CHAPTER

9

Euclid explaining geometry (from The School of Athens, *painted by Raphael around 1510).*

Operators on Real Vector Spaces

In the last chapter we learned about the structure of an operator on a finite-dimensional complex vector space. In this chapter, we will use our results about operators on complex vector spaces to learn about operators on real vector spaces.

Our assumptions for this chapter are as follows:

9.1 **Notation** $\mathbf{F}$, V

- $\mathbf{F}$ denotes $\mathbf{R}$ or $\mathbf{C}$.
- V denotes a finite-dimensional nonzero vector space over $\mathbf{F}$.

LEARNING OBJECTIVES FOR THIS CHAPTER

- complexification of a real vector space
- complexification of an operator on a real vector space
- operators on finite-dimensional real vector spaces have an eigenvalue or a 2-dimensional invariant subspace
- characteristic polynomial and the Cayley–Hamilton Theorem
- description of normal operators on a real inner product space
- description of isometries on a real inner product space

S. Axler, *Linear Algebra Done Right*, Undergraduate Texts in Mathematics,
DOI 10.1007/978-3-319-11080-6_9

9.A Complexification

Complexification of a Vector Space

As we will soon see, a real vector space V can be embedded, in a natural way, in a complex vector space called the complexification of V. Each operator on V can be extended to an operator on the complexification of V. Our results about operators on complex vector spaces can then be translated to information about operators on real vector spaces.

We begin by defining the complexification of a real vector space.

9.2 Definition ***complexification of*** V, $V_{\mathbf{C}}$

Suppose V is a real vector space.

- The ***complexification*** of V, denoted $V_{\mathbf{C}}$, equals $V \times V$. An element of $V_{\mathbf{C}}$ is an ordered pair (u, v), where $u, v \in V$, but we will write this as $u + iv$.
- Addition on $V_{\mathbf{C}}$ is defined by
$$(u_1 + iv_1) + (u_2 + iv_2) = (u_1 + u_2) + i(v_1 + v_2)$$
for $u_1, v_1, u_2, v_2 \in V$.
- Complex scalar multiplication on $V_{\mathbf{C}}$ is defined by
$$(a + bi)(u + iv) = (au - bv) + i(av + bu)$$
for $a, b \in \mathbf{R}$ and $u, v \in V$.

Motivation for the definition above of complex scalar multiplication comes from usual algebraic properties and the identity $i^2 = -1$. If you remember the motivation, then you do not need to memorize the definition above.

We think of V as a subset of $V_{\mathbf{C}}$ by identifying $u \in V$ with $u + i0$. The construction of $V_{\mathbf{C}}$ from V can then be thought of as generalizing the construction of $\mathbf{C}^n$ from $\mathbf{R}^n$.

9.3 $V_{\mathbf{C}}$ **is a complex vector space.**

Suppose V is a real vector space. Then with the definitions of addition and scalar multiplication as above, $V_{\mathbf{C}}$ is a complex vector space.

The proof of the result above is left as an exercise for the reader. Note that the additive identity of $V_{\mathbf{C}}$ is $0 + i0$, which we write as just 0.

Probably everything that you think should work concerning complexification does work, usually with a straightforward verification, as illustrated by the next result.

9.4 Basis of V is basis of $V_{\mathbf{C}}$

Suppose V is a real vector space.

(a) If $v_1, \dots, v_n$ is a basis of V (as a real vector space), then $v_1, \dots, v_n$ is a basis of $V_{\mathbf{C}}$ (as a complex vector space).

(b) The dimension of $V_{\mathbf{C}}$ (as a complex vector space) equals the dimension of V (as a real vector space).

Proof To prove (a), suppose $v_1, \dots, v_n$ is a basis of the real vector space V. Then $\operatorname{span}(v_1, \dots, v_n)$ in the complex vector space $V_{\mathbf{C}}$ contains all the vectors $v_1, \dots, v_n, iv_1, \dots, iv_n$. Thus $v_1, \dots, v_n$ spans the complex vector space $V_{\mathbf{C}}$.

To show that $v_1, \dots, v_n$ is linearly independent in the complex vector space $V_{\mathbf{C}}$, suppose $\lambda_1, \dots, \lambda_n \in \mathbf{C}$ and

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

Then the equation above and our definitions imply that

$$(\operatorname{Re}\lambda_1)v_1 + \cdots + (\operatorname{Re}\lambda_n)v_n = 0 \quad \text{and} \quad (\operatorname{Im}\lambda_1)v_1 + \cdots + (\operatorname{Im}\lambda_n)v_n = 0.$$

Because $v_1, \dots, v_n$ is linearly independent in V, the equations above imply $\operatorname{Re}\lambda_1 = \cdots = \operatorname{Re}\lambda_n = 0$ and $\operatorname{Im}\lambda_1 = \cdots = \operatorname{Im}_n = 0$. Thus we have $\lambda_1 = \cdots = \lambda_n = 0$. Hence $v_1, \dots, v_n$ is linearly independent in $V_{\mathbf{C}}$, completing the proof of (a).

Clearly (b) follows immediately from (a). ■

Complexification of an Operator

Now we can define the complexification of an operator.

9.5 Definition ***complexification of T, $T_{\mathbf{C}}$***

Suppose V is a real vector space and $T \in \mathcal{L}(V)$. The ***complexification*** of T, denoted $T_{\mathbf{C}}$, is the operator $T_{\mathbf{C}} \in \mathcal{L}(V_{\mathbf{C}})$ defined by

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

for $u, v \in V$.

You should verify that if V is a real vector space and $T \in \mathcal{L}(V)$, then $T_{\mathbf{C}}$ is indeed in $\mathcal{L}(V_{\mathbf{C}})$. The key point here is that our definition of complex scalar multiplication can be used to show that $T_{\mathbf{C}}\big(\lambda(u + iv)\big) = \lambda T_{\mathbf{C}}(u + iv)$ for all $u, v \in V$ and all **complex** numbers λ.

The next example gives a good way to think about the complexification of a typical operator.

9.6 **Example** Suppose A is an n-by-n matrix of real numbers. Define $T \in \mathcal{L}(\mathbf{R}^n)$ by $Tx = Ax$, where elements of $\mathbf{R}^n$ are thought of as n-by-1 column vectors. Identifying the complexification of $\mathbf{R}^n$ with $\mathbf{C}^n$, we then have $T_{\mathbf{C}}z = Az$ for each $z \in \mathbf{C}^n$, where again elements of $\mathbf{C}^n$ are thought of as n-by-1 column vectors.

In other words, if T is the operator of matrix multiplication by A on $\mathbf{R}^n$, then the complexification $T_{\mathbf{C}}$ is also matrix multiplication by A but now acting on the larger domain $\mathbf{C}^n$.

The next result makes sense because 9.4 tells us that a basis of a real vector space is also a basis of its complexification. The proof of the next result follows immediately from the definitions.

9.7 Matrix of $T_{\mathbf{C}}$ equals matrix of T

Suppose V is a real vector space with basis $v_1, \dots, v_n$ and $T \in \mathcal{L}(V)$. Then $\mathcal{M}(T) = \mathcal{M}(T_{\mathbf{C}})$, where both matrices are with respect to the basis $v_1, \dots, v_n$.

The result above and Example 9.6 provide complete insight into complexification, because once a basis is chosen, every operator essentially looks like Example 9.6. Complexification of an operator could have been defined using matrices, but the approach taken here is more natural because it does not depend on the choice of a basis.

We know that every operator on a nonzero finite-dimensional complex vector space has an eigenvalue (see 5.21) and thus has a 1-dimensional invariant subspace. We have seen an example [5.8(a)] of an operator on a nonzero finite-dimensional real vector space with no eigenvalues and thus no 1-dimensional invariant subspaces. However, we now show that an invariant subspace of dimension 1 or 2 always exists. Notice how complexification leads to a simple proof of this result.

9.8 Every operator has an invariant subspace of dimension 1 or 2

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

Proof Every operator on a nonzero finite-dimensional complex vector space has an eigenvalue (5.21) and thus has a 1-dimensional invariant subspace.

Hence assume V is a real vector space and $T \in \mathcal{L}(V)$. The complexification $T_{\mathbf{C}}$ has an eigenvalue $a + bi$ (by 5.21), where $a, b \in \mathbf{R}$. Thus there exist $u, v \in V$, not both 0, such that $T_{\mathbf{C}}(u + iv) = (a + bi)(u + iv)$. Using the definition of $T_{\mathbf{C}}$, the last equation can be rewritten as

$$Tu + iTv = (au - bv) + (av + bu)i.$$

Thus

$$Tu = au - bv \quad \text{and} \quad Tv = av + bu.$$

Let U equal the span in V of the list u, v. Then U is a subspace of V with dimension 1 or 2. The equations above show that U is invariant under T, completing the proof. ■

The Minimal Polynomial of the Complexification

Suppose V is a real vector space and $T \in \mathcal{L}(V)$. Repeated application of the definition of $T_{\mathbf{C}}$ shows that

9.9
$$(T_{\mathbf{C}})^n(u + iv) = T^n u + iT^n v$$

for every positive integer n and all $u, v \in V$.

Notice that the next result implies that the minimal polynomial of $T_{\mathbf{C}}$ has real coefficients.

9.10 Minimal polynomial of $T_{\mathbf{C}}$ equals minimal polynomial of T

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

Proof Let $p \in \mathcal{P}(\mathbf{R})$ denote the minimal polynomial of T. From 9.9 it is easy to see that $p(T_{\mathbf{C}}) = \big(p(T)\big)_{\mathbf{C}}$, and thus $p(T_{\mathbf{C}}) = 0$.

Suppose $q \in \mathcal{P}(\mathbf{C})$ is a monic polynomial such that $q(T_{\mathbf{C}}) = 0$. Then $\big(q(T_{\mathbf{C}})\big)(u) = 0$ for every $u \in V$. Letting r denote the polynomial whose j^{th} coefficient is the real part of the j^{th} coefficient of q, we see that r is a monic polynomial and $r(T) = 0$. Thus $\deg q = \deg r \geq \deg p$.

The conclusions of the two previous paragraphs imply that p is the minimal polynomial of $T_{\mathbf{C}}$, as desired. ■

Eigenvalues of the Complexification

Now we turn to questions about the eigenvalues of the complexification of an operator. Again, everything that we expect to work indeed works easily.

We begin with a result showing that the real eigenvalues of $T_{\mathbf{C}}$ are precisely the eigenvalues of T. We give two different proofs of this result. The first proof is more elementary, but the second proof is shorter and gives some useful insight.

9.11 Real eigenvalues of $T_{\mathbf{C}}$

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

Proof 1 First suppose λ is an eigenvalue of T. Then there exists $v \in V$ with $v \neq 0$ such that $Tv = \lambda v$. Thus $T_{\mathbf{C}} v = \lambda v$, which shows that λ is an eigenvalue of $T_{\mathbf{C}}$, completing one direction of the proof.

To prove the other direction, suppose now that λ is an eigenvalue of $T_{\mathbf{C}}$. Then there exist $u, v \in V$ with $u + iv \neq 0$ such that

$$T_{\mathbf{C}}(u + iv) = \lambda(u + iv).$$

The equation above implies that $Tu = \lambda u$ and $Tv = \lambda v$. Because $u \neq 0$ or $v \neq 0$, this implies that λ is an eigenvalue of T, completing the proof. ■

Proof 2 The (real) eigenvalues of T are the (real) zeros of the minimal polynomial of T (by 8.49). The real eigenvalues of $T_{\mathbf{C}}$ are the real zeros of the minimal polynomial of $T_{\mathbf{C}}$ (again by 8.49). These two minimal polynomials are the same (by 9.10). Thus the eigenvalues of T are precisely the real eigenvalues of $T_{\mathbf{C}}$, as desired. ■

Our next result shows that $T_{\mathbf{C}}$ behaves symmetrically with respect to an eigenvalue λ and its complex conjugate $\bar{\lambda}$.

9.12 $T_{\mathbf{C}} - \lambda I$ and $T_{\mathbf{C}} - \bar{\lambda} I$

Suppose V is a real vector space, $T \in \mathcal{L}(V)$, $\lambda \in \mathbf{C}$, j is a nonnegative integer, and $u, v \in V$. Then

$$(T_{\mathbf{C}} - \lambda I)^j (u + iv) = 0 \quad \text{if and only if} \quad (T_{\mathbf{C}} - \bar{\lambda} I)^j (u - iv) = 0.$$

Proof We will prove this result by induction on j. To get started, note that if $j = 0$ then (because an operator raised to the power 0 equals the identity operator) the result claims that $u + iv = 0$ if and only if $u - iv = 0$, which is clearly true.

Thus assume by induction that $j \geq 1$ and the desired result holds for $j - 1$. Suppose $(T_{\mathbf{C}} - \lambda I)^j (u + iv) = 0$. Then

9.13 $$(T_{\mathbf{C}} - \lambda I)^{j-1}\big((T_{\mathbf{C}} - \lambda I)(u + iv)\big) = 0.$$

Writing $\lambda = a + bi$, where $a, b \in \mathbf{R}$, we have

9.14 $$(T_{\mathbf{C}} - \lambda I)(u + iv) = (Tu - au + bv) + i(Tv - av - bu)$$

and

9.15 $$(T_{\mathbf{C}} - \bar{\lambda} I)(u - iv) = (Tu - au + bv) - i(Tv - av - bu).$$

Our induction hypothesis, 9.13, and 9.14 imply that

$$(T_{\mathbf{C}} - \bar{\lambda} I)^{j-1}\big((Tu - au + bv) - i(Tv - av - bu)\big) = 0.$$

Now the equation above and 9.15 imply that $(T_{\mathbf{C}} - \bar{\lambda} I)^j (u - iv) = 0$, completing the proof in one direction.

The other direction is proved by replacing λ with $\bar{\lambda}$, replacing v with $-v$, and then using the first direction. ■

An important consequence of the result above is the next result, which states that if a number is an eigenvalue of $T_{\mathbf{C}}$, then its complex conjugate is also an eigenvalue of $T_{\mathbf{C}}$.

9.16 Nonreal eigenvalues of $T_{\mathbf{C}}$ come in pairs

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

Proof Take $j = 1$ in 9.12. ■

By definition, the eigenvalues of an operator on a real vector space are real numbers. Thus 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.

Recall that the multiplicity of an eigenvalue is defined to be the dimension of the generalized eigenspace corresponding to that eigenvalue (see 8.24). The next result states that the multiplicity of an eigenvalue of a complexification equals the multiplicity of its complex conjugate.

9.17 Multiplicity of λ equals multiplicity of $\bar{\lambda}$

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

Proof Suppose $u_1 + iv_1, \dots, u_m + iv_m$ is a basis of the generalized eigenspace $G(\lambda, T_{\mathbf{C}})$, where $u_1, \dots, u_m, v_1, \dots, v_m \in V$. Then using 9.12, routine arguments show that $u_1 - iv_1, \dots, u_m - iv_m$ is a basis of the generalized eigenspace $G(\bar{\lambda}, T_{\mathbf{C}})$. Thus both λ and $\bar{\lambda}$ have multiplicity m as eigenvalues of $T_{\mathbf{C}}$. ■

9.18 Example Suppose $T \in \mathcal{L}(\mathbf{R}^3)$ is defined by

$$T(x_1, x_2, x_3) = (2x_1, x_2 - x_3, x_2 + x_3).$$

The matrix of T with respect to the standard basis of $\mathbf{R}^3$ is $\begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 1 & 1 \end{pmatrix}$.

As you can verify, 2 is an eigenvalue of T with multiplicity 1 and T has no other eigenvalues.

If we identify the complexification of $\mathbf{R}^3$ with $\mathbf{C}^3$, then the matrix of $T_{\mathbf{C}}$ with respect to the standard basis of $\mathbf{C}^3$ is the matrix above. As you can verify, the eigenvalues of $T_{\mathbf{C}}$ are 2, $1 + i$, and $1 - i$, each with multiplicity 1. Thus the nonreal eigenvalues of $T_{\mathbf{C}}$ come as a pair, with each the complex conjugate of the other and with the same multiplicity, as expected by 9.17.

We have seen an example [5.8(a)] of an operator on $\mathbf{R}^2$ with no eigenvalues. The next result shows that no such example exists on $\mathbf{R}^3$.

9.19 Operator on odd-dimensional vector space has eigenvalue

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

Proof Suppose V is a real vector space with odd dimension and $T \in \mathcal{L}(V)$. Because the nonreal eigenvalues of $T_{\mathbf{C}}$ come in pairs with equal multiplicity (by 9.17), the sum of the multiplicities of all the nonreal eigenvalues of $T_{\mathbf{C}}$ is an even number.

Because the sum of the multiplicities of all the eigenvalues of $T_{\mathbf{C}}$ equals the (complex) dimension of $V_{\mathbf{C}}$ (by Theorem 8.26), the conclusion of the paragraph above implies that $T_{\mathbf{C}}$ has a real eigenvalue. Every real eigenvalue of $T_{\mathbf{C}}$ is also an eigenvalue of T (by 9.11), giving the desired result. ■

Characteristic Polynomial of the Complexification

In the previous chapter we defined the characteristic polynomial of an operator on a finite-dimensional complex vector space (see 8.34). The next result is a key step toward defining the characteristic polynomial for operators on finite-dimensional real vector spaces.

9.20 Characteristic polynomial of $T_{\mathbf{C}}$

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

Proof Suppose λ is a nonreal eigenvalue of $T_{\mathbf{C}}$ with multiplicity m. Then $\bar{\lambda}$ is also an eigenvalue of $T_{\mathbf{C}}$ with multiplicity m (by 9.17). Thus the characteristic polynomial of $T_{\mathbf{C}}$ includes factors of $(z-\lambda)^m$ and $(z-\bar{\lambda})^m$. Multiplying together these two factors, we have

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

The polynomial above on the right has real coefficients.

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

Now we can define the characteristic polynomial of an operator on a finite-dimensional real vector space to be the characteristic polynomial of its complexification.

9.21 Definition ***Characteristic polynomial***

Suppose V is a real vector space and $T \in \mathcal{L}(V)$. Then the ***characteristic polynomial*** of T is defined to be the characteristic polynomial of $T_{\mathbf{C}}$.

9.22 Example Suppose $T \in \mathcal{L}(\mathbf{R}^3)$ is defined by

$$T(x_1, x_2, x_3) = (2x_1, x_2 - x_3, x_2 + x_3).$$

As we noted in 9.18, the eigenvalues of $T_{\mathbf{C}}$ are 2, $1+i$, and $1-i$, each with multiplicity 1. Thus the characteristic polynomial of the complexification $T_{\mathbf{C}}$ is $(z-2)\big(z-(1+i)\big)\big(z-(1-i)\big)$, which equals $z^3 - 4z^2 + 6z - 4$. Hence the characteristic polynomial of T is also $z^3 - 4z^2 + 6z - 4$.

In the next result, the eigenvalues of T are all real (because T is an operator on a real vector space).

9.23 Degree and zeros of characteristic polynomial

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;

(b) the characteristic polynomial of T has degree $\dim V$;

(c) the eigenvalues of T are precisely the real zeros of the characteristic polynomial of T.

Proof Part (a) holds because of 9.20.

Part (b) follows from 8.36(a).

Part (c) holds because the real zeros of the characteristic polynomial of T are the real eigenvalues of $T_{\mathbf{C}}$ [by 8.36(a)], which are the eigenvalues of T (by 9.11). ■

In the previous chapter, we proved the Cayley–Hamilton Theorem (8.37) for complex vector spaces. Now we can also prove it for real vector spaces.

9.24 Cayley–Hamilton Theorem

Suppose $T \in \mathcal{L}(V)$. Let q denote the characteristic polynomial of T. Then $q(T) = 0$.

Proof We have already proved this result when V is a complex vector space. Thus assume that V is a real vector space.

The complex case of the Cayley–Hamilton Theorem (8.37) implies that $q(T_{\mathbf{C}}) = 0$. Thus we also have $q(T) = 0$, as desired. ■

9.25 **Example** Suppose $T \in \mathcal{L}(\mathbf{R}^3)$ is defined by

$$T(x_1, x_2, x_3) = (2x_1, x_2 - x_3, x_2 + x_3).$$

As we saw in 9.22, the characteristic polynomial of T is $z^3 - 4z^2 + 6z - 4$. Thus the Cayley–Hamilton Theorem implies that $T^3 - 4T^2 + 6T - 4I = 0$, which can also be verified by direct calculation.

We can now prove another result that we previously knew only in the complex case.

9.26 Characteristic polynomial is a multiple of minimal polynomial

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

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

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

Proof Part (a) follows immediately from the Cayley–Hamilton Theorem. Part (b) follows from the Cayley–Hamilton Theorem and 8.46. ■

EXERCISES 9.A

1 Prove 9.3.

2 Verify that if V is a real vector space and $T \in \mathcal{L}(V)$, then $T_{\mathbf{C}} \in \mathcal{L}(V_{\mathbf{C}})$.

3 Suppose V is a real vector space and $v_1, \dots, v_m \in V$. Prove that $v_1, \dots, v_m$ is linearly independent in $V_{\mathbf{C}}$ if and only if $v_1, \dots, v_m$ is linearly independent in V.

4 Suppose V is a real vector space and $v_1, \dots, v_m \in V$. Prove that $v_1, \dots, v_m$ spans $V_{\mathbf{C}}$ if and only if $v_1, \dots, v_m$ spans V.

5 Suppose that V is a real vector space and $S, T \in \mathcal{L}(V)$. Show that $(S + T)_{\mathbf{C}} = S_{\mathbf{C}} + T_{\mathbf{C}}$ and that $(\lambda T)_{\mathbf{C}} = \lambda T_{\mathbf{C}}$ for every $\lambda \in \mathbf{R}$.

6 Suppose V is a real vector space and $T \in \mathcal{L}(V)$. Prove that $T_{\mathbf{C}}$ is invertible if and only if T is invertible.

7 Suppose V is a real vector space and $N \in \mathcal{L}(V)$. Prove that $N_{\mathbf{C}}$ is nilpotent if and only if N is nilpotent.

8 Suppose $T \in \mathcal{L}(\mathbf{R}^3)$ and $5, 7$ are eigenvalues of T. Prove that $T_{\mathbf{C}}$ has no nonreal eigenvalues.

9 Prove there does not exist an operator $T \in \mathcal{L}(\mathbf{R}^7)$ such that $T^2 + T + I$ is nilpotent.

10 Give an example of an operator $T \in \mathcal{L}(\mathbf{C}^7)$ such that $T^2 + T + I$ is nilpotent.

11 Suppose V is a real vector space and $T \in \mathcal{L}(V)$. Suppose there exist $b, c \in \mathbf{R}$ such that $T^2 + bT + cI = 0$. Prove that T has an eigenvalue if and only if $b^2 \geq 4c$.

12 Suppose V is a real vector space and $T \in \mathcal{L}(V)$. Suppose there exist $b, c \in \mathbf{R}$ such that $b^2 < 4c$ and $T^2 + bT + cI$ is nilpotent. Prove that T has no eigenvalues.

13 Suppose V is a real vector space, $T \in \mathcal{L}(V)$, and $b, c \in \mathbf{R}$ are such that $b^2 < 4c$. Prove that $\operatorname{null}(T^2 + bT + cI)^j$ has even dimension for every positive integer j.

14 Suppose V is a real vector space with $\dim V = 8$. Suppose $T \in \mathcal{L}(V)$ is such that $T^2 + T + I$ is nilpotent. Prove that $(T^2 + T + I)^4 = 0$.

15 Suppose V is a real vector space and $T \in \mathcal{L}(V)$ has no eigenvalues. Prove that every subspace of V invariant under T has even dimension.

16 Suppose V is a real vector space. Prove that there exists $T \in \mathcal{L}(V)$ such that $T^2 = -I$ if and only if V has even dimension.

17 Suppose V is a real vector space and $T \in \mathcal{L}(V)$ satisfies $T^2 = -I$. Define complex scalar multiplication on V as follows: if $a, b \in \mathbf{R}$, then

$$(a + bi)v = av + bTv.$$

(a) Show that the complex scalar multiplication on V defined above and the addition on V makes V into a complex vector space.

(b) Show that the dimension of V as a complex vector space is half the dimension of V as a real vector space.

18 Suppose V is a real vector space and $T \in \mathcal{L}(V)$. Prove that the following are equivalent:

(a) All the eigenvalues of $T_{\mathbf{C}}$ are real.

(b) There exists a basis of V with respect to which T has an upper-triangular matrix.

(c) There exists a basis of V consisting of generalized eigenvectors of T.

19 Suppose V is a real vector space with $\dim V = n$ and $T \in \mathcal{L}(V)$ is such that $\operatorname{null} T^{n-2} \neq \operatorname{null} T^{n-1}$. Prove that T has at most two distinct eigenvalues and that $T_{\mathbf{C}}$ has no nonreal eigenvalues.

9.B Operators on Real Inner Product Spaces

We now switch our focus to the context of inner product spaces. We will give a complete description of normal operators on real inner product spaces; a key step in the proof of this result (9.34) requires the result from the previous section that an operator on a finite-dimensional real vector space has an invariant subspace of dimension 1 or 2 (9.8).

After describing the normal operators on real inner product spaces, we will use that result to give a complete description of isometries on such spaces.

Normal Operators on Real Inner Product Spaces

The Complex Spectral Theorem (7.24) gives a complete description of normal operators on complex inner product spaces. In this subsection we will give a complete description of normal operators on real inner product spaces.

We begin with a description of the operators on 2-dimensional real inner product spaces that are normal but not self-adjoint.

9.27 Normal but not self-adjoint operators

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 First suppose (a) holds, so that T is normal but not self-adjoint. Let e_1, e_2 be an orthonormal basis of V. Suppose

9.28 $$\mathcal{M}(T, (e_1, e_2)) = \begin{pmatrix} a & c \\ b & d \end{pmatrix}.$$

Then $\|Te_1\|^2 = a^2 + b^2$ and $\|T^*e_1\|^2 = a^2 + c^2$. Because T is normal, $\|Te_1\| = \|T^*e_1\|$ (see 7.20); thus these equations imply that $b^2 = c^2$. Thus $c = b$ or $c = -b$. But $c \neq b$, because otherwise T would be self-adjoint, as can be seen from the matrix in 9.28. Hence $c = -b$, so

9.29
$$\mathcal{M}\big(T, (e_1, e_2)\big) = \begin{pmatrix} a & -b \\ b & d \end{pmatrix}.$$

The matrix of T^* is the transpose of the matrix above. Use matrix multiplication to compute the matrices of TT^* and T^*T (do it now). Because T is normal, these two matrices are equal. Equating the entries in the upper-right corner of the two matrices you computed, you will discover that $bd = ab$. Now $b \neq 0$, because otherwise T would be self-adjoint, as can be seen from the matrix in 9.29. Thus $d = a$, completing the proof that (a) implies (b).

Now suppose (b) holds. We want to prove that (c) holds. Choose an orthonormal basis e_1, e_2 of V. We know that the matrix of T with respect to this basis has the form given by (b), with $b \neq 0$. If $b > 0$, then (c) holds and we have proved that (b) implies (c). If $b < 0$, then, as you should verify, the matrix of T with respect to the orthonormal basis $e_1, -e_2$ equals $\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$, where $-b > 0$; thus in this case we also see that (b) implies (c).

Now suppose (c) holds, so that the matrix of T with respect to some orthonormal basis has the form given in (c) with $b > 0$. Clearly the matrix of T is not equal to its transpose (because $b \neq 0$). Hence T is not self-adjoint. Now use matrix multiplication to verify that the matrices of TT^* and T^*T are equal. We conclude that $TT^* = T^*T$. Hence T is normal. Thus (c) implies (a), completing the proof. ■

The next result tells us that a normal operator restricted to an invariant subspace is normal. This will allow us to use induction on $\dim V$ when we prove our description of normal operators (9.34).

9.30 Normal operators and invariant subspaces

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;

(b) U is invariant under T^*;

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

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

Proof First we will prove (a). Let $e_1, \dots, e_m$ be an orthonormal basis of U. Extend to an orthonormal basis $e_1, \dots, e_m, f_1, \dots, f_n$ of V (this is possible by 6.35). Because U is invariant under T, each Te_j is a linear combination of $e_1, \dots, e_m$. Thus the matrix of T with respect to the basis $e_1, \dots, e_m, f_1, \dots, f_n$ is of the form

$$\mathcal{M}(T) = \begin{array}{c} \\ e_1 \\ \vdots \\ e_m \\ f_1 \\ \vdots \\ f_n \end{array} \begin{array}{c} \begin{array}{cccccc} e_1 & \dots & e_m & f_1 & \dots & f_n \end{array} \\ \left(\begin{array}{cc} A & B \\ & \\ 0 & C \end{array} \right) \end{array};$$

here A denotes an m-by-m matrix, 0 denotes the n-by-m matrix of all 0's, B denotes an m-by-n matrix, C denotes an n-by-n matrix, and for convenience the basis has been listed along the top and left sides of the matrix.

For each $j \in \{1, \dots, m\}$, $\|Te_j\|^2$ equals the sum of the squares of the absolute values of the entries in the j^{th} column of A (see 6.25). Hence

9.31 $$\sum_{j=1}^{m} \|Te_j\|^2 = \begin{array}{l}\text{the sum of the squares of the absolute} \\ \text{values of the entries of } A.\end{array}$$

For each $j \in \{1, \dots, m\}$, $\|T^*e_j\|^2$ equals the sum of the squares of the absolute values of the entries in the j^{th} rows of A and B. Hence

9.32 $$\sum_{j=1}^{m} \|T^*e_j\|^2 = \begin{array}{l}\text{the sum of the squares of the absolute} \\ \text{values of the entries of } A \text{ and } B.\end{array}$$

Because T is normal, $\|Te_j\| = \|T^*e_j\|$ for each j (see 7.20); thus

$$\sum_{j=1}^{m} \|Te_j\|^2 = \sum_{j=1}^{m} \|T^*e_j\|^2.$$

This equation, along with 9.31 and 9.32, implies that the sum of the squares of the absolute values of the entries of B equals 0. In other words, B is the matrix of all 0's. Thus

9.33
$$\mathcal{M}(T) = \begin{array}{c} e_1 \\ \vdots \\ e_m \\ f_1 \\ \vdots \\ f_n \end{array} \overset{\begin{array}{cccccc} e_1 & \dots & e_m & f_1 & \dots & f_n \end{array}}{\left(\begin{array}{cc} A & 0 \\ & \\ 0 & C \end{array} \right)}.$$

This representation shows that Tf_k is in the span of $f_1, \dots, f_n$ for each k. Because $f_1, \dots, f_n$ is a basis of $U^{\perp}$, this implies that $Tv \in U^{\perp}$ whenever $v \in U^{\perp}$. In other words, $U^{\perp}$ is invariant under T, completing the proof of (a).

To prove (b), note that $\mathcal{M}(T^*)$, which is the conjugate transpose of $\mathcal{M}(T)$, has a block of 0's in the lower left corner (because $\mathcal{M}(T)$, as given above, has a block of 0's in the upper right corner). In other words, each T^*e_j can be written as a linear combination of $e_1, \dots, e_m$. Thus U is invariant under T^*, completing the proof of (b).

To prove (c), let $S = T|_U \in \mathcal{L}(U)$. Fix $v \in U$. Then

$$\begin{aligned} \langle Su, v \rangle &= \langle Tu, v \rangle \\ &= \langle u, T^*v \rangle \end{aligned}$$

for all $u \in U$. Because $T^*v \in U$ [by (b)], the equation above shows that $S^*v = T^*v$. In other words, $(T|_U)^* = (T^*)|_U$, completing the proof of (c).

To prove (d), note that T commutes with T^* (because T is normal) and that $(T|_U)^* = (T^*)|_U$ [by (c)]. Thus $T|_U$ commutes with its adjoint and hence is normal. Interchanging the roles of U and $U^{\perp}$, which is justified by (a), shows that $T|_{U^{\perp}}$ is also normal, completing the proof of (d). ■

Note that if an operator T has a block diagonal matrix with respect to some basis, then the entry in each 1*-by-*1 *block on the diagonal of this matrix is an eigenvalue of T.*

Our next result shows that normal operators on real inner product spaces come close to having diagonal matrices. Specifically, we get block diagonal matrices, with each block having size at most 2-by-2.

We cannot expect to do better than the next result, because on a real inner product space there exist normal operators that do not have a diagonal matrix with respect to any basis. For example, the operator $T \in \mathcal{L}(\mathbf{R}^2)$ defined by $T(x, y) = (-y, x)$ is normal (as you should verify) but has no eigenvalues; thus this particular T does not have even an upper-triangular matrix with respect to any basis of $\mathbf{R}^2$.

9.34 Characterization of normal operators when $\mathbf{F} = \mathbf{R}$

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-by-1 matrix or a 2-by-2 matrix of the form

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

with $b > 0$.

Proof First suppose (b) holds. With respect to the basis given by (b), the matrix of T commutes with the matrix of T^* (which is the transpose of the matrix of T), as you should verify (use Exercise 9 in Section 8.B for the product of two block diagonal matrices). Thus T commutes with T^*, which means that T is normal, completing the proof that (b) implies (a).

Now suppose (a) holds, so T is normal. We will prove that (b) holds by induction on $\dim V$. To get started, note that our desired result holds if $\dim V = 1$ (trivially) or if $\dim V = 2$ [if T is self-adjoint, use the Real Spectral Theorem (7.29); if T is not self-adjoint, use 9.27].

Now assume that $\dim V > 2$ and that the desired result holds on vector spaces of smaller dimension. Let U be a subspace of V of dimension 1 that is invariant under T if such a subspace exists (in other words, if T has an eigenvector, let U be the span of this eigenvector). If no such subspace exists, let U be a subspace of V of dimension 2 that is invariant under T (an invariant subspace of dimension 1 or 2 always exists by 9.8).

If $\dim U = 1$, choose a vector in U with norm 1; this vector will be an orthonormal basis of U, and of course the matrix of $T|_U \in \mathcal{L}(U)$ is a 1-by-1 matrix. If $\dim U = 2$, then $T|_U \in \mathcal{L}(U)$ is normal (by 9.30) but not self-adjoint (otherwise $T|_U$, and hence T, would have an eigenvector by 7.27). Thus we can choose an orthonormal basis of U with respect to which the matrix of $T|_U \in \mathcal{L}(U)$ has the required form (see 9.27).

Now $U^\perp$ is invariant under T and $T|_{U^\perp}$ is a normal operator on $U^\perp$ (by 9.30). Thus by our induction hypothesis, there is an orthonormal basis of $U^\perp$ with respect to which the matrix of $T|_{U^\perp}$ has the desired form. Adjoining this basis to the basis of U gives an orthonormal basis of V with respect to which the matrix of T has the desired form. Thus (b) holds. ■

Isometries on Real Inner Product Spaces

As we will see, the next example is a key building block for isometries on real inner product spaces. Also, note that the next example shows that an isometry on $\mathbf{R}^2$ may have no eigenvalues.

9.35 **Example** Let $\theta \in \mathbf{R}$. Then the operator on $\mathbf{R}^2$ of counterclockwise rotation (centered at the origin) by an angle of θ is an isometry, as is geometrically obvious. The matrix of this operator with respect to the standard basis is

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

If θ is not an integer multiple of π, then no nonzero vector of $\mathbf{R}^2$ gets mapped to a scalar multiple of itself, and hence the operator has no eigenvalues.

The next result shows that every isometry on a real inner product space is composed of pieces that are rotations on 2-dimensional subspaces, pieces that equal the identity operator, and pieces that equal multiplication by -1.

9.36 Description of isometries when $\mathbf{F} = \mathbf{R}$

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

(a) S is an isometry.

(b) There is an orthonormal basis of V with respect to which S has a block diagonal matrix such that each block on the diagonal is a 1-by-1 matrix containing 1 or -1 or is a 2-by-2 matrix of the form

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix},$$

with $\theta \in (0, \pi)$.

Proof First suppose (a) holds, so S is an isometry. Because S is normal, there is an orthonormal basis of V with respect to which S has a block diagonal matrix such that each block is a 1-by-1 matrix or a 2-by-2 matrix of the form

9.37
$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix},$$

with $b > 0$ (by 9.34).

If λ is an entry in a 1-by-1 matrix along the diagonal of the matrix of S (with respect to the basis mentioned above), then there is a basis vector e_j such that $Se_j = \lambda e_j$. Because S is an isometry, this implies that $|\lambda| = 1$. Thus $\lambda = 1$ or $\lambda = -1$, because these are the only real numbers with absolute value 1.

Now consider a 2-by-2 matrix of the form 9.37 along the diagonal of the matrix of S. There are basis vectors e_j, e_{j+1} such that

$$Se_j = ae_j + be_{j+1}.$$

Thus

$$1 = \|e_j\|^2 = \|Se_j\|^2 = a^2 + b^2.$$

The equation above, along with the condition $b > 0$, implies that there exists a number $\theta \in (0, \pi)$ such that $a = \cos\theta$ and $b = \sin\theta$. Thus the matrix 9.37 has the required form, completing the proof in this direction.

Conversely, now suppose (b) holds, so there is an orthonormal basis of V with respect to which the matrix of S has the form required by the theorem. Thus there is a direct sum decomposition

$$V = U_1 \oplus \cdots \oplus U_m,$$

where each U_j is a subspace of V of dimension 1 or 2. Furthermore, any two vectors belonging to distinct U's are orthogonal, and each $S|_{U_j}$ is an isometry mapping U_j into U_j. If $v \in V$, we can write

$$v = u_1 + \cdots + u_m,$$

where each u_j is in U_j. Applying S to the equation above and then taking norms gives

$$\begin{aligned}\|Sv\|^2 &= \|Su_1 + \cdots + Su_m\|^2 \\ &= \|Su_1\|^2 + \cdots + \|Su_m\|^2 \\ &= \|u_1\|^2 + \cdots + \|u_m\|^2 \\ &= \|v\|^2.\end{aligned}$$

Thus S is an isometry, and hence (a) holds. ■

EXERCISES 9.B

1 Suppose $S \in \mathcal{L}(\mathbf{R}^3)$ is an isometry. Prove that there exists a nonzero vector $x \in \mathbf{R}^3$ such that $S^2 x = x$.

2 Prove that every isometry on an odd-dimensional real inner product space has 1 or -1 as an eigenvalue.

3 Suppose V is a real inner product space. Show that

$$\langle u + iv, x + iy \rangle = \langle u, x \rangle + \langle v, y \rangle + \big(\langle v, x \rangle - \langle u, y \rangle\big)i$$

for $u, v, x, y \in V$ defines a complex inner product on $V_{\mathbf{C}}$.

4 Suppose V is a real inner product space and $T \in \mathcal{L}(V)$ is self-adjoint. Show that $T_{\mathbf{C}}$ is a self-adjoint operator on the inner product space $V_{\mathbf{C}}$ defined by the previous exercise.

5 Use the previous exercise to give a proof of the Real Spectral Theorem (7.29) via complexification and the Complex Spectral Theorem (7.24).

6 Give an example of an operator T on an inner product space such that T has an invariant subspace whose orthogonal complement is not invariant under T.
[*The exercise above shows that 9.30 can fail without the hypothesis that T is normal.*]

7 Suppose $T \in \mathcal{L}(V)$ and T has a block diagonal matrix

$$\begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_m \end{pmatrix}$$

with respect to some basis of V. For $j = 1, \dots, m$, let T_j be the operator on V whose matrix with respect to the same basis is a block diagonal matrix with blocks the same size as in the matrix above, with A_j in the j^{th} block, and with all the other blocks on the diagonal equal to identity matrices (of the appropriate size). Prove that $T = T_1 \cdots T_m$.

8 Suppose D is the differentiation operator on the vector space V in Exercise 21 in Section 7.A. Find an orthonormal basis of V such that the matrix of the normal operator D has the form promised by 9.34.