

Solutions to Chapter 3

Solution to Exercise 3.18. If we identify 1×1 matrices over F with elements of F , it is obvious that $a, b \in F$ are conjugate if and only if they are equal (the equation $b = x^{-1}ax$ collapses to $b = a$). So every element of F , when thought of as a 1×1 matrix, is in its own conjugacy class. By the equivalence of Problems 3.6 and 3.7, we deduce that every 1-dimensional $F[x]$ -module is isomorphic to F_a for a unique $a \in F$, where F_a means the vector space F on which x acts as multiplication by a .

Solution to Exercise 3.19. The statement about characteristic polynomials remains true if A is block-lower-triangular (or block-upper-triangular), because the determinant of a block-triangular matrix is the product of the determinants of the diagonal blocks. (Considering Definition 2.36, one needs to observe that if a permutation σ uses any of the entries outside the diagonal blocks, it must use some entries above as well as some entries below.)

The statement about minimal polynomials, however, is not necessarily true in the block-triangular case; the argument in the proof of Proposition 3.16 still works for the diagonal blocks, which shows that $m_A(x)$ is a common multiple of $m_{A_1}(x), \dots, m_{A_k}(x)$, but it is not necessarily the least common multiple. For example, the minimal polynomial of $\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$ is x^2 , whereas both 1×1 diagonal blocks have minimal polynomial x .

Solution to Exercise 3.40.

- (i) The characteristic polynomial is $x^2(x-2)$ (easy to calculate, since the matrix is upper-triangular). So the torsion invariants are either x and

$x(x-2)$, or $x^2(x-2)$ alone. Either by computing the invariant factors of $xI - A$ (which are $1, 1, x^2(x-2)$) or by computing that $A(A-2) \neq 0$ (which shows that the minimal polynomial is not $x(x-2)$) or by computing the 0-eigenspace of A (which is one-dimensional, showing that A is not diagonalizable), we see that the sole torsion invariant is $x^2(x-2)$. So the primary invariants are x^2 and $x-2$, and the rational and primary rational canonical forms are

$$\begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 2 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

- (ii) The characteristic polynomial is $x(x-1)^2$ (easiest to calculate by expanding along the middle row). So the torsion invariants are either $x-1$ and $x(x-1)$, or $x(x-1)^2$ alone. Either by computing the invariant factors of $xI - A$ (which are $1, x-1, x(x-1)$) or by computing that $A(A-1) = 0$ (which shows that the minimal polynomial is $x(x-1)$) or by computing the 1-eigenspace of A (which is two-dimensional, showing that A is diagonalizable), we see that the torsion invariants are $x-1$ and $x(x-1)$. So the primary invariants are $x-1, x-1$, and x , and the rational and primary rational canonical forms are

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

- (iii) Here the characteristic polynomial is $x^2(x+3)$. (We can conjugate the matrix to upper-triangular form easily, by swapping first and second rows and first and second columns.) As in (i), we find that the sole torsion invariant is $x^2(x+3)$, and the primary invariants are x^2 and $x+3$. The rational and primary rational canonical forms are

$$\begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & -3 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -3 \end{pmatrix}.$$

- (iv) The characteristic polynomial is $x^4 - 3x^2 + 1$, which factorizes over \mathbb{Q} as $(x^2 - x - 1)(x^2 + x - 1)$. Both of these quadratic factors are irreducible over \mathbb{Q} , because they have irrational roots in \mathbb{R} , namely τ and $1 - \tau$ for the first and $-\tau$ and $\tau - 1$ for the second, where $\tau = \frac{1+\sqrt{5}}{2}$ is the

golden ratio. (One way to arrive at the factorization of $x^4 - 3x^2 + 1$ is to find these roots first by applying the quadratic formula, and then see how they can be combined to give rational polynomials.) So the sole torsion invariant is $x^4 - 3x^2 + 1$, and the primary invariants are $x^2 - x - 1$ and $x^2 + x - 1$. The rational and primary rational canonical forms are

$$\begin{pmatrix} 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 0 \end{pmatrix} \text{ and } \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & -1 \end{pmatrix}.$$

- (v) This matrix is block upper-triangular where the diagonal blocks are two copies of $C(x^2 + 1)$. So the characteristic polynomial is $(x^2 + 1)^2$. Of course, $x^2 + 1$ is irreducible in $\mathbb{Q}[x]$, so the torsion invariants of A are either $x^2 + 1$ and $x^2 + 1$, or $(x^2 + 1)^2$ alone. Computing just the bottom-left entry of A^2 shows that $A^2 + 1 \neq 0$, so the minimal polynomial cannot be $x^2 + 1$, and the sole torsion invariant is $(x^2 + 1)^2$, which equals the sole primary invariant. The rational and primary rational canonical forms both equal

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

Solution to Exercise 3.41. We just need two different collections of torsion invariants such that the product (the characteristic polynomial) is the same and the last torsion invariant (the minimal polynomial) is the same. One example is x , x , and x^2 as against x^2 and x^2 . The corresponding rational canonical form matrices are

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \text{ and } \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

So these or any matrices conjugate to them provide an example.

Solution to Exercise 3.42. If A is a 2×2 matrix, the torsion invariants of A are either $x - \lambda$ and $x - \lambda$ for some $\lambda \in F$, or $x^2 + px + q$ alone for

some $p, q \in F$. The corresponding rational canonical form matrices are:

$$\begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}, \lambda \in F, \quad \text{and} \quad \begin{pmatrix} 0 & -q \\ 1 & -p \end{pmatrix}, p, q \in F.$$

So every 2×2 matrix is conjugate to a unique one on this list. An equally good answer would be provided by the primary rational canonical form matrices in these conjugacy classes. The only difference would be in the classes whose torsion invariant $x^2 + px + q$ factorized as $(x - \lambda_1)(x - \lambda_2)$ for some distinct elements $\lambda_1 \neq \lambda_2$ of F : for these the primary invariants would be $x - \lambda_1$ and $x - \lambda_2$, so we could replace the rational canonical form matrix by the primary rational canonical form matrix $\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$.

Solution to Exercise 3.59. The Jordan canonical form is the same as the primary rational canonical form in cases where the primary invariants all either have degree 1 (leading to a 1×1 block) or are powers of x (because the companion matrix $C(x^\ell)$ equals the Jordan block $J_\ell(0)$). So in parts (i)–(iii) we found the answer already in Exercise 3.40.

To find the Jordan canonical form in part (iv), we need to extend \mathbb{Q} to a field containing $\sqrt{5}$, so that the characteristic polynomial factorizes as $(x - \tau)(x - 1 + \tau)(x + \tau)(x + 1 - \tau)$. The primary invariants are then these degree-1 factors, so the matrix becomes diagonalizable, and its Jordan canonical form is $\text{diag}(\tau, 1 - \tau, -\tau, \tau - 1)$.

To find the Jordan canonical form in part (v), we need to extend \mathbb{Q} to a field containing $i = \sqrt{-1}$, so that the characteristic polynomial factorizes as $(x - i)^2(x + i)^2$. Since the torsion invariant remains unchanged, the primary invariants are $(x - i)^2$ and $(x + i)^2$, and the Jordan canonical form is

$$\begin{pmatrix} i & 0 & 0 & 0 \\ 1 & i & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & 0 & 1 & -i \end{pmatrix}.$$

Solution to Exercise 3.60. The characteristic polynomial is $x(x^2 + x + 1)$, and $x^2 + x + 1$ is irreducible in $\mathbb{Z}_2[x]$ (because it has no roots in \mathbb{Z}_2 , so can't

be the product of two degree-1 factors). So the sole torsion invariant is $x(x^2 + x + 1)$, and the rational canonical form of A is

$$C(x^3 + x^2 + x) = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}.$$

To find the Jordan canonical form, we need to extend \mathbb{Z}_2 to a field in which $x^2 + x + 1$ factorizes: for example, we can take $E = \mathbb{Z}_2[x]/\mathbb{Z}_2[x](x^2 + x + 1)$, which is a field with four elements $0, 1, \alpha, \alpha + 1$, where α denotes the image of x in E , which satisfies $\alpha^2 = \alpha + 1$. Then $\chi_A(x)$ factorizes completely as $x(x + \alpha)(x + \alpha + 1)$. So A is diagonalizable over E , and its Jordan canonical form is $\text{diag}(0, \alpha, \alpha + 1)$.

Solution to Exercise 3.61.

- (i) If $A \in M_n$ and $B = X^{-1}AX$ for invertible X , then $B^n = X^{-1}A^nX = 1$, so $B \in M_n$. This shows that M_n contains the whole conjugacy class of each of its elements, i.e. it is a union of conjugacy classes.
- (ii) In Examples 3.52 and 3.53, we saw that every 2×2 real matrix is either conjugate to $\begin{pmatrix} \lambda & 0 \\ 1 & \lambda \end{pmatrix}$ for some $\lambda \in \mathbb{R}$, or diagonalizable over \mathbb{C} . Now $\begin{pmatrix} \lambda & 0 \\ 1 & \lambda \end{pmatrix}^2 = \begin{pmatrix} \lambda^2 & 0 \\ 2\lambda & \lambda^2 \end{pmatrix}$ can never be the identity, so none of the first kind of conjugacy class occurs in M_2 . Moreover, $\text{diag}(\lambda_1, \lambda_2)^2 = 1$ exactly when $\lambda_1, \lambda_2 \in \{\pm 1\}$. So M_2 consists of three conjugacy classes: the identity $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and the half-turn about the origin $\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ each form a singleton conjugacy class, and the other class, say W , consists of all real matrices whose eigenvalues are 1 and -1 . (Alternatively, W consists of all real matrices whose trace is 0 and whose determinant is -1 .) In terms of linear transformations of \mathbb{R}^2 , W contains all reflections in lines through the origin (the reflection axis is the 1-eigenspace, and the perpendicular line is the (-1) -eigenspace): in particular, W contains the reflection $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ in the x -axis and the reflection $\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$ in the y -axis. It also contains other transformations which are not isometries of the plane (i.e. do not preserve distance): for these, there are still two lines through the origin which are the eigenspaces for the eigenvalues 1 and -1 , but they are not perpendicular.
- (iii) Again, we rule out the case of a 2×2 Jordan block by observing that $\begin{pmatrix} \lambda & 0 \\ 1 & \lambda \end{pmatrix}^3 = \begin{pmatrix} \lambda^3 & 0 \\ 3\lambda^2 & \lambda^3 \end{pmatrix}$ can never be the identity. Now $\text{diag}(\lambda_1, \lambda_2)^3 = 1$ exactly when $\lambda_1, \lambda_2 \in \{1, \omega, \omega^2\}$ where $\omega = e^{\frac{2\pi i}{3}}$ is the usual complex

cube root of 1. Since a real 2×2 matrix either has real eigenvalues or complex-conjugate eigenvalues (see Example 3.53), M_3 consists of two conjugacy classes: the identity forms a class by itself, and the other class, say Z , consists of all real matrices whose eigenvalues are ω and ω^2 . (Alternatively, Z consists of all real matrices whose trace is -1 and whose determinant is 1.) The rational canonical form matrix in Z is $C(x^2 + x + 1) = \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix}$, which is not an isometry of the plane. In fact, the only two elements of Z which are isometries are the anti-clockwise rotation through $\frac{2\pi}{3}$ about the origin, namely $\begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}$, and the clockwise rotation through $\frac{2\pi}{3}$ about the origin, namely $\begin{pmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix}$.

Solution to Exercise 3.62. The property of having a square root is conjugation-invariant, because $B^2 = A$ implies $(X^{-1}BX)^2 = X^{-1}AX$ for any invertible X . So it suffices to prove the result when A runs over a set of representatives for the conjugacy classes of invertible complex matrices, namely the matrices in Jordan canonical form with no zero eigenvalues. It is enough to show that a single Jordan block $J_\ell(\lambda)$ with $\lambda \neq 0$ has a square root, because we can then find the square root of a general JCF matrix ‘block by block’. The naive guess for the square root of $J_\ell(\lambda)$ is $J_\ell(\sqrt{\lambda})$, where $\sqrt{\lambda}$ denotes either of the two complex square roots of λ ; this works for $\ell = 1$ but not in general, since

$$J_\ell(\sqrt{\lambda})^2 = \begin{pmatrix} \lambda & 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 2\sqrt{\lambda} & \lambda & 0 & 0 & \cdots & 0 & 0 & 0 \\ 1 & 2\sqrt{\lambda} & \lambda & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & 2\sqrt{\lambda} & \lambda & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \lambda & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 2\sqrt{\lambda} & \lambda & 0 \\ 0 & 0 & 0 & 0 & \cdots & 1 & 2\sqrt{\lambda} & \lambda \end{pmatrix},$$

which is obviously different from $J_\ell(\lambda)$. But invoking once more the principle that having a square root is conjugation-invariant, it suffices to show that this matrix $J_\ell(\sqrt{\lambda})^2$ is conjugate to $J_\ell(\lambda)$. It certainly has the right characteristic polynomial, namely $(x - \lambda)^\ell$. All we have to show is that its minimal polynomial is also $(x - \lambda)^\ell$, because this forces the sole torsion

invariant to be $(x - \lambda)^\ell$. So assume for a contradiction that the minimal polynomial of $J_\ell(\sqrt{\lambda})^2$ was $(x - \lambda)^m$ for some $m < \ell$. But then we have that $(J_\ell(\sqrt{\lambda})^2 - \lambda)^m = 0$, and we also know that the minimal polynomial of $J_\ell(\sqrt{\lambda})$ is $(x - \sqrt{\lambda})^\ell$, so we deduce that the polynomial $(x - \sqrt{\lambda})^\ell$ divides the polynomial $(x^2 - \lambda)^m$. This is a contradiction, because the irreducible factorization of $(x^2 - \lambda)^m$ is $(x - \sqrt{\lambda})^m(x + \sqrt{\lambda})^m$. The proof is finished.

Solution to Exercise 3.63. We know that V can be written as a direct sum of indecomposable cyclic $F[x]$ -modules $F[x]/F[x]p_i(x)^{\ell_i}$, where the polynomials $p_i(x)^{\ell_i}$ are the primary invariants of (the matrix of) T . From Exercises 1.113 and 1.114 we see that V is a semisimple $F[x]$ -module if and only if each of these modules $F[x]/F[x]p_i(x)^{\ell_i}$ is semisimple; from the description of the submodules of such a module given in the proof of Proposition 2.116, we see that this happens if and only if each $\ell_i = 1$, i.e. all the primary invariants are irreducible.

On the other hand, it follows from Proposition 3.55 that T is diagonalizable over an extension field E of F if and only if the primary invariants of T over E all have degree 1. The primary invariants over E are the irreducible-power factors of the previous primary invariants $p_i(x)^{\ell_i}$, now regarded as polynomials in $E[x]$. So if these all have degree 1, it must indeed be the case that all $\ell_i = 1$, as required.

The converse is not true for arbitrary fields, since there are cases where an irreducible polynomial in $F[x]$ has repeated factors when factorized over an extension field. The classic example is when $F = \mathbb{Z}_p(t)$, the field of rational functions in the variable t with coefficients in the field \mathbb{Z}_p ; the polynomial $x^p - t$ is irreducible in $F[x]$, but over a splitting field factorizes as $(x - \zeta)^p$.

However, any field F of characteristic 0 (such as \mathbb{Q} or \mathbb{R}) has the property known as separability, namely that any irreducible polynomial $p(x) \in F[x]$ factorizes into distinct degree-1 factors over some extension field E of F , i.e. $p(x) = (x - \alpha_1)(x - \alpha_2) \cdots (x - \alpha_k)$ where the elements $\alpha_i \in E$ are all different. (The proof that $\alpha_i \neq \alpha_j$ uses the fact that any repeated factor of $p(x)$ would also be a factor of the derivative $p'(x)$.) If F is separable, it is indeed true that V is a semisimple $F[x]$ -module if and only if T is diagonalizable over some extension field of F .