

9.8 Primary Decomposition

Chapter 6 ended with a lemma about any polynomial that kills an operator: if it factors into coprime pieces, the space splits into the kernels of the pieces. The previous sections produced the best polynomial to feed it, the minimal polynomial, which factors into powers of distinct irreducibles. This section runs the lemma on that factorization. The result, the primary decomposition, cuts V into invariant pieces, one for each irreducible factor of m_T , and on each piece the operator has a minimal polynomial that is a single prime power. Two answers follow at once. An operator is diagonalizable exactly when its minimal polynomial is a product of distinct linear factors. And every complex matrix with $\mathbf{A}^k = \mathbf{I}$ is diagonalizable, a promise from Section 4. Finally, when m_T splits, the pieces are the spaces on which Chapter 10 builds the Jordan form.

9.8.1 Splitting the space along the minimal polynomial

Here is the question. Diagonalization splits V into eigenspaces, and it can fail. What splitting survives for **every** operator? The minimal polynomial suggests one. By Theorem 6.3.5, if $V \neq \{0\}$, then m_T (monic, of degree at least 1 by Theorem 9.5.1) factors as

$$m_T = p_1^{e_1} p_2^{e_2} \cdots p_k^{e_k},$$

with $p_1, \dots, p_k$ **distinct** monic irreducible polynomials and exponents $e_i \geq 1$. Powers of distinct irreducibles have no common factor, so the kernel splitting lemma of Chapter 6 (Theorem 6.7.5) applies to them, peeling off one factor at a time. It gives a direct sum. The new information is what happens on the pieces.

The primary decomposition splits V into one invariant piece for each irreducible factor of m_T , and on that piece the operator is killed by exactly that factor's power.

Theorem 9.8.1. Primary Decomposition Theorem

Let $V \neq \{0\}$ be a finite-dimensional vector space over F , let $T \in \mathcal{L}(V)$, and let $m_T = p_1^{e_1} \cdots p_k^{e_k}$, where $p_1, \dots, p_k$ are distinct monic irreducible polynomials and $e_i \geq 1$. Put

$$V_i := \ker p_i(T)^{e_i} \quad (i = 1, \dots, k).$$

Then:

- (a) $V = V_1 \oplus V_2 \oplus \cdots \oplus V_k$;
- (b) each V_i is T -invariant and nonzero;
- (c) the restriction $T|_{V_i}$ has minimal polynomial $m_{T|_{V_i}} = p_i^{e_i}$;
- (d) for each i there is $h_i \in F[x]$ such that $h_i(T)$ is the projection of V onto V_i along $\bigoplus_{j \neq i} V_j$.

In words: (a) every vector is **uniquely** a sum of pieces, one from each V_i . (b) T maps each piece into itself, so T is determined by the k restrictions, and in a basis made of bases of the V_i its matrix is block diagonal (Theorem 9.1.5). (c) Nothing is wasted: on V_i the operator is killed by $p_i^{e_i}$, and by no proper divisor of it. (d) The pieces can be computed with polynomials in T alone. The word *primary* refers to the prime powers $p_i^{e_i}$.

Idea. ① The splitting is Chapter 6's lemma: distinct irreducible powers are pairwise coprime, and their product m_T kills T . ② Invariance holds for every kernel of a polynomial in T . ③ For (c), each restriction is killed by $p_i^{e_i}$, so its minimal polynomial is some $p_i^{f_i}$ with $f_i \leq e_i$. Conversely, the product of these minimal polynomials kills T piece by piece, hence on all of V , so it is a multiple of m_T . A degree count then forces $f_i = e_i$.

Proof. Coprime. Let $i \neq j$ and $d = \gcd(p_i^{e_i}, p_j^{e_j})$, which is monic. By Lemma 9.5.2, $d = p_i^f$ and $d = p_j^g$ for some $f, g \geq 0$. By the uniqueness in Theorem 6.3.5, and since $p_i \neq p_j$, this forces $f = g = 0$, so $d = 1$.

(a) and (d). These follow from the kernel splitting lemma, applied one factor at a time.

Claim. Let U be a vector space over F , let $R \in \mathcal{L}(U)$, and let $q_1, \dots, q_r \in F[x]$ be pairwise coprime with $(q_1 \cdots q_r)(R) = 0$. Then $U = \ker q_1(R) \oplus \cdots \oplus \ker q_r(R)$, and for each i the projection of U onto $\ker q_i(R)$ along the sum of the other summands is $g_i(R)$ for some $g_i \in F[x]$.

Proof. Induction on r , with U and R allowed to vary at each step. For $r = 1$, $q_1(R) = 0$ gives $\ker q_1(R) = U$, and $g_1 = 1$ works.

Let $r \geq 2$ and put $q = q_2 \cdots q_r$. Since q_1 is coprime to each of $q_2, \dots, q_r$, it is coprime to q , and $aq_1 + bq = 1$ for some $a, b \in F[x]$ (Lemma 6.6.3 (a)). As $(q_1 q)(R) = 0$, Theorem 6.7.5 (c) gives

$$U = \ker q_1(R) \oplus W, \quad W := \ker q(R),$$

with the projections onto the two summands along each other equal to $(bq)(R)$ and $(aq_1)(R)$.

The subspace W is R -invariant (Theorem 6.7.4 (b)); write $S = R|_W$. Because W is R -invariant, $S^j w = R^j w$ for every $j \geq 0$ and $w \in W$, hence $g(S)w = g(R)w$ for every $g \in F[x]$. In particular $q(S) = 0$. By the induction hypothesis applied to W , $S \in \mathcal{L}(W)$, and $q_2, \dots, q_r$,

$$W = \ker q_2(S) \oplus \cdots \oplus \ker q_r(S),$$

with the projection of W onto its i -th summand equal to $h_i(S)$ for some $h_i \in F[x]$. For $i \geq 2$ we have $q_i \mid q$, so $\ker q_i(R) \subseteq \ker q(R) = W$, and therefore

$$\ker q_i(S) = \{w \in W : q_i(R)w = 0\} = \ker q_i(R) \cap W = \ker q_i(R).$$

So $U = \ker q_1(R) + \cdots + \ker q_r(R)$, and the sum is direct: if $0 = u_1 + \cdots + u_r$ with $u_i \in \ker q_i(R)$, then $u_2 + \cdots + u_r \in W$, so $u_1 = 0$ and $u_2 + \cdots + u_r = 0$ by the directness of $U = \ker q_1(R) \oplus W$, and then every $u_i = 0$ by the directness of W (Theorem 1.7.5, (a) $\Rightarrow$ (b) and back).

Finally the projections. For $v \in U$, the component in $\ker q_1(R)$ is $(bq)(R)v$. For $i \geq 2$, the component of v in $\ker q_i(R)$ is the component in $\ker q_i(S)$ of the W -part $(aq_1)(R)v$, namely $h_i(S)(aq_1)(R)v = (h_i a q_1)(R)v$, the middle step because $g(S)$ and $g(R)$ agree on W . Each projection is therefore a polynomial in R . ■

The polynomials $p_1^{e_1}, \dots, p_k^{e_k}$ are pairwise coprime by the paragraph above, and their product m_T annihilates T . The Claim, applied to V , T , and these polynomials, gives $V = V_1 \oplus \cdots \oplus V_k$ with each projection $h_i(T)$ a polynomial in T . This proves (a) and (d).

- (b) Each V_i is the kernel of a polynomial in T , so it is T -invariant by Theorem 6.7.4 (b). That $V_i \neq \{0\}$ will follow from (c): the minimal polynomial of an operator on the zero space is 1, and $p_i^{e_i} \neq 1$.
- (c) Each V_i is T -invariant by (b), so $(T|_{V_i})^j v = T^j v$ for every $j \geq 0$ and $v \in V_i$, and hence

$$g(T|_{V_i}) = g(T)|_{V_i} \quad \text{for every } g \in F[x]. \quad (*)$$

Let $q_i = m_{T|_{V_i}}$. By (*), $p_i^{e_i}(T|_{V_i}) = p_i^{e_i}(T)|_{V_i} = 0$, so $q_i \mid p_i^{e_i}$ (Theorem 9.5.1), and $q_i = p_i^{f_i}$ with $0 \leq f_i \leq e_i$ by Lemma 9.5.2. Put $q = q_1 \cdots q_k$. For $v_i \in V_i$, polynomials in T commute (Theorem 6.7.1 (c)), so

$$q(T)v_i = \left(\prod_{j \neq i} q_j \right)(T) q_i(T)v_i = \left(\prod_{j \neq i} q_j \right)(T) q_i(T|_{V_i})v_i = 0,$$

using (*) again. By (a), every $v \in V$ is a sum of such v_i , so $q(T) = 0$, and $m_T \mid q$ by Theorem 9.5.1. Comparing degrees (Proposition 6.1.2 (c), Theorem 0.8.3),

$$\sum_{i=1}^k e_i \deg p_i = \deg m_T \leq \deg q = \sum_{i=1}^k f_i \deg p_i.$$

Since $f_i \leq e_i$ and $\deg p_i \geq 1$ for every i , each term on the right is at most the matching term on the left, so equality holds term by term, and $f_i = e_i$. Hence $m_{T|_{V_i}} = p_i^{e_i}$. This proves the theorem. ■

The theorem reduces the study of T to the study of operators whose minimal polynomial is a power of **one** irreducible. For a matrix, choose a basis of each V_i and put them together: $\mathbf{A}$ is similar to a block diagonal matrix $\mathbf{A}_1 \oplus \cdots \oplus \mathbf{A}_k$ with $m_{\mathbf{A}_i} = p_i^{e_i}$, and then $m_{\mathbf{A}} = \text{lcm}(m_{\mathbf{A}_1}, \dots, m_{\mathbf{A}_k})$ by Proposition 9.5.6, applied $k-1$ times, consistent with (c).

Example. A Primary Decomposition with an Irreducible Quadratic

Let $\mathbf{C} = \begin{pmatrix} 0 & 0 & 2 \\ 1 & 0 & -1 \\ 0 & 1 & 2 \end{pmatrix} \in M_3(\mathbb{R})$. Chapter 6 showed that its minimal polynomial is $m_{\mathbf{C}} = (x-2)(x^2+1)$ (Exercise (B2: An annihilating polynomial by elimination)). Find the primary decomposition of $\mathbb{R}^3$, the projections as polynomials in $\mathbf{C}$, and the matrix of $\mathbf{C}$ in an adapted basis.

Solution. The factorization. $x-2$ is irreducible (degree 1), and x^2+1 is irreducible over $\mathbb{R}$ because it has no real root (Theorem 6.3.1 (b)). So $k=2$, $p_1 = x-2$, $p_2 = x^2+1$, $e_1 = e_2 = 1$.

The pieces. $V_1 = \ker(\mathbf{C} - 2\mathbf{I})$: the matrix $\mathbf{C} - 2\mathbf{I} = \begin{pmatrix} -2 & 0 & 2 \\ 1 & -2 & -1 \\ 0 & 1 & 0 \end{pmatrix}$ gives $y=0$ and $x=z$,

so $V_1 = \text{span}((1, 0, 1))$. For $V_2 = \ker(\mathbf{C}^2 + \mathbf{I})$, multiply out:

$$\mathbf{C}^2 = \begin{pmatrix} 0 & 2 & 4 \\ 0 & -1 & 0 \\ 1 & 2 & 3 \end{pmatrix}, \quad \mathbf{C}^2 + \mathbf{I} = \begin{pmatrix} 1 & 2 & 4 \\ 0 & 0 & 0 \\ 1 & 2 & 4 \end{pmatrix}.$$

So V_2 is the plane $x + 2y + 4z = 0$, with basis $(-2, 1, 0)$, $(-4, 0, 1)$. As the theorem predicts, $\dim V_1 + \dim V_2 = 3$, and $(1, 0, 1) \notin V_2$ since $1 + 0 + 4 \neq 0$.

The projections. Bézout's identity for p_1, p_2 : since $x^2 + 1 = (x + 2)(x - 2) + 5$,

$$\frac{1}{5}(x^2 + 1) - \frac{1}{5}(x + 2)(x - 2) = 1.$$

Following Theorem 6.7.5 (c) with $p = x - 2$, $q = x^2 + 1$, $a = -\frac{1}{5}(x + 2)$, and $b = \frac{1}{5}$, the projection onto V_1 along V_2 is $h_1(\mathbf{C}) = b(\mathbf{C})q(\mathbf{C}) = \frac{1}{5}(\mathbf{C}^2 + \mathbf{I})$, and the projection onto V_2 along V_1 is $h_2(\mathbf{C}) = \mathbf{I} - h_1(\mathbf{C})$. Concretely,

$$h_1(\mathbf{C}) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \frac{x + 2y + 4z}{5} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix},$$

which lies in V_1 , is $\mathbf{0}$ on V_2 , and fixes $(1, 0, 1)$.

The adapted basis. $\mathbf{C}(1, 0, 1) = (2, 0, 2)$. On V_2 , $\mathbf{C}(-2, 1, 0) = (0, -2, 1) = -2(-2, 1, 0) + (-4, 0, 1)$ and $\mathbf{C}(-4, 0, 1) = (2, -5, 2) = -5(-2, 1, 0) + 2(-4, 0, 1)$. So in the basis $\mathcal{B} = ((1, 0, 1), (-2, 1, 0), (-4, 0, 1))$,

$$[T_{\mathbf{C}}]_{\mathcal{B}}^{\mathcal{B}} = (2) \oplus \begin{pmatrix} -2 & -5 \\ 1 & 2 \end{pmatrix}.$$

The second block has trace 0 and determinant $-4 + 5 = 1$, so its characteristic polynomial is $x^2 + 1$; it is not a scalar matrix, so its minimal polynomial is also $x^2 + 1$, as (c) says. Over $\mathbb{R}$, no further splitting of V_2 is possible, since $\mathbf{C}$ has no eigenvector in V_2 . Over $\mathbb{C}$, the factor $x^2 + 1 = (x - i)(x + i)$ splits, and so does the plane.

Warning. The pieces are kernels of the powers $p_i^{e_i}$, not of the p_i . Replacing $\ker p_i(T)^{e_i}$ by $\ker p_i(T)$ can lose most of the space. For $\mathbf{J} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, $m_{\mathbf{J}} = (x - 1)^2$, and the primary decomposition has the single piece $\ker(\mathbf{J} - \mathbf{I})^2 = F^2$; but $\ker(\mathbf{J} - \mathbf{I}) = \text{span}(\mathbf{e}_1)$ is only a line. The sum of the kernels $\ker p_i(T)$ is still direct, but it equals V only when every $e_i = 1$, which is the subject of the next subsection.

Quick check. Let $T \in \mathcal{L}(V)$, $V \neq \{0\}$ finite-dimensional over $\mathbb{Q}$, satisfy $T^3 = T^2$ but $T^2 \neq T$. Which polynomials can m_T be? What are the pieces of the primary decomposition?

Solution. $x^3 - x^2 = x^2(x - 1)$ annihilates T , so by Lemma 9.5.2 $m_T = x^a(x - 1)^b$ with $a \leq 2$, $b \leq 1$, not both 0. Since $x(x - 1)$ does not annihilate T , and neither do its divisors x and

$x - 1$ (they would give $T = 0$ or $T = \text{id}_V$, both satisfying $T^2 = T$), we need $a = 2$. So $m_T = x^2$ or $m_T = x^2(x - 1)$. In the first case the only piece is $\ker T^2 = V$; in the second, $V = \ker T^2 \oplus \ker(T - \text{id}_V)$.

9.8.2 Diagonalizability and the minimal polynomial

When every p_i is linear and every $e_i = 1$, the pieces are kernels of $T - \lambda_i \text{id}_V$, that is, eigenspaces. That is exactly diagonalizability, and it gives the criterion Chapter 6 promised.

Theorem 9.8.2. Diagonalizability via the Minimal Polynomial

Let V be a finite-dimensional vector space over F with $\dim V \geq 1$, and let $T \in \mathcal{L}(V)$. The following are equivalent:

- (a) T is diagonalizable;
- (b) $m_T = (x - \lambda_1)(x - \lambda_2) \cdots (x - \lambda_k)$ for some **distinct** $\lambda_1, \dots, \lambda_k \in F$, that is, m_T splits over F with no repeated root;
- (c) some polynomial $(x - \mu_1) \cdots (x - \mu_r)$ with **distinct** $\mu_1, \dots, \mu_r \in F$ annihilates T .

In that case $\lambda_1, \dots, \lambda_k$ are the distinct eigenvalues of T . The same holds for $\mathbf{A} \in M_n(F)$.

Idea.

- (a) $\Rightarrow$ (b): the polynomial $\prod_i (x - \lambda_i)$ over the distinct eigenvalues kills each vector of an eigenbasis, hence kills T ; and m_T must contain every factor $x - \lambda_i$, because its roots are the eigenvalues. (b) $\Rightarrow$ (c) is free. (c) $\Rightarrow$ (a): m_T divides a product of distinct linear factors, so it is one, and the primary decomposition with every $e_i = 1$ splits V into eigenspaces.

Proof.

- (a) $\Rightarrow$ (b). Let $\lambda_1, \dots, \lambda_k$ be the distinct eigenvalues of T , let $\mathcal{B}$ be a basis of eigenvectors (Theorem 9.4.1 (b)), and put $q = (x - \lambda_1) \cdots (x - \lambda_k)$. For $v \in \mathcal{B}$ with eigenvalue λ_j , induction on i gives $T^i v = \lambda_j^i v$, hence $q(T)v = q(\lambda_j)v = 0$, as $x - \lambda_j$ is a factor of q . So $q(T)$ and the zero operator agree on a basis, and $q(T) = 0$ by the uniqueness in Theorem 2.1.5. Hence $m_T \mid q$, and by Lemma 9.5.2 m_T is a product of some of the factors $x - \lambda_i$. Each λ_i is a root of m_T by Theorem 9.5.3, so no factor is missing, and $m_T = q$.
- (b) $\Rightarrow$ (c). Take $q = m_T$.
- (c) $\Rightarrow$ (a). Let $q = (x - \mu_1) \cdots (x - \mu_r)$ with distinct μ_j annihilate T . Each factor $x - \mu_j$ is monic irreducible, being of degree 1 (Theorem 0.8.3: a factorization would leave one factor constant). By Theorem 9.5.1 and Lemma 9.5.2, $m_T = (x - \lambda_1) \cdots (x - \lambda_k)$ for some distinct λ_i among the μ_j , and $k \geq 1$ since $V \neq \{0\}$. By Theorem 9.5.3, $\lambda_1, \dots, \lambda_k$ are exactly the eigenvalues of T . Apply Theorem 9.8.1 with $p_i = x - \lambda_i$ and $e_i = 1$:

$$V = \ker(T - \lambda_1 \text{id}_V) \oplus \cdots \oplus \ker(T - \lambda_k \text{id}_V) = E_{\lambda_1}(T) \oplus \cdots \oplus E_{\lambda_k}(T).$$

This is condition (c) of Theorem 9.4.1, so T is diagonalizable. The statement for matrices follows by applying this to $T_{\mathbf{A}}$, since $m_{\mathbf{A}} = m_{T_{\mathbf{A}}}$. ■

This is the promise of Chapter 6 kept: **diagonalizable means the minimal polynomial splits with distinct roots**. The criterion is often the fastest test, because it needs no eigenvectors, only one polynomial identity. Chapter 6 also supplied a way to detect repeated factors without finding roots: if $\gcd(m_T, m'_T) = 1$, then no irreducible factor of m_T is repeated (Corollary 6.4.6 (a)). So T is diagonalizable as soon as m_T splits and is coprime to its derivative.

Condition (c) is especially useful because it does not require knowing m_T at all.

- **Projections.** If $P^2 = P$, then $x(x-1)$ annihilates P , and $0 \neq 1$ in every field. So every projection on a finite-dimensional space is diagonalizable, over every field.
- **Involutions.** If $T^2 = \text{id}_V$ and the characteristic of F is not 2, then $(x-1)(x+1)$ annihilates T , and $1 \neq -1$ because $2 \neq 0$. So T is diagonalizable. In characteristic 2 this fails: over $\mathbb{F}_2$, $\mathbf{J} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ has $\mathbf{J}^2 = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} = \mathbf{I}$, but $m_{\mathbf{J}} = (x-1)^2$.
- **Too few eigenvectors.** $\mathbf{J}$ over any field has $m_{\mathbf{J}} = (x-1)^2$, with a repeated root, so it is not diagonalizable, confirming Section 4 without any rank computation. Likewise the matrix of Example (The Minimal Polynomial from the Eigenvalues), with $m = (x-1)(x-3)^2$, is not diagonalizable, while that of Example (The Minimal Polynomial from the First Dependence), with $m = (x-1)(x-3)$, is.

The most striking use settles the question left open in Section 4 (Exercise (C1: Matrices with $\mathbf{A}^k = \mathbf{I}$)).

Example. Complex Matrices of Finite Order Are Diagonalizable

Let $\mathbf{A} \in M_n(\mathbb{C})$, $n \geq 1$, satisfy $\mathbf{A}^k = \mathbf{I}$ for some $k \geq 1$. Show that $\mathbf{A}$ is diagonalizable, with eigenvalues among the k -th roots of unity. Show that both “over $\mathbb{C}$ ” and “characteristic 0” matter.

Solution. Diagonalizable. The polynomial $q = x^k - 1$ annihilates $\mathbf{A}$. By Corollary 6.5.5, $q = (x - \zeta_1) \cdots (x - \zeta_k)$ for some $\zeta_j \in \mathbb{C}$. The ζ_j are distinct: if c were a multiple root, then $q(c) = 0$ and $q'(c) = kc^{k-1} = 0$ by Theorem 6.4.5 (a). But $q(c) = 0$ means $c^k = 1$, so $c \neq 0$; and $k \neq 0$ in $\mathbb{C}$; so $kc^{k-1} \neq 0$, a contradiction. By Theorem 9.8.2 ((c) $\Rightarrow$ (a)), $\mathbf{A}$ is diagonalizable. Its eigenvalues are roots of $m_{\mathbf{A}}$, which divides q , so each eigenvalue λ satisfies $\lambda^k = 1$. (Explicitly, the ζ_j are $e^{2\pi ij/k}$, $j = 0, \dots, k-1$.)

Over $\mathbb{R}$ it fails. The rotation $\mathbf{R} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ has $\mathbf{R}^4 = \mathbf{I}$ but is not diagonalizable over $\mathbb{R}$, since $m_{\mathbf{R}} = x^2 + 1$ does not split. The step that breaks is the splitting of $x^4 - 1 = (x-1)(x+1)(x^2+1)$.

In characteristic p it fails. Over $\mathbb{F}_p$, $\mathbf{J} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ has $\mathbf{J}^p = \begin{pmatrix} 1 & p \\ 0 & 1 \end{pmatrix} = \mathbf{I}$ (the powers of $\mathbf{J}$ computed in Section 4), but $m_{\mathbf{J}} = (x-1)^2$. The step that breaks is distinctness: the derivative of $x^p - 1$ is $px^{p-1} = 0$ in $\mathbb{F}_p$, and indeed 1 is a multiple root of $x^p - 1$.

9.8.3 Restrictions and polynomials of diagonalizable operators

The criterion in terms of a single annihilating polynomial passes easily to smaller spaces.

Corollary 9.8.3. Restrictions of Diagonalizable Operators

Let V be finite-dimensional, $T \in \mathcal{L}(V)$ diagonalizable, and U a T -invariant subspace. Then $T|_U$ is diagonalizable. Moreover $U = \bigoplus_{\lambda \in \text{spec}(T)} (U \cap E_\lambda(T))$.

Proof. If $U = \{0\}$, the empty basis gives the empty (diagonal) matrix, and both statements are trivial. Let $U \neq \{0\}$. By Theorem 9.8.2 ((a) $\Rightarrow$ (b)), $m_T = (x - \lambda_1) \cdots (x - \lambda_k)$ with distinct λ_i , the eigenvalues of T . Since U is T -invariant, $(T|_U)^j u = T^j u$ for every $j \geq 0$ and $u \in U$, so $m_T(T|_U) = m_T(T)|_U = 0$ and m_T is a product of distinct linear factors annihilating $T|_U$. By Theorem 9.8.2 ((c) $\Rightarrow$ (a)), $T|_U$ is diagonalizable.

For the decomposition, Theorem 9.4.1 (c) applied to $T|_U$ gives $U = \bigoplus_\mu E_\mu(T|_U)$, over the eigenvalues μ of $T|_U$. An eigenvector of $T|_U$ is an eigenvector of T lying in U , so $E_\mu(T|_U) = U \cap E_\mu(T)$, and every such μ is in $\text{spec}(T)$. For $\lambda \in \text{spec}(T)$ that is not an eigenvalue of $T|_U$, $U \cap E_\lambda(T) = \{0\}$, and adding zero summands does not change a direct sum. ■

The second statement says that an invariant subspace of a diagonalizable operator is spanned by the eigenvectors it contains. It will be the key step in Section 9, where two commuting diagonalizable operators are diagonalized by a single basis.

Diagonalizability is also preserved by polynomials: if T is diagonalizable and $p \in F[x]$, then every eigenvector of T for λ is an eigenvector of $p(T)$ for $p(\lambda)$ (Exercise (B3: Eigenvalues of powers and inverses) (a)), so a basis of eigenvectors of T is also one for $p(T)$. This is the operator form of Exercise (C2: Polynomials of diagonalizable matrices). The converse fails, as that exercise shows.

9.8.4 When the minimal polynomial splits: a first look at Chapter 10

Suppose p_T splits, $p_T = (x - \lambda_1)^{a_1} \cdots (x - \lambda_k)^{a_k}$ with distinct λ_i ; this always happens over $\mathbb{C}$. By Corollary 9.6.2, $m_T = (x - \lambda_1)^{s_1} \cdots (x - \lambda_k)^{s_k}$ with $1 \leq s_i \leq a_i$, and the primary decomposition reads

$$V = \ker(T - \lambda_1 \text{id}_V)^{s_1} \oplus \cdots \oplus \ker(T - \lambda_k \text{id}_V)^{s_k}.$$

On the i -th piece, $T|_{V_i} = \lambda_i \text{id}_{V_i} + N_i$, where $N_i = (T - \lambda_i \text{id}_V)|_{V_i}$ satisfies $N_i^{s_i} = 0$, because $V_i = \ker(T - \lambda_i \text{id}_V)^{s_i}$. So every operator whose characteristic polynomial splits is, piece by piece, a scalar plus a nilpotent operator. Chapter 10 calls the pieces **generalized eigenspaces** and finds the best basis for each nilpotent part, which leads to the Jordan form. Here we record only their dimensions.

Corollary 9.8.4. Dimensions of the Primary Components

Let V be finite-dimensional with $\dim V \geq 1$, $T \in \mathcal{L}(V)$, and suppose p_T splits over F , with distinct eigenvalues $\lambda_1, \dots, \lambda_k$ and $m_T = \prod_i (x - \lambda_i)^{s_i}$. Let $V_i = \ker(T - \lambda_i \text{id}_V)^{s_i}$. Then

$$p_{T|_{V_i}} = (x - \lambda_i)^{\dim V_i} \quad \text{and} \quad \dim V_i = a_T(\lambda_i).$$

Moreover $E_{\lambda_i}(T) \subseteq V_i$, with equality if and only if $s_i = 1$.

Idea. The decomposition makes p_T the product of the characteristic polynomials of the restrictions. Each of those factors is a monic divisor of p_T , and its roots are the eigenvalues of the restriction, which the minimal polynomial $(x - \lambda_i)^{s_i}$ pins down to the single value λ_i . So the i -th factor is $(x - \lambda_i)^{\dim V_i}$, and multiplying the factors back together identifies $\dim V_i$ as the multiplicity of λ_i in p_T .

Proof. By Theorem 9.8.1, $V = V_1 \oplus \dots \oplus V_k$ with each V_i invariant, nonzero, and $m_{T|_{V_i}} = (x - \lambda_i)^{s_i}$. By Theorem 9.1.5, $p_T = p_{T|_{V_1}} \cdots p_{T|_{V_k}}$. So each $p_{T|_{V_i}}$ is a monic divisor of p_T , and by Lemma 9.5.2 it is a product of powers of the $x - \lambda_j$. Its roots are the eigenvalues of $T|_{V_i}$ (Theorem 9.2.1), which are the roots of $m_{T|_{V_i}} = (x - \lambda_i)^{s_i}$ (Theorem 9.5.3), that is, only λ_i . So $p_{T|_{V_i}} = (x - \lambda_i)^{d_i}$, with $d_i = \deg p_{T|_{V_i}} = \dim V_i$ (Theorem 7.10.2). Then $p_T = \prod_i (x - \lambda_i)^{d_i}$, and since the λ_i are distinct, Lemma 6.4.1 gives $a_T(\lambda_i) = d_i$.

If $Tv = \lambda_i v$, then $(T - \lambda_i \text{id}_V)^{s_i} v = 0$ since $s_i \geq 1$, so $E_{\lambda_i}(T) \subseteq V_i$. If $s_i = 1$, the two spaces are equal by definition. Conversely, if $E_{\lambda_i}(T) = V_i$, then $x - \lambda_i$ annihilates $T|_{V_i}$, so $m_{T|_{V_i}} = (x - \lambda_i)^{s_i}$ divides $x - \lambda_i$, and $s_i = 1$. ■

So, when p_T splits, the dimension of the i -th piece is the algebraic multiplicity $a(\lambda_i)$, and inside it the eigenspace has dimension $g(\lambda_i)$. The operator is diagonalizable exactly when all these pieces are eigenspaces, which ties together the three criteria of this chapter: $g(\lambda_i) = a(\lambda_i)$ for all i (Section 4), every $s_i = 1$ (this section), and every piece $\ker(T - \lambda_i \text{id}_V)^{s_i}$ equal to $\ker(T - \lambda_i \text{id}_V)$.

Warning. A single eigenvalue gives a single piece, whatever the operator looks like.

If $p_T = (x - \lambda)^n$, then the primary decomposition is $V = \ker(T - \lambda \text{id}_V)^n$, all of V , and says

nothing new. For instance $2\mathbf{I}_3$, $\begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} \oplus (2)$, and the matrix $\mathbf{A}$ of Example (Triangularizing

a Matrix with a Single Eigenvector Line) all have $p = (x - 2)^3$ and a primary decomposition with one piece, yet they are pairwise nonsimilar: their minimal polynomials are $x - 2$, $(x - 2)^2$

(Proposition 9.5.6), and $(x - 2)^3$, the last because $(\mathbf{A} - 2\mathbf{I})^2 = \begin{pmatrix} 1 & -1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \neq \mathbf{O}$, and equal

minimal polynomials are necessary for similarity (Theorem 9.5.5). Distinguishing operators within one piece is the work of Chapter 10.

9.8.5 Exercises

9.8.5.1 A. Check your understanding

Exercise. A1

- (a) State the Primary Decomposition Theorem.
- (b) State the criterion for diagonalizability in terms of the minimal polynomial.
- (c) True or false: if $m_{\mathbf{A}} = x(x-1)(x+1)$ over $\mathbb{Q}$, then $\mathbf{A}$ is diagonalizable over $\mathbb{Q}$. Justify your answer.
- (d) True or false: if $\mathbf{A} \in M_n(\mathbb{R})$ and $\mathbf{A}^k = \mathbf{I}$ for some $k \geq 1$, then $\mathbf{A}$ is diagonalizable over $\mathbb{R}$. Justify your answer.
- (e) Let $m_T = (x-2)^2(x+1)$. What are the summands of the primary decomposition, and is T diagonalizable?
- (f) True or false: the restriction of a diagonalizable operator to an invariant subspace is diagonalizable. Justify your answer.

Solution.

- (a) If $V \neq \{0\}$ is finite-dimensional and $m_T = p_1^{e_1} \cdots p_k^{e_k}$ with distinct monic irreducible p_i , then $V = \bigoplus_i \ker p_i(T)^{e_i}$; the summands are invariant and nonzero, T restricted to the i -th has minimal polynomial $p_i^{e_i}$, and the projections are polynomials in T (Theorem 9.8.1).
- (b) T is diagonalizable if and only if m_T is a product of distinct monic linear factors over F (Theorem 9.8.2).
- (c) True: $m_{\mathbf{A}}$ splits over $\mathbb{Q}$ with distinct roots $0, 1, -1$.
- (d) False. The rotation by a right angle satisfies $\mathbf{R}^4 = \mathbf{I}$ but has no real eigenvalue (Example (Complex Matrices of Finite Order Are Diagonalizable)).
- (e) $\ker(T - 2\text{id}_V)^2$ and $\ker(T + \text{id}_V)$. T is not diagonalizable, since m_T has the repeated root 2.
- (f) True, by Corollary 9.8.3.

9.8.5.2 B. Practice

Exercise. B1: Diagonalizable or not?

Determine whether each matrix is diagonalizable over the given field. Justify your answer.

- (a) $\mathbf{A}_1 = \begin{pmatrix} 2 & 1 & 1 \\ -1 & 0 & -1 \\ 1 & 1 & 2 \end{pmatrix}$ over $\mathbb{R}$ (see Exercise (B1: Computing minimal polynomials)).
- (b) $\mathbf{M}_1 = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ -1 & 1 & 3 \end{pmatrix}$ over $\mathbb{R}$ (see Exercise (B3: Minimal polynomial from one product)).
- (c) Any $\mathbf{A} \in M_5(\mathbb{Q})$ with $\mathbf{A}^3 = 4\mathbf{A}$.
- (d) Any $\mathbf{A} \in M_2(\mathbb{R})$ with $\mathbf{A}^2 + \mathbf{I} = \mathbf{O}$, over $\mathbb{R}$ and over $\mathbb{C}$.

Solution.

- (a) By Exercise (B1: Computing minimal polynomials) (d), $m_{\mathbf{A}_1} = (x-1)(x-2)$, which has distinct linear factors. Diagonalizable, by [Theorem 9.8.2](#).
- (b) By Exercise (B3: Minimal polynomial from one product), $m_{\mathbf{M}_1} = (x-2)^2(x-3)$, with a repeated root. Not diagonalizable.
- (c) $x^3 - 4x = x(x-2)(x+2)$ annihilates $\mathbf{A}$, and $0, 2, -2$ are distinct in $\mathbb{Q}$. Diagonalizable, by condition (c) of [Theorem 9.8.2](#).
- (d) $m_{\mathbf{A}} \mid x^2 + 1$. Over $\mathbb{R}$, $x^2 + 1$ has no root, so it is irreducible (Theorem 6.3.1), and $m_{\mathbf{A}} = x^2 + 1$ (it is not 1 since $\mathbf{A}$ acts on $\mathbb{R}^2 \neq \{\mathbf{0}\}$). It does not split over $\mathbb{R}$, so $\mathbf{A}$ is not diagonalizable over $\mathbb{R}$. Over $\mathbb{C}$, $x^2 + 1 = (x-i)(x+i)$ with $i \neq -i$, so condition (c) holds and $\mathbf{A}$ is diagonalizable over $\mathbb{C}$.

Exercise. B2: Decomposing $\mathbb{R}^3$

Let $\mathbf{A} = \begin{pmatrix} 2 & 1 & 1 \\ 0 & 1 & -1 \\ -1 & -1 & 1 \end{pmatrix} \in M_3(\mathbb{R})$, which has $m_{\mathbf{A}} = (x-1)^2(x-2)$ (Exercise (B1: Computing minimal polynomials) (c)).

- (a) Find bases of $V_1 = \ker(\mathbf{A} - \mathbf{I})^2$ and $V_2 = \ker(\mathbf{A} - 2\mathbf{I})$, and check that $\mathbb{R}^3 = V_1 \oplus V_2$.
- (b) Compare V_1 with $E_1(\mathbf{A})$.
- (c) Using $(x-1)^2 - x(x-2) = 1$, write the two projections as polynomials in $\mathbf{A}$, and compute them.
- (d) Write the matrix of $T_{\mathbf{A}}$ in the basis formed by your bases of V_1 and V_2 .

Solution.

$$(a) (\mathbf{A} - \mathbf{I})^2 = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & -1 \\ -1 & -1 & 0 \end{pmatrix}^2 = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 1 & 0 \\ -1 & -1 & 0 \end{pmatrix}, \text{ whose kernel is } x + y = 0, \text{ with basis}$$

$$(-1, 1, 0), (0, 0, 1). \quad \mathbf{A} - 2\mathbf{I} = \begin{pmatrix} 0 & 1 & 1 \\ 0 & -1 & -1 \\ -1 & -1 & -1 \end{pmatrix} \text{ gives } y + z = 0 \text{ and } x = -y - z = 0,$$

so $V_2 = \text{span}((0, -1, 1))$. The three vectors form a basis of $\mathbb{R}^3$: the matrix with these columns has determinant $-1 \cdot (0 \cdot 1 - 1 \cdot (-1)) - 0 + 0 = -1 \neq 0$, expanding along the first row. So $\mathbb{R}^3 = V_1 \oplus V_2$, as [Theorem 9.8.1](#) guarantees.

- (b) $\mathbf{A} - \mathbf{I}$ has rank 2 (its third row is minus the sum of the first two, and the first two are independent), so $E_1(\mathbf{A})$ is a line, $\text{span}((-1, 1, 0))$, strictly inside the plane V_1 . This matches [Corollary 9.8.4](#): $\dim V_1 = a(1) = 2$, $g(1) = 1$, $s_1 = 2$.
- (c) With $p = (x-1)^2$, $q = x-2$, $a = 1$, $b = -x$ in Theorem 6.7.5 (c), the projection onto V_1 along V_2 is $b(\mathbf{A})q(\mathbf{A}) = -\mathbf{A}(\mathbf{A} - 2\mathbf{I})$, and onto V_2 along V_1 it is $a(\mathbf{A})p(\mathbf{A}) = (\mathbf{A} - \mathbf{I})^2$.

Computing,

$$-\mathbf{A}(\mathbf{A} - 2\mathbf{I}) = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix}, \quad (\mathbf{A} - \mathbf{I})^2 = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 1 & 0 \\ -1 & -1 & 0 \end{pmatrix}.$$

They add up to $\mathbf{I}$. The columns of the first satisfy $x + y = 0$, so they lie in V_1 ; the columns of the second are multiples of $(0, -1, 1)$, so they lie in V_2 .

- (d) $\mathbf{A}(-1, 1, 0) = (-1, 1, 0)$, $\mathbf{A}(0, 0, 1) = (1, -1, 1) = -(-1, 1, 0) + (0, 0, 1)$, and $\mathbf{A}(0, -1, 1) = (0, -2, 2) = 2(0, -1, 1)$. So the matrix is

$$\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \oplus (2),$$

block diagonal as Theorem 9.1.5 predicts, with first block of minimal polynomial $(x - 1)^2$.

Exercise. B3: Matrices with $\mathbf{A}^3 = \mathbf{A}$

Let F have characteristic not 2, and let $\mathbf{A} \in M_n(F)$, $n \geq 1$, satisfy $\mathbf{A}^3 = \mathbf{A}$.

- (a) Prove that $\mathbf{A}$ is diagonalizable, with $\text{spec}(\mathbf{A}) \subseteq \{0, 1, -1\}$.
- (b) Now let $F = \mathbb{R}$. Prove that $\text{rank } \mathbf{A} = \text{tr}(\mathbf{A}^2)$.

Solution.

- (a) $x^3 - x = x(x - 1)(x + 1)$ annihilates $\mathbf{A}$. The roots 0, 1, -1 are distinct: $1 \neq 0$, $-1 \neq 0$, and $1 \neq -1$ because $2 \neq 0$. By Theorem 9.8.2 ((c) $\Rightarrow$ (a)), $\mathbf{A}$ is diagonalizable, and its eigenvalues are roots of $m_{\mathbf{A}} \mid x^3 - x$, so they lie in $\{0, 1, -1\}$.
- (b) Write $\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{-1}$ with $\mathbf{D}$ diagonal, entries in $\{0, 1, -1\}$. Then $\text{rank } \mathbf{A} = \text{rank } \mathbf{D}$ (Proposition 4.2.6 (a)), which is the number r of nonzero diagonal entries. Also $\mathbf{A}^2 = \mathbf{P}\mathbf{D}^2\mathbf{P}^{-1}$ (Theorem 9.4.3), and $\mathbf{D}^2$ is diagonal with r entries 1 and the rest 0. By Theorem 4.2.5, $\text{tr } \mathbf{A}^2 = \text{tr } \mathbf{D}^2 = r$.

9.8.5.3 C. Going deeper

Exercise. C1: $\mathbf{A}^p = \mathbf{A}$ over $\mathbb{F}_p$

Let p be a prime and $\mathbf{A} \in M_n(\mathbb{F}_p)$, $n \geq 1$, with $\mathbf{A}^p = \mathbf{A}$.

- (a) Let $a \in \mathbb{F}_p$ be nonzero. Show that $b \mapsto ab$ is a bijection of $\mathbb{F}_p \setminus \{0\}$, and deduce $a^{p-1} = 1$. Conclude that $a^p = a$ for **every** $a \in \mathbb{F}_p$.
- (b) Deduce that $x^p - x = \prod_{a \in \mathbb{F}_p} (x - a)$ in $\mathbb{F}_p[x]$.
- (c) Prove that $\mathbf{A}$ is diagonalizable over $\mathbb{F}_p$.
- (d) Show that $\mathbf{J} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ satisfies $\mathbf{J}^p = \mathbf{I}$ over $\mathbb{F}_p$ but is not diagonalizable, and explain why this does not contradict (c).

Hint: for (a), compare the product of all nonzero elements with the product of all their multiples by a .

Solution.

- (a) The map $b \mapsto ab$ sends nonzero elements to nonzero elements (Theorem 0.7.1), and it is injective, since $ab = ab'$ gives $b = b'$ after multiplying by a^{-1} . An injective map from the finite set $\mathbb{F}_p \setminus \{0\}$ to itself is bijective (Theorem 0.4.9). So the elements ab , for $b \neq 0$, are the nonzero elements in another order. Let c be the product of all nonzero elements; it is nonzero. Multiplying the $p-1$ elements ab gives $a^{p-1}c = c$, by commutativity, so $a^{p-1} = 1$. Multiplying by a gives $a^p = a$, which also holds for $a = 0$.
- (b) By (a), each of the p distinct elements of $\mathbb{F}_p$ is a root of $f = x^p - x$. By Theorem 6.4.3, the multiplicities of these roots, each at least 1, add up to at most $\deg f = p$. So each multiplicity is 1, the sum equals p , and f splits; as f is monic, $f = \prod_{a \in \mathbb{F}_p} (x - a)$.
- (c) By (b), $x^p - x$ is a product of distinct monic linear factors, and it annihilates $\mathbf{A}$. By Theorem 9.8.2 ((c) $\Rightarrow$ (a)), $\mathbf{A}$ is diagonalizable, with eigenvalues in $\mathbb{F}_p$.
- (d) As in Example (Complex Matrices of Finite Order Are Diagonalizable), $\mathbf{J}^p = \begin{pmatrix} 1 & p \\ 0 & 1 \end{pmatrix} = \mathbf{I}$ in $\mathbb{F}_p$, while $m_{\mathbf{J}} = (x-1)^2$, so $\mathbf{J}$ is not diagonalizable. There is no contradiction: $\mathbf{J}$ satisfies $\mathbf{J}^p = \mathbf{I}$, not $\mathbf{J}^p = \mathbf{J}$ (for $p \geq 2$, $\mathbf{J}^p = \mathbf{I} \neq \mathbf{J}$). The polynomial $x^p - 1$ has the multiple root 1, because its derivative px^{p-1} is 0 in $\mathbb{F}_p$ (Theorem 6.4.5 (a)).

Exercise. C2: Irreducible minimal polynomials

Let $V \neq \{0\}$ be finite-dimensional, $T \in \mathcal{L}(V)$, and suppose m_T is **irreducible**, of degree d . For $v \in V$, let $W_v = \{q(T)v : q \in F[x]\}$.

- (a) Show that W_v is the smallest T -invariant subspace of V containing v .
- (b) Let $v \neq 0$. Prove that $(v, Tv, \dots, T^{d-1}v)$ is a basis of W_v . In particular $\dim W_v = d$.
- (c) Deduce that the only T -invariant subspaces of W_v are $\{0\}$ and W_v .
- (d) Prove that V is a direct sum of subspaces of the form W_v , and hence that d divides $\dim V$.
- (e) Illustrate with the rotation $\mathbf{R}$ of $\mathbb{R}^2$ and with $\mathbf{R} \oplus \mathbf{R}$ on $\mathbb{R}^4$.

Hint: for (b), use the gcd of m_T with a polynomial of degree less than d . For (d), take a direct sum $W_{v_1} \oplus \dots \oplus W_{v_s}$ that is as large as possible.

Solution.

- (a) W_v is the image of the linear map $F[x] \rightarrow V$, $q \mapsto q(T)v$ (linear by Theorem 6.7.1 (a)), so it is a subspace (Theorem 2.2.2). It contains $v = 1(T)v$, and $T(q(T)v) = (xq)(T)v \in W_v$, so it is invariant. If W is any invariant subspace containing v , then $T^k v \in W$ for all k by induction, so every $q(T)v \in W$, and $W_v \subseteq W$.
- (b) *Independence.* Suppose $q(T)v = 0$ for some nonzero q of degree less than d . Let $g = \gcd(q, m_T)$. It is a monic divisor of the irreducible m_T , so $g = 1$ or $g = m_T$ (Definition 6.3.1). It is not m_T , since $m_T \nmid q$ for degree reasons (Proposition 6.1.2 (c)). So $g = 1$, and $aq + bm_T = 1$ for some a, b (Corollary 6.2.4). Then $v = a(T)q(T)v +$

$b(T)m_T(T)v = 0$, a contradiction. So no nontrivial combination of $v, \dots, T^{d-1}v$ is 0. *Spanning.* For $q \in F[x]$, divide: $q = sm_T + r$ with $\deg r < d$ (Theorem 6.1.1). Then $q(T)v = r(T)v$, a combination of $v, \dots, T^{d-1}v$.

- (c) Let $W \subseteq W_v$ be invariant and nonzero, and take $w \in W$ with $w \neq 0$. By (a), $W_w \subseteq W$. By (b), $\dim W_w = d = \dim W_v$, so $W_w = W_v$ (Theorem 1.6.8), and hence $W = W_v$.
- (d) Among all families $v_1, \dots, v_s$ of nonzero vectors for which the sum $S = W_{v_1} + \dots + W_{v_s}$ is direct, choose one with s as large as possible; this is possible since $\dim S = sd \leq \dim V$ by (b) and Theorem 1.7.5. Suppose $S \neq V$, and take $v \in V \setminus S$. The subspace $W_v \cap S$ is invariant (Proposition 9.1.4) and lies in W_v , but does not contain v ; by (c) it is $\{0\}$. Then the sum $S + W_v$ is direct: if $u_1 + \dots + u_s + w = 0$ with $u_j \in W_{v_j}$ and $w \in W_v$, then $w \in W_v \cap S = \{0\}$, and directness of S gives all $u_j = 0$. This contradicts the maximality of s . So $V = W_{v_1} \oplus \dots \oplus W_{v_s}$, and $\dim V = sd$.
- (e) For the rotation $\mathbf{R}$, $m_{\mathbf{R}} = x^2 + 1$ is irreducible over $\mathbb{R}$, $d = 2$, and $W_{\mathbf{v}} = \mathbb{R}^2$ for every $\mathbf{v} \neq \mathbf{0}$: no line is invariant (Example (A Rotation Has No Invariant Lines)). For $\mathbf{R} \oplus \mathbf{R}$, $m = \text{lcm}(x^2 + 1, x^2 + 1) = x^2 + 1$ (Proposition 9.5.6), and $\mathbb{R}^4 = W_{\mathbf{e}_1} \oplus W_{\mathbf{e}_3}$ with $W_{\mathbf{e}_1} = \text{span}(\mathbf{e}_1, \mathbf{e}_2)$ and $W_{\mathbf{e}_3} = \text{span}(\mathbf{e}_3, \mathbf{e}_4)$; indeed $2 \mid 4$.