

# 12.8 Spectral Resolutions and Functions of a Normal Operator

The last two sections produce an orthonormal basis of eigenvectors. A basis, though, is a choice: permute it, or rotate inside the eigenspace of a repeated eigenvalue, and the basis changes while the operator does not. This section rewrites the same information with nothing chosen — a list of eigenvalues, and the orthogonal projections onto their eigenspaces — and then spends it. Once an operator is written as  $\lambda_1 P_1 + \cdots + \lambda_k P_k$ , any function whatsoever on those  $k$  numbers produces a new operator, and powers, square roots, and exponentials all become one line of arithmetic on the spectrum.

Throughout,  $V$  is a **nonzero** finite-dimensional inner product space over  $F = \mathbb{R}$  or  $F = \mathbb{C}$ , and  $T \in \mathcal{L}(V)$ . We say that  $T$  satisfies

$$(*) \quad V \text{ has an orthonormal basis of eigenvectors of } T.$$

By Theorem 12.4.2 this holds over  $\mathbb{C}$  exactly for the **normal** operators, and by Theorem 12.5.1 it holds over  $\mathbb{R}$  exactly for the **self-adjoint** ones. Stating everything for an operator satisfying  $(*)$  covers both fields at once, and keeps in view the only property we ever use.

## 12.8.1 Grouping an eigenbasis by eigenvalue

Here is the observation the section runs on. Let  $(e_1, \dots, e_n)$  be an orthonormal eigenbasis of  $T$ , with  $Te_j = \mu_j e_j$ . The list  $\mu_1, \dots, \mu_n$  repeats: an eigenvalue of geometric multiplicity 3 appears three times. Sort the basis vectors into bundles by the eigenvalue they carry, and each bundle spans one eigenspace. Summing the coordinates of a vector inside one bundle is exactly projecting it onto that eigenspace — orthogonally, because the basis is orthonormal. So  $T$  is a weighted sum of perpendicular projections, one weight per **distinct** eigenvalue, and the weights and the projections no longer remember the basis.

*A normal operator is a sum of perpendicular projections, weighted by its distinct eigenvalues.*

### Definition 12.8.1. Spectral Resolution

Let  $T \in \mathcal{L}(V)$  satisfy  $(*)$ , let  $\lambda_1, \dots, \lambda_k$  be the **distinct** eigenvalues of  $T$ , and write  $E_i = E_{\lambda_i}(T)$  for the corresponding eigenspaces. Let

$$P_i := P_{E_i} \in \mathcal{L}(V)$$

be the **orthogonal** projection onto  $E_i$  (Definition 11.3.2). The identity

$$T = \lambda_1 P_1 + \lambda_2 P_2 + \cdots + \lambda_k P_k$$

is the **spectral resolution** of  $T$ , and the family  $(P_1, \dots, P_k)$  is the **resolution of the identity** attached to  $T$ .

Two clauses carry weight. The eigenvalues are listed **once each**, so  $k$  is the number of distinct eigenvalues and not  $\dim V$ ; and the projections are **orthogonal**, which is the whole content of the definition, since every diagonalizable operator has *some* such expression with oblique projections. That the displayed identity is true, and that the  $P_i$  fit together as neatly as one

could want, is the next theorem.

**Theorem 12.8.1. The Spectral Resolution and Its Projections**

Let  $T \in \mathcal{L}(V)$  satisfy (\*), with distinct eigenvalues  $\lambda_1, \dots, \lambda_k$ , eigenspaces  $E_i = E_{\lambda_i}(T)$ , and  $P_i = P_{E_i}$ . Then:

- (a)  $V = E_1 \oplus \dots \oplus E_k$ , and  $E_i \perp E_j$  whenever  $i \neq j$ ;
- (b)  $P_i^* = P_i$  and  $P_i^2 = P_i$  for every  $i$ , and  $P_i P_j = 0$  whenever  $i \neq j$ ;
- (c)  $P_1 + \dots + P_k = \text{id}_V$ ;
- (d)  $T = \lambda_1 P_1 + \dots + \lambda_k P_k$ ;
- (e)  $T P_i = P_i T = \lambda_i P_i$  for every  $i$ .

*Idea.* Everything is read off one orthonormal eigenbasis. Sorting its vectors by eigenvalue cuts it into  $k$  bundles; the only thing to check is that the  $i$ -th bundle spans the **whole** of  $E_i$ , not just part of it, and that is a coordinate comparison. After that, (a) is the orthonormality of the basis, (c) and (d) are the coordinate formula  $v = \sum_j \langle v, e_j \rangle e_j$  with its terms grouped, and (b) and (e) are bookkeeping.

*Proof.* Let  $\mathcal{B} = (e_1, \dots, e_n)$  be an orthonormal basis of  $V$  with  $T e_j = \mu_j e_j$ , which exists by (\*). Each  $e_j$  is nonzero, so each  $\mu_j$  is an eigenvalue of  $T$  and therefore equals exactly one  $\lambda_i$ . Put

$$\mathcal{S}_i := \{j : \mu_j = \lambda_i\},$$

so that  $\{1, \dots, n\}$  is the disjoint union of  $\mathcal{S}_1, \dots, \mathcal{S}_k$ .

**Step 1.**  $E_i = \text{span}(e_j : j \in \mathcal{S}_i)$ . The inclusion  $(\supseteq)$  is immediate, since  $T e_j = \lambda_i e_j$  for  $j \in \mathcal{S}_i$ . For  $(\subseteq)$ , let  $v \in E_i$  and write  $v = \sum_j c_j e_j$  in the basis  $\mathcal{B}$ . Applying  $T$ ,

$$\sum_j c_j \mu_j e_j = T v = \lambda_i v = \sum_j c_j \lambda_i e_j.$$

Coordinates in a basis are unique, so  $c_j(\mu_j - \lambda_i) = 0$  for every  $j$ , and hence  $c_j = 0$  unless  $j \in \mathcal{S}_i$ . In particular  $\mathcal{S}_i \neq \emptyset$ , because  $E_i \neq \{0\}$ .

**Step 2 (a).** By Step 1,  $V = \text{span}(\mathcal{B}) = E_1 + \dots + E_k$ . If  $i \neq j$ , then  $\mathcal{S}_i$  and  $\mathcal{S}_j$  are disjoint and  $\mathcal{B}$  is orthonormal, so every spanning vector of  $E_i$  is orthogonal to every spanning vector of  $E_j$ ; hence  $E_i \perp E_j$ . Pairwise orthogonal subspaces are independent: if  $u_1 + \dots + u_k = 0$  with  $u_i \in E_i$ , then pairing with  $u_l$  kills every term but one and leaves  $\|u_l\|^2 = 0$ , so  $u_l = 0$ . By Theorem 1.7.4 the sum is direct.

**Step 3 (b).** Each  $P_i$  is self-adjoint by Theorem 11.3.3 (b), and satisfies  $P_i^2 = P_i$  with  $\text{im } P_i = E_i$  and  $\ker P_i = E_i^\perp$  (Definition 11.3.2 and the discussion following it). For  $i \neq j$ , Step 2 gives  $E_j \subseteq E_i^\perp$ , so  $\text{im } P_j \subseteq \ker P_i$  and  $P_i P_j = 0$ .

**Step 4 (c) and (d).** By Step 1,  $(e_j : j \in \mathcal{S}_i)$  is an orthonormal basis of  $E_i$ , so Theorem 11.3.3 (a) gives

$$P_i v = \sum_{j \in \mathcal{S}_i} \langle v, e_j \rangle e_j \quad (v \in V).$$

Summing over  $i$  collects every index  $j$  exactly once, so by Theorem 11.2.1,

$$\sum_{i=1}^k P_i v = \sum_{j=1}^n \langle v, e_j \rangle e_j = v,$$

which is (c). Multiplying the  $i$ -th group by  $\lambda_i$  instead,

$$\begin{aligned} \sum_{i=1}^k \lambda_i P_i v &= \sum_{i=1}^k \sum_{j \in \mathcal{S}_i} \langle v, e_j \rangle \lambda_i e_j \\ &= \sum_{j=1}^n \langle v, e_j \rangle \mu_j e_j = T \left( \sum_{j=1}^n \langle v, e_j \rangle e_j \right) = Tv, \end{aligned}$$

where the second equality uses  $\mu_j = \lambda_i$  for  $j \in \mathcal{S}_i$ , and the third the linearity of  $T$  together with  $Te_j = \mu_j e_j$ . This is (d).

**Step 5 (e).** For  $v \in V$  we have  $P_i v \in E_i$ , so  $T(P_i v) = \lambda_i P_i v$ ; thus  $TP_i = \lambda_i P_i$ . In the other order, (d) and (b) give

$$P_i T = P_i \left( \sum_{j=1}^k \lambda_j P_j \right) = \lambda_i P_i^2 = \lambda_i P_i.$$

This proves the theorem. ■

For matrices the picture is the familiar one, packaged differently. If  $\mathbf{A} \in M_n(\mathbb{C})$  is normal, write  $\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{U}^*$  with  $\mathbf{U}$  unitary and  $\mathbf{D}$  diagonal (Corollary 12.4.3). Grouping the columns  $\mathbf{q}_1, \dots, \mathbf{q}_n$  of  $\mathbf{U}$  by their eigenvalue turns this into

$$\mathbf{A} = \sum_{i=1}^k \lambda_i \mathbf{P}_i, \quad \mathbf{P}_i = \sum_{j \in \mathcal{S}_i} \mathbf{q}_j \mathbf{q}_j^*,$$

so each  $\mathbf{P}_i$  is a sum of rank-one projections and  $\text{rank } \mathbf{P}_i = \dim E_i$ .

#### Example. A Symmetric $3 \times 3$ Matrix

Let  $\mathbf{A} = \begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{pmatrix} \in M_3(\mathbb{R})$ . Find the spectral resolution of  $\mathbf{A}$ .

*Solution.* Write  $\mathbf{A} = 2\mathbf{J} - \mathbf{I}$ , where  $\mathbf{J}$  is the all-ones matrix. Since  $\mathbf{J}\mathbf{x} = (x_1 + x_2 + x_3)\mathbf{1}$ , we get  $\mathbf{J}\mathbf{1} = 3\mathbf{1}$  and  $\mathbf{J}\mathbf{x} = \mathbf{0}$  exactly when  $x_1 + x_2 + x_3 = 0$ . So  $\mathbf{J}$  has eigenvalues 3 and 0, and  $\mathbf{A}$  has eigenvalues

$$\lambda_1 = 2 \cdot 3 - 1 = 5, \quad \lambda_2 = 2 \cdot 0 - 1 = -1,$$

with  $E_5 = \text{span}(\mathbf{1})$  of dimension 1 and  $E_{-1} = \mathbf{1}^\perp$  of dimension 2. Since  $\|\mathbf{1}\|^2 = 3$ , the orthogonal projection onto  $E_5$  is  $\mathbf{P}_1 = \frac{1}{3}\mathbf{1}\mathbf{1}^\top = \frac{1}{3}\mathbf{J}$ , and  $\mathbf{P}_2 = \mathbf{I} - \mathbf{P}_1$  by Theorem 12.8.1 (c). Concretely,

$$\mathbf{P}_1 = \frac{1}{3} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}, \quad \mathbf{P}_2 = \frac{1}{3} \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}.$$

*Check.* Both are symmetric, both satisfy  $\mathbf{P}_i^2 = \mathbf{P}_i$ , their product is  $\mathbf{O}$ , they sum to  $\mathbf{I}$ , and

$$5\mathbf{P}_1 - \mathbf{P}_2 = \frac{5}{3}\mathbf{J} - \mathbf{I} + \frac{1}{3}\mathbf{J} = 2\mathbf{J} - \mathbf{I} = \mathbf{A}.$$

Note that the resolution has two terms, not three:  $k = 2$  counts **distinct** eigenvalues, while  $\text{rank } \mathbf{P}_2 = 2$  records the multiplicity of  $-1$ .

## 12.8.2 Orthogonal projections versus oblique ones

An expression  $T = \sum \lambda_i \pi_i$  with idempotents summing to the identity is not by itself news. In the proof of Theorem 10.8.3 the semisimple part of  $T$  is written as  $\sum_i \lambda_i \pi_i$ , where  $\pi_i = h_i(T)$  projects onto the generalized eigenspace  $G_{\lambda_i}(T)$  along the sum of the others (Theorem 9.8.1 (d)); when  $T$  is diagonalizable, that is a decomposition of  $T$  itself, and those  $\pi_i$  satisfy  $\pi_i^2 = \pi_i$ ,  $\pi_i \pi_j = 0$  for  $i \neq j$ , and  $\sum_i \pi_i = \text{id}_V$ , exactly as in Theorem 12.8.1 (b) and (c). What they do **not** satisfy is  $\pi_i^* = \pi_i$ , and that single missing condition is the whole difference between Chapter 10 and this section.

### Example. Two Projections That Are Not Orthogonal

Let  $\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} \in M_2(\mathbb{R})$ , which is diagonalizable with eigenvalues 1 and 2. Compute the projections  $\pi_1, \pi_2$  of Theorem 9.8.1 for  $\mathbf{A}$ , verify that  $\mathbf{A} = \pi_1 + 2\pi_2$ , and decide whether they are orthogonal projections.

*Solution.* The eigenspaces are  $E_1 = \text{span}(\mathbf{e}_1)$  and  $E_2 = \text{span}((1, 1))$ , read off from  $\mathbf{A} - \mathbf{I} = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}$  and  $\mathbf{A} - 2\mathbf{I} = \begin{pmatrix} -1 & 1 \\ 0 & 0 \end{pmatrix}$ . Since the eigenvalues are distinct, the generalized eigenspaces are the eigenspaces. With  $\mathbf{P} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$  we have  $\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \text{diag}(1, 2)$ , so

$$\pi_1 = \mathbf{P} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \mathbf{P}^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}, \quad \pi_2 = \mathbf{I} - \pi_1 = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}.$$

Both square to themselves, their products vanish in both orders, they sum to  $\mathbf{I}$ , and

$$1 \cdot \pi_1 + 2 \cdot \pi_2 = \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 2 \\ 0 & 2 \end{pmatrix} = \mathbf{A}.$$

They are **not** orthogonal projections:  $\pi_1^\top = \begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix} \neq \pi_1$ . Geometrically,  $\pi_1$  projects onto  $\text{span}(\mathbf{e}_1)$  **along**  $\text{span}((1, 1))$ , and those two lines meet at  $45^\circ$ , not at a right angle. Consistently,  $\mathbf{A}$  is not normal:

$$\mathbf{A}^\top \mathbf{A} = \begin{pmatrix} 1 & 1 \\ 1 & 5 \end{pmatrix} \neq \begin{pmatrix} 2 & 2 \\ 2 & 4 \end{pmatrix} = \mathbf{A}\mathbf{A}^\top.$$

*Warning.* “**Diagonalizable**” gives projections; only “**normal**” makes them **orthogonal**. Every diagonalizable operator splits as  $\sum \lambda_i \pi_i$  with idempotents that annihilate each other and sum to the identity, and [Example \(Two Projections That Are Not Orthogonal\)](#) is such a splitting for a matrix that is not normal. The projections there are not self-adjoint, so they are not the projections of [Definition 12.8.1](#), and the eigenspaces they project onto are not perpendicular. Whenever an argument needs  $\langle P_i u, v \rangle = \langle u, P_i v \rangle$  — and most of this section does — normality is what pays for it.

### 12.8.3 What makes a resolution the resolution

The projections in [Definition 12.8.1](#) were built from the eigenspaces, so they are visibly determined by  $T$ . It is much more useful to know the converse: that *any* decomposition with the right formal properties has to be that one. This is what lets us recognize a spectral resolution when it turns up by accident, and it is the engine of the uniqueness proof for square roots below.

#### Theorem 12.8.2. Uniqueness of the Spectral Resolution

Let  $V$  be a nonzero finite-dimensional inner product space over  $F$  and let  $T \in \mathcal{L}(V)$ . Suppose

$$T = \mu_1 Q_1 + \mu_2 Q_2 + \cdots + \mu_m Q_m,$$

where  $\mu_1, \dots, \mu_m \in F$  are **distinct** and  $Q_1, \dots, Q_m \in \mathcal{L}(V)$  are **nonzero** operators satisfying

(R1)  $Q_i^* = Q_i$  and  $Q_i^2 = Q_i$  for every  $i$ ;

(R2)  $Q_i Q_j = 0$  whenever  $i \neq j$ ;

(R3)  $Q_1 + \cdots + Q_m = \text{id}_V$ .

Then  $T$  satisfies (\*),  $\text{spec}(T) = \{\mu_1, \dots, \mu_m\}$ , and  $Q_i$  is the orthogonal projection onto  $E_{\mu_i}(T)$  for every  $i$ . In particular  $m$  is the number of distinct eigenvalues of  $T$ , and the given decomposition is the spectral resolution of  $T$ .

*Idea.* Three moves. ① A self-adjoint idempotent is the orthogonal projection onto its own image: self-adjointness is used here and nowhere else, which is exactly why [Example \(Two Projections That Are Not Orthogonal\)](#) escapes the theorem. ② Conditions (R2) and (R3) then break  $V$  into the orthogonal direct sum of the images  $U_i = \text{im } Q_i$ . ③ On  $U_i$  the operator  $T$  is multiplication by  $\mu_i$ , so  $U_i \subseteq E_{\mu_i}(T)$ ; and because the  $\mu_i$  are distinct, an eigenvector of  $T$  cannot spread across two pieces, which upgrades the inclusion to an equality.

*Proof.* Write  $U_i = \text{im } Q_i$ , which is nonzero because  $Q_i \neq 0$ .

**Step 1.**  $Q_i = P_{U_i}$ . Fix  $i$  and drop the subscript. If  $u \in U$ , say  $u = Qx$ , then  $Qu = Q^2x = Qx = u$ , so  $Q$  is the identity on  $U$ . Next,  $\ker Q = U^\perp$ . Indeed, if  $Qw = 0$ , then for every  $x$ ,

$$\langle w, Qx \rangle = \langle Qw, x \rangle = 0$$

by Definition 11.6.1 and (R1), so  $w \in U^\perp$ ; conversely if  $w \in U^\perp$ , then  $Qw \in U$  and

$$\|Qw\|^2 = \langle Qw, Qw \rangle = \langle w, Q^2w \rangle = \langle w, Qw \rangle = 0,$$

using (R1) twice and then  $Qw \in U \perp w$ ; so  $Qw = 0$ . Finally, for any  $v \in V$ ,

$$v = Qv + (v - Qv), \quad Qv \in U, \quad Q(v - Qv) = Qv - Q^2v = 0,$$

so  $v - Qv \in \ker Q = U^\perp$ . This is the splitting of  $v$  in  $U \oplus U^\perp$  (Theorem 11.3.2), so  $Qv = P_U v$  by Definition 11.3.2.

**Step 2.**  $V = U_1 \oplus \cdots \oplus U_m$ , *orthogonally*. By (R3), every  $v$  equals  $\sum_i Q_i v \in \sum_i U_i$ . For  $i \neq j$ ,  $u = Q_i x$  and  $w = Q_j y$  satisfy

$$\langle u, w \rangle = \langle Q_i x, Q_j y \rangle = \langle Q_j Q_i x, y \rangle = 0$$

by (R1) and (R2), so  $U_i \perp U_j$ . As in Step 2 of the proof of [Theorem 12.8.1](#), pairwise orthogonal subspaces are independent, so the sum is direct.

**Step 3.**  $T$  acts on  $U_i$  as multiplication by  $\mu_i$ . Let  $u \in U_i$ . By Step 1,  $Q_i u = u$ , and for  $j \neq i$ ,  $Q_j u = Q_j Q_i u = 0$  by (R2). Hence

$$Tu = \sum_{l=1}^m \mu_l Q_l u = \mu_i u.$$

Since  $U_i \neq \{0\}$ , this makes  $\mu_i$  an eigenvalue of  $T$ , with  $U_i \subseteq E_{\mu_i}(T)$ .

**Step 4.**  $\text{spec}(T) = \{\mu_1, \dots, \mu_m\}$  and  $U_i = E_{\mu_i}(T)$ . Let  $\lambda$  be an eigenvalue of  $T$  with eigenvector  $v \neq 0$ . By (R3),  $v = \sum_i Q_i v$  with  $Q_i v \in U_i$ , and by Step 3,

$$\sum_{i=1}^m \mu_i Q_i v = Tv = \lambda v = \sum_{i=1}^m \lambda Q_i v.$$

Subtracting,  $\sum_i (\mu_i - \lambda) Q_i v = 0$  with the  $i$ -th term in  $U_i$ , so by the directness of Step 2 every term vanishes:  $Q_i v = 0$  whenever  $\mu_i \neq \lambda$ . As  $v \neq 0$ , some  $Q_i v \neq 0$ , and for that  $i$  we get  $\lambda = \mu_i$ . The  $\mu_l$  being distinct,  $i$  is the only index with  $\mu_i = \lambda$ , so  $v = Q_i v \in U_i$ . Therefore  $\text{spec}(T) \subseteq \{\mu_1, \dots, \mu_m\}$  and  $E_{\mu_i}(T) \subseteq U_i$ ; with Step 3 both are equalities.

**Step 5.** By Steps 1 and 4,  $Q_i = P_{U_i} = P_{E_{\mu_i}(T)}$ . Choosing an orthonormal basis of each  $U_i$  (Theorem 11.2.2) and concatenating gives an orthonormal basis of  $V$  by Step 2, made of eigenvectors of  $T$  by Step 3; so  $T$  satisfies (\*). This proves the theorem. ■

Read the other way round, the theorem is the complex spectral theorem again with no basis in sight: over  $\mathbb{C}$ , an operator is normal **if and only if** it can be written as  $\sum \mu_i Q_i$  with distinct  $\mu_i$  and nonzero self-adjoint idempotents satisfying (R2) and (R3) (Theorem 12.4.2). Conditions (R1)–(R3) are the entire content of the theorem, compressed into three lines of algebra.

*Quick check.* The matrices  $\pi_1, \pi_2$  of [Example \(Two Projections That Are Not Orthogonal\)](#) satisfy  $\pi_i^2 = \pi_i$ ,  $\pi_1 \pi_2 = \pi_2 \pi_1 = \mathbf{O}$ , and  $\pi_1 + \pi_2 = \mathbf{I}$ , and the coefficients 1 and 2 are distinct. Which hypothesis of [Theorem 12.8.2](#) fails, and what does its failure cost?

*Solution.* Self-adjointness in (R1):  $\pi_1^\top \neq \pi_1$ . Everything else holds. The cost is Step 1 of the proof:  $\pi_1$  is a projection onto  $\text{span}(\mathbf{e}_1)$ , but along  $\text{span}((1, 1))$  rather than along  $\text{span}(\mathbf{e}_1)^\perp =$

$\text{span}(\mathbf{e}_2)$ , so it is not  $P_{\text{span}(\mathbf{e}_1)}$ . Consequently the two images are not perpendicular, and the conclusion fails at its strongest point:  $\begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$  does not satisfy  $(*)$ , since it is not normal.

## 12.8.4 Functions of an operator

Powers are now trivial. Multiplying the resolution by itself and using  $P_i P_j = 0$  for  $i \neq j$  leaves only the diagonal terms:

$$T^2 = \left( \sum_i \lambda_i P_i \right) \left( \sum_j \lambda_j P_j \right) = \sum_i \lambda_i^2 P_i,$$

and the same computation repeated gives  $T^m = \sum_i \lambda_i^m P_i$ . So  $p(T) = \sum_i p(\lambda_i) P_i$  for every polynomial  $p$ . Stare at that formula and notice what it does **not** use: nothing about  $p$  except its  $k$  values on the spectrum. Any function at all supplies those.

### Definition 12.8.2. Function of an Operator

Let  $T \in \mathcal{L}(V)$  satisfy  $(*)$ , with distinct eigenvalues  $\lambda_1, \dots, \lambda_k$  and spectral resolution  $T = \sum_i \lambda_i P_i$ . For **any** function  $f: \text{spec}(T) \rightarrow F$ , define

$$f(T) := \sum_{i=1}^k f(\lambda_i) P_i \in \mathcal{L}(V).$$

In words: apply  $f$  to each eigenvalue and reweight the same projections. The function needs no continuity, no differentiability, and no formula — it is  $k$  numbers. And the definition is unambiguous: the  $\lambda_i$  and the  $P_i$  are determined by  $T$  alone, by [Theorem 12.8.2](#), so no choice of basis or of eigenvectors enters.

Four instances show the range. Taking  $f(\lambda) = \lambda$  returns  $T$ . Taking  $f \equiv 1$  returns  $\text{id}_V$ , by [Theorem 12.8.1](#) (c). Taking  $f(\lambda) = \bar{\lambda}$  returns  $T^*$ , since  $P_i^* = P_i$ . And if  $0 \notin \text{spec}(T)$ , taking  $f(\lambda) = \lambda^{-1}$  returns  $T^{-1}$ , because

$$\left( \sum_i \lambda_i P_i \right) \left( \sum_j \lambda_j^{-1} P_j \right) = \sum_i P_i = \text{id}_V.$$

The notation is, on its face, dangerous. For a polynomial  $p$  the symbol  $p(T)$  already means something — the operator  $a_0 \text{id}_V + a_1 T + \dots + a_N T^N$  of Definition 2.4.2, which Chapter 6 studied at length. If [Definition 12.8.2](#) disagreed with it, the whole notation would have to be abandoned. Part (a) below says it does not, and that is the reason the rest of the theorem is safe to use.

### Theorem 12.8.3. Properties of the Functional Calculus

Let  $T \in \mathcal{L}(V)$  satisfy  $(*)$ , with distinct eigenvalues  $\lambda_1, \dots, \lambda_k$  and resolution of the identity  $(P_1, \dots, P_k)$ . Let  $f, g: \text{spec}(T) \rightarrow F$  and  $c \in F$ , with  $f + g$ ,  $cf$ , and  $fg$  taken pointwise.

- (a) **(Compatibility)** For every  $p \in F[x]$ , the operator  $p(T)$  of Definition 2.4.2 equals  $\sum_i p(\lambda_i) P_i$ ; that is, it agrees with [Definition 12.8.2](#) applied to the restriction of  $p$  to  $\text{spec}(T)$ .
- (b) **(Algebra homomorphism)**  $(f + g)(T) = f(T) + g(T)$ ,  $(cf)(T) = c f(T)$ ,  $(fg)(T) =$

$f(T)g(T) = g(T)f(T)$ , and  $1(T) = \text{id}_V$ .

- (c) **(Adjoint)**  $f(T)^* = \overline{f}(T)$ , where  $\overline{f}(\lambda) = \overline{f(\lambda)}$ . In particular  $f(T)$  is self-adjoint when  $f$  is real-valued, and unitary when  $|f(\lambda_i)| = 1$  for every  $i$ .
- (d) **(Spectral mapping)**  $f(T)$  satisfies  $(*)$ , with the same orthonormal eigenbasis as  $T$ , and  $\text{spec}(f(T)) = f(\text{spec}(T))$ .
- (e) **(Commutation)** If  $S \in \mathcal{L}(V)$  commutes with  $T$ , then  $S$  commutes with every  $P_i$  and hence with  $f(T)$ . In particular  $T$  and  $f(T)$  commute.

*Idea.* Everything rests on one multiplication table,  $P_i P_j = \delta_{ij} P_i$ . It says that the operators  $\sum_i c_i P_i$  multiply exactly the way the coefficient lists  $(c_1, \dots, c_k)$  multiply coordinatewise, so  $\mathcal{L}(V)$  contains a copy of the algebra of all  $F$ -valued functions on  $\text{spec}(T)$ , and  $f \mapsto f(T)$  is the obvious map into it. That makes (b) and (c) one-line computations. The one statement needing an argument is (a) — that the copy sitting inside is the *same* copy Chapter 2 built by substituting  $T$  into polynomials — and the argument is to compute the powers of  $T$ . For (e), a commuting  $S$  preserves each eigenspace, and an operator preserving all the summands of an orthogonal direct sum commutes with the projections onto them.

*Proof.*

- (a) We first show that

$$T^m = \sum_{i=1}^k \lambda_i^m P_i \quad \text{for every integer } m \geq 0,$$

with the usual convention  $\lambda^0 = 1$ . For  $m = 0$  both sides equal  $\text{id}_V$ , by [Theorem 12.8.1](#) (c). Assuming the formula for  $m$ , [Theorem 12.8.1](#) (d) and (b) give

$$\begin{aligned} T^{m+1} &= T \cdot T^m = \left( \sum_i \lambda_i P_i \right) \left( \sum_j \lambda_j^m P_j \right) \\ &= \sum_{i,j} \lambda_i \lambda_j^m P_i P_j = \sum_i \lambda_i^{m+1} P_i, \end{aligned}$$

since  $P_i P_j = 0$  for  $i \neq j$  and  $P_i^2 = P_i$ . Now let  $p = a_0 + a_1 x + \dots + a_N x^N$ . By [Definition 2.4.2](#) and the formula just proved,

$$\begin{aligned} p(T) &= \sum_{m=0}^N a_m T^m = \sum_{m=0}^N a_m \sum_{i=1}^k \lambda_i^m P_i \\ &= \sum_{i=1}^k \left( \sum_{m=0}^N a_m \lambda_i^m \right) P_i = \sum_{i=1}^k p(\lambda_i) P_i, \end{aligned}$$

the middle step being a rearrangement of a finite double sum.

- (b) The first two identities are immediate from [Definition 12.8.2](#). For the product, [Theorem 12.8.1](#) (b) gives

$$f(T)g(T) = \sum_{i,j} f(\lambda_i)g(\lambda_j)P_i P_j = \sum_i f(\lambda_i)g(\lambda_i)P_i,$$

which is  $(fg)(T)$ . The right-hand side is symmetric in  $f$  and  $g$ , so  $f(T)g(T) = g(T)f(T)$ . Finally  $1(T) = \sum_i P_i = \text{id}_V$  by [Theorem 12.8.1 \(c\)](#).

- (c) By [Theorem 11.6.2 \(a\)](#) and (b), and  $P_i^* = P_i$ ,

$$f(T)^* = \sum_{i=1}^k \overline{f(\lambda_i)} P_i^* = \overline{f}(T).$$

If  $f$  is real-valued then  $\overline{f} = f$  and  $f(T)^* = f(T)$ . If  $|f(\lambda_i)| = 1$  for all  $i$ , then  $\overline{f}f \equiv 1$ , so by (b)  $f(T)^*f(T) = (\overline{f}f)(T) = \text{id}_V$ , and likewise  $f(T)f(T)^* = \text{id}_V$ ; so  $f(T)$  is unitary ([Definition 11.7.2](#)).

- (d) Let  $\mathcal{B} = (e_1, \dots, e_n)$  be an orthonormal eigenbasis of  $T$ , with  $Te_j = \mu_j e_j$ , and fix  $j$ . Let  $i$  be the index with  $\mu_j = \lambda_i$ . Then  $e_j \in E_i$ , so  $P_i e_j = e_j$ , while  $P_l e_j = 0$  for  $l \neq i$  because  $E_i \subseteq E_l^\perp = \ker P_l$ , by [Theorem 12.8.1 \(a\)](#) and [Definition 11.3.2](#). Hence

$$f(T)e_j = f(\lambda_i)e_j = f(\mu_j)e_j.$$

So  $\mathcal{B}$  is an orthonormal basis of eigenvectors of  $f(T)$ , which is  $(*)$  for  $f(T)$ . The matrix of  $f(T)$  in  $\mathcal{B}$  is  $\text{diag}(f(\mu_1), \dots, f(\mu_n))$ , so its eigenvalues are exactly the numbers  $f(\mu_j)$ , that is the numbers  $f(\lambda_i)$ ; hence  $\text{spec}(f(T)) = f(\text{spec}(T))$ .

- (e) Suppose  $ST = TS$ . Each  $E_i$  is  $S$ -invariant: if  $Tv = \lambda_i v$ , then  $T(Sv) = S(Tv) = \lambda_i Sv$ . Let  $v \in V$  and write  $v = \sum_i P_i v$  with  $P_i v \in E_i$  ([Theorem 12.8.1 \(c\)](#)). Applying  $S$  gives  $Sv = \sum_i SP_i v$  with  $SP_i v \in E_i$ , and this is the decomposition of  $Sv$  in  $E_1 \oplus \dots \oplus E_k$ , which is unique. Comparing it with  $Sv = \sum_i P_i(Sv)$  gives  $P_i Sv = SP_i v$  for every  $i$ , that is  $SP_i = P_i S$ . Hence  $S$  commutes with  $f(T) = \sum_i f(\lambda_i)P_i$ . Since  $T$  commutes with itself, the last claim follows. This proves the theorem. ■

Part (a) is what makes the notation honest. Writing  $\sqrt{T}$ ,  $e^T$ , and  $p(T)$  in the same paragraph is legitimate precisely because the polynomial case is common ground: the new definition extends the old one instead of competing with it. It also reconciles this section with Chapter 10. An operator satisfying  $(*)$  is diagonalizable, so every Jordan block of its matrix has size 1, and [Definition 10.9.1](#) consults no derivatives at all — it too returns  $\sum_i f(\lambda_i)\mathbf{P}_i$ . The two constructions of  $f(\mathbf{A})$  agree for normal  $\mathbf{A}$ , and Chapter 10's extra machinery is precisely the price of dropping normality.

**Warning.** Outside the normal case,  $f$  on the spectrum is not enough data. For  $\mathbf{A} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$  the only eigenvalue is 0, so the recipe “ $f(\mathbf{A}) = \sum_i f(\lambda_i)\mathbf{P}_i$ ” would read  $f(\mathbf{A}) = f(0)\mathbf{I}$  for every  $f$  — and with  $f(x) = x$  it would announce  $\mathbf{A} = \mathbf{O}$ . The recipe is not wrong, it is inapplicable:  $\mathbf{A}$  is not normal, so there is no resolution of the identity to expand in. Chapter 10's definition of  $f(\mathbf{A})$  has to consult  $f'(0)$  as well ([Definition 10.9.1](#)), and that derivative is exactly what the missing orthogonality would have spared us.

*Quick check.* Let  $T \in \mathcal{L}(V)$  satisfy  $(*)$  with  $\text{spec}(T) = \{0, 1\}$ . Identify  $T$ .

*Solution.*  $T$  is the orthogonal projection onto  $E_1(T)$ . Indeed the spectral resolution reads  $T = 0 \cdot P_1 + 1 \cdot P_2 = P_2 = P_{E_1(T)}$ . So the operators satisfying  $(*)$  with spectrum in  $\{0, 1\}$  are exactly the orthogonal projections, together with 0 and  $\text{id}_V$  for the degenerate spectra  $\{0\}$  and  $\{1\}$ .

## 12.8.5 Square roots

The function  $\lambda \mapsto \sqrt{\lambda}$  is defined on  $[0, \infty)$ , so [Definition 12.8.2](#) produces a square root of  $T$  as soon as the spectrum lies there. The interesting part is that the square root is **unique** once we ask it to be of the same kind, and the uniqueness proof is a direct application of [Theorem 12.8.2](#): a square root's own resolution, squared, has to be  $T$ 's.

Note first that over  $\mathbb{C}$  the hypothesis is stronger than it looks. If  $T$  satisfies  $(*)$  and  $\text{spec}(T) \subseteq \mathbb{R}$ , then  $f(\lambda) = \lambda$  is real-valued on the spectrum, so  $T = f(T)$  is self-adjoint by [Theorem 12.8.3](#) (c). A normal operator with real spectrum is self-adjoint.

### Theorem 12.8.4. The Square Root of a Nonnegative Normal Operator

Let  $V$  be a nonzero finite-dimensional inner product space over  $F$  and let  $T \in \mathcal{L}(V)$  satisfy  $(*)$  with  $\text{spec}(T) \subseteq [0, \infty)$ . Then there is **exactly one** operator  $S \in \mathcal{L}(V)$  such that

- (i)  $S$  satisfies  $(*)$ ;
- (ii)  $\text{spec}(S) \subseteq [0, \infty)$ ;
- (iii)  $S^2 = T$ .

It is  $S = \sum_i \sqrt{\lambda_i} P_i$ , where  $T = \sum_i \lambda_i P_i$  is the spectral resolution of  $T$ . This  $S$  is self-adjoint, and it commutes with every operator that commutes with  $T$ .

*Idea.* Existence is the functional calculus applied to  $\sqrt{\cdot}$ : squaring a function squares its values. Uniqueness is where the work is, and the trick is not to compare  $S$  with the candidate directly but to *square  $S$ 's own resolution*. If  $S = \sum_j \mu_j Q_j$ , then  $S^2 = \sum_j \mu_j^2 Q_j$ , and the  $\mu_j^2$  are still distinct because the  $\mu_j$  are distinct **and nonnegative**. So  $\sum_j \mu_j^2 Q_j$  meets every condition of [Theorem 12.8.2](#) for  $T$ , which forces it to be  $T$ 's spectral resolution term by term.

*Proof. Existence.* Let  $f(\lambda) = \sqrt{\lambda}$  on  $\text{spec}(T) \subseteq [0, \infty)$ , which takes values in  $[0, \infty) \subseteq F$ , and put  $S = f(T) = \sum_i \sqrt{\lambda_i} P_i$ . Then  $f^2$  is the function  $\lambda \mapsto \lambda$ , so by [Theorem 12.8.3](#) (b) and (a),

$$S^2 = f(T)f(T) = (f^2)(T) = T,$$

giving (iii). By [Theorem 12.8.3](#) (d),  $S$  satisfies  $(*)$  and  $\text{spec}(S) = \{\sqrt{\lambda_1}, \dots, \sqrt{\lambda_k}\} \subseteq [0, \infty)$ , giving (i) and (ii). Since  $f$  is real-valued,  $S$  is self-adjoint by [Theorem 12.8.3](#) (c), and it commutes with everything commuting with  $T$  by (e) of the same theorem.

*Uniqueness.* Let  $S$  satisfy (i), (ii), and (iii), and let

$$S = \mu_1 Q_1 + \dots + \mu_l Q_l$$

be its spectral resolution ([Theorem 12.8.1](#)): the  $\mu_j$  are the distinct eigenvalues of  $S$ , and the

$Q_j$  are nonzero self-adjoint idempotents with  $Q_j Q_{j'} = 0$  for  $j \neq j'$  and  $\sum_j Q_j = \text{id}_V$ . By [Theorem 12.8.3](#) (a) applied to  $S$  and  $p = x^2$ ,

$$T = S^2 = \sum_{j=1}^l \mu_j^2 Q_j.$$

The scalars  $\mu_j^2$  are **distinct**: the  $\mu_j$  are distinct by construction, and nonnegative by (ii), so  $\mu_j^2 = \mu_{j'}^2$  would force  $\mu_j = \mu_{j'}$ . The family  $(Q_1, \dots, Q_l)$  satisfies (R1), (R2), and (R3) of [Theorem 12.8.2](#), and every  $Q_j$  is nonzero. That theorem therefore identifies the displayed expression as the spectral resolution of  $T$ : after renumbering,  $l = k$ ,  $\mu_j^2 = \lambda_j$ , and  $Q_j = P_j$  for every  $j$ . Since  $\mu_j \geq 0$  and  $\mu_j^2 = \lambda_j$ , we get  $\mu_j = \sqrt{\lambda_j}$ , and hence

$$S = \sum_{j=1}^k \sqrt{\lambda_j} P_j,$$

which is the operator constructed above. This proves the theorem. ■

The square root is written  $\sqrt{T}$ , or  $T^{1/2}$ . Chapter 13 takes the hypothesis  $\text{spec}(T) \subseteq [0, \infty)$  as the definition of a **positive** operator, proves that it is equivalent to  $\langle Tv, v \rangle \geq 0$  for all  $v$ , and makes this theorem the tool behind polar decomposition and the singular value decomposition; the notation  $\mathbf{A}^{1/2}$  in the notation table refers to exactly the operator built here.

### Example. A Square Root, and Three Impostors

Let  $\mathbf{A} = \begin{pmatrix} 10 & 6 \\ 6 & 10 \end{pmatrix} \in M_2(\mathbb{R})$ . Find its spectral resolution, compute  $\sqrt{\mathbf{A}}$  and  $\mathbf{A}^m$ , and find a symmetric square root of  $\mathbf{A}$  other than  $\sqrt{\mathbf{A}}$ .

*Solution. The resolution.*  $\mathbf{A}$  is symmetric, hence self-adjoint on  $\mathbb{R}^2$ . Since  $\mathbf{A}(1, 1) = (16, 16)$  and  $\mathbf{A}(1, -1) = (4, -4)$ , the eigenvalues are 16 and 4, with eigenlines  $\text{span}((1, 1))$  and  $\text{span}((1, -1))$ . Both spanning vectors have squared length 2, so

$$\mathbf{P}_1 = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad \mathbf{P}_2 = \frac{1}{2} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}, \quad \mathbf{A} = 16\mathbf{P}_1 + 4\mathbf{P}_2.$$

*The square root.* Both eigenvalues are positive, so [Theorem 12.8.4](#) applies:

$$\sqrt{\mathbf{A}} = 4\mathbf{P}_1 + 2\mathbf{P}_2 = \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix} + \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix},$$

and indeed  $\begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}^2 = \begin{pmatrix} 10 & 6 \\ 6 & 10 \end{pmatrix}$ .

*Powers.* By [Theorem 12.8.3](#) (a),  $\mathbf{A}^m = 16^m \mathbf{P}_1 + 4^m \mathbf{P}_2$ , that is

$$\mathbf{A}^m = \frac{1}{2} \begin{pmatrix} 16^m + 4^m & 16^m - 4^m \\ 16^m - 4^m & 16^m + 4^m \end{pmatrix}.$$

For  $m = 3$  this reads  $\begin{pmatrix} 2080 & 2016 \\ 2016 & 2080 \end{pmatrix}$ , which is  $\mathbf{A}^3$ .

*An impostor.* Changing one sign gives  $\mathbf{B} = 4\mathbf{P}_1 - 2\mathbf{P}_2 = \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix}$ , and  $\mathbf{B}^2 = \mathbf{A}$  as well.

So  $\mathbf{A}$  has more than one symmetric square root;  $\mathbf{B}$  fails clause (ii) of [Theorem 12.8.4](#), since  $\text{spec}(\mathbf{B}) = \{4, -2\}$ . Flipping the other sign, and both, gives  $-\mathbf{B}$  and  $-\sqrt{\mathbf{A}}$ : four symmetric square roots, and exactly one of them has nonnegative spectrum.

*Warning.* “**A square root**” and “**the square root**” are different things. Uniqueness in [Theorem 12.8.4](#) needs all three clauses. Dropping (ii) admits the three sign-flipped roots of [Example \(A Square Root, and Three Impostors\)](#). Dropping (i) admits more still, and the cheapest witness is the zero operator on  $\mathbb{C}^2$ . Take

$$\mathbf{S} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \mathbf{S}^2 = \mathbf{O}.$$

Here  $\text{spec}(\mathbf{S}) = \{0\} \subseteq [0, \infty)$ , so (ii) holds, and (iii) holds by the display, yet  $\mathbf{S} \neq \mathbf{O} = \sqrt{\mathbf{O}}$ . Only (i) fails:  $\mathbf{S}^*\mathbf{S} = \text{diag}(0, 1)$  while  $\mathbf{S}\mathbf{S}^* = \text{diag}(1, 0)$ . The phrase “**the square root of T**” always means the one satisfying all three.

## 12.8.6 The exponential and the unitary group

Chapter 10 defined  $e^{\mathbf{A}}$  for every  $\mathbf{A} \in M_n(\mathbb{C})$  as the sum of the series  $\sum_m \mathbf{A}^m/m!$  (Definition 10.9.2) and proved that the series converges (Theorem 10.9.3). We do not redevelop any of that. What the spectral resolution adds is that for a normal matrix the series can be summed by inspection, and that the answer identifies the unitary matrices as the exponentials of the Hermitian ones — the matrix version of  $|e^{i\theta}| = 1$ .

### Theorem 12.8.5. Unitary Matrices Are Exponentials of Hermitian Matrices

Let  $n \geq 1$ .

- (a) If  $\mathbf{A} \in M_n(\mathbb{C})$  is normal with spectral resolution  $\mathbf{A} = \sum_i \lambda_i \mathbf{P}_i$ , then  $e^{\mathbf{A}} = \sum_i e^{\lambda_i} \mathbf{P}_i$ .
- (b) If  $\mathbf{H} \in M_n(\mathbb{C})$  is Hermitian, then  $e^{i\mathbf{H}} \in \text{U}(n)$ , and  $\det e^{i\mathbf{H}} = e^{i \text{tr } \mathbf{H}}$ .
- (c) Conversely, every  $\mathbf{U} \in \text{U}(n)$  equals  $e^{i\mathbf{H}}$  for some Hermitian  $\mathbf{H}$ , which may be chosen with  $\text{spec}(\mathbf{H}) \subseteq [0, 2\pi)$ .

*Idea.* For (a), the partial sums of the series collapse:  $\mathbf{A}^m = \sum_i \lambda_i^m \mathbf{P}_i$ , so each partial sum is a fixed combination of the  $k$  scalar partial sums of  $\sum_m \lambda_i^m/m!$ , whose limits are the numbers  $e^{\lambda_i}$ . For (b) and (c), the resolution turns everything into a statement about  $k$  complex numbers:  $\mathbf{H}$  Hermitian means real eigenvalues,  $e^{i\mathbf{H}}$  then has eigenvalues on the unit circle, and a unitary matrix has eigenvalues on the unit circle, each of which is  $e^{i\theta}$  for a unique  $\theta \in [0, 2\pi)$ . Reading the last sentence backwards constructs  $\mathbf{H}$ .

*Proof.*

- (a) By Theorem 12.8.3 (a),  $\mathbf{A}^m = \sum_i \lambda_i^m \mathbf{P}_i$  for every  $m \geq 0$ , so for every  $N \geq 0$ ,

$$\sum_{m=0}^N \frac{\mathbf{A}^m}{m!} = \sum_{i=1}^k \left( \sum_{m=0}^N \frac{\lambda_i^m}{m!} \right) \mathbf{P}_i.$$

Fix a position  $(r, s)$ . Its entry on the right is  $\sum_i c_{N,i} (\mathbf{P}_i)_{rs}$ , where  $c_{N,i} = \sum_{m \leq N} \lambda_i^m / m!$  and the coefficients  $(\mathbf{P}_i)_{rs}$  do not depend on  $N$ . Each  $c_{N,i} \rightarrow e^{\lambda_i}$  as  $N \rightarrow \infty$  by the scalar exponential series, so the entry converges to  $\sum_i e^{\lambda_i} (\mathbf{P}_i)_{rs}$ . Hence the series converges entrywise to  $\sum_i e^{\lambda_i} \mathbf{P}_i$ , and by Definition 10.9.2 this matrix is  $e^{\mathbf{A}}$ .

- (b)  $\mathbf{H}$  is normal, so it has a spectral resolution  $\mathbf{H} = \sum_j \theta_j \mathbf{P}_j$ , and its eigenvalues  $\theta_j$  are real by Theorem 12.2.1. The scalars  $i\theta_j$  are distinct, so  $i\mathbf{H} = \sum_j i\theta_j \mathbf{P}_j$  is the spectral resolution of  $i\mathbf{H}$  by Theorem 12.8.2, and (a) gives

$$e^{i\mathbf{H}} = \sum_j e^{i\theta_j} \mathbf{P}_j.$$

Each  $\theta_j$  is real, so  $\overline{e^{i\theta_j}} e^{i\theta_j} = |e^{i\theta_j}|^2 = 1$ . Hence, using  $\mathbf{P}_j^* = \mathbf{P}_j$  and  $\mathbf{P}_j \mathbf{P}_{j'} = \delta_{jj'} \mathbf{P}_j$  from Theorem 12.8.1 (b),

$$(e^{i\mathbf{H}})^* e^{i\mathbf{H}} = \sum_j \overline{e^{i\theta_j}} e^{i\theta_j} \mathbf{P}_j = \sum_j \mathbf{P}_j = \mathbf{I},$$

so  $e^{i\mathbf{H}} \in \mathbf{U}(n)$  by Definition 11.7.3. For the determinant, Theorem 10.9.4 (e) gives  $\det e^{i\mathbf{H}} = e^{\text{tr}(i\mathbf{H})} = e^{i \text{tr} \mathbf{H}}$ .

- (c) Let  $\mathbf{U} \in \mathbf{U}(n)$ . Then  $\mathbf{U}^* \mathbf{U} = \mathbf{U} \mathbf{U}^* = \mathbf{I}$ , so  $\mathbf{U}$  is normal; let  $\mathbf{U} = \sum_j \nu_j \mathbf{P}_j$  be its spectral resolution. Every eigenvalue has  $|\nu_j| = 1$ : if  $\mathbf{U}\mathbf{x} = \nu_j \mathbf{x}$  with  $\mathbf{x} \neq \mathbf{0}$ , then  $\|\mathbf{x}\| = \|\mathbf{U}\mathbf{x}\| = |\nu_j| \|\mathbf{x}\|$  because  $\mathbf{U}$  is an isometry (Theorem 11.7.3), and  $\|\mathbf{x}\| \neq 0$ . So there is a unique  $\theta_j \in [0, 2\pi)$  with  $\nu_j = e^{i\theta_j}$ , and the  $\theta_j$  are distinct because the  $\nu_j$  are. Put

$$\mathbf{H} := \sum_j \theta_j \mathbf{P}_j.$$

The  $\theta_j$  are real, so  $\mathbf{H}^* = \sum_j \theta_j \mathbf{P}_j^* = \mathbf{H}$  by Theorem 11.6.2 (a), (b) and Theorem 12.8.1 (b):  $\mathbf{H}$  is Hermitian, with  $\text{spec}(\mathbf{H}) = \{\theta_j\} \subseteq [0, 2\pi)$  by Theorem 12.8.2. By (a) applied to  $i\mathbf{H}$ ,

$$e^{i\mathbf{H}} = \sum_j e^{i\theta_j} \mathbf{P}_j = \sum_j \nu_j \mathbf{P}_j = \mathbf{U}.$$

This proves the theorem. ■

Part (c) says that the exponential map from Hermitian matrices to  $\mathbf{U}(n)$  is **surjective**. One consequence belongs to topology rather than algebra, so we only record it: the path  $t \mapsto e^{it\mathbf{H}}$ ,  $t \in [0, 1]$ , stays inside  $\mathbf{U}(n)$  and runs from  $\mathbf{I}$  to  $\mathbf{U}$ , so every unitary matrix can be deformed continuously to the identity. Nothing of the sort holds in  $\mathbf{O}(n)$ , where the determinant is  $\pm 1$  and cannot change sign along a path.

**Example. Exponentiating a Hermitian Matrix, and Undoing It**

Let  $\mathbf{H} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ . Compute  $e^{i\theta\mathbf{H}}$  for  $\theta \in \mathbb{R}$ . Then find a Hermitian  $\mathbf{K}$  with  $e^{i\mathbf{K}} = \mathbf{Q}$ , where  $\mathbf{Q} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$ .

*Solution. Forwards.*  $\mathbf{H}$  is symmetric with  $\mathbf{H}(1, 1) = (1, 1)$  and  $\mathbf{H}(1, -1) = -(1, -1)$ , so its eigenvalues are  $\pm 1$  and

$$\mathbf{H} = \mathbf{P}_+ - \mathbf{P}_-, \quad \mathbf{P}_{\pm} = \frac{1}{2} \begin{pmatrix} 1 & \pm 1 \\ \pm 1 & 1 \end{pmatrix}.$$

Hence  $i\theta\mathbf{H}$  has resolution  $i\theta\mathbf{P}_+ - i\theta\mathbf{P}_-$  (for  $\theta \neq 0$ ), and [Theorem 12.8.5](#) (a) gives

$$\begin{aligned} e^{i\theta\mathbf{H}} &= e^{i\theta\mathbf{P}_+} + e^{-i\theta\mathbf{P}_-} \\ &= \begin{pmatrix} \cos \theta & i \sin \theta \\ i \sin \theta & \cos \theta \end{pmatrix}, \end{aligned}$$

using  $e^{\pm i\theta} = \cos \theta \pm i \sin \theta$ . Each row has squared length  $\cos^2 \theta + \sin^2 \theta = 1$  and the two rows are orthogonal, confirming that the result is unitary.

*Backwards.*  $\mathbf{Q}$  is the rotation by  $\pi/4$ , hence unitary, with  $\mathbf{Q}(1, -i) = e^{i\pi/4}(1, -i)$  and  $\mathbf{Q}(1, i) = e^{-i\pi/4}(1, i)$ . Both eigenvectors have squared length 2, so the spectral projections are

$$\mathbf{R}_+ = \frac{1}{2} \begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix}, \quad \mathbf{R}_- = \frac{1}{2} \begin{pmatrix} 1 & -i \\ i & 1 \end{pmatrix}.$$

Writing  $e^{-i\pi/4} = e^{i7\pi/4}$  to land in  $[0, 2\pi)$ , the recipe of [Theorem 12.8.5](#) (c) gives

$$\mathbf{K} = \frac{\pi}{4}\mathbf{R}_+ + \frac{7\pi}{4}\mathbf{R}_- = \begin{pmatrix} \pi & -\frac{3\pi i}{4} \\ \frac{3\pi i}{4} & \pi \end{pmatrix},$$

Hermitian with spectrum  $\{\pi/4, 7\pi/4\}$ . Leaving  $[0, 2\pi)$ , the smaller choice  $\frac{\pi}{4}(\mathbf{R}_+ - \mathbf{R}_-) = \frac{\pi}{4} \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}$  works just as well: the Hermitian logarithm is far from unique.

## 12.8.7 Exercises

### 12.8.7.1 A. Check your understanding

#### Exercise. A1

- Define the spectral resolution of  $T$ , and say which operators have one over  $\mathbb{C}$  and which over  $\mathbb{R}$ .
- List the properties of the family  $(P_1, \dots, P_k)$  given by [Theorem 12.8.1](#), and say how many terms the resolution of a self-adjoint  $T \in \mathcal{L}(\mathbb{R}^5)$  with eigenvalues  $2, 2, 2, 7, 7$  has.
- True or false: if  $T = \mu_1 Q_1 + \mu_2 Q_2$  with  $\mu_1 \neq \mu_2$  and  $Q_1, Q_2$  nonzero operators satisfying  $Q_i^2 = Q_i$ ,  $Q_1 Q_2 = Q_2 Q_1 = 0$ , and  $Q_1 + Q_2 = \text{id}_V$ , then  $T$  is normal. Justify your answer.

- (d) Define  $f(T)$  for a function  $f$  on  $\text{spec}(T)$ , and say exactly how much information about  $f$  it uses.
- (e) State [Theorem 12.8.3](#) (a) and explain in one sentence why the rest of the section would be unusable without it.
- (f) Give two different self-adjoint square roots of  $\text{id}_V$  on  $\mathbb{R}^2$ , and say which one [Theorem 12.8.4](#) singles out and why.

*Solution.*

- (a) For  $T$  with an orthonormal basis of eigenvectors, distinct eigenvalues  $\lambda_1, \dots, \lambda_k$ , and  $P_i = P_{E_{\lambda_i}(T)}$  the orthogonal projection onto the  $i$ -th eigenspace, the spectral resolution is  $T = \sum_i \lambda_i P_i$  ([Definition 12.8.1](#)). Over  $\mathbb{C}$  the operators with one are exactly the normal operators ([Theorem 12.4.2](#)); over  $\mathbb{R}$ , exactly the self-adjoint ones ([Theorem 12.5.1](#)).
- (b) Each  $P_i$  is self-adjoint and idempotent,  $P_i P_j = 0$  for  $i \neq j$ ,  $\sum_i P_i = \text{id}_V$ , and  $T P_i = P_i T = \lambda_i P_i$ . The resolution has **two** terms,  $T = 2P_1 + 7P_2$ , with  $\text{rank } P_1 = 3$  and  $\text{rank } P_2 = 2$ : the resolution counts distinct eigenvalues, not multiplicities.
- (c) False. Self-adjointness of the  $Q_i$  is missing, and it is the one hypothesis of [Theorem 12.8.2](#) that does the work. The matrices  $\pi_1, \pi_2$  of [Example \(Two Projections That Are Not Orthogonal\)](#) satisfy everything listed, with  $\mu_1 = 1$  and  $\mu_2 = 2$ , and  $\pi_1 + 2\pi_2 = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$ , which is not normal.
- (d)  $f(T) = \sum_i f(\lambda_i) P_i$  ([Definition 12.8.2](#)). It uses exactly the  $k$  values  $f(\lambda_1), \dots, f(\lambda_k)$ ; two functions agreeing on  $\text{spec}(T)$  give the same operator, and  $f$  need not be defined, let alone continuous, anywhere else.
- (e) For  $p \in F[x]$ , the operator  $\sum_i p(\lambda_i) P_i$  coincides with the operator  $p(T)$  of [Definition 2.4.2](#). Without it, the symbol  $p(T)$  would carry two unrelated meanings at once, and no computation mixing polynomials with other functions — for instance  $(\sqrt{T})^2 = T$  — could be trusted.
- (f)  $\text{id}$  and the reflection  $T_{\mathbf{D}}$  with  $\mathbf{D} = \text{diag}(1, -1)$ , both self-adjoint with square  $\text{id}$ . [Theorem 12.8.4](#) singles out  $\text{id}$ , because  $\text{spec}(\text{id}) = \{1\} \subseteq [0, \infty)$  while  $\text{spec}(T_{\mathbf{D}}) = \{1, -1\}$  is not. (A third root is  $-\text{id}$ , and a fourth is  $T_{\mathbf{D}'}$  with  $\mathbf{D}' = \text{diag}(-1, 1)$ .)

### 12.8.7.2 B. Practice

**Exercise. B1:** Write these as  $\sum \lambda_i P_i$

Find the spectral resolution of each of the following matrices, and check in each case that the projections are self-adjoint, idempotent, mutually annihilating, and sum to  $\mathbf{I}$ .

(a)  $\mathbf{A} = \begin{pmatrix} 2 & 1-i \\ 1+i & 3 \end{pmatrix} \in M_2(\mathbb{C})$ .

(b)  $\mathbf{S} = \begin{pmatrix} 2 & 0 & 1 \\ 0 & 3 & 0 \\ 1 & 0 & 2 \end{pmatrix} \in M_3(\mathbb{R})$ .

(c)  $\mathbf{U} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \in M_2(\mathbb{C})$ .

*Solution.*

- (a)  $\mathbf{A}^* = \mathbf{A}$ ,  $\text{tr } \mathbf{A} = 5$ , and  $\det \mathbf{A} = 6 - (1-i)(1+i) = 4$ , so  $p_{\mathbf{A}} = x^2 - 5x + 4$  and  $\text{spec}(\mathbf{A}) = \{1, 4\}$ . From  $\mathbf{A} - 4\mathbf{I} = \begin{pmatrix} -2 & 1-i \\ 1+i & -1 \end{pmatrix}$  the first row gives the eigenvector  $\mathbf{v}_4 = (1-i, 2)$ , with  $\|\mathbf{v}_4\|^2 = 2 + 4 = 6$ ; from  $\mathbf{A} - \mathbf{I}$  likewise  $\mathbf{v}_1 = (-1+i, 1)$ , with  $\|\mathbf{v}_1\|^2 = 3$ . Hence

$$\mathbf{P}_4 = \frac{\mathbf{v}_4 \mathbf{v}_4^*}{6} = \frac{1}{3} \begin{pmatrix} 1 & 1-i \\ 1+i & 2 \end{pmatrix}, \quad \mathbf{P}_1 = \frac{1}{3} \begin{pmatrix} 2 & -1+i \\ -1-i & 1 \end{pmatrix},$$

and  $\mathbf{A} = 1 \cdot \mathbf{P}_1 + 4 \cdot \mathbf{P}_4$ . Both are Hermitian, both square to themselves,  $\mathbf{P}_1 \mathbf{P}_4 = \mathbf{O}$ , and  $\mathbf{P}_1 + \mathbf{P}_4 = \mathbf{I}$ .

- (b)  $\mathbf{S}$  is symmetric, and  $\mathbf{S}(1, 0, 1) = 3(1, 0, 1)$ ,  $\mathbf{S}(0, 1, 0) = 3(0, 1, 0)$ ,  $\mathbf{S}(1, 0, -1) = (1, 0, -1)$ , so  $\text{spec}(\mathbf{S}) = \{3, 1\}$  with  $E_1 = \text{span}((1, 0, -1))$  and  $E_3 = \text{span}((1, 0, 1), (0, 1, 0))$ . Hence

$$\mathbf{P}_1 = \frac{1}{2} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{pmatrix}, \quad \mathbf{P}_3 = \mathbf{I} - \mathbf{P}_1 = \frac{1}{2} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix},$$

and  $\mathbf{S} = \mathbf{P}_1 + 3\mathbf{P}_3$ . The stated properties follow from  $\mathbf{P}_1^\top = \mathbf{P}_1$ ,  $\mathbf{P}_1^2 = \mathbf{P}_1$ , and  $\mathbf{P}_3 = \mathbf{I} - \mathbf{P}_1$ .

- (c)  $\mathbf{U}$  is real orthogonal, hence unitary and normal. Since  $\mathbf{U}(1, -i) = \frac{1}{\sqrt{2}}(1+i, 1-i) = e^{i\pi/4}(1, -i)$ , and similarly  $\mathbf{U}(1, i) = e^{-i\pi/4}(1, i)$ , we get  $\text{spec}(\mathbf{U}) = \{e^{\pm i\pi/4}\}$ . Both eigenvectors have squared length 2, so with  $\mathbf{v}_{\pm} = (1, \mp i)$ ,

$$\mathbf{P}_{\pm} = \frac{\mathbf{v}_{\pm} \mathbf{v}_{\pm}^*}{2} = \frac{1}{2} \begin{pmatrix} 1 & \pm i \\ \mp i & 1 \end{pmatrix}, \quad \mathbf{U} = e^{i\pi/4} \mathbf{P}_+ + e^{-i\pi/4} \mathbf{P}_-.$$

Both are Hermitian and idempotent,  $\mathbf{P}_+ \mathbf{P}_- = \mathbf{O}$  and  $\mathbf{P}_+ + \mathbf{P}_- = \mathbf{I}$ .

### Exercise. B2: Computing through the resolution

Let  $\mathbf{C} = \begin{pmatrix} 13 & 12 \\ 12 & 13 \end{pmatrix} \in M_2(\mathbb{R})$ .

- Find the spectral resolution of  $\mathbf{C}$ .
- Hence compute  $\sqrt{\mathbf{C}}$ , and verify by squaring.
- Hence give a closed formula for  $\mathbf{C}^m$  and write down  $\mathbf{C}^{100}$ .
- Hence compute  $e^{\mathbf{C}}$ .

*Solution.*

- (a)  $\mathbf{C}(1, 1) = 25(1, 1)$  and  $\mathbf{C}(1, -1) = (1, -1)$ , so  $\text{spec}(\mathbf{C}) = \{25, 1\}$  and, both spanning vectors having squared length 2,

$$\mathbf{P}_1 = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad \mathbf{P}_2 = \frac{1}{2} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}, \quad \mathbf{C} = 25\mathbf{P}_1 + \mathbf{P}_2.$$

(b) Both eigenvalues are positive, so by [Theorem 12.8.4](#),

$$\sqrt{\mathbf{C}} = 5\mathbf{P}_1 + 1 \cdot \mathbf{P}_2 = \begin{pmatrix} 3 & 2 \\ 2 & 3 \end{pmatrix},$$

$$\text{and } \begin{pmatrix} 3 & 2 \\ 2 & 3 \end{pmatrix}^2 = \begin{pmatrix} 13 & 12 \\ 12 & 13 \end{pmatrix} = \mathbf{C}.$$

(c) By [Theorem 12.8.3](#) (a),  $\mathbf{C}^m = 25^m\mathbf{P}_1 + 1^m\mathbf{P}_2$ , that is

$$\mathbf{C}^m = \frac{1}{2} \begin{pmatrix} 25^m + 1 & 25^m - 1 \\ 25^m - 1 & 25^m + 1 \end{pmatrix},$$

so  $\mathbf{C}^{100} = \frac{1}{2} \begin{pmatrix} 25^{100} + 1 & 25^{100} - 1 \\ 25^{100} - 1 & 25^{100} + 1 \end{pmatrix}$ . (As a check,  $m = 2$  gives  $\frac{1}{2} \begin{pmatrix} 626 & 624 \\ 624 & 626 \end{pmatrix} = \begin{pmatrix} 313 & 312 \\ 312 & 313 \end{pmatrix}$ , which is  $\mathbf{C}^2$ .)

(d) By [Theorem 12.8.5](#) (a),

$$e^{\mathbf{C}} = e^{25}\mathbf{P}_1 + e\mathbf{P}_2 = \frac{1}{2} \begin{pmatrix} e^{25} + e & e^{25} - e \\ e^{25} - e & e^{25} + e \end{pmatrix}.$$

### Exercise. B3: Is it a spectral resolution?

Determine which of the following displays the spectral resolution of the operator  $T_{\mathbf{A}}$  on the stated space. Justify your answer; for those that do not, name the condition of [Theorem 12.8.2](#) that fails and say what the correct resolution is.

- (a) On  $\mathbb{R}^2$ :  $\mathbf{A} = \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}$ , with  $\mu_1 = 2$ ,  $\mathbf{Q}_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ ,  $\mu_2 = -1$ ,  $\mathbf{Q}_2 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ .
- (b) On  $\mathbb{R}^2$ :  $\mathbf{A} = 3\mathbf{I}$ , with  $\mu_1 = \mu_2 = 3$ ,  $\mathbf{Q}_1 = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ ,  $\mathbf{Q}_2 = \frac{1}{2} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$ .
- (c) On  $\mathbb{R}^2$ :  $\mathbf{A} = \begin{pmatrix} 3 & 1 \\ 0 & 1 \end{pmatrix}$ , with  $\mu_1 = 3$ ,  $\mathbf{Q}_1 = \frac{1}{2} \begin{pmatrix} 2 & 1 \\ 0 & 0 \end{pmatrix}$ ,  $\mu_2 = 1$ ,  $\mathbf{Q}_2 = \frac{1}{2} \begin{pmatrix} 0 & -1 \\ 0 & 2 \end{pmatrix}$ .

*Solution.*

- (a) Yes. The two matrices are symmetric, idempotent, nonzero, annihilate each other, and sum to  $\mathbf{I}$ , and  $2\mathbf{Q}_1 - \mathbf{Q}_2 = \mathbf{A}$ ; the coefficients 2 and  $-1$  are distinct. By [Theorem 12.8.2](#) this is the spectral resolution, and indeed  $E_2 = \text{span}(\mathbf{e}_1)$ ,  $E_{-1} = \text{span}(\mathbf{e}_2)$ .
- (b) No: the scalars  $\mu_1, \mu_2$  are not **distinct**, so the hypothesis of [Theorem 12.8.2](#) fails at once. The identity  $3\mathbf{Q}_1 + 3\mathbf{Q}_2 = 3\mathbf{I} = \mathbf{A}$  is true but uninformative — the eigenvalue 3 has been split artificially across two projections. The spectral resolution of  $3\mathbf{I}$  has a single term,  $3\mathbf{I} = 3 \cdot \text{id}$ , since  $\text{spec}(3\mathbf{I}) = \{3\}$  and  $E_3 = \mathbb{R}^2$ .
- (c) No:  $\mathbf{Q}_1$  is not self-adjoint, since  $\mathbf{Q}_1^\top = \frac{1}{2} \begin{pmatrix} 2 & 0 \\ 1 & 0 \end{pmatrix} \neq \mathbf{Q}_1$ , so (R1) fails. Everything else holds:  $\mathbf{Q}_1^2 = \mathbf{Q}_1$ ,  $\mathbf{Q}_2 = \mathbf{I} - \mathbf{Q}_1$ ,  $\mathbf{Q}_1\mathbf{Q}_2 = \mathbf{Q}_2\mathbf{Q}_1 = \mathbf{O}$ , and  $3\mathbf{Q}_1 + \mathbf{Q}_2 = \mathbf{A}$ . There is no correct resolution to offer, because  $\mathbf{A}$  is not normal:  $\mathbf{A}^\top\mathbf{A} = \begin{pmatrix} 9 & 3 \\ 3 & 2 \end{pmatrix}$  while  $\mathbf{A}\mathbf{A}^\top =$

$\begin{pmatrix} 10 & 1 \\ 1 & 1 \end{pmatrix}$ . The projections here are onto  $E_3 = \text{span}(\mathbf{e}_1)$  along  $E_1 = \text{span}((-1, 2))$ , two lines that are not perpendicular.

### 12.8.7.3 C. Going deeper

#### Exercise. C1: The projections are polynomials in the operator

Let  $T \in \mathcal{L}(V)$  satisfy (\*), with distinct eigenvalues  $\lambda_1, \dots, \lambda_k \in F$  and resolution of the identity  $(P_1, \dots, P_k)$ .

- (a) Prove that  $P_i = \ell_i(T)$ , where  $\ell_i \in F[x]$  is the  $i$ -th Lagrange basis polynomial for the nodes  $\lambda_1, \dots, \lambda_k$ .
- (b) Deduce that for every  $f: \text{spec}(T) \rightarrow F$  there is a polynomial  $p$  of degree less than  $k$  with  $f(T) = p(T)$ .
- (c) Deduce that every operator commuting with  $T$  commutes with each  $P_i$ , re-proving [Theorem 12.8.3](#) (e) by a different route.

*Hint for (a): Definition 6.6.1 produces a polynomial taking prescribed values at prescribed nodes.*

*Solution.*

- (a) The  $\lambda_i$  are  $k$  distinct elements of  $F$ , so Definition 6.6.1 gives  $\ell_1, \dots, \ell_k \in F[x]$  of degree  $k - 1$ , and Theorem 6.6.1 (a) gives  $\ell_i(\lambda_j) = \delta_{ij}$ . By [Theorem 12.8.3](#) (a),

$$\ell_i(T) = \sum_{j=1}^k \ell_i(\lambda_j) P_j = \sum_{j=1}^k \delta_{ij} P_j = P_i.$$

- (b) Let  $p$  be the interpolating polynomial of degree less than  $k$  with  $p(\lambda_i) = f(\lambda_i)$  for every  $i$ , which exists by Theorem 6.6.1. By [Theorem 12.8.3](#) (a),

$$p(T) = \sum_i p(\lambda_i) P_i = \sum_i f(\lambda_i) P_i = f(T).$$

Explicitly,  $p = \sum_i f(\lambda_i) \ell_i$ .

- (c) If  $ST = TS$ , then  $S$  commutes with every polynomial in  $T$  by Theorem 6.7.1 (d), in particular with  $\ell_i(T) = P_i$ .

#### Exercise. C2: When the projections have rank one

Let  $T \in \mathcal{L}(V)$  satisfy (\*), with  $\dim V = n$  and resolution of the identity  $(P_1, \dots, P_k)$ .

- (a) Prove that  $\text{rank } P_i = \dim E_{\lambda_i}(T)$  and that  $\sum_i \text{rank } P_i = n$ .
- (b) Deduce that every  $P_i$  has rank 1 if and only if  $T$  has  $n$  distinct eigenvalues.
- (c) Suppose every  $P_i$  has rank 1. Prove that every  $S \in \mathcal{L}(V)$  commuting with  $T$  is a polynomial in  $T$ .

*Solution.*

- (a)  $\text{im } P_i = E_{\lambda_i}(T)$  by Definition 11.3.2, so  $\text{rank } P_i = \dim E_{\lambda_i}(T)$ . By Theorem 12.8.1 (a),  $V = E_1 \oplus \cdots \oplus E_k$ , so the dimensions add to  $n$ .
- (b) If every  $\text{rank } P_i = 1$ , then (a) gives  $k = n$ , so  $T$  has  $n$  distinct eigenvalues. Conversely, if  $T$  has  $n$  distinct eigenvalues, then  $k = n$  and (a) forces each of the  $n$  positive integers  $\text{rank } P_i$  to be 1.
- (c) By (b),  $k = n$  and each  $E_i$  is a line, say  $E_i = \text{span}(e_i)$  with  $(e_1, \dots, e_n)$  an orthonormal eigenbasis. Let  $ST = TS$ . As in the proof of Theorem 12.8.3 (e),  $S$  maps  $E_i$  into itself, so  $Se_i = c_i e_i$  for some  $c_i \in F$ . Define  $f: \text{spec}(T) \rightarrow F$  by  $f(\lambda_i) = c_i$ . Then  $f(T)e_i = c_i e_i = Se_i$  for every  $i$ , and two operators agreeing on a basis are equal, so  $S = f(T)$ . By Exercise (C1: The projections are polynomials in the operator) (b),  $f(T)$  is a polynomial in  $T$ .

**Exercise. C3: What the functional calculus forgets**

Let  $T \in \mathcal{L}(V)$  satisfy (\*), with distinct eigenvalues  $\lambda_1, \dots, \lambda_k$ , and let  $f, g: \text{spec}(T) \rightarrow F$ .

- (a) Prove that  $f(T) = 0$  if and only if  $f(\lambda_i) = 0$  for every  $i$ . Deduce that  $f(T) = g(T)$  if and only if  $f = g$ .
- (b) Deduce that  $\{p(T) : p \in F[x]\} = \text{span}(P_1, \dots, P_k)$ , a subspace of  $\mathcal{L}(V)$  of dimension exactly  $k$ .
- (c) Deduce that  $m_T = (x - \lambda_1)(x - \lambda_2) \cdots (x - \lambda_k)$ , so that the minimal polynomial of an operator satisfying (\*) has no repeated root.

*Solution.*

- (a)  $(\Leftarrow)$  is immediate from Definition 12.8.2. For  $(\Rightarrow)$ , fix  $i$  and pick  $v \in E_{\lambda_i}(T)$  with  $v \neq 0$ , which is possible because  $\lambda_i$  is an eigenvalue. Then  $P_i v = v$  and  $P_j v = 0$  for  $j \neq i$ , since  $E_i \subseteq \ker P_j$  (Theorem 12.8.1 (a), (b)). Hence

$$0 = f(T)v = \sum_j f(\lambda_j)P_j v = f(\lambda_i)v,$$

and  $v \neq 0$  forces  $f(\lambda_i) = 0$ . For the second claim, apply this to  $f - g$ , using Theorem 12.8.3 (b).

- (b) By Theorem 12.8.3 (a), every  $p(T)$  equals  $\sum_i p(\lambda_i)P_i$ , so the left-hand side is contained in  $\text{span}(P_1, \dots, P_k)$ ; conversely every  $\sum_i c_i P_i$  is  $f(T)$  for the function  $f(\lambda_i) = c_i$ , hence is a polynomial in  $T$  by Exercise (C1: The projections are polynomials in the operator) (b). The  $P_i$  are linearly independent: if  $\sum_i c_i P_i = 0$ , that operator is  $f(T)$  for  $f(\lambda_i) = c_i$ , so every  $c_i = 0$  by (a). Hence the dimension is  $k$ .
- (c) Put  $p = \prod_i (x - \lambda_i)$ , monic of degree  $k$ . Since  $p(\lambda_i) = 0$  for every  $i$ , part (a) gives  $p(T) = 0$ , so  $m_T \mid p$  by Theorem 9.5.1. Conversely  $m_T(T) = 0$ , so  $m_T(\lambda_i) = 0$  for every  $i$  by (a), and hence each  $x - \lambda_i$  divides  $m_T$  (Theorem 6.1.3 (b)). Distinct monic linear polynomials are coprime, since a common divisor would divide their difference, a nonzero constant; so  $p = \prod_i (x - \lambda_i)$  divides  $m_T$  by Lemma 6.6.3 (b). Two monic polynomials dividing each other are equal, so  $m_T = p$ , a product of distinct linear factors.