

9.2 Eigenvalues and Eigenvectors

The previous section ended with an observation: an operator has a diagonal matrix exactly when the space splits into invariant lines, and a line $\text{span}(v)$ is invariant exactly when Tv is a multiple of v . This section turns that observation into the central notion of the chapter. We name the vectors that an operator only stretches, and the stretching factors. We find the factors as roots of the characteristic polynomial of Chapter 7, and we prove the two facts that everything later rests on: eigenvectors for different factors are independent, and over $\mathbb{C}$ every operator has at least one eigenvector.

9.2.1 When is a matrix diagonal?

Chapter 4 asked how simple $[T]_{\mathcal{B}}^{\mathcal{B}}$ can be made by choosing $\mathcal{B}$. The simplest square matrices are the diagonal ones, so start with the wish and unwind it.

Let V be finite-dimensional with basis $\mathcal{B} = (v_1, \dots, v_n)$, and $T \in \mathcal{L}(V)$. By Definition 4.1.1, column j of $[T]_{\mathcal{B}}^{\mathcal{B}}$ holds the coordinates of Tv_j . The matrix is $\text{diag}(\lambda_1, \dots, \lambda_n)$ exactly when column j is $\lambda_j \mathbf{e}_j$ for each j , that is,

$$Tv_j = \lambda_j v_j \quad (j = 1, \dots, n).$$

So the wish forces an equation: every basis vector must be sent to a multiple of itself. We met this in Chapter 2, where the reflection $R(x, y) = (y, x)$ became $\text{diag}(1, -1)$ in the basis $((1, 1), (1, -1))$ (Example (The reflection in a mirror-adapted basis)), precisely because $R(1, 1) = (1, 1)$ and $R(1, -1) = -(1, -1)$.

Where should we look for vectors with $Tv = \lambda v$? Rewrite λv as $\lambda \text{id}_V(v)$. Then the equation says $(T - \lambda \text{id}_V)v = 0$: the vector lies in a **kernel**. Kernels are subspaces we know how to compute, so this rewriting is a plan, not just a restatement.

An eigenvector of T is a nonzero vector that T only stretches, and the eigenvalue is the stretching factor.

Definition 9.2.1. Eigenvalue and Eigenvector

Let V be a vector space over F and $T \in \mathcal{L}(V)$. A scalar $\lambda \in F$ is an **eigenvalue** of T if there **exists** a **nonzero** vector $v \in V$ with

$$Tv = \lambda v.$$

Every **nonzero** $v \in V$ with $Tv = \lambda v$ is called an **eigenvector** of T for the eigenvalue λ .

For a matrix $\mathbf{A} \in M_n(F)$, the eigenvalues and eigenvectors of $\mathbf{A}$ are those of $T_{\mathbf{A}}: F^n \rightarrow F^n$: a scalar $\lambda \in F$ and a nonzero $\mathbf{v} \in F^n$ with $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$.

In words: an eigenvalue is a scalar **in the field** F over which V is a vector space. It qualifies if **some** nonzero vector is sent to λ times itself. The vector must be nonzero; the scalar may be anything, including 0. An eigenvector is a witness for its eigenvalue, and there are always many witnesses: if v is one, so is cv for every $c \neq 0$, since $T(cv) = cTv = \lambda(cv)$.

Well-definedness. An eigenvector determines its eigenvalue. If $Tv = \lambda v$ and $Tv = \mu v$ with $v \neq 0$, then $(\lambda - \mu)v = 0$, and Theorem 1.1.9 gives $\lambda = \mu$. So “the eigenvalue of v ” makes sense, while “the eigenvector of λ ” does not.

The eigenvectors for one eigenvalue, together with 0, are exactly the kernel found above. That set is a subspace, and it deserves its own name.

Definition 9.2.2. Eigenspace

Let $T \in \mathcal{L}(V)$ and $\lambda \in F$. The **eigenspace** of T for λ is

$$E_\lambda(T) := \ker(T - \lambda \operatorname{id}_V) = \{v \in V : Tv = \lambda v\}.$$

For $\mathbf{A} \in M_n(F)$ we write $E_\lambda(\mathbf{A}) = E_\lambda(T_{\mathbf{A}}) = \operatorname{nul}(\mathbf{A} - \lambda \mathbf{I}_n)$.

$E_\lambda(T)$ is a subspace by Theorem 2.2.1, and it is T -invariant by Theorem 6.7.4 (b) with $p = x - \lambda$. It is defined for **every** $\lambda \in F$, and λ is an eigenvalue exactly when $E_\lambda(T) \neq \{0\}$; in that case the eigenvectors for λ are the nonzero vectors of $E_\lambda(T)$. The eigenspace is the home of all eigenvectors for λ at once, and it is one of the invariant subspaces of the previous section: on it, T acts as $\lambda \operatorname{id}$.

Examples.

- **Diagonal matrices.** Let $\mathbf{A} = \operatorname{diag}(d_1, \dots, d_n)$. Then $\mathbf{A}\mathbf{e}_j = d_j\mathbf{e}_j$ and $\mathbf{e}_j \neq \mathbf{0}$, so each d_j is an eigenvalue with eigenvector $\mathbf{e}_j$. The standard basis is a basis of eigenvectors, which is the observation we started from, read backwards.
- **Degenerate cases.** On $V \neq \{0\}$, the identity has the eigenvalue 1 with $E_1(\operatorname{id}_V) = V$: **every** nonzero vector is an eigenvector. The zero operator has the eigenvalue 0 with $E_0 = V$. On $V = \{0\}$, no operator has any eigenvalue, because there is no nonzero vector to be a witness. That last case matters: it is why the existence theorem later in this section must assume $V \neq \{0\}$.
- **Projections.** Let $P \in \mathcal{L}(V)$ be a projection, $P^2 = P$ (Definition 4.5.1). If $Pv = \lambda v$ with $v \neq 0$, then $\lambda v = Pv = P^2v = P(\lambda v) = \lambda^2 v$, so $(\lambda^2 - \lambda)v = 0$ and $\lambda(\lambda - 1) = 0$ by Theorem 1.1.9. So the only possible eigenvalues are 0 and 1. By definition $E_0(P) = \ker P$. Also $E_1(P) = \operatorname{im} P$: a vector with $Pv = v$ is in the image, and conversely P is the identity on $\operatorname{im} P$ by Theorem 4.5.1 (a). So a vector is fixed by P exactly when it is in the image, and killed exactly when it is in the kernel.
- **Differentiation of polynomials.** Let D act on $V = \mathbb{R}[x]_{\leq n}$ with $n \geq 1$. Suppose $Dp = \lambda p$ with $p \neq 0$. If $\lambda \neq 0$, then $\deg(\lambda p) = \deg p$, while Dp is either 0 or of degree $\deg p - 1$; neither equals $\deg p$ for $p \neq 0$. So $\lambda = 0$, and $E_0(D) = \ker D$ is the space of constant polynomials. Differentiation on $\mathbb{R}[x]_{\leq n}$ has exactly one eigenvalue, and its eigenvectors span only a line.

Nonexample by minimal change. For $\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 0 & 3 \end{pmatrix}$, the vector $\mathbf{e}_1$ is an eigenvector: $\mathbf{A}\mathbf{e}_1 = 2\mathbf{e}_1$. Move to the neighboring basis vector $\mathbf{e}_2$. It is still nonzero, and $\mathbf{A}\mathbf{e}_2 = (1, 3)$ is still a vector of $\mathbb{R}^2$. What fails is the equation $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$: no λ makes $(1, 3) = (0, \lambda)$. So $\mathbf{e}_2$ is not an eigenvector, and the line $\operatorname{span}(\mathbf{e}_2)$ is not invariant, as we saw in the previous section.

Why this definition. The clause “**nonzero**” is what gives the notion content. The zero vector satisfies $T0 = \lambda 0$ for **every** $\lambda \in F$, so if we allowed $v = 0$, every scalar would be an eigenvalue of every operator. On the other hand we do **not** exclude $\lambda = 0$: the equation $Tv = 0v$ says $v \in \ker T$, and a nonzero kernel vector is exactly the kind of information we want to record. Requiring $\lambda \in F$, rather than in some larger field, makes the notion depend on the field, and this is deliberate: a real matrix may have no real eigenvalues, and then no real basis diagonalizes it. The prefix *eigen* is German for “own”: an eigenvector is a direction that belongs to T , which T maps into itself.

Warning. The eigenvalue may be 0; the eigenvector may not. For $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$, the vector $(2, -1)$ satisfies $\mathbf{A}(2, -1) = (0, 0) = 0 \cdot (2, -1)$, so 0 is an eigenvalue. By contrast, $\mathbf{A}\mathbf{0} = 7 \cdot \mathbf{0}$ does **not** make 7 an eigenvalue. And say “**an** eigenvector for 0”, never “the”: $(2, -1)$, $(-4, 2)$, and $(200, -100)$ are all eigenvectors for 0.

Quick check. Let $T(x, y, z) = (x, z, y)$ on $\mathbb{R}^3$. Which of $\mathbf{e}_1$, $(0, 1, 1)$, $(0, 1, -1)$, $\mathbf{e}_2$ are eigenvectors of T , and for which eigenvalues?

Solution. $T\mathbf{e}_1 = \mathbf{e}_1$, so $\mathbf{e}_1$ is an eigenvector for 1. $T(0, 1, 1) = (0, 1, 1)$, an eigenvector for 1 as well. $T(0, 1, -1) = (0, -1, 1) = -(0, 1, -1)$, an eigenvector for -1 . $T\mathbf{e}_2 = \mathbf{e}_3$, which is not a multiple of $\mathbf{e}_2$, so $\mathbf{e}_2$ is **not** an eigenvector. The three eigenvectors form a basis $\mathcal{B}$ of $\mathbb{R}^3$, and $[T]_{\mathcal{B}}^{\mathcal{B}} = \text{diag}(1, 1, -1)$.

9.2.2 Finding eigenvalues

The definition asks for a nonzero vector in $\ker(T - \lambda \text{id}_V)$, a question about one λ at a time. In finite dimension, Chapters 2 and 7 turn it into a test that handles all λ at once. The last item is the one we compute with.

Theorem 9.2.1. Characterizations of Eigenvalues

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

- (a) λ is an eigenvalue of T ;
- (b) $E_{\lambda}(T) \neq \{0\}$, that is, $T - \lambda \text{id}_V$ is **not injective**;
- (c) $T - \lambda \text{id}_V$ is **not invertible**;
- (d) $\det(\lambda \text{id}_V - T) = 0$;
- (e) $p_T(\lambda) = 0$.

In particular T has at most n eigenvalues. For $\mathbf{A} \in M_n(F)$: λ is an eigenvalue of $\mathbf{A}$ if and only if $p_{\mathbf{A}}(\lambda) = 0$, if and only if $\lambda \mathbf{I} - \mathbf{A}$ is not invertible.

Proof.

- (a) $\Leftrightarrow$ (b). By Definition 9.2.1 and Definition 9.2.2, λ is an eigenvalue exactly when $E_{\lambda}(T) = \ker(T - \lambda \text{id}_V)$ contains a nonzero vector. By Theorem 2.2.3, this happens exactly when $T - \lambda \text{id}_V$ is not injective.
- (b) $\Leftrightarrow$ (c). V is finite-dimensional, so by Theorem 2.5.7 ((a) $\Leftrightarrow$ (b)) an operator on V is invertible if and only if it is injective.
- (c) $\Leftrightarrow$ (d). The operators $\lambda \text{id}_V - T$ and $T - \lambda \text{id}_V$ differ by the factor -1 , so one is invertible exactly when the other is. By Theorem 7.5.8, $\lambda \text{id}_V - T$ is invertible if and only if $\det(\lambda \text{id}_V - T) \neq 0$.

- (d) $\Leftrightarrow$ (e). Fix a basis $\mathcal{B}$ of V and put $\mathbf{A} = [T]_{\mathcal{B}}^{\mathcal{B}}$. By Corollary 4.1.3, $[\lambda \text{id}_V - T]_{\mathcal{B}}^{\mathcal{B}} = \lambda \mathbf{I}_n - \mathbf{A}$, and by Theorem 4.1.5 (b), $\lambda \text{id}_V - T$ is invertible if and only if $\lambda \mathbf{I}_n - \mathbf{A}$ is. By Theorem 7.10.4 ((a), (i) $\Leftrightarrow$ (ii)), $\lambda \mathbf{I}_n - \mathbf{A}$ is not invertible if and only if $p_{\mathbf{A}}(\lambda) = 0$, and $p_{\mathbf{A}} = p_T$ by Definition 7.10.2.

The bound on the number of eigenvalues is Theorem 7.10.4 (b) applied to $\mathbf{A}$, since by (e) the eigenvalues of T are the roots of $p_{\mathbf{A}}$ in F . The statement for matrices is the case $T = T_{\mathbf{A}}$, $\mathcal{B}$ standard, $[T_{\mathbf{A}}]_{\mathcal{B}}^{\mathcal{B}} = \mathbf{A}$. ■

For matrices, this is Theorem 7.10.4 read in the new language: Chapter 7 found the scalars c with $c\mathbf{I} - \mathbf{A}$ singular, and those are the eigenvalues. It gives a procedure with two steps.

1. Compute $p_{\mathbf{A}}(x) = \det(x\mathbf{I} - \mathbf{A})$ and find its roots **in** F . These are the eigenvalues.
2. For each eigenvalue λ , solve the homogeneous system $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$ by elimination. Its solution set is $E_{\lambda}(\mathbf{A})$, and a basis of it lists “all” eigenvectors for λ (Theorem 3.5.4).

Example. Eigenvalues and Eigenspaces of a 2×2 Matrix

Find the eigenvalues and eigenspaces of $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix} \in M_2(\mathbb{R})$.

Solution. Step 1. For a 2×2 matrix, $p_{\mathbf{A}}(x) = x^2 - (\text{tr } \mathbf{A})x + \det \mathbf{A}$ (the examples after Definition 7.10.1). Here $\text{tr } \mathbf{A} = 3$ and $\det \mathbf{A} = 2 - 6 = -4$, so

$$p_{\mathbf{A}}(x) = x^2 - 3x - 4 = (x - 4)(x + 1).$$

By Theorem 9.2.1, the eigenvalues are 4 and -1 .

Step 2. For $\lambda = 4$,

$$\mathbf{A} - 4\mathbf{I} = \begin{pmatrix} -3 & 2 \\ 3 & -2 \end{pmatrix}.$$

The second row is -1 times the first, so the system reduces to $-3x + 2y = 0$, with solutions $(x, y) = t(2, 3)$. Hence $E_4(\mathbf{A}) = \text{span}((2, 3))$. For $\lambda = -1$,

$$\mathbf{A} + \mathbf{I} = \begin{pmatrix} 2 & 2 \\ 3 & 3 \end{pmatrix},$$

which reduces to $x + y = 0$, so $E_{-1}(\mathbf{A}) = \text{span}((1, -1))$.

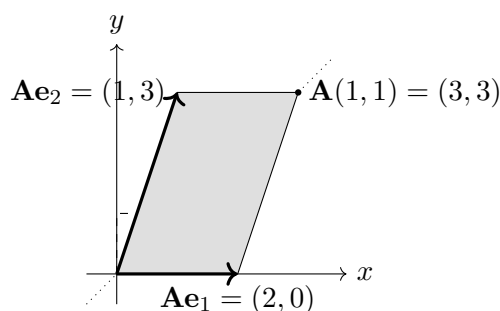
Check. $\mathbf{A}(2, 3) = (2 + 6, 6 + 6) = (8, 12) = 4(2, 3)$ and $\mathbf{A}(1, -1) = (1 - 2, 3 - 2) = (-1, 1) = -1(1, -1)$. In the basis $((2, 3), (1, -1))$, the operator $T_{\mathbf{A}}$ has matrix $\text{diag}(4, -1)$, by the computation at the start of this section.

In Step 2, the matrix $\mathbf{A} - \lambda\mathbf{I}$ **must** be singular; if elimination produces only the trivial solution, then λ was not a root of $p_{\mathbf{A}}$, and Step 1 contains an error. This is a built-in check.

Triangular matrices. If $\mathbf{A}$ is upper or lower triangular, then $p_{\mathbf{A}}(x) = (x - a_{11}) \cdots (x - a_{nn})$ (the examples after Definition 7.10.1), so the eigenvalues of a triangular matrix are its diagonal

entries. For $\begin{pmatrix} 2 & 1 \\ 0 & 3 \end{pmatrix}$ they are 2 and 3, with eigenvectors $\mathbf{e}_1$ and $(1, 1)$, found in the previous section.

You can watch eigenvectors in the widget below. It shows the unit square and its image under $\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 0 & 3 \end{pmatrix}$. The image $\mathbf{A}\mathbf{e}_1$ lies on the x -axis, which says that $\mathbf{e}_1$ is an eigenvector. The far corner of the image parallelogram is $\mathbf{A}(1, 1) = \mathbf{A}\mathbf{e}_1 + \mathbf{A}\mathbf{e}_2 = (3, 3)$, and it lies on the diagonal line through $(1, 1)$: so $(1, 1)$ is an eigenvector too, for the eigenvalue 3. Drag $\mathbf{A}\mathbf{e}_2$ to $(1, 2)$, which changes the matrix to $\begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$. Now the far corner $(3, 2)$ leaves the diagonal, and only the x -axis is still mapped into itself. The next section explains what was lost.



The unit square (dashed) and its image under $\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 0 & 3 \end{pmatrix}$. The image of $\mathbf{e}_1$ stays on the x -axis (eigenvalue 2), and the image of $(1, 1)$ stays on the dotted diagonal (eigenvalue 3). The image of $\mathbf{e}_2$ leaves the y -axis, so $\mathbf{e}_2$ is not an eigenvector.

The set of all eigenvalues is used so often that it gets a symbol.

Definition 9.2.3. Spectrum

Let $T \in \mathcal{L}(V)$. The **spectrum** of T is the set $\text{spec}(T) \subseteq F$ of all eigenvalues of T . For $\mathbf{A} \in M_n(F)$, $\text{spec}(\mathbf{A}) := \text{spec}(T_{\mathbf{A}})$.

For instance $\text{spec}\begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix} = \{4, -1\}$, $\text{spec}(\text{id}_V) = \{1\}$ for $V \neq \{0\}$, and the spectrum of a projection is contained in $\{0, 1\}$. By [Theorem 9.2.1](#), if $\dim V = n \geq 1$, then $\text{spec}(T)$ is the set of roots of p_T in F , and it has at most n elements. It may be empty: the characteristic polynomial of the rotation $\mathbf{R} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ is $x^2 + 1$, which has no real root, so $\text{spec}(\mathbf{R}) = \emptyset$ over $\mathbb{R}$. This matches Example (A Rotation Has No Invariant Lines): no invariant lines, no eigenvectors.

Infinite dimension. Without a characteristic polynomial, we go back to the definition, and anything can happen.

- **Differentiation of smooth functions.** Let V be the space of infinitely differentiable functions $\mathbb{R} \rightarrow \mathbb{R}$ (a subspace of $\mathbb{R}^{\mathbb{R}}$, by the argument of Example (Differentiable Functions) applied to all derivatives), and $Df = f'$. Here we take three facts from calculus as known: $e^{\lambda x}$ is smooth with derivative $\lambda e^{\lambda x}$; the product rule; and a function on $\mathbb{R}$ with derivative 0 everywhere is constant. For **every** $\lambda \in \mathbb{R}$, the function $f(x) = e^{\lambda x}$ is nonzero and satisfies $Df = \lambda f$. So $\text{spec}(D) = \mathbb{R}$, an infinite set. In fact $E_{\lambda}(D) = \text{span}(e^{\lambda x})$: if $f' = \lambda f$, then by the product rule $g(x) = e^{-\lambda x} f(x)$ has $g' = e^{-\lambda x}(f' - \lambda f) = 0$, so g is a constant c and

$$f = ce^{\lambda x}.$$

- **The right shift.** Let $R(s_0, s_1, s_2, \dots) = (0, s_0, s_1, \dots)$ on $F^{\mathbb{N}}$ (Example (The Shift Maps on Sequences)). If $Rs = \lambda s$, compare entries: $0 = \lambda s_0$, and $s_k = \lambda s_{k+1}$ for all $k \geq 0$. If $\lambda = 0$, the second equation gives $s_k = 0$ for all k . If $\lambda \neq 0$, the first gives $s_0 = 0$, and then $s_{k+1} = \lambda^{-1}s_k$ gives $s_1 = 0$, $s_2 = 0$, and so on by induction. Either way $s = 0$. So $\text{spec}(R) = \emptyset$, over **every** field.

Warning. In infinite dimension, “not invertible” does not produce an eigenvalue. The step (b) $\Leftrightarrow$ (c) of Theorem 9.2.1 used finite dimension. The right shift R is injective but not surjective (Example (The Shifts: Injective but Not Surjective, and Conversely)), so $R - 0 \cdot \text{id}$ is not invertible, yet 0 is not an eigenvalue of R : nothing nonzero is sent to 0.

9.2.3 Distinct eigenvalues give independent eigenvectors

In Example (Eigenvalues and Eigenspaces of a 2×2 Matrix) the eigenvectors $(2, 3)$ and $(1, -1)$ for the eigenvalues 4 and -1 formed a basis. That is no accident. Eigenvectors for different eigenvalues can never be linearly dependent.

Theorem 9.2.2. Distinct Eigenvalues Give Independent Eigenvectors

Let $T \in \mathcal{L}(V)$, let $\lambda_1, \dots, \lambda_k \in F$ be **distinct** eigenvalues of T , and for each i let v_i be an eigenvector of T for λ_i . Then $(v_1, \dots, v_k)$ is linearly independent.

Idea. Try two vectors first. From $av_1 + bv_2 = 0$, we want to kill one term, and the operator $T - \lambda_2 \text{id}_V$ does exactly that: it sends v_2 to 0 and v_1 to $(\lambda_1 - \lambda_2)v_1$. So $a(\lambda_1 - \lambda_2)v_1 = 0$, and since $\lambda_1 \neq \lambda_2$ and $v_1 \neq 0$, $a = 0$; then $b = 0$. In general, applying $T - \lambda_j \text{id}_V$ to a dependence kills the j -th vector and rescales the others by nonzero factors, producing a **shorter** dependence. So a shortest dependence cannot exist. The Linear Dependence Lemma lets us pick that shortest one cleanly.

Proof. Suppose, for a contradiction, that $(v_1, \dots, v_k)$ is linearly dependent. By Theorem 1.4.1 (1), there is an index j with $v_j \in \text{span}(v_1, \dots, v_{j-1})$; choose the **smallest** such j . Since $v_1 \neq 0$, $j \geq 2$. The list $(v_1, \dots, v_{j-1})$ is linearly independent: otherwise Theorem 1.4.1 (1), applied to it, would give a smaller index with the same property. Write

$$v_j = a_1 v_1 + \dots + a_{j-1} v_{j-1} \quad (a_i \in F).$$

Apply $T - \lambda_j \text{id}_V$ to both sides. Since $(T - \lambda_j \text{id}_V)v_i = (\lambda_i - \lambda_j)v_i$ for each i , linearity gives

$$0 = a_1(\lambda_1 - \lambda_j)v_1 + \dots + a_{j-1}(\lambda_{j-1} - \lambda_j)v_{j-1}.$$

By independence of $(v_1, \dots, v_{j-1})$, every coefficient $a_i(\lambda_i - \lambda_j)$ is 0. The eigenvalues are distinct, so $\lambda_i - \lambda_j \neq 0$ for $i < j$, and hence $a_i = 0$ for all $i < j$. Then $v_j = 0$, which contradicts the fact that an eigenvector is nonzero. This proves that $(v_1, \dots, v_k)$ is linearly independent.

The theorem works in any vector space, finite-dimensional or not. For instance, for distinct real numbers $\lambda_1, \dots, \lambda_k$, the functions $e^{\lambda_1 x}, \dots, e^{\lambda_k x}$ are eigenvectors of differentiation for distinct eigenvalues, so they are linearly independent, with no computation at all. Two consequences follow at once, and the second is the first half of a promise made in Chapter 1, §7: eigenspaces sit inside V as a direct sum.

Corollary 9.2.3. Eigenspaces Form a Direct Sum

Let $T \in \mathcal{L}(V)$ and let $\lambda_1, \dots, \lambda_k \in F$ be distinct. Then the sum $E_{\lambda_1}(T) + \dots + E_{\lambda_k}(T)$ is direct. If V is finite-dimensional, then

$$\dim E_{\lambda_1}(T) + \dots + \dim E_{\lambda_k}(T) \leq \dim V,$$

and T has at most $\dim V$ distinct eigenvalues.

Proof. By Theorem 1.7.5 ((b) $\Rightarrow$ (a)), it suffices to show: if $u_1 + \dots + u_k = 0$ with $u_i \in E_{\lambda_i}(T)$, then every $u_i = 0$. Suppose not, and let S be the nonempty set of indices i with $u_i \neq 0$. For $i \in S$, u_i is an eigenvector of T for λ_i , and these eigenvalues are distinct. By Theorem 9.2.2, the list $(u_i)_{i \in S}$ is linearly independent. But $\sum_{i \in S} 1 \cdot u_i = \sum_{i=1}^k u_i = 0$ is a relation with all coefficients $1 \neq 0$, a contradiction. So the sum is direct.

If V is finite-dimensional, Theorem 1.7.5 ((a) $\Rightarrow$ (e)) gives $\dim(E_{\lambda_1}(T) \oplus \dots \oplus E_{\lambda_k}(T)) = \sum_i \dim E_{\lambda_i}(T)$, and this is at most $\dim V$ by Theorem 1.6.7. If the λ_i are eigenvalues, each $\dim E_{\lambda_i}(T) \geq 1$, so $k \leq \dim V$.

The corollary gives a second, polynomial-free proof that an operator on an n -dimensional space has at most n eigenvalues. Section 4 shows that T has a diagonal matrix exactly when the inequality is an equality.

9.2.4 Existence over the complex numbers

We have seen operators without eigenvalues: the rotation of $\mathbb{R}^2$, and the right shift. The rotation fails because $\mathbb{R}$ is missing the roots of $x^2 + 1$; the shift fails because its space is infinite-dimensional. Remove both obstacles and eigenvalues always exist.

Theorem 9.2.4. Complex Operators Have Eigenvalues

Let V be a finite-dimensional vector space over $\mathbb{C}$ with $V \neq \{0\}$, and let $T \in \mathcal{L}(V)$. Then T has an eigenvalue. The same holds for every $\mathbf{A} \in M_n(\mathbb{C})$ with $n \geq 1$.

Proof. Let $n = \dim V \geq 1$. By Theorem 7.10.2 and Definition 7.10.2, $p_T \in \mathbb{C}[x]$ is monic of degree $n \geq 1$, so it is not constant. By the Fundamental Theorem of Algebra (Theorem 6.5.1), $p_T(\lambda) = 0$ for some $\lambda \in \mathbb{C}$. By Theorem 9.2.1 ((e) $\Rightarrow$ (a)), λ is an eigenvalue of T . For $\mathbf{A} \in M_n(\mathbb{C})$, apply this to $T_{\mathbf{A}}$.

Alternatively, determinants can be avoided altogether, and the proof then shows where the eigenvector comes from. By Theorem 6.7.2, there is a nonzero $p \in \mathbb{C}[x]$ with $p(T) = 0$. It is not a nonzero constant c , since $c \operatorname{id}_V \neq 0$ on $V \neq \{0\}$. By Corollary 6.5.5, $p = c(x - z_1) \cdots (x - z_m)$ with $c \neq 0$ and $m \geq 1$, so by Theorem 6.7.1 (a) and (b),

$$0 = p(T) = c(T - z_1 \operatorname{id}_V)(T - z_2 \operatorname{id}_V) \cdots (T - z_m \operatorname{id}_V).$$

Since $c \neq 0$, multiplying by c^{-1} shows that the composite $(T - z_1 \operatorname{id}_V) \cdots (T - z_m \operatorname{id}_V)$ is the zero operator, which is not injective on $V \neq \{0\}$. A composition of injective maps is injective (Theorem 0.4.3 (a), applied $m - 1$ times), so at least one factor $T - z_j \operatorname{id}_V$ is not injective, and z_j is an eigenvalue by Theorem 9.2.1 ((b) $\Rightarrow$ (a)).

Both routes use the Fundamental Theorem of Algebra, and both need all three hypotheses. Over $\mathbb{R}$ the rotation has no eigenvalue. In infinite dimension the right shift on $\mathbb{C}^{\mathbb{N}}$ has none, and indeed no nonzero polynomial annihilates it: for $p = a_0 + a_1x + \cdots + a_Nx^N \neq 0$, the sequence $p(R)(1, 0, 0, \dots) = (a_0, a_1, \dots, a_N, 0, \dots)$ is not zero. On $V = \{0\}$ there is nothing to be an eigenvector. This theorem pays off the promise made in Chapter 6, and it is the base step of every structure theorem over $\mathbb{C}$ in this book: find one eigenvector, then look at what is left.

Quick check. Does every operator on $\mathbb{C}^3$ have an invariant line? Does every operator on $\mathbb{R}^2$?

Solution. Every operator on $\mathbb{C}^3$ has one: by Theorem 9.2.4 it has an eigenvector $\mathbf{v}$, and $\operatorname{span}(\mathbf{v})$ is invariant by Proposition 9.1.2. Not every operator on $\mathbb{R}^2$: the rotation by a right angle has no invariant line (Example (A Rotation Has No Invariant Lines)).

9.2.5 Real matrices and complex eigenvalues

A real matrix may have no real eigenvalues, but it is also a complex matrix, and as such it has eigenvalues by Theorem 9.2.4. This is the situation Chapter 1, §9 prepared for. Let $\mathbf{A} \in M_n(\mathbb{R})$. The complexification of $\mathbb{R}^n$ is $\mathbb{C}^n$: the pair $(\mathbf{x}, \mathbf{y})$ is the vector $\mathbf{x} + i\mathbf{y}$ (Example (The Complexification of $\mathbb{R}^n$ Is $\mathbb{C}^n$)). The real operator $T_{\mathbf{A}}$ extends to the complexification by the only rule compatible with complex scalars,

$$\mathbf{x} + i\mathbf{y} \mapsto \mathbf{A}\mathbf{x} + i\mathbf{A}\mathbf{y} = \mathbf{A}(\mathbf{x} + i\mathbf{y}),$$

which is just $T_{\mathbf{A}}$ for $\mathbf{A}$ regarded as a matrix in $M_n(\mathbb{C})$. By Theorem 1.9.3, the real standard basis is also a complex basis of $\mathbb{C}^n$, and in it the extended operator still has matrix $\mathbf{A}$; in particular it has the same characteristic polynomial $p_{\mathbf{A}} \in \mathbb{R}[x] \subseteq \mathbb{C}[x]$. So the **complex eigenvalues** of a real matrix are the roots of $p_{\mathbf{A}}$ in $\mathbb{C}$, with eigenvectors in $\mathbb{C}^n$, and the real eigenvalues are among them. The complex ones cannot appear alone.

Theorem 9.2.5. Complex Eigenvalues of Real Matrices Come in Conjugate Pairs

Let $\mathbf{A} \in M_n(\mathbb{R})$, regarded as an element of $M_n(\mathbb{C})$. If $\lambda \in \mathbb{C}$ is an eigenvalue of $\mathbf{A}$ with eigenvector $\mathbf{v} \in \mathbb{C}^n$, then $\bar{\lambda}$ is an eigenvalue of $\mathbf{A}$ with eigenvector $\bar{\mathbf{v}}$, the vector of conjugated entries. Moreover:

- (a) $\text{mult}_\lambda(p_{\mathbf{A}}) = \text{mult}_{\bar{\lambda}}(p_{\mathbf{A}})$;
- (b) if $\lambda \notin \mathbb{R}$, then $(\mathbf{v}, \bar{\mathbf{v}})$ is linearly independent over $\mathbb{C}$.

Proof. Let $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$ with $\mathbf{v} = (v_1, \dots, v_n) \neq \mathbf{0}$. For each k , the k -th entry of $\mathbf{A}\mathbf{v}$ is $\sum_j a_{kj}v_j$ (Definition 0.9.9). By Theorem 0.6.2 (a) and (b), and since $\overline{a_{kj}} = a_{kj}$ for the real entries of $\mathbf{A}$,

$$\overline{\sum_j a_{kj}v_j} = \sum_j a_{kj}\bar{v}_j, \quad \overline{\lambda v_k} = \bar{\lambda} \bar{v}_k.$$

Conjugating the equation $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$ entry by entry therefore gives $\mathbf{A}\bar{\mathbf{v}} = \bar{\lambda}\bar{\mathbf{v}}$. Since $\overline{\bar{v}_k} = v_k$, $\bar{\mathbf{v}} = \mathbf{0}$ would force $\mathbf{v} = \mathbf{0}$; so $\bar{\mathbf{v}} \neq \mathbf{0}$, and $\bar{\mathbf{v}}$ is an eigenvector for $\bar{\lambda}$.

- (a) The entries of $x\mathbf{I} - \mathbf{A}$ lie in $\mathbb{R}[x]$, so its determinant $p_{\mathbf{A}}$ lies in $\mathbb{R}[x]$ (Definition 7.10.1). By Theorem 6.5.6, $\text{mult}_{\bar{\lambda}}(p_{\mathbf{A}}) = \text{mult}_\lambda(p_{\mathbf{A}})$.
- (b) If $\lambda \notin \mathbb{R}$, then $\bar{\lambda} \neq \lambda$, so $\mathbf{v}$ and $\bar{\mathbf{v}}$ are eigenvectors for distinct eigenvalues, and they are independent by Theorem 9.2.2. ■

For the rotation $\mathbf{R} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, the vector $(1, -i)$ satisfies $\mathbf{R}(1, -i) = (i, 1) = i(1, -i)$, so i is an eigenvalue. Conjugating, $\mathbf{R}(1, i) = (-i, 1) = -i(1, i)$. Over $\mathbb{C}$, the two eigenvectors form a basis of $\mathbb{C}^2$, and the rotation has a diagonal matrix $\text{diag}(i, -i)$ in it, although no real basis diagonalizes it. So the field matters for eigenvalues, and it will matter for diagonalization.

Example. A Real Matrix with Non-Real Eigenvalues

Find the eigenvalues of $\mathbf{A} = \begin{pmatrix} 3 & -2 \\ 4 & -1 \end{pmatrix}$ over $\mathbb{C}$, with an eigenvector for each.

Solution. Here $\text{tr } \mathbf{A} = 2$ and $\det \mathbf{A} = -3 + 8 = 5$, so $p_{\mathbf{A}}(x) = x^2 - 2x + 5 = (x - 1)^2 + 4$. Over $\mathbb{R}$ this has no root, so $\text{spec}(\mathbf{A}) = \emptyset$ over $\mathbb{R}$. Over $\mathbb{C}$, $(x - 1)^2 = -4$ gives the roots $1 + 2i$ and $1 - 2i$.

For $\lambda = 1 + 2i$,

$$\mathbf{A} - \lambda\mathbf{I} = \begin{pmatrix} 2 - 2i & -2 \\ 4 & -2 - 2i \end{pmatrix}.$$

The first row gives $(2 - 2i)x = 2y$, that is, $y = (1 - i)x$; take $\mathbf{v} = (1, 1 - i)$. Check the second row: $4 \cdot 1 + (-2 - 2i)(1 - i) = 4 + (-2 + 2i - 2i + 2i^2) = 4 - 4 = 0$. So $\mathbf{v}$ is an eigenvector for $1 + 2i$. By Theorem 9.2.5, $\bar{\mathbf{v}} = (1, 1 + i)$ is an eigenvector for $1 - 2i$, with no further computation.

9.2.6 The invertible matrix theorem grows

Chapter 2 collected the conditions equivalent to invertibility for an operator (Theorem 2.5.7), Chapter 3 gave the matrix form, the eight conditions of the Invertible Matrix Theorem (Theorem 3.4.4), and Chapter 7 added the determinant (Theorem 7.5.3). Eigenvalues give one more:

a matrix is singular exactly when it kills a nonzero vector, which is to say, when it stretches some nonzero vector by the factor 0.

Theorem 9.2.6. Invertible Matrix Theorem, with Eigenvalues

Let $\mathbf{A} \in M_n(F)$ with $n \geq 1$. The following are equivalent, and each is equivalent to each of the conditions of Theorem 7.5.3 and Theorem 3.4.4.

- (a) $\mathbf{A}$ is invertible.
- (b) 0 is **not** an eigenvalue of $\mathbf{A}$, that is, $0 \notin \text{spec}(\mathbf{A})$.

Likewise, an operator T on a finite-dimensional space $V \neq \{0\}$ is invertible if and only if $0 \notin \text{spec}(T)$.

Proof. By Definition 9.2.2, $E_0(\mathbf{A}) = \ker(T_{\mathbf{A}} - 0 \text{ id}) = \ker T_{\mathbf{A}}$. So $0 \notin \text{spec}(\mathbf{A})$ if and only if $\ker T_{\mathbf{A}} = \{0\}$, if and only if $T_{\mathbf{A}}$ is injective (Theorem 2.2.3), if and only if $\mathbf{A}$ is invertible (Theorem 7.5.3, (a) $\Leftrightarrow$ (b)). For an operator, $E_0(T) = \ker T$, and Theorem 2.5.7 ((a) $\Leftrightarrow$ (c)) finishes the same way. ■

For example, $\text{spec} \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix} = \{4, -1\}$ does not contain 0, so that matrix is invertible, while $\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$ has the eigenvalue 0 and is singular.

Eigenvalues are easy to compute for one matrix, and it is tempting to assume they behave well under the operations we use on matrices. Two such assumptions are false.

Warning. Eigenvalues do not add or multiply. Let $\mathbf{A} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$. Both are triangular with zero diagonal, so $\text{spec}(\mathbf{A}) = \text{spec}(\mathbf{B}) = \{0\}$. But $\mathbf{A} + \mathbf{B} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ has $p(x) = x^2 - 1$ and eigenvalues ± 1 , which are not sums of eigenvalues of $\mathbf{A}$ and $\mathbf{B}$; and $\mathbf{AB} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ has the eigenvalue 1, which is not a product of them.

Warning. Row operations change eigenvalues. Elimination preserves the solution set of $\mathbf{Ax} = \mathbf{0}$ (Theorem 3.2.2), not the solution sets of $\mathbf{Ax} = \lambda\mathbf{x}$. For $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}$, with eigenvalues 4 and -1 , subtracting 3 times row 1 from row 2 gives the triangular matrix $\begin{pmatrix} 1 & 2 \\ 0 & -4 \end{pmatrix}$, whose eigenvalues are 1 and -4 . So never row reduce $\mathbf{A}$ to find its eigenvalues; row reduce $\mathbf{A} - \lambda\mathbf{I}$ to find eigenvectors, once λ is known.

9.2.7 Exercises

9.2.7.1 A. Check your understanding

Exercise. A1

- (a) Define eigenvalue, eigenvector, and eigenspace of an operator $T \in \mathcal{L}(V)$.
- (b) True or false: 0 is an eigenvector of every operator. Justify your answer.
- (c) True or false: if 0 is an eigenvalue of $\mathbf{A} \in M_n(F)$, then $\mathbf{A}$ is not invertible. Justify your answer.
- (d) State four conditions equivalent to “ λ is an eigenvalue of T ” for an operator on a finite-dimensional space $V \neq \{0\}$.
- (e) True or false: every operator on $\mathbb{R}^2$ has an eigenvalue. Justify your answer.
- (f) True or false: the eigenvalues of $\mathbf{A}$ are the diagonal entries of a row echelon form of $\mathbf{A}$. Justify your answer.

Solution.

- (a) $\lambda \in F$ is an eigenvalue of T if $Tv = \lambda v$ for some **nonzero** $v \in V$; each such nonzero v is an eigenvector for λ ([Definition 9.2.1](#)). The eigenspace is $E_\lambda(T) = \ker(T - \lambda \text{id}_V)$ ([Definition 9.2.2](#)).
- (b) False. Eigenvectors are nonzero by definition. (The equation $T0 = \lambda 0$ holds for every λ , which is why 0 is excluded.)
- (c) True, by [Theorem 9.2.6](#).
- (d) By [Theorem 9.2.1](#): $T - \lambda \text{id}_V$ is not injective; $T - \lambda \text{id}_V$ is not invertible; $\det(\lambda \text{id}_V - T) = 0$; $p_T(\lambda) = 0$.
- (e) False. The rotation $R(x, y) = (-y, x)$ has $p_R = x^2 + 1$, with no real root, so it has no real eigenvalue.
- (f) False. Row operations change eigenvalues: $\begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix}$ has eigenvalues $4, -1$, but the echelon form $\begin{pmatrix} 1 & 2 \\ 0 & -4 \end{pmatrix}$ has diagonal entries $1, -4$.

9.2.7.2 B. Practice

Exercise. B1: Computing eigenvalues and eigenspaces

Find the eigenvalues and a basis of each eigenspace of the following real matrices.

- (a) $\mathbf{A} = \begin{pmatrix} 5 & 4 \\ -2 & -1 \end{pmatrix}$.
- (b) $\mathbf{B} = \begin{pmatrix} 2 & 0 & 1 \\ -1 & 3 & 1 \\ 2 & -2 & 1 \end{pmatrix}$.

Hence write down a basis of $\mathbb{R}^2$ consisting of eigenvectors of $\mathbf{A}$, and a basis of $\mathbb{R}^3$ consisting of eigenvectors of $\mathbf{B}$.

Solution.

- (a) $\operatorname{tr} \mathbf{A} = 4$ and $\det \mathbf{A} = -5 + 8 = 3$, so $p_{\mathbf{A}}(x) = x^2 - 4x + 3 = (x - 1)(x - 3)$, and $\operatorname{spec}(\mathbf{A}) = \{1, 3\}$ by Theorem 9.2.1. For $\lambda = 1$, $\mathbf{A} - \mathbf{I} = \begin{pmatrix} 4 & 4 \\ -2 & -2 \end{pmatrix}$ reduces to $x + y = 0$, so $E_1(\mathbf{A}) = \operatorname{span}((1, -1))$. For $\lambda = 3$, $\mathbf{A} - 3\mathbf{I} = \begin{pmatrix} 2 & 4 \\ -2 & -4 \end{pmatrix}$ reduces to $x + 2y = 0$, so $E_3(\mathbf{A}) = \operatorname{span}((2, -1))$. Check: $\mathbf{A}(1, -1) = (1, -1)$ and $\mathbf{A}(2, -1) = (6, -3) = 3(2, -1)$.
- (b) Expanding $\det(x\mathbf{I} - \mathbf{B})$ along the first row,

$$\begin{aligned} p_{\mathbf{B}}(x) &= \det \begin{pmatrix} x-2 & 0 & -1 \\ 1 & x-3 & -1 \\ -2 & 2 & x-1 \end{pmatrix} \\ &= (x-2)[(x-3)(x-1) + 2] + (-1)[1 \cdot 2 - (x-3)(-2)]. \end{aligned}$$

The first bracket is $x^2 - 4x + 5$ and the second is $2x - 4 = 2(x - 2)$, so

$$\begin{aligned} p_{\mathbf{B}}(x) &= (x-2)(x^2 - 4x + 5) - 2(x-2) \\ &= (x-2)(x^2 - 4x + 3) = (x-1)(x-2)(x-3). \end{aligned}$$

So $\operatorname{spec}(\mathbf{B}) = \{1, 2, 3\}$. Solving $(\mathbf{B} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$:

- $\lambda = 1$: $\mathbf{B} - \mathbf{I} = \begin{pmatrix} 1 & 0 & 1 \\ -1 & 2 & 1 \\ 2 & -2 & 0 \end{pmatrix}$. The first row gives $z = -x$, the third gives $y = x$, and the second is then $-x + 2x - x = 0$. So $E_1(\mathbf{B}) = \operatorname{span}((1, 1, -1))$.
- $\lambda = 2$: $\mathbf{B} - 2\mathbf{I} = \begin{pmatrix} 0 & 0 & 1 \\ -1 & 1 & 1 \\ 2 & -2 & -1 \end{pmatrix}$. The first row gives $z = 0$, and then the second gives $y = x$. So $E_2(\mathbf{B}) = \operatorname{span}((1, 1, 0))$.
- $\lambda = 3$: $\mathbf{B} - 3\mathbf{I} = \begin{pmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ 2 & -2 & -2 \end{pmatrix}$. The first row gives $z = x$, and the third gives $2x - 2y - 2x = 0$, so $y = 0$. So $E_3(\mathbf{B}) = \operatorname{span}((1, 0, 1))$.

Check: $\mathbf{B}(1, 1, -1) = (1, 1, -1)$, $\mathbf{B}(1, 1, 0) = (2, 2, 0)$, $\mathbf{B}(1, 0, 1) = (3, 0, 3)$.

Hence: the eigenvalues of $\mathbf{A}$ are distinct, so by Theorem 9.2.2 the eigenvectors $((1, -1), (2, -1))$ are independent, and two independent vectors in $\mathbb{R}^2$ form a basis (Theorem 1.6.4). In the same way $((1, 1, -1), (1, 1, 0), (1, 0, 1))$ is a basis of $\mathbb{R}^3$ consisting of eigenvectors of $\mathbf{B}$.

Exercise. B2: The transpose map

Let F be a field of characteristic not 2, and let $\tau: M_2(F) \rightarrow M_2(F)$, $\tau(\mathbf{X}) = \mathbf{X}^\top$.

- (a) Show that every eigenvalue λ of τ satisfies $\lambda^2 = 1$.

- (b) Find $\text{spec}(\tau)$, and describe $E_1(\tau)$ and $E_{-1}(\tau)$ with a basis of each.

Solution.

- (a) Let $\tau(\mathbf{X}) = \lambda\mathbf{X}$ with $\mathbf{X} \neq \mathbf{O}$. Since $(\mathbf{X}^\top)^\top = \mathbf{X}$ (Theorem 0.9.5), $\mathbf{X} = \tau(\tau(\mathbf{X})) = \tau(\lambda\mathbf{X}) = \lambda\tau(\mathbf{X}) = \lambda^2\mathbf{X}$. So $(\lambda^2 - 1)\mathbf{X} = \mathbf{O}$ with $\mathbf{X} \neq \mathbf{O}$, and $\lambda^2 = 1$ by Theorem 1.1.9.
- (b) From $\lambda^2 - 1 = (\lambda - 1)(\lambda + 1) = 0$, the only candidates are 1 and -1 , which are different because the characteristic is not 2. $E_1(\tau) = \{\mathbf{X} : \mathbf{X}^\top = \mathbf{X}\}$ is the space of symmetric matrices. Writing $\mathbf{X} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, the condition is $b = c$, so $E_1(\tau) = \text{span}(\mathbf{E}_{11}, \mathbf{E}_{22}, \mathbf{E}_{12} + \mathbf{E}_{21})$, and these three matrices are independent (compare entries). $E_{-1}(\tau) = \{\mathbf{X} : \mathbf{X}^\top = -\mathbf{X}\}$ is the space of skew-symmetric matrices. The condition $\mathbf{X}^\top = -\mathbf{X}$ says $a = -a$, $d = -d$, and $c = -b$; since $2 \neq 0$, $a = d = 0$, so $E_{-1}(\tau) = \text{span}(\mathbf{E}_{12} - \mathbf{E}_{21})$. Both eigenspaces are nonzero, so $\text{spec}(\tau) = \{1, -1\}$. This agrees with Theorem 4.5.5: $\tau^2 = \text{id}$, and $M_2(F) = E_1(\tau) \oplus E_{-1}(\tau)$, with dimensions $3 + 1 = 4$.

Exercise. B3: Eigenvalues of powers and inverses

Let $T \in \mathcal{L}(V)$ and let v be an eigenvector of T for λ .

- (a) Prove that v is an eigenvector of T^2 for λ^2 , and more generally of $p(T)$ for $p(\lambda)$, for every $p \in F[x]$.
- (b) Suppose T is invertible. Prove that $\lambda \neq 0$ and that v is an eigenvector of T^{-1} for λ^{-1} .
- (c) Show that the converse of the first statement in (a) fails: give $T \in \mathcal{L}(\mathbb{R}^2)$ such that -1 is an eigenvalue of T^2 but T has no eigenvalue.

Solution.

- (a) $T^2v = T(Tv) = T(\lambda v) = \lambda Tv = \lambda^2v$, and $v \neq 0$. By induction on k , $T^k v = \lambda^k v$: if $T^k v = \lambda^k v$, then $T^{k+1}v = T(\lambda^k v) = \lambda^{k+1}v$. For $p = \sum_k a_k x^k$, Definition 2.4.2 gives $p(T)v = \sum_k a_k T^k v = \sum_k a_k \lambda^k v = p(\lambda)v$.
- (b) An invertible operator is injective, so $\ker T = \{0\}$ and $Tv \neq 0$; as $Tv = \lambda v$, $\lambda \neq 0$ by Theorem 1.1.9. Applying T^{-1} to $Tv = \lambda v$ gives $v = \lambda T^{-1}v$, so $T^{-1}v = \lambda^{-1}v$.
- (c) Take the rotation $R(x, y) = (-y, x)$. It has no real eigenvalue (Example (A Rotation Has No Invariant Lines)). But $R^2(x, y) = R(-y, x) = (-x, -y)$, so $R^2 = -\text{id}$, and every nonzero vector is an eigenvector of R^2 for -1 . Part (a) would need a real λ with $\lambda^2 = -1$, and there is none.

9.2.7.3 C. Going deeper

Exercise. C1: Every vector an eigenvector

Let V be a vector space over F and $T \in \mathcal{L}(V)$. Suppose every nonzero vector of V is an eigenvector of T . Prove that $T = c \text{id}_V$ for some $c \in F$.

Hint: compare the eigenvalues of v , w , and $v + w$.

Solution. If $V = \{0\}$, then $T = 0 = 1 \cdot \text{id}_V$. Otherwise, for each $v \neq 0$ let λ_v be its eigenvalue, which is unique by the well-definedness remark after [Definition 9.2.1](#). We show that $\lambda_v = \lambda_w$ for all nonzero v, w .

Case 1: (v, w) is linearly dependent. Then $w = av$ for some $a \neq 0$, and $Tw = aTv = \lambda_v(av) = \lambda_v w$. By uniqueness of the eigenvalue of w , $\lambda_w = \lambda_v$.

Case 2: (v, w) is linearly independent. Then $v + w \neq 0$, and

$$\lambda_{v+w}v + \lambda_{v+w}w = T(v+w) = Tv + Tw = \lambda_v v + \lambda_w w.$$

So $(\lambda_{v+w} - \lambda_v)v + (\lambda_{v+w} - \lambda_w)w = 0$, and by independence $\lambda_v = \lambda_{v+w} = \lambda_w$.

Fix $v_0 \neq 0$ and put $c = \lambda_{v_0}$. Then $Tv = cv$ for every $v \neq 0$, and also for $v = 0$. Hence $T = c \text{id}_V$.

Exercise. C2: Nilpotent operators

An operator $T \in \mathcal{L}(V)$ is **nilpotent** if $T^k = 0$ for some integer $k \geq 1$.

- (a) Let $V \neq \{0\}$ and T nilpotent. Prove that $\text{spec}(T) = \{0\}$. (No finite dimension is needed.)
- (b) Let $\mathbf{A} \in M_n(\mathbb{C})$, $n \geq 1$, be nilpotent. Deduce that $p_{\mathbf{A}}(x) = x^n$.
- (c) Let $\mathbf{A} \in M_n(F)$, $n \geq 1$, be nilpotent. Prove that $\mathbf{I} + \mathbf{A}$ is invertible.

Solution.

- (a) Let $Tv = \lambda v$ with $v \neq 0$. By [Exercise \(B3: Eigenvalues of powers and inverses\)](#) (a), $0 = T^k v = \lambda^k v$, so $\lambda^k = 0$ by Theorem 1.1.9, and hence $\lambda = 0$ (a product of nonzero elements of a field is nonzero, Theorem 0.7.1). So $\text{spec}(T) \subseteq \{0\}$. Conversely, $T^k = 0$ is not injective on $V \neq \{0\}$; if T were injective, so would be the composite T^k (Theorem 0.4.3). So $\ker T \neq \{0\}$, and $0 \in \text{spec}(T)$.
- (b) By Corollary 6.5.5 and Theorem 7.10.2, $p_{\mathbf{A}} = (x - z_1) \cdots (x - z_n)$ for some $z_i \in \mathbb{C}$. Each z_i is a root of $p_{\mathbf{A}}$, hence an eigenvalue of $\mathbf{A}$ by [Theorem 9.2.1](#), hence $z_i = 0$ by (a). So $p_{\mathbf{A}} = x^n$.
- (c) By [Theorem 9.2.6](#), it suffices to show $0 \notin \text{spec}(\mathbf{I} + \mathbf{A})$. Suppose $(\mathbf{I} + \mathbf{A})\mathbf{v} = \mathbf{0}$ with $\mathbf{v} \neq \mathbf{0}$. Then $\mathbf{A}\mathbf{v} = -\mathbf{v}$, so $-1 \in \text{spec}(\mathbf{A})$. By (a), $\text{spec}(\mathbf{A}) = \{0\}$, but $-1 \neq 0$ in every field. This contradiction shows that $\mathbf{I} + \mathbf{A}$ is invertible.

Exercise. C3: Real matrices of odd size

- (a) Prove that every $\mathbf{A} \in M_n(\mathbb{R})$ with n odd has a real eigenvalue.
- (b) Deduce that every operator on a real vector space of dimension 3 has an invariant line.
- (c) Show that (a) fails for $n = 4$: find $\mathbf{A} \in M_4(\mathbb{R})$ with no real eigenvalue.

Solution.

- (a) By Theorem 7.10.2, $p_{\mathbf{A}} \in \mathbb{R}[x]$ is monic of degree n , which is odd. By Corollary 6.5.9, $p_{\mathbf{A}}(\lambda) = 0$ for some $\lambda \in \mathbb{R}$, and λ is an eigenvalue of $\mathbf{A}$ by Theorem 9.2.1.
- (b) Let $\dim V = 3$ over $\mathbb{R}$ and $T \in \mathcal{L}(V)$. Choose a basis $\mathcal{B}$. By (a), the real matrix $[T]_{\mathcal{B}}^{\mathcal{B}}$ has a real eigenvalue λ , which is a root of $p_{[T]_{\mathcal{B}}^{\mathcal{B}}} = p_T$ (Theorem 9.2.1, Definition 7.10.2). So λ is an eigenvalue of T , again by Theorem 9.2.1, and T has an eigenvector v for it. By Proposition 9.1.2, $\text{span}(v)$ is invariant.
- (c) Let $\mathbf{R} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ and $\mathbf{A} = \mathbf{R} \oplus \mathbf{R} \in M_4(\mathbb{R})$. Then $x\mathbf{I}_4 - \mathbf{A} = (x\mathbf{I}_2 - \mathbf{R}) \oplus (x\mathbf{I}_2 - \mathbf{R})$ is block diagonal, so by the block triangular determinant over $\mathbb{R}[x]$ (the remark after Theorem 7.8.4), $p_{\mathbf{A}} = p_{\mathbf{R}}^2 = (x^2 + 1)^2$. For real λ , $(\lambda^2 + 1)^2 \geq 1 > 0$, so $p_{\mathbf{A}}$ has no real root, and $\mathbf{A}$ has no real eigenvalue. The argument of (a) needs odd degree.