Chapter 9

Operators on Real Vector Spaces

9.A Complexification

10/24:

- **Complexification** (of V): The set $V \times V$, where V is a real vector space. *Denoted by* $\boldsymbol{V}_{\mathbb{C}}$.
  - The complexification of V allows us to embed a real vector space in a complex vector space so that our results concerning operators on complex vector spaces can be translated into information about operators on real vector spaces.
- An element of $V_{\mathbb{C}}$ is an ordered pair (u, v), where $u, v \in V$, but we will write this as $u + iv$.
- Addition on $V_{\mathbb{C}}$ is defined by

$$(u_1 + iv_1) + (u_2 + iv_2) = (u_1 + u_2) + i(v_1 + v_2)$$

  for all $u_1, v_1, u_2, v_2 \in V$.
- Scalar multiplication on $V_{\mathbb{C}}$ is defined by

$$(a + bi)(u + iv) = (au - bv) + i(av + bu)$$

  for all $a, b \in \mathbb{R}$ and $u, v \in V$.
- Thus, we can prove that $V \times V$ is a vector space.
- If we identify $u \in V$ with $u + 0i \in V_{\mathbb{C}}$, then we can think of V as a subset of $V_{\mathbb{C}}$.
  - Basically, the construction of $V_{\mathbb{C}}$ from V generalizes the construction of $\mathbb{C}^n$ from $\mathbb{R}^n$.
- Many things transfer nicely from V to $V_{\mathbb{C}}$, as exemplified by the following.

  Theorem 9.1. *Suppose V is a real vector space.*

  (a) *If $v_1, \dots, v_n$ is a basis of V, then $v_1, \dots, v_n$ is a basis of $V_{\mathbb{C}}$.*

  Proof. Let $v_1, \dots, v_n$ be a basis of V.
  To prove that $v_1, \dots, v_n$ spans $V_{\mathbb{C}}$, we will prove an inclusion in both directions. Clearly, $\operatorname{span}(v_1, \dots, v_n) \subset V_{\mathbb{C}}$. In the other direction, let $u + iv \in V_{\mathbb{C}}$ be arbitrary. Since $u, v \in V$,

$$u = \alpha_1 v_1 + \cdots + \alpha_n v_n \qquad\qquad v = \beta_1 v_1 + \cdots + \beta_n v_n$$

  for some $\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_n \in \mathbb{R}$. It follows that

$$\begin{aligned} u + iv &= \alpha_1 v_1 + \cdots + \alpha_n v_n + i(\beta_1 v_1 + \cdots + \beta_n v_n) \\ &= (\alpha_1 + i\beta_1) v_1 + \cdots + (\alpha_n + i\beta_n) v_n \end{aligned}$$

so $u + iv \in \text{span}(v_1, \dots, v_n)$, as desired.

To prove that $v_1, \dots, v_n$ is linearly independent, suppose $\lambda_1, \dots, \lambda_n \in C$ make

$$\lambda_1 v_1 + \cdots + \lambda_n v_n = 0$$

Then naturally

$$(\operatorname{Re}\lambda_1)v_1 + \cdots + (\operatorname{Re}\lambda_n)v_n = 0 \qquad (\operatorname{Im}\lambda_1)v_1 + \cdots + (\operatorname{Im}\lambda_n)v_n = 0$$

so we must have that $\operatorname{Re}\lambda_j = \operatorname{Im}\lambda_j = \lambda_j = 0$ for each $j = 1, \dots, n$ since $v_1, \dots, v_n$ is linearly independent in V. ∎

(*b*) *The dimension of $V_{\mathbb{C}}$ equals the dimension of V.*

Proof. This follows from part (a) by the definition of dimension. ∎

- **Complexification** (of $T \in \mathcal{L}(V)$): The operator $T_{\mathbb{C}} \in \mathcal{L}(V_{\mathbb{C}})$ defined by

$$T_{\mathbb{C}}(u + iv) = Tu + iTv$$

for all $u, v \in V$ a real vector space.

  - Note that technically, we must prove that $T_{\mathbb{C}}$ is *actually* in $\mathcal{L}(V_{\mathbb{C}})$ as defined.
- If V is a real vector space with basis $v_1, \dots, v_n$ and $T \in \mathcal{L}(V)$, then $\mathcal{M}(T, (v_1, \dots, v_n)) = \mathcal{M}(T_{\mathbb{C}}, (v_1, \dots, v_n))$.
  - The proof of this claim follows immediately from the definitions.
- We now apply complexification to answer a question about invariant subspaces.

Theorem 9.2. *Every operator on a nonzero finite-dimensional vector space has an invariant subspace of dimension 1 or 2.*

Proof. Let V be a nonzero finite-dimensional vector space, and let $T \in \mathcal{L}(V)$. We divide into two cases (V is complex and V is real).

If V is complex, then by Theorem 5.5, T has an eigenvalue and hence a corresponding eigenvector v. Thus, T has a 1-dimensional invariant subspace (namely, span(v)).

If V is real, then by Theorem 5.5, $T_{\mathbb{C}}$ has an eigenvalue $a + bi$ and a corresponding eigenvector $u + iv$. It follows that

$$Tu + iTv = T_{\mathbb{C}}(u + iv) = (a + ib)(u + iv) = (au - bv) + (av + bu)i$$

i.e., that

$$Tu = au - bv \qquad Tv = av + bu$$

The above two equations prove that span(u, v) is an invariant subspace of V under T of dimension ≤ 2. ∎

- Relating the minimal polynomials of $T_{\mathbb{C}}$ and T.

Theorem 9.3. *Suppose V is a real vector space and $T \in \mathcal{L}(V)$. Then the minimal polynomial of $T_{\mathbb{C}}$ equals the minimal polynomial of T.*

Proof. Let $p \in \mathcal{P}(\mathbb{R})$ be the minimal polynomial of T. To prove that p is the minimal polynomial of $T_{\mathbb{C}}$, it will suffice to show that $p(T_{\mathbb{C}}) = 0$ and that if $q \in \mathcal{P}(\mathbb{C})$ is a monic polynomial such that $q(T_{\mathbb{C}}) = 0$, then $\deg q \geq \deg p$ (see Theorem 8.18 for the second claim). Let's begin.

For the first part, since $(T_{\mathbb{C}})^n(u + iv) = T^n u + iT^n v$ by the definition of $T_{\mathbb{C}}$, we have that $p(T_{\mathbb{C}}) = (p(T))_{\mathbb{C}} = 0_{\mathbb{C}} = 0$, as desired.

For the second part, suppose $q \in \mathcal{P}(\mathbb{C})$ is a monic polynomial such that $q(T_{\mathbb{C}}) = 0$. Then $(q(T_{\mathbb{C}}))u = 0$ for all $u = u + 0i \in V$. It follows that if $r \in \mathcal{P}(\mathbb{R})$ is the polynomial with j^{th} coefficient equal to the real part of the j^{th} coefficient of q, then $r(T) = 0$. Therefore, $\deg q = \deg r \geq \deg p$, as desired. ∎

- An interesting corollary to the previous result is that the coefficients of the minimal polynomial of $T_{\mathbb{C}}$ are real.
- We now show that the *real* eigenvalues of the complexification of T are exactly the eigenvalues of T.

Theorem 9.4. *Suppose V is a real vector space, $T \in \mathcal{L}(V)$, and $\lambda \in \mathbb{R}$. Then λ is an eigenvalue of $T_{\mathbb{C}}$ if and only if λ is an eigenvalue of T.*

Proof[1]. Let $p \in \mathcal{P}(\mathbb{C})$ be the minimal polynomial of $T_{\mathbb{C}}$. Then we have that

$$\begin{aligned} \lambda \text{ is a real eigenvalue of } T_{\mathbb{C}} &\iff \lambda \text{ is a real zero of } p(T_{\mathbb{C}}) && \text{Theorem 8.21} \\ &\iff \lambda \text{ is a zero of } p(T) && \text{Theorem 9.3} \\ &\iff \lambda \text{ is an eigenvalue of } T && \text{Theorem 8.21} \end{aligned}$$

as desired. ■

- We now show that $T_{\mathbb{C}}$ treats $\lambda, \bar{\lambda}$ the same way.

Theorem 9.5. *Suppose V is a real vector space, $T \in \mathcal{L}(V)$, $\lambda \in \mathbb{C}$, j is a nonnegative integer, and $u, v \in V$. Then $(T_{\mathbb{C}} - \lambda I)^j(u + iv) = 0$ if and only if $(T_{\mathbb{C}} - \bar{\lambda} I)^j(u - iv) = 0$.*

Proof. We induct on j.

For the base case $j = 0$, suppose first that $(T_{\mathbb{C}} - \lambda I)^0(u + iv) = u + iv = 0$. Then $u, v = 0$. It follows that $T_{\mathbb{C}} - \bar{\lambda} I)^0(u - iv) = u - iv = 0$, as desired. The proof is symmetric in the reverse direction.

Now suppose inductively that we have proven the claim for $j-1$; we now seek to prove it for j. Suppose first that $(T_{\mathbb{C}} - \lambda I)^j(u + iv) = 0$. Then

$$\begin{aligned} 0 &= (T_{\mathbb{C}} - \lambda I)^{j-1}((T_{\mathbb{C}} - \lambda I)(u + iv)) \\ &= (T_{\mathbb{C}} - \lambda I)^{j-1}((Tu - au + bv) + i(Tv - av - bu)) \end{aligned}$$

It follows by the inductive hypothesis that

$$\begin{aligned} 0 &= (T_{\mathbb{C}} - \lambda I)^{j-1}((Tu - au + bv) - i(Tv - av - bu)) \\ &= (T_{\mathbb{C}} - \lambda I)^{j-1}((T_{\mathbb{C}} - \lambda I)(u - iv)) \\ &= (T_{\mathbb{C}} - \bar{\lambda} I)^j(u - iv) \end{aligned}$$

as desired. The proof is symmetric in the other direction. ■

- We can now prove that having one complex number be an eigenvalue of $T_{\mathbb{C}}$ necessitates that its complex conjugate is an eigenvalue of $T_{\mathbb{C}}$.

Theorem 9.6. *Suppose V is a real vector space, $T \in \mathcal{L}(V)$, and $\lambda \in \mathbb{C}$. Then λ is an eigenvalue of $T_{\mathbb{C}}$ if and only if $\bar{\lambda}$ is an eigenvalue of $T_{\mathbb{C}}$.*

Proof. Take $j = 1$ in Theorem 9.5. ■

- Since a real operator can naturally only have real eigenvalues, "when mathematicians sometimes informally mention the complex eigenvalues of an operator on a real vector space, what they have in mind is the eigenvalues of the complexification of the operator" (Axler, 2015, p. 281)..
- We now prove that the multiplicities of complex conjugate eigenvalues coincide.

Theorem 9.7. *Suppose V is a real vector space, $T \in \mathcal{L}(V)$, and $\lambda \in \mathbb{C}$ is an eigenvalue of $T_{\mathbb{C}}$. Then the multiplicity of λ as an eigenvalue of $T_{\mathbb{C}}$ equals the multiplicity of $\bar{\lambda}$ as an eigenvalue of $T_{\mathbb{C}}$.*

Proof. Suppose $u_1 + iv_1, \dots, u_m + iv_m$ is a basis of the generalized eigenspace $G(\lambda, T_{\mathbb{C}})$. Then with the help of Theorem 9.5, we can easily show that $u_1 - iv_1, \dots, u_m - iv_m$ is a basis of the generalized eigenspace $G(\bar{\lambda}, T_{\mathbb{C}})$. Therefore, the multiplicities coincide at m. ■

- Although there exist operators on $\mathbb{R}^2$ (for example) with no eigenvalues, this is not true for every real vector space.

Theorem 9.8. *Every operator on an odd-dimensional real vector space has an eigenvalue.*

Proof. Suppose V is a real vector space with odd dimension, and let $T \in \mathcal{L}(V)$. Since every complex eigenvalue of $T_{\mathbb{C}}$ comes paired with its conjugate (see Theorem 9.6) and the members of each conjugate pair have the same multiplicity (see Theorem 9.7), the sum of multiplicities of the complex eigenvalues will be an even number. However, by Theorem 8.12, the sum of all of the multiplicities (counting the complex *and* real) of the eigenvalues of $T_{\mathbb{C}}$ will equal $\dim V_{\mathbb{C}}$, an odd number. Thus, there must be at least one additional eigenvalue λ of $T_{\mathbb{C}}$ that is not complex, i.e., is real. It follows by Theorem 9.4 that λ is an eigenvalue of T, as desired. ■

- We now build up to defining the characteristic polynomial for real operators.

Theorem 9.9. *Suppose V is a real vector space and $T \in \mathcal{L}(V)$. Then the coefficients of the characteristic polynomial of $T_{\mathbb{C}}$ are all real.*

Proof. Suppose λ is a nonreal eigenvalue of $T_{\mathbb{C}}$ with multiplicity m. Then by Theorems 9.6 and 9.7, $\bar{\lambda}$ is also an eigenvalue of $T_{\mathbb{C}}$ with multiplicity m. Thus, the characteristic polynomial of $T_{\mathbb{C}}$ includes the term

$$(z-\lambda)^m (z-\bar{\lambda})^m = (z^2 - 2(\operatorname{Re}\lambda)z + |\lambda|^2)^m$$

which has only real coefficients.

Since the characteristic polynomial of $T_{\mathbb{C}}$ is the product of terms with the above form and terms of the form $(z-t)^d$, where t is a real eigenvalue of $T_{\mathbb{C}}$ with multiplicity d, the coefficients of the characteristic polynomial of $T_{\mathbb{C}}$ are all real. ■

- The above result allows for the following definition.
- **Characteristic polynomial** (of $T \in \mathcal{L}(V)$, V real): The characteristic polynomial of $T_{\mathbb{C}}$.
- Properties of the characteristic polynomial.

Theorem 9.10. *Suppose V is a real vector space and $T \in \mathcal{L}(V)$. Then*

(*a*) *The coefficients of the characteristic polynomial of T are all real.*

Proof. See Theorem 9.9. ■

(*b*) *The characteristic polynomial of T has degree* $\dim V$.

Proof. See Theorem 8.16a. ■

(*c*) *The eigenvalues of T are precisely the real zeroes of the characteristic polynomial of T.*

Proof. By Theorem 9.3, the real zeroes of the characteristic polynomial of T are the real zeroes of the characteristic polynomial of $T_{\mathbb{C}}$. These are, in turn, the real eigenvalues of $T_{\mathbb{C}}$ (by Theorem 8.16b). These, lastly, are in turn the eigenvalues of T (by Theorem 9.4). ■

- We can now prove the complete Cayley-Hamilton Theorem.

Theorem 9.11 (Cayley-Hamilton Theorem)**.** *Suppose $T \in \mathcal{L}(V)$. Let q denote the characteristic polynomial of T. Then $q(T) = 0$.*

Proof. We divide into two cases (V is complex and V is real).

If V is complex, then apply the Complex Cayley-Hamilton Theorem.

If V is real, then by the Complex Cayley-Hamilton Theorem, $q(T_\mathbb{C}) = 0$. It follows by the definition of the characteristic polynomial of T that $q(T) = q(T_\mathbb{C}) = 0$, as desired. ■

- We now extend one last result from the complex to the real case.

Theorem 9.12. *Suppose $T \in \mathcal{L}(V)$. Then*

(a) *The degree of the minimal polynomial of T is at most* $\dim V$*.*

Proof. Let $p \in \mathcal{P}(\mathbb{R})$ be the minimal polynomial of T, and let $q \in \mathcal{P}(\mathbb{R})$ be the characteristic polynomial of T. By Theorem 9.10b, $\deg q = \dim V$. By the Cayley-Hamilton Theorem, $q(T) = 0$. Thus, by the definition of the minimal polynomial

$$\dim p \leq \dim q = \dim V$$

as desired. ■

(b) *The characteristic polynomial of T is a polynomial multiple of the minimal polynomial of T.*

Proof. Let $p \in \mathcal{P}(\mathbb{R})$ be the minimal polynomial of T, and let $q \in \mathcal{P}(\mathbb{R})$ be the characteristic polynomial of T. By the Complex Cayley-Hamilton Theorem, $q(T_\mathbb{C}) = 0$. Thus, by Theorem 8.19 $q(T) = q(T_\mathbb{C})$ is a polynomial multiple of $p(T_\mathbb{C}) = p(T)$ (where the last equality follows from Theorem 9.3). ■

9.B Operators on Real Inner Product Spaces

10/25: • In this section, we characterize normal operators and isometries on real inner product spaces.

- First off, we describe normal but not self-adjoint operators on 2-dimensional real inner product spaces.

Theorem 9.13. *Suppose V is a 2-dimensional real inner product space and $T \in \mathcal{L}(V)$. Then the following are equivalent.*

(a) *T is normal but not self-adjoint.*

(b) *The matrix of T with respect to every orthonormal basis of V has the form*

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

with $b \neq 0$.

(c) *The matrix of T with respect to some orthonormal basis of V has the form*

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

with $b > 0$.

Proof. We will show that (a) $\Rightarrow$ (b) $\Rightarrow$ (c) $\Rightarrow$ (a). Let's begin.

First, suppose that T is normal but not self-adjoint, and let e_1, e_2 be an arbitrary orthonormal basis of V. We have from the definitions that

$$\mathcal{M}(T, (e_1, e_2)) = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$$

for some $a, b, c, d \in \mathbb{R}$. We now look to express some of these variable in terms of others using known constraints on normal but not self-adjoint matrices. Since T is normal, Theorem 7.9 implies that $\|Te_1\|^2 = \|T^*e_1\|^2$. Thus, we have that

$$\begin{aligned} a^2 + b^2 = \|Te_1\|^2 = \|T^*e_1\|^2 &= a^2 + c^2 \\ b^2 &= c^2 \\ c &= \pm b \end{aligned}$$

Since T is not self-adjoint, we must choose $c = -b$ instead of $c = b$. Thus, we have that

$$\mathcal{M}(T, (e_1, e_2)) = \begin{pmatrix} a & -b \\ b & d \end{pmatrix}$$

Additionally, since T is normal, we have by definition that

$$\begin{aligned} TT^* &= T^*T \\ \begin{pmatrix} a & -b \\ b & d \end{pmatrix}\begin{pmatrix} a & b \\ -b & d \end{pmatrix} &= \begin{pmatrix} a & b \\ -b & d \end{pmatrix}\begin{pmatrix} a & -b \\ b & d \end{pmatrix} \\ \begin{pmatrix} a^2 + b^2 & ab - bd \\ ab - bd & b^2 + d^2 \end{pmatrix} &= \begin{pmatrix} a^2 + b^2 & -ab + bd \\ -ab + bd & b^2 + d^2 \end{pmatrix} \end{aligned}$$

Thus, we have that $ab = bd$. But since $b \neq 0$ (otherwise, T would be self-adjoint), we must have $d = a$, as desired.

Second, suppose that (b) holds. Since $b \neq 0$, either $b > 0$ or $b < 0$ in $\mathcal{M}(T)$ with respect to any orthonormal basis of V. We now divide into two cases. If $b > 0$, then we are done. On the other hand, if $b < 0$ in $\mathcal{M}(T, (e_1, e_2))$, we will have $b > 0$ in $\mathcal{M}(T, (e_1, -e_2))$, as desired.

Third, suppose that (c) holds. Since $b > 0$ in $\mathcal{M}(T)$ with respect to some orthonormal basis, the upper right and lower left entries in $\mathcal{M}(T)$ are distinct. Thus, T is not self-adjoint. However, we can use matrix multiplication to verify that the matrices of TT^* and T^*T are equal with respect to this basis, proving that T is normal, as desired. ■

- We now prove a lemma to our main description of normal operators.

Theorem 9.14. *Suppose V is an inner product space, $T \in \mathcal{L}(V)$ is normal, and U is a subspace of V that is invariant under T. Then*

(a) *$U^\perp$ is invariant under T.*

Proof. Let $e_1, \dots, e_m$ be an orthonormal basis of U. Extend it to an orthonormal basis $e_1, \dots, e_m, f_1, \dots, f_m$ of V (see Theorem 6.15). Since U is invariant under T each Te_j is a linear combination of $e_1, \dots, e_m$. Thus, $\mathcal{M}(T)$ is of the following form.

$$\mathcal{M}(T) = \begin{array}{c} \\ e_1 \\ \vdots \\ e_m \\ f_1 \\ \vdots \\ f_m \end{array} \begin{array}{c} \begin{array}{cccccc} e_1 & \cdots & e_m & f_1 & \cdots & f_m \end{array} \\ \begin{pmatrix} \\ & A & & & B & \\ \\ \\ & 0 & & & C & \\ \\ \end{pmatrix} \end{array}$$

For each $j = 1, \dots, m$, Theorem 6.9 asserts that $\|Te_j\|^2$ is the sum of the squares of the absolute values of the entries in the j^{th} column of A. Similarly, for each $j = 1, \dots, m$, Theorem 6.9 asserts that $\|T^*e_j\|^2$ is the sum of the squares of the absolute values of the entries in the j^{th} columns of A and B. But since these two values are equal by Theorem 7.9, we must have that all of the values in the j^{th} column of B are 0 for each $j = 1, \dots, n$, i.e., that $B = 0$. Thus, $Tf_k \in \operatorname{span}(f_1, \dots, f_n)$ for each $k = 1, \dots, n$. It follows since $f_1, \dots, f_n$ is a basis of $U^\perp$ that $Tv \in U^\perp$ for any $v \in U^\perp$, implying that $U^\perp$ is invariant under T, as desired. ■

(*b*) *U is invariant under T^*.*

Proof. By part (a), $\mathcal{M}(T^*)$ is of the same form as $\mathcal{M}(T)$. Thus, by a similar argument, T^* is invariant under U. ■

(*c*) $(T|_U)^* = (T^*)|_U$.

Proof. For every $u, v \in U$, we have that

$$\langle u, (T|_U)^* v\rangle = \langle (T|_U)u, v\rangle = \langle Tu, v\rangle = \langle u, T^* v\rangle = \langle u, ((T^*)|_U)v\rangle$$

Therefore, $(T|_U)^* = (T^*)|_U$, as desired. ■

(*d*) *$T|_U \in \mathcal{L}(U)$ and $T|_{U^\perp} \in \mathcal{L}(U^\perp)$ are normal operators.*

Proof. We have from the above results that

$$(T|_U)(T|_U)^* = T|_U(T^*)|_U = (TT^*)|_U = (T^*T)|_U = (T^*)|_U T|_U = (T|_U)^*(T|_U)$$

and similarly for $U^\perp$. ■

- Note that if an operator has a block diagonal matrix with respect to some basis, then the entry in any 1×1 block on the diagonal of this matrix is an eigenvalue of T.
- We now prove that all normal operators on real inner product spaces have block-diagonal matrices with blocks of size at most 2×2.

Theorem 9.15. *Suppose V is a real inner product space and $T \in \mathcal{L}(V)$. Then the following are equivalent:*

(*a*) *T is normal.*

(*b*) *There is an orthonormal basis of V with respect to which T has a block diagonal matrix such that each block is a 1×1 matrix or a 2×2 matrix of the form*

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

with $b > 0$.

Proof. Suppose first that T is normal. We induct on $n = \dim V$. For the base case $n = 1$, T will trivially have a block diagonal matrix with a single 1×1 block for any basis of V, in particular any orthonormal one we can pick by Theorem 6.14. For the base case $n = 2$, we divide into two cases (T is self-adjoint and T is not self-adjoint). If T is self-adjoint, then by the Real Spectral Theorem, there is an orthonormal basis of V with respect to which T has a diagonal matrix (equivalently, a block diagonal matrix where each [the one] block is a 1×1 matrix), as desired. If T is not self-adjoint, then since T is also normal by hypothesis, Theorem 9.13 asserts that there exists an orthonormal basis of V with respect to which T has a matrix of the form $\left(\begin{smallmatrix} a & -b \\ b & a \end{smallmatrix}\right)$ with $b > 0$ (equivalently, a block-diagonal matrix where each [the one] block is a 2×2 matrix of the desired form), as desired.

Now suppose using strong induction that we have proven the claim for $n - 1$; we now seek to prove it for n. By Theorem 9.2, T has an invariant subspace U of dimension 1 or 2. We now divide into two cases. If $\dim U = 1$, choose $u \in U$ such that $\|u\| = 1$; this vector forms an orthonormal basis of U. If $\dim U = 2$, then by Theorem 9.14, $T|_U \in \mathcal{L}(U)$ is normal. Additionally, $T|_U$ is not self-adjoint (otherwise, Theorem 7.14 would imply that T has an eigenvalue and hence a corresponding eigenvector, making $\dim U = 1$). Thus, by Theorem 9.13, we can choose an orthonormal basis of U with respect to which $\mathcal{M}(T|_U)$ has the required form. Either way, we now have $\dim U^\perp < \dim V$. This combined with the facts that $U^\perp$ is invariant under T and $T|_{U^\perp}$ is normal (see Theorem 9.14) allows us to apply our induction hypothesis. Doing so reveals that there is an orthonormal basis of $U^\perp$ with respect to which

$\mathcal{M}(T|_{U^\perp})$ has the desired form. Concatenating this basis to the previously found basis of U gives an orthonormal basis of V with respect to which the matrix of T has the desired form overall.

Now suppose that (b) holds. Then using Exercise 8.B.9, to show that $TT^* = T^*T$, it will suffice to show that all submatrices along the diagonal commute, too. The 1×1 submatrices obviously commute, and the 2×2 ones commute since

$$\begin{aligned} T_j T_j^* &= \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \\ &= \begin{pmatrix} a^2+b^2 & 0 \\ 0 & a^2+b^2 \end{pmatrix} \\ &= \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \\ &= T_j^* T_j \end{aligned}$$

as desired. ■