

12.4 The Complex Spectral Theorem

Two things are now on the table. Section 1 put every complex operator into upper triangular form in an **orthonormal** basis, and Section 3 identified the operators that commute with their adjoints. Put them together and the chapter's main theorem falls out in a few lines: a triangular matrix that is normal has nothing above the diagonal. This section proves that, reads off the consequences, and says exactly how much of the resulting picture is unique.

The field is $\mathbb{C}$ throughout this section, and every inner product space is finite-dimensional. The reason for insisting on $\mathbb{C}$ is Schur's theorem, and behind it the Fundamental Theorem of Algebra (Theorem 6.5.1): over $\mathbb{R}$ an operator can fail to have a single eigenvector, and Section 5 shows what has to change.

12.4.1 Triangular plus normal forces diagonal

Start with the smallest case, where the mechanism is visible at a glance. Let

$$\mathbf{A} = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \in M_2(\mathbb{C})$$

be upper triangular. The $(1, 1)$ entry of $\mathbf{A}^*\mathbf{A}$ is the squared length of the first **column** of $\mathbf{A}$, and the $(1, 1)$ entry of $\mathbf{A}\mathbf{A}^*$ is the squared length of the first **row**:

$$(\mathbf{A}^*\mathbf{A})_{11} = |a|^2, \quad (\mathbf{A}\mathbf{A}^*)_{11} = |a|^2 + |b|^2.$$

If $\mathbf{A}$ is normal these are equal, so $|b|^2 = 0$ and $b = 0$. One entry of one product did all the work. (This is Exercise (B3: Triangular and normal, by hand), which you may already have done.)

The general case is the same comparison run down the diagonal, one row at a time. Nothing is clever; the only care needed is to use, at step j , the zeros that steps $1, \dots, j-1$ have already produced.

Lemma 12.4.1. A Normal Triangular Matrix Is Diagonal

Let $\mathbf{T} = (t_{ij}) \in M_n(\mathbb{C})$ be upper triangular and normal. Then $\mathbf{T}$ is diagonal.

Idea. For any square matrix, the j -th diagonal entry of $\mathbf{T}^*\mathbf{T}$ is the squared length of column j , and the j -th diagonal entry of $\mathbf{T}\mathbf{T}^*$ is the squared length of row j . Normality says those two lengths agree for every j . For an upper triangular matrix, column j reaches only down to the diagonal and row j only rightwards from it, so the equation compares “what sits above t_{jj} ” with “what sits to the right of t_{jj} ”. Row 1 has nothing above it, so the right side must be empty too; that clears row 1, which empties the column sums one step further down, and the argument walks down the diagonal.

Proof. For each j , computing the (j, j) entries of the two products entry by entry,

$$(\mathbf{T}^*\mathbf{T})_{jj} = \sum_{i=1}^n \overline{t_{ij}} t_{ij} = \sum_{i=1}^n |t_{ij}|^2, \quad (\mathbf{T}\mathbf{T}^*)_{jj} = \sum_{k=1}^n |t_{jk}|^2.$$

Since $\mathbf{T}$ is upper triangular, $t_{ij} = 0$ whenever $i > j$, and $t_{jk} = 0$ whenever $k < j$. So normality of $\mathbf{T}$ says

$$\sum_{i \leq j} |t_{ij}|^2 = \sum_{k \geq j} |t_{jk}|^2 \quad (j = 1, \dots, n). \quad (12.4.1)$$

We prove by induction on j that $t_{jk} = 0$ for every $k > j$.

Base case $j = 1$. The left-hand side of Eq. 12.4.1 has the single term $|t_{11}|^2$, so

$$|t_{11}|^2 = |t_{11}|^2 + \sum_{k > 1} |t_{1k}|^2.$$

Hence $\sum_{k > 1} |t_{1k}|^2 = 0$, a sum of nonnegative reals, so $t_{1k} = 0$ for every $k > 1$.

Inductive step. Let $1 < j \leq n$ and suppose $t_{mk} = 0$ whenever $m < j$ and $k > m$. In the left-hand side of Eq. 12.4.1, every term with $i < j$ is $|t_{ij}|^2$ with $i < j$, hence with $j > i$; by the inductive hypothesis, taking $m = i$ and $k = j$, each such t_{ij} is 0. So the left-hand side reduces to $|t_{jj}|^2$ and Eq. 12.4.1 becomes

$$|t_{jj}|^2 = |t_{jj}|^2 + \sum_{k > j} |t_{jk}|^2,$$

whence $t_{jk} = 0$ for every $k > j$, again because a sum of nonnegative reals vanishes only when every term does.

By induction, every entry strictly above the diagonal is zero. Every entry strictly below it is zero because $\mathbf{T}$ is upper triangular. So $\mathbf{T}$ is diagonal, as claimed. ■

Note what the proof did **not** need: no eigenvector was chosen, no dimension was inducted on, no subspace was split off. The whole argument is an identity between two lists of nonnegative numbers. That is why it also survives, unchanged in spirit, into the block version used in Section 5 and the quantitative version in Section 11.

12.4.2 The theorem

Theorem 12.4.2. Complex Spectral Theorem

Let V be a finite-dimensional inner product space over $\mathbb{C}$ and let $T \in \mathcal{L}(V)$. The following are equivalent.

- (a) T is normal.
- (b) V has an orthonormal basis consisting of eigenvectors of T .

Idea. $(\Leftarrow)$ is a one-line check: in such a basis the matrix of T is diagonal, the matrix of T^* is its conjugate transpose, and two diagonal matrices commute. $(\Rightarrow)$ is Schur plus the lemma: take an orthonormal basis making the matrix triangular, observe that normality of T is normality of that matrix because the basis is orthonormal, and apply Lemma 12.4.1. A diagonal matrix in the basis $\mathcal{B}$ says precisely that each vector of $\mathcal{B}$ is an eigenvector.

Proof. ($\Leftarrow$) Suppose $\mathcal{B} = (e_1, \dots, e_n)$ is an orthonormal basis with $Te_j = \lambda_j e_j$ for each j . Then $\mathbf{D} := [T]_{\mathcal{B}}^{\mathcal{B}} = \text{diag}(\lambda_1, \dots, \lambda_n)$, and since $\mathcal{B}$ is orthonormal, $[T^*]_{\mathcal{B}}^{\mathcal{B}} = \mathbf{D}^* = \text{diag}(\overline{\lambda_1}, \dots, \overline{\lambda_n})$ by Theorem 11.6.3. Both products are the same diagonal matrix:

$$\mathbf{D}^* \mathbf{D} = \mathbf{D} \mathbf{D}^* = \text{diag}(|\lambda_1|^2, \dots, |\lambda_n|^2).$$

Taking the matrix of an operator in a fixed basis turns composition into matrix multiplication and is injective (Theorem 4.1.2, Theorem 4.1.4), so $T^*T = TT^*$.

($\Rightarrow$) Suppose T is normal. By Schur's theorem (Theorem 12.1.2) there is an orthonormal basis $\mathcal{B} = (e_1, \dots, e_n)$ of V in which $\mathbf{T} := [T]_{\mathcal{B}}^{\mathcal{B}}$ is upper triangular. Because $\mathcal{B}$ is orthonormal, Theorem 11.6.3 gives $[T^*]_{\mathcal{B}}^{\mathcal{B}} = \mathbf{T}^*$, so the equation $T^*T = TT^*$ reads $\mathbf{T}^* \mathbf{T} = \mathbf{T} \mathbf{T}^*$: the matrix $\mathbf{T}$ is normal. By Lemma 12.4.1, $\mathbf{T}$ is diagonal, say $\mathbf{T} = \text{diag}(\lambda_1, \dots, \lambda_n)$. Reading off the j -th column (Definition 4.1.1), $Te_j = \lambda_j e_j$, and $e_j \neq 0$ because $\|e_j\| = 1$. So $\mathcal{B}$ is an orthonormal basis of eigenvectors of T . This proves the theorem. ■

Corollary 12.4.3. Unitary Diagonalization

Let $\mathbf{A} \in M_n(\mathbb{C})$. Then $\mathbf{A}$ is normal if and only if there exist a unitary $\mathbf{U} \in U(n)$ and a diagonal $\mathbf{D} \in M_n(\mathbb{C})$ with

$$\mathbf{A} = \mathbf{U} \mathbf{D} \mathbf{U}^*.$$

A matrix admitting such a factorization is called **unitarily diagonalizable**. The diagonal of $\mathbf{D}$ is the list of eigenvalues of $\mathbf{A}$ with multiplicity, in an order that may be prescribed in advance, and the columns of $\mathbf{U}$ are a corresponding orthonormal basis of eigenvectors.

Proof. Apply Theorem 12.4.2 to the operator $T_{\mathbf{A}}$ on $\mathbb{C}^n$ with its standard inner product, for which the standard basis is orthonormal.

($\Rightarrow$) Let $(\mathbf{u}_1, \dots, \mathbf{u}_n)$ be an orthonormal basis of eigenvectors, $\mathbf{A} \mathbf{u}_j = \lambda_j \mathbf{u}_j$, and let $\mathbf{U}$ be the matrix with these columns; it is unitary because its columns are orthonormal (Theorem 11.7.3 (f)). Then $\mathbf{A} \mathbf{U}$ has j -th column $\lambda_j \mathbf{u}_j$, which is the j -th column of $\mathbf{U} \mathbf{D}$ for $\mathbf{D} = \text{diag}(\lambda_1, \dots, \lambda_n)$. So $\mathbf{A} \mathbf{U} = \mathbf{U} \mathbf{D}$ and $\mathbf{A} = \mathbf{U} \mathbf{D} \mathbf{U}^{-1} = \mathbf{U} \mathbf{D} \mathbf{U}^*$.

($\Leftarrow$) If $\mathbf{A} = \mathbf{U} \mathbf{D} \mathbf{U}^*$ then $\mathbf{A}^* = \mathbf{U} \mathbf{D}^* \mathbf{U}^*$, and

$$\mathbf{A}^* \mathbf{A} = \mathbf{U} \mathbf{D}^* \mathbf{D} \mathbf{U}^* = \mathbf{U} \mathbf{D} \mathbf{D}^* \mathbf{U}^* = \mathbf{A} \mathbf{A}^*,$$

using $\mathbf{U}^* \mathbf{U} = \mathbf{I}$ twice and the fact that diagonal matrices commute.

Finally, $\mathbf{A}$ is similar to $\mathbf{D}$, so they have the same characteristic polynomial (Theorem 7.10.3), and the roots of $\det(x\mathbf{I} - \mathbf{D}) = \prod_j (x - \lambda_j)$ are the diagonal entries with multiplicity. Reordering the eigenvectors reorders the columns of $\mathbf{U}$ and the diagonal of $\mathbf{D}$ in the same way, so any order may be prescribed. ■

Example. Unitarily diagonalizing a Hermitian matrix

Find a unitary $\mathbf{U}$ and a diagonal $\mathbf{D}$ with $\mathbf{U}^* \mathbf{A} \mathbf{U} = \mathbf{D}$, for

$$\mathbf{A} = \begin{pmatrix} 2 & 1-i \\ 1+i & 3 \end{pmatrix}.$$

Solution. $\mathbf{A}^* = \mathbf{A}$, so $\mathbf{A}$ is Hermitian, hence normal, and [Corollary 12.4.3](#) applies. Its characteristic polynomial is

$$\begin{aligned} x^2 - (\operatorname{tr} \mathbf{A})x + \det \mathbf{A} &= x^2 - 5x + (6 - (1-i)(1+i)) \\ &= x^2 - 5x + 4, \end{aligned}$$

since $(1-i)(1+i) = 1 - i^2 = 2$. So the eigenvalues are 4 and 1, both real as Theorem 12.2.1 promises.

For $\lambda = 4$, the first row of $\mathbf{A} - 4\mathbf{I} = \begin{pmatrix} -2 & 1-i \\ 1+i & -1 \end{pmatrix}$ gives $-2z_1 + (1-i)z_2 = 0$; taking $z_2 = 2$ gives $\mathbf{u} = (1-i, 2)$, with $\|\mathbf{u}\|^2 = 2 + 4 = 6$.

For $\lambda = 1$, the first row of $\mathbf{A} - \mathbf{I} = \begin{pmatrix} 1 & 1-i \\ 1+i & 2 \end{pmatrix}$ gives $z_1 = -(1-i)z_2$; taking $z_2 = 1$ gives $\mathbf{w} = (-1+i, 1)$, with $\|\mathbf{w}\|^2 = 2 + 1 = 3$.

The two are orthogonal, as Corollary 12.3.4 predicts:

$$\begin{aligned} \langle \mathbf{u}, \mathbf{w} \rangle &= (1-i)\overline{(-1+i)} + 2 \cdot \bar{1} \\ &= (1-i)(-1-i) + 2 = -2 + 2 = 0. \end{aligned}$$

Normalizing and using them as columns,

$$\mathbf{U} = \begin{pmatrix} \frac{1-i}{\sqrt{6}} & \frac{-1+i}{\sqrt{3}} \\ \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} \end{pmatrix}, \quad \mathbf{D} = \begin{pmatrix} 4 & 0 \\ 0 & 1 \end{pmatrix},$$

and $\mathbf{U}^* \mathbf{U} = \mathbf{I}$, $\mathbf{U}^* \mathbf{A} \mathbf{U} = \mathbf{D}$. As a check on the arithmetic, $\operatorname{tr} \mathbf{D} = 5 = \operatorname{tr} \mathbf{A}$ and $\det \mathbf{D} = 4 = \det \mathbf{A}$.

12.4.3 Reading the class off the spectrum

Once the eigenvalues sit on a diagonal, every question about T that is invariant under a unitary change of basis becomes a question about a list of complex numbers. Here are three answers.

Corollary 12.4.4. The Spectrum Decides

Let $T \in \mathcal{L}(V)$ be normal on a finite-dimensional complex inner product space.

- (a) T is self-adjoint if and only if every eigenvalue of T is real.
- (b) T is unitary if and only if every eigenvalue of T has modulus 1.
- (c) $\langle Tv, v \rangle$ is a nonnegative real number for every $v \in V$ if and only if every eigenvalue of

T is a nonnegative real.

Idea. By [Theorem 12.4.2](#) there is an orthonormal basis $(e_1, \dots, e_n)$ with $Te_j = \lambda_j e_j$, and then $T^*e_j = \overline{\lambda_j}e_j$ by [Theorem 12.3.3](#). Two operators that agree on a basis are equal, so each of the three conditions can be tested one basis vector at a time, where it becomes a statement about the single number λ_j .

Proof. Fix an orthonormal basis $(e_1, \dots, e_n)$ of eigenvectors, $Te_j = \lambda_j e_j$ ([Theorem 12.4.2](#)), and recall $T^*e_j = \overline{\lambda_j}e_j$ ([Theorem 12.3.3](#)). Every eigenvalue of T occurs among $\lambda_1, \dots, \lambda_n$, because $\text{diag}(\lambda_1, \dots, \lambda_n)$ is the matrix of T and its characteristic polynomial is $\prod_j (x - \lambda_j)$.

- (a) $T = T^*$ holds if and only if $Te_j = T^*e_j$ for every j ([Theorem 2.1.5](#)), that is $\lambda_j e_j = \overline{\lambda_j} e_j$, that is $\lambda_j = \overline{\lambda_j}$ since $e_j \neq 0$. A complex number equals its conjugate exactly when it is real.
- (b) $T^*Te_j = T^*(\lambda_j e_j) = \lambda_j \overline{\lambda_j} e_j = |\lambda_j|^2 e_j$, so $T^*T = \text{id}_V$ if and only if $|\lambda_j|^2 = 1$ for every j . That is what it means for T to be unitary ([Definition 11.7.2](#)): the equation $T^*T = \text{id}_V$ makes T injective, hence invertible on a finite-dimensional space ([Theorem 2.5.7](#)), with $T^{-1} = T^*$ and therefore $TT^* = \text{id}_V$ as well.
- (c) Write $v = \sum_j c_j e_j$. Expanding both slots and using $\langle e_j, e_i \rangle = \delta_{ij}$,

$$\langle Tv, v \rangle = \left\langle \sum_j c_j \lambda_j e_j, \sum_i c_i e_i \right\rangle = \sum_j \lambda_j |c_j|^2.$$

($\Leftarrow$) If every λ_j is a nonnegative real, this is a sum of nonnegative reals. ($\Rightarrow$) Taking $v = e_j$ gives $\langle Te_j, e_j \rangle = \lambda_j$, which is then a nonnegative real by hypothesis.

■

Part (c) names a class that Chapter 13 takes as its subject: the **positive** operators. Over $\mathbb{C}$ the hypothesis of normality is not really needed in (c), because $\langle Tv, v \rangle$ real for every v already forces $T = T^*$: apply [Theorem 12.2.6](#) to the operator $T - T^*$, for which $\langle (T - T^*)v, v \rangle = \langle Tv, v \rangle - \langle T^*v, v \rangle = 0$. Over $\mathbb{R}$ that shortcut is unavailable, and [Section 5](#) shows the quarter turn defeating it.

The next corollary is a small shock. The adjoint is defined through the inner product and has nothing to do with polynomials. For a normal operator it is one anyway.

Corollary 12.4.5. The Adjoint of a Normal Operator Is a Polynomial in It

Let $T \in \mathcal{L}(V)$ be normal on a finite-dimensional complex inner product space, with distinct eigenvalues $\lambda_1, \dots, \lambda_k$. Then there is a polynomial $p \in \mathbb{C}[x]$ of degree less than k with $T^* = p(T)$.

Proof. The scalars $\lambda_1, \dots, \lambda_k$ are distinct, so by Lagrange interpolation ([Theorem 6.6.1](#)) there is a polynomial $p \in \mathbb{C}[x]$ of degree less than k with

$$p(\lambda_i) = \overline{\lambda_i} \quad (i = 1, \dots, k).$$

Let $(e_1, \dots, e_n)$ be an orthonormal basis of eigenvectors (Theorem 12.4.2), with $Te_j = \mu_j e_j$ where each μ_j is one of the λ_i . Then $p(T)e_j = p(\mu_j)e_j = \overline{\mu_j}e_j = T^*e_j$, the first equality because $T^i e_j = \mu_j^i e_j$ for every $i \geq 0$, by induction from $Te_j = \mu_j e_j$, and the last by Theorem 12.3.3. The operators $p(T)$ and T^* agree on a basis, hence are equal (Theorem 2.1.5). ■

One consequence is worth stating on its own. Any operator S that commutes with T also commutes with every polynomial in T , hence with T^* . So for normal operators the apparently stronger requirement “commutes with T and with T^* ” is no requirement at all, and Section 7 will use this to diagonalize whole commuting families at once. The polynomial depends on T and not only on its size: for $T = \text{id}_V$ one may take the constant $p = 1$, while for the matrix $\text{diag}(1, i)$ the conditions $p(1) = 1$ and $p(i) = -i$ are met by

$$p(x) = ix + (1 - i),$$

as one checks by substituting $x = 1$ and $x = i$.

12.4.4 Exactly what is unique

The factorization $\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{U}^*$ is not unique, and it is worth being precise about which parts of it are.

Unique: the **set** of eigenvalues $\text{spec}(T)$, and for each $\lambda \in \text{spec}(T)$ the eigenspace $E_\lambda(T)$ with its dimension. These are defined from T alone, without any choice. Consequently the diagonal of $\mathbf{D}$ is unique **as a list with multiplicities up to reordering**: each λ appears exactly $\dim E_\lambda(T)$ times.

Not unique: the order of the diagonal, and the matrix $\mathbf{U}$. Two freedoms create all the others. First, permuting the eigenvectors permutes the columns of $\mathbf{U}$ and the diagonal of $\mathbf{D}$ in step. Second, within a single eigenspace $E_\lambda(T)$ of dimension m , **any** orthonormal basis of that eigenspace will do, and there are many when $m \geq 2$: the choices are related by the unitary maps of $E_\lambda(T)$. For $\mathbf{A} = \mathbf{I}_n$, for instance, every unitary $\mathbf{U}$ satisfies $\mathbf{U}^*\mathbf{A}\mathbf{U} = \mathbf{I}_n$. What no freedom can change is that eigenvectors for distinct eigenvalues stay orthogonal (Corollary 12.3.4), so the decomposition

$$V = E_{\lambda_1}(T) \oplus \cdots \oplus E_{\lambda_k}(T)$$

into mutually orthogonal eigenspaces is itself unique. Section 8 turns that decomposition into the spectral resolution and shows it determines T completely.

Quick check. The matrix $\mathbf{A} = \text{diag}(1, 1, 2)$ is normal. Write down two different unitary matrices $\mathbf{U}$ with $\mathbf{U}^*\mathbf{A}\mathbf{U} = \mathbf{A}$, and explain why, for $\text{diag}(1, 2, 3)$ with the order of the diagonal fixed, the freedom shrinks to a single unit-modulus scalar on each column.

Solution. Take $\mathbf{U}_1 = \mathbf{I}_3$ and $\mathbf{U}_2 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$; more generally $\mathbf{R} \oplus (1)$ works for **any**

$\mathbf{R} \in \text{U}(2)$, because $\mathbf{A} = \mathbf{I}_2 \oplus (2)$ and $\mathbf{R}^*\mathbf{I}_2\mathbf{R} = \mathbf{I}_2$. The freedom comes from the two-dimensional eigenspace $E_1(\mathbf{A}) = \text{span}(\mathbf{e}_1, \mathbf{e}_2)$, inside which every orthonormal basis is as good as any other. For $\text{diag}(1, 2, 3)$ each eigenspace is a line, and a line has exactly two unit vectors over $\mathbb{R}$ and a circle of them over $\mathbb{C}$; the residual freedom is only the scalar factor of

modulus 1 in each column, which changes $\mathbf{U}$ but not the decomposition into eigenspaces.

Warning. “**Diagonalizable**” and “**unitarily diagonalizable**” are different conditions. Every unitarily diagonalizable matrix is diagonalizable, because $\mathbf{U}$ is in particular invertible. The converse fails, and it fails for matrices as small and as ordinary as $\begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$. Having n distinct eigenvalues guarantees diagonalizability (Corollary 9.4.2) and guarantees nothing about orthogonality.

Example. Diagonalizable, and no orthonormal eigenbasis

For $\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} \in M_2(\mathbb{R})$, find an invertible $\mathbf{P}$ with $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ diagonal, and show that no orthogonal $\mathbf{Q}$ makes $\mathbf{Q}^\top \mathbf{A} \mathbf{Q}$ diagonal.

Solution. The matrix is upper triangular, so its eigenvalues are the diagonal entries 1 and 2 (Theorem 7.4.6). For $\lambda = 1$, $\mathbf{A} - \mathbf{I} = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}$ has kernel spanned by $(1, 0)$. For $\lambda = 2$, $\mathbf{A} - 2\mathbf{I} = \begin{pmatrix} -1 & 1 \\ 0 & 0 \end{pmatrix}$ has kernel spanned by $(1, 1)$. Two distinct eigenvalues in dimension 2 give a basis of eigenvectors (Corollary 9.4.2), so with

$$\mathbf{P} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{we get} \quad \mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}.$$

Now suppose $\mathbf{Q}^\top \mathbf{A} \mathbf{Q} = \mathbf{D}$ were diagonal with $\mathbf{Q}$ orthogonal. Then $\mathbf{A} = \mathbf{Q}\mathbf{D}\mathbf{Q}^\top$ would be normal by Corollary 12.4.3. But

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

so $\mathbf{A}$ is not normal and no such $\mathbf{Q}$ exists. The obstruction is visible without any theory: the eigenvectors are $(1, 0)$ and $(1, 1)$, the eigenspaces are lines and therefore the *only* candidates for eigenvector directions, and $(1, 0) \cdot (1, 1) = 1 \neq 0$. There simply is no orthogonal pair of eigenvectors to build $\mathbf{Q}$ from.

12.4.5 Exercises

12.4.5.1 A. Check your understanding

Exercise. A1

- (a) State the complex spectral theorem, including every hypothesis, and state its matrix form.
- (b) What goes wrong with the proof if the field is $\mathbb{R}$? Name the exact step.

- (c) True or false: a matrix in $M_n(\mathbb{C})$ with n distinct eigenvalues is unitarily diagonalizable. Justify your answer.
- (d) Let $\mathbf{A} \in M_3(\mathbb{C})$ be normal with $\text{spec}(\mathbf{A}) = \{2, 2, -1\}$ as a list with multiplicity. Is the unitary $\mathbf{U}$ in $\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{U}^*$ determined by $\mathbf{A}$ and the order of $\mathbf{D}$? What is determined?
- (e) A normal operator T satisfies $T^* = T^3$. What can you say about its eigenvalues?
- (f) Explain why [Lemma 12.4.1](#) cannot be weakened to “a normal matrix with some of its below-diagonal entries zero is diagonal”.

Solution.

- (a) For a finite-dimensional inner product space V over $\mathbb{C}$ and $T \in \mathcal{L}(V)$: T is normal if and only if V has an orthonormal basis of eigenvectors of T ([Theorem 12.4.2](#)). Matrix form: $\mathbf{A} \in M_n(\mathbb{C})$ is normal if and only if $\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{U}^*$ with $\mathbf{U}$ unitary and $\mathbf{D}$ diagonal ([Corollary 12.4.3](#)).
- (b) Schur’s theorem. It needs an eigenvalue at every stage of its induction, which over $\mathbb{C}$ comes from Theorem 6.5.1; over $\mathbb{R}$ an operator may have none at all, and then there is no triangular form to start from.
- (c) False. $\begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$ has the distinct eigenvalues 1 and 2 and is not normal, hence is not unitarily diagonalizable ([Example \(Diagonalizable, and no orthonormal eigenbasis\)](#)).
- (d) No. The eigenspace $E_2(\mathbf{A})$ is two-dimensional, and replacing its orthonormal basis by any other one changes $\mathbf{U}$ while leaving $\mathbf{D}$ alone. What is determined: $\text{spec}(\mathbf{A}) = \{2, -1\}$, the eigenspaces $E_2(\mathbf{A})$ and $E_{-1}(\mathbf{A})$, their dimensions 2 and 1, and their mutual orthogonality.
- (e) By Theorem 12.3.3, $Tv = \lambda v$ gives $T^*v = \bar{\lambda}v$, while $T^* = T^3$ gives $T^*v = \lambda^3v$. Hence $\bar{\lambda} = \lambda^3$. Taking moduli, $|\lambda| = |\lambda|^3$, so $|\lambda|$ is 0 or 1. If $\lambda \neq 0$, multiplying $\bar{\lambda} = \lambda^3$ by λ gives $|\lambda|^2 = \lambda^4$, that is $\lambda^4 = 1$. So every eigenvalue lies in $\{0, 1, -1, i, -i\}$.
- (f) The lemma uses triangularity at **every** index, not at a few places: the comparison of row and column lengths at step j is informative only because the whole of column j below the diagonal is already known to vanish. Zeros in some of those positions buy nothing. The circulant $\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}$ is normal and has zeros in positions $(2, 1)$ and $(3, 2)$, two of the three places below its diagonal, and it is not diagonal.

12.4.5.2 B. Practice

Exercise. B1: Unitarily diagonalize

For each matrix, find a unitary $\mathbf{U}$ and a diagonal $\mathbf{D}$ with $\mathbf{U}^*\mathbf{A}\mathbf{U} = \mathbf{D}$.

- (a) $\mathbf{A} = \begin{pmatrix} 3 & -i \\ i & 3 \end{pmatrix}$.
- (b) $\mathbf{A} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}$.

Solution.

- (a) $\mathbf{A}^* = \mathbf{A}$, so $\mathbf{A}$ is Hermitian and normal. Its characteristic polynomial is $x^2 - 6x + (9 - (-i)(i)) = x^2 - 6x + 8$, with roots 4 and 2.

For $\lambda = 4$: $\mathbf{A} - 4\mathbf{I} = \begin{pmatrix} -1 & -i \\ i & -1 \end{pmatrix}$, and the first row gives $z_1 = -iz_2$; take $\mathbf{u} = (-i, 1)$,

$\|\mathbf{u}\|^2 = 2$. For $\lambda = 2$: $\mathbf{A} - 2\mathbf{I} = \begin{pmatrix} 1 & -i \\ i & 1 \end{pmatrix}$, and the first row gives $z_1 = iz_2$; take $\mathbf{w} = (i, 1)$, $\|\mathbf{w}\|^2 = 2$. They are orthogonal: $\langle \mathbf{u}, \mathbf{w} \rangle = (-i)\bar{i} + 1 = (-i)(-i) + 1 = -1 + 1 = 0$. Hence

$$\mathbf{U} = \frac{1}{\sqrt{2}} \begin{pmatrix} -i & i \\ 1 & 1 \end{pmatrix}, \quad \mathbf{D} = \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix}.$$

- (b) $\mathbf{A}$ is unitary (Example (Which of these are normal?) (c)), hence normal. Here $\text{tr } \mathbf{A} = \sqrt{2}$ and $\det \mathbf{A} = \frac{1}{2}(1 - i^2) = 1$, so the characteristic polynomial is $x^2 - \sqrt{2}x + 1$, with roots

$$x = \frac{\sqrt{2} \pm \sqrt{2-4}}{2} = \frac{1 \pm i}{\sqrt{2}},$$

both of modulus 1, as [Corollary 12.4.4](#) (b) requires. For $\lambda = (1+i)/\sqrt{2}$, the equation $\mathbf{A}\mathbf{z} = \lambda\mathbf{z}$ reads $\frac{1}{\sqrt{2}}(z_1 + iz_2) = \frac{1+i}{\sqrt{2}}z_1$, that is $iz_2 = iz_1$, so $\mathbf{u} = (1, 1)$. For $\lambda = (1-i)/\sqrt{2}$ the same line gives $iz_2 = -iz_1$, so $\mathbf{w} = (1, -1)$. These are orthogonal, so

$$\mathbf{U} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad \mathbf{D} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1+i & 0 \\ 0 & 1-i \end{pmatrix}.$$

Here $\mathbf{U}$ is real orthogonal even though $\mathbf{A}$ is not real.

Exercise. B2: A circulant, diagonalized

Let $\omega = e^{2\pi i/3}$, so $\omega^3 = 1$ and $1 + \omega + \omega^2 = 0$, and let

$$\mathbf{C} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \in M_3(\mathbb{C}).$$

- (a) Verify that $\mathbf{C}$ is normal.
 (b) Show that $\mathbf{f}_j = \frac{1}{\sqrt{3}}(1, \omega^j, \omega^{2j})$ is an eigenvector of $\mathbf{C}$ for $j = 0, 1, 2$, and find the eigenvalue.
 (c) Deduce a unitary $\mathbf{F}$ with $\mathbf{F}^*\mathbf{C}\mathbf{F}$ diagonal, and write down the diagonal explicitly.

Solution.

- (a) Both products are the same matrix:

$$\mathbf{C}^*\mathbf{C} = \mathbf{C}\mathbf{C}^* = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}.$$

(The entries are dot products of pairs of rows, respectively columns, of $\mathbf{C}$; cyclic shifting makes the two lists agree.)

- (b) Write $\mathbf{f} = (1, \omega^j, \omega^{2j})$ before normalizing. Then

$$\mathbf{C}\mathbf{f} = (1 + \omega^j, \omega^j + \omega^{2j}, 1 + \omega^{2j}).$$

Using $\omega^{3j} = 1$, the second entry is $\omega^j(1 + \omega^j)$ and the third is $\omega^{2j}(\omega^{-2j} + 1) = \omega^{2j}(\omega^j + 1)$, since $\omega^{-2j} = \omega^j$. So $\mathbf{C}\mathbf{f} = (1 + \omega^j)\mathbf{f}$, and the eigenvalue is $1 + \omega^j$.

- (c) The three vectors $\mathbf{f}_0, \mathbf{f}_1, \mathbf{f}_2$ are orthonormal: each has norm $\frac{1}{\sqrt{3}}\sqrt{3} = 1$, and for $j \neq l$,

$$\langle \mathbf{f}_j, \mathbf{f}_l \rangle = \frac{1}{3} \sum_{m=0}^2 \omega^{mj} \overline{\omega^{ml}} = \frac{1}{3} \sum_{m=0}^2 (\omega^{j-l})^m = 0,$$

because $\omega^{j-l} \neq 1$ is a cube root of unity and $1 + z + z^2 = 0$ for such z . Taking $\mathbf{F}$ with columns $\mathbf{f}_0, \mathbf{f}_1, \mathbf{f}_2$,

$$\begin{aligned} \mathbf{F}^* \mathbf{C} \mathbf{F} &= \text{diag}(2, 1 + \omega, 1 + \omega^2) \\ &= \text{diag}\left(2, \frac{1}{2} + \frac{\sqrt{3}}{2}i, \frac{1}{2} - \frac{\sqrt{3}}{2}i\right), \end{aligned}$$

using $\omega = -\frac{1}{2} + \frac{\sqrt{3}}{2}i$. Section 9 shows that this one matrix $\mathbf{F}$ diagonalizes **every** 3×3 circulant.

Exercise. B3: Decide unitary diagonalizability

Determine which of the following matrices are unitarily diagonalizable over $\mathbb{C}$, and which are merely diagonalizable. Justify your answer.

(a) $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$

(b) $\begin{pmatrix} 2 & 3 \\ 0 & 2 \end{pmatrix}.$

(c) $\begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}.$

(d) $\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}.$

Solution.

- (a) Unitarily diagonalizable: the matrix is skew-symmetric, hence skew-adjoint over $\mathbb{C}$ and normal, so [Corollary 12.4.3](#) applies. (Its eigenvalues are $\pm i$.)
- (b) Neither. Comparing $(1, 1)$ entries, $(\mathbf{A}^* \mathbf{A})_{11} = 4$ and $(\mathbf{A} \mathbf{A}^*)_{11} = 4 + 9 = 13$, so $\mathbf{A}$ is not normal. It is not even diagonalizable: the only eigenvalue is 2 and $\mathbf{A} - 2\mathbf{I} = 3\mathbf{J}_2(0) \neq \mathbf{O}$ has rank 1, so $\dim E_2(\mathbf{A}) = 1 < 2$ (Theorem 9.4.1).
- (c) Diagonalizable but not unitarily: the eigenvalues 1 and 3 are distinct (Corollary 9.4.2), while $(\mathbf{A}^* \mathbf{A})_{11} = 1 \neq 5 = (\mathbf{A} \mathbf{A}^*)_{11}$, so $\mathbf{A}$ is not normal.
- (d) Unitarily diagonalizable: the matrix is a permutation matrix, so its columns are or-

thonormal and it is unitary, hence normal. (Its eigenvalues are the cube roots of unity.)

12.4.5.3 C. Going deeper

Exercise. C1: Operators of finite order

Let $T \in \mathcal{L}(V)$ be normal on a finite-dimensional complex inner product space, and suppose $T^k = \text{id}_V$ for some integer $k \geq 1$.

- (a) Prove that every eigenvalue of T is a k -th root of unity.
- (b) Deduce that T is unitary and unitarily diagonalizable.
- (c) Show that the hypothesis “normal” cannot be dropped from (b) by exhibiting a T with $T^2 = \text{id}_V$ that is not unitary.

Solution.

- (a) If $Tv = \lambda v$ with $v \neq 0$, then $T^k v = \lambda^k v$ (Exercise (B3: Eigenvalues of powers and inverses) (a)) and $T^k v = v$, so $(\lambda^k - 1)v = 0$ and $\lambda^k = 1$.
- (b) Every k -th root of unity has modulus 1, so T is unitary by [Corollary 12.4.4](#) (b), and it is unitarily diagonalizable by [Theorem 12.4.2](#) because it is normal.
- (c) On $\mathbb{C}^2$ take $\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix}$. Then $\mathbf{A}^2 = \mathbf{I}$, so $T_{\mathbf{A}}$ has order 2; but its second column $(1, -1)$ has norm $\sqrt{2} \neq 1$, so $\mathbf{A}$ is not unitary (Theorem 11.7.3 (f)). It is still diagonalizable, with eigenvalues 1 and -1 ; what fails is only the orthogonality of the eigenvectors.

Exercise. C2: Same spectrum, same multiplicities

Let $\mathbf{A}, \mathbf{B} \in M_n(\mathbb{C})$ be normal, and suppose they have the same eigenvalues with the same multiplicities, that is, the same characteristic polynomial.

- (a) Prove that there is a unitary $\mathbf{W}$ with $\mathbf{B} = \mathbf{W}\mathbf{A}\mathbf{W}^*$ (one says $\mathbf{A}$ and $\mathbf{B}$ are **unitarily similar**).
- (b) Give two matrices in $M_2(\mathbb{C})$ with the same characteristic polynomial that are not unitarily similar, and say which hypothesis of (a) they violate.
- (c) Deduce that two normal matrices are unitarily similar if and only if they are similar.

Solution.

- (a) By [Corollary 12.4.3](#) write $\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{U}^*$ and $\mathbf{B} = \mathbf{V}\mathbf{D}'\mathbf{V}^*$ with $\mathbf{U}, \mathbf{V}$ unitary and $\mathbf{D}, \mathbf{D}'$ diagonal. The diagonals of $\mathbf{D}$ and $\mathbf{D}'$ are the eigenvalue lists of $\mathbf{A}$ and $\mathbf{B}$ with multiplicity, which agree as multisets; and the order on the diagonal may be prescribed, so we may choose the eigenvector orderings to make $\mathbf{D}' = \mathbf{D}$. Then

$$\mathbf{B} = \mathbf{V}\mathbf{D}\mathbf{V}^* = \mathbf{V}\mathbf{U}^*\mathbf{A}\mathbf{U}\mathbf{V}^* = \mathbf{W}\mathbf{A}\mathbf{W}^* \quad \text{with} \quad \mathbf{W} = \mathbf{V}\mathbf{U}^*,$$

and $\mathbf{W}$ is unitary because $U(n)$ is closed under products and inverses (Proposi-

tion 11.7.4).

- (b) $\mathbf{A} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ and $\mathbf{B} = \mathbf{O}$ both have characteristic polynomial x^2 ; they are not even similar, since similar matrices have equal rank. A pair that is similar but not unitarily similar: $\begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$ and $\text{diag}(1, 2)$. In both cases the violated hypothesis is normality of one of the two matrices.
- (c) $(\Rightarrow)$ A unitary similarity is a similarity. $(\Leftarrow)$ Similar matrices have the same characteristic polynomial (Theorem 7.10.3), so if both are normal, (a) upgrades the similarity to a unitary one.

Exercise. C3: Normality cannot be dropped

Corollary 12.4.4 reads the class of a **normal** operator off its spectrum. This exercise shows that each part fails without that hypothesis.

- (a) Give $\mathbf{A} \in M_2(\mathbb{C})$ with both eigenvalues real that is not Hermitian.
- (b) Give $\mathbf{A} \in M_2(\mathbb{C})$ with both eigenvalues of modulus 1 that is not unitary.
- (c) Over $\mathbb{R}$, give $\mathbf{A} \in M_2(\mathbb{R})$ with $\langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle \geq 0$ for every $\mathbf{x} \in \mathbb{R}^2$ whose eigenvalues are not nonnegative reals. Why does no such example exist over $\mathbb{C}$?

Solution.

- (a) $\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix}$ has eigenvalues 1 and 2 and is not Hermitian, since $\mathbf{A}^* = \begin{pmatrix} 1 & 0 \\ 1 & 2 \end{pmatrix} \neq \mathbf{A}$. It is not normal (Example (Diagonalizable, and no orthonormal eigenbasis)).

- (b) $\mathbf{A} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ has 1 as its only eigenvalue, of modulus 1, and is not unitary: its second column has norm $\sqrt{2}$. Again it is not normal, the $(1, 1)$ entries of $\mathbf{A}^*\mathbf{A}$ and $\mathbf{A}\mathbf{A}^*$ being 1 and 2.

- (c) Take the quarter turn $\mathbf{A} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. For every $\mathbf{x} = (x_1, x_2) \in \mathbb{R}^2$,

$$\langle \mathbf{A}\mathbf{x}, \mathbf{x} \rangle = -x_2x_1 + x_1x_2 = 0 \geq 0,$$

yet $\mathbf{A}$ has no real eigenvalue at all; over $\mathbb{C}$ its eigenvalues are $\pm i$, which are not nonnegative reals. Over $\mathbb{C}$ no such example exists: if $\langle T\mathbf{v}, \mathbf{v} \rangle = 0$ for every $\mathbf{v}$ then $T = 0$ by Theorem 12.2.6, and more generally $\langle T\mathbf{v}, \mathbf{v} \rangle$ real for every $\mathbf{v}$ forces T to be self-adjoint, hence normal, by applying Theorem 12.2.6 to $T - T^*$. Real scalars are not enough to detect an operator over $\mathbb{R}$, and they are over $\mathbb{C}$.