2(x)^2$  having  $|\text{Aut } M(A)| = 72$ ,  
 24 classes  $(x-a)p_3(x)$  having  $|\text{Aut } M(A)| = 2 \times 26 = 52$ ,  
 18 classes  $p_4(x)$  with  $|\text{Aut } M(A)| = 80$ .

Using (6.29) the numbers of  $4 \times 4$  matrices over  $\mathbb{Z}_3$  in the similarity class are:

1 (3 classes), 1040 (3 classes), 1080 (6 classes), 74880 (3 classes), 84240 (6 classes),  
 112320 (6 classes), 126360 (3 classes), 63180 (9 classes), 6240 (3 classes),  
 10530 (3 classes), 4212 (3 classes), 449280 (3 classes), 673920 (9 classes),  
 1010880 (3 classes), 505440 (9 classes), 758160 (9 classes), 379080 (3 classes),  
 336960 (3 classes), 466560 (24 classes), 303264 (18 classes).

Selecting the similarity (conjugacy) classes in  $GL_4(\mathbb{Z}_3)$ , i.e. the classes in  $\mathfrak{M}_4(\mathbb{Z}_3)$  having no invariant factor divisible by  $x$ , gives:

1 (2 classes), 1040 (2 classes), 1080 (2 classes), 74880 (2 classes), 84240 (2 classes),  
 112320 (2 classes), 126360 (0 classes), 63180 (6 classes), 6240 (2 classes),  
 10530 (1 class), 4212 (3 classes), 449280 (2 classes), 673920 (3 classes),  
 1010880 (0 classes), 505440 (6 classes), 758160 (3 classes), 379080 (3 classes),  
 336960 (3 classes), 466560 (16 classes), 303264 (18 classes).

Finally we check that all matrices in  $\mathfrak{M}_4(\mathbb{Z}_3)$ , respectively  $GL_4(\mathbb{Z}_3)$ , have been accounted for by noting:

$$\begin{aligned}
 &1 \times 3 + 1040 \times 3 + 1080 \times 6 + 74880 \times 3 + 84240 \times 6 + 112320 \times 6 + 126360 \times 3 + \\
 &63180 \times 9 + 6240 \times 3 + 10530 \times 3 + 4212 \times 3 + 449280 \times 3 + 673920 \times 9 + 1010880 \times 3 + \\
 &505440 \times 9 + 758160 \times 9 + 379080 \times 3 + 336960 \times 3 + 466560 \times 24 + 303264 \times 18 = \\
 &43046721 = 3^{16} = |\mathfrak{M}_4(\mathbb{Z}_3)|, \\
 &1 \times 2 + 1040 \times 2 + 1080 \times 2 + 74880 \times 2 + 84240 \times 2 + 112320 \times 2 + 126360 \times 0 + \\
 &63180 \times 6 + 6240 \times 2 + 10530 \times 1 + 4212 \times 3 + 449280 \times 2 + 673920 \times 3 + 1010880 \times 0 + \\
 &505440 \times 6 + 758160 \times 3 + 379080 \times 3 + 336960 \times 3 + 466560 \times 16 + 303264 \times 18 = \\
 &24261120 = |GL_4(\mathbb{Z}_3)|.
 \end{aligned}$$

Finitely Generated Abelian Groups and Similarity of  
Matrices over a Field

Norman, C.

2012, XII, 381 p., Softcover

ISBN: 978-1-4471-2729-1