

4.5 Projections and the Trace of an Operator

Chapter 1 ended its discussion of complements with a promise: a splitting $V = U \oplus W$ should give a projection of V onto U along W , and different complements should give different projections. This section keeps that promise. Projections are the operators that are simplest in a well-chosen basis, and computing their matrices there leads to a second topic: the trace of an operator, a number attached to $T \in \mathcal{L}(V)$ that no choice of basis can change. We end with involutions, which split a space into two pieces in the same way.

4.5.1 Projections

Suppose $V = U \oplus W$. By Definition 1.7.2, every $v \in V$ can be written as $v = u + w$ with $u \in U$ and $w \in W$ in **exactly one** way. So the rule “ $v \mapsto u$ ”, keep the U -piece and throw away the W -piece, is a well-defined function $V \rightarrow V$. In $\mathbb{R}^3 = \{\text{the } xy\text{-plane}\} \oplus \{\text{the } z\text{-axis}\}$, it sends (a, b, c) to $(a, b, 0)$: the shadow of a point on the floor when light falls straight down.

What property does such a map have that singles it out? Apply it twice. The output u already lies in U , and its splitting is $u = u + 0$, so applying the rule again returns u . Doing it twice is the same as doing it once. We take this as the definition, because it can be checked without knowing U and W in advance.

A projection is an operator that leaves alone everything it has already produced.

Definition 4.5.1. Projection

Let V be a vector space over F . An operator $P \in \mathcal{L}(V)$ is a **projection** if

$$P^2 = P,$$

that is, $P(Pv) = Pv$ **for every** $v \in V$.

In words: P^2 is the composite PP (Definition 2.4.2), so the condition says that P fixes each of its own outputs. Equivalently, $Px = x$ **for every** $x \in \text{im } P$: if $x = Pv$, then $Px = P^2(v) = Pv = x$, and conversely this for $x = Pv$ is $P^2(v) = Pv$. No inner product and no notion of angle is involved. The projections of this section are often called **oblique** projections, to distinguish them from the orthogonal projections met later, in the chapters on inner product spaces.

Examples.

- **Onto an axis.** $P(x, y) = (x, 0)$ on F^2 . Then $P^2(x, y) = P(x, 0) = (x, 0)$, so $P^2 = P$.
- **A rank-one projection.** $T(x, y) = (2x - y, 2x - y)$ on $\mathbb{R}^2$ from Example (Making a matrix diagonal with a well-chosen basis). We saw there that its matrix $\mathbf{A}$ satisfies $\mathbf{A}^2 = \mathbf{A}$; directly, $T(2x - y, 2x - y) = (4x - 2y - 2x + y, \dots) = (2x - y, 2x - y)$.
- **Constant term.** On $F[x]_{\leq n}$, $P(p) = p(0)$, the constant polynomial with value $p(0)$. The constant polynomial c has value c at 0, so $P(P(p)) = P(p)$.
- **Degenerate cases.** The identity id_V and the zero operator are projections, since $\text{id}_V^2 = \text{id}_V$ and $0^2 = 0$. They correspond to keeping everything and keeping nothing.

Nonexample by minimal change. Change the first example to $T(x, y) = (2x, 0)$ on $\mathbb{R}^2$. Its image is still the x -axis and its kernel is still the y -axis. But $T^2(x, y) = (4x, 0) \neq T(x, y)$ when $x \neq 0$: T does not fix the vectors of its own image, since $T(1, 0) = (2, 0)$. So having “the right

image and kernel” is not enough; the clause $P^2 = P$ is what forces P to be the identity on its image. Likewise the reflection $R(x, y) = (y, x)$ is not a projection: $R^2 = \text{id} \neq R$.

Here is the promised correspondence. It says that projections and direct sum decompositions are two descriptions of the same thing.

Theorem 4.5.1. Projections and Direct Sums

Let V be a vector space.

- (a) If $P \in \mathcal{L}(V)$ is a projection, then

$$V = \text{im } P \oplus \ker P,$$

P is the identity on $\text{im } P$, and P is zero on $\ker P$.

- (b) Conversely, suppose $V = U \oplus W$. Then there is **exactly one** operator $P \in \mathcal{L}(V)$ with $Pu = u$ for all $u \in U$ and $Pw = 0$ for all $w \in W$, namely $P(u + w) = u$. It is a projection, with $\text{im } P = U$ and $\ker P = W$. It is called the **projection onto U along W** .

Idea. For (a), every vector splits as “what P keeps” plus “what P removes”: $v = Pv + (v - Pv)$. The first piece is in the image by definition, and $P^2 = P$ is exactly what puts the second piece in the kernel. For (b), uniqueness of the splitting makes the rule a function, and uniqueness again makes it linear.

Proof.

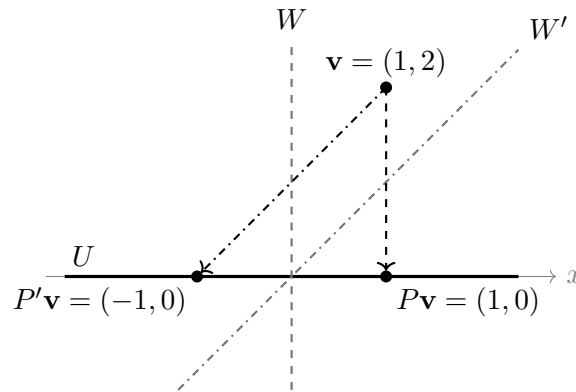
- (a) Let $v \in V$. Then $v = Pv + (v - Pv)$, where $Pv \in \text{im } P$ and $P(v - Pv) = Pv - P^2(v) = 0$ because $P^2 = P$. Hence $V = \text{im } P + \ker P$. If $x \in \text{im } P \cap \ker P$, then $x = Px$, as shown after Definition 4.5.1, and $Px = 0$; so $x = 0$. By Theorem 1.7.4, $V = \text{im } P \oplus \ker P$. We have also seen that P fixes $\text{im } P$, and P is zero on $\ker P$ by definition.
- (b) *Existence.* By Definition 1.7.2, each $v \in V$ has exactly one splitting $v = u + w$, so $Pv := u$ defines a function $V \rightarrow V$. It is linear: if $v = u + w$ and $v' = u' + w'$, then $v + v' = (u + u') + (w + w')$ and $av = au + aw$ are splittings, because U and W are subspaces, so by uniqueness $P(v + v') = u + u' = Pv + Pv'$ and $P(av) = au = aPv$. For $u \in U$ the splitting is $u = u + 0$, so $Pu = u$; for $w \in W$ it is $w = 0 + w$, so $Pw = 0$.

Properties. Every value Pv lies in U , and every $u \in U$ is Pu ; so $\text{im } P = U$. Then $P(Pv) = Pv$ because P fixes U ; so $P^2 = P$. Finally $P(u + w) = u$ is 0 exactly when $u = 0$, that is, when $v = w \in W$; so $\ker P = W$.

Uniqueness. If $P' \in \mathcal{L}(V)$ is identity on U and zero on W , then $P'(u + w) = P'u + P'w = u$ by linearity, so $P' = P$. ■

So a projection **is** a direct sum decomposition $V = U \oplus W$, recorded as an operator: U is its image and W its kernel. Given P , the decomposition is $\text{im } P \oplus \ker P$; given the decomposition, P is the projection onto U along W . The two constructions undo each other, by the “exactly one” in (b).

Now the promise from Chapter 1. There, the x -axis U in $\mathbb{R}^2$ had many complements, for instance the y -axis W and the line W' spanned by $(1, 1)$. Each gives a projection onto the same line U .



To project $\mathbf{v}$ onto U along a complement, slide $\mathbf{v}$ parallel to that complement until it hits U . Along the y -axis, $(1, 2) = (1, 0) + (0, 2)$ gives $P\mathbf{v} = (1, 0)$ (dashed arrow). Along W' , $(1, 2) = (-1, 0) + (2, 2)$ gives $P'\mathbf{v} = (-1, 0)$ (dash-dotted arrow). In general $P(x, y) = (x, 0)$ and $P'(x, y) = (x - y, 0)$, with standard matrices $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ and $\begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}$.

Warning. “**Projection onto U** ” is incomplete: the direction “along W ” matters. The two operators above are both projections, both have image U , and they are different: $P(1, 2) = (1, 0)$ but $P'(1, 2) = (-1, 0)$. Since U has infinitely many complements in $\mathbb{R}^2$, it is the image of infinitely many projections. A projection is determined by its image **and** its kernel together.

In finite dimension, the direct sum $V = \text{im } P \oplus \ker P$ hands us a basis in which P is as simple as possible.

Corollary 4.5.2. The matrix of a projection

Let V be finite-dimensional and $P \in \mathcal{L}(V)$ a projection with $\text{rank } P = r$. Let $(u_1, \dots, u_r)$ be a basis of $\text{im } P$ and $(w_1, \dots, w_s)$ a basis of $\ker P$. Then $\mathcal{B} = (u_1, \dots, u_r, w_1, \dots, w_s)$ is a basis of V , and

$$[P]_{\mathcal{B}}^{\mathcal{B}} = \begin{pmatrix} \mathbf{I}_r & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{pmatrix} = \text{diag}(1, \dots, 1, 0, \dots, 0),$$

with r ones.

Proof. By Theorem 4.5.1 (a), $V = \text{im } P \oplus \ker P$, so the concatenated list $\mathcal{B}$ is a basis of V by Theorem 1.7.5 ((a) $\Rightarrow$ (d)). Since $Pu_i = u_i$ and $Pw_j = 0$, the columns of $[P]_{\mathcal{B}}^{\mathcal{B}}$ are $\mathbf{e}_1, \dots, \mathbf{e}_r$ followed by s zero columns (Definition 4.1.1). ■

Conversely, an operator whose matrix in some basis is $\text{diag}(1, \dots, 1, 0, \dots, 0)$ is a projection, since that matrix squares to itself. With the change-of-basis square, this gives a way to write down any projection in standard coordinates.

Example. Projecting onto a plane along a line

In $\mathbb{R}^3$, let $U = \{(x, y, z) : x + y + z = 0\}$ and $L = \text{span}((1, 1, 1))$, so that $\mathbb{R}^3 = U \oplus L$ by Example (A plane and a line in $\mathbb{R}^3$). Find the standard matrix of the projection P onto U along L , and check that it squares to itself.

Solution. Direct route. By Example (A plane and a line in $\mathbb{R}^3$), (a, b, c) splits as $(a - m, b - m, c - m) + m(1, 1, 1)$ with $m = \frac{a+b+c}{3}$, and the first piece lies in U . So

$$P(a, b, c) = \left(\frac{2a-b-c}{3}, \frac{-a+2b-c}{3}, \frac{-a-b+2c}{3} \right), \quad [P]_{\mathcal{E}}^{\mathcal{E}} = \frac{1}{3} \begin{pmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{pmatrix}.$$

Adapted basis. $(1, -1, 0)$ and $(0, -1, 1)$ lie in U and are independent (look at the first and third entries), so they form a basis of the 2-dimensional plane U . With $\mathcal{B} = ((1, -1, 0), (0, -1, 1), (1, 1, 1))$, Corollary 4.5.2 gives $[P]_{\mathcal{B}}^{\mathcal{B}} = \text{diag}(1, 1, 0)$, and by Theorem 4.2.3 $[P]_{\mathcal{E}}^{\mathcal{E}} = M \text{diag}(1, 1, 0) M^{-1}$ with $M = [\text{id}]_{\mathcal{B}}^{\mathcal{E}}$, whose columns are the vectors of $\mathcal{B}$. Multiplying out gives the same matrix as the direct route.

Check. Let $\mathbf{A}$ be the matrix found. The first column of $\mathbf{A}^2$ is $\mathbf{A}$ applied to $\frac{1}{3}(2, -1, -1)$, which is $\frac{1}{9}(4 + 1 + 1, -2 - 2 + 1, -2 + 1 - 2) = \frac{1}{3}(2, -1, -1)$, the first column of $\mathbf{A}$; the other columns work the same way by symmetry of the entries. Also $\mathbf{A}(1, 1, 1) = \mathbf{0}$ and $\mathbf{A}(1, -1, 0) = (1, -1, 0)$, as a projection onto U along L must satisfy. The division by 3 needs $3 \neq 0$ in the field, which holds over $\mathbb{R}$.

Quick check. Is $P = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$ (acting on $\mathbb{R}^2$) a projection? If so, onto which line, and along which line?

Solution. $P^2 = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} = P$, so yes. Its image is spanned by its columns, $\text{span}((1, 0))$, the x -axis. Its kernel is $\{x + y = 0\} = \text{span}((1, -1))$. By Theorem 4.5.1, it is the projection onto the x -axis along the line $y = -x$.

4.5.2 The trace of an operator

For a projection of rank r , the diagonal matrix of Corollary 4.5.2 has trace $1 + \cdots + 1 = r \cdot 1$. But traces are defined for matrices, and the same projection has many other matrices, such as the 3×3 matrix of Example (Projecting onto a plane along a line), whose trace is $\frac{2+2+2}{3} = 2$. The agreement is no accident. In Theorem 4.2.5 we proved that similar matrices have the same trace, and by Theorem 4.2.4 all matrices of one operator are similar. So the trace belongs to the operator.

The trace of an operator is the trace of its matrix in any basis; the basis does not matter.

Definition 4.5.2. Trace of an operator

Let V be a finite-dimensional vector space over F with $\dim V \geq 1$, and let $T \in \mathcal{L}(V)$. The **trace** of T is

$$\operatorname{tr} T := \operatorname{tr}[T]_{\mathcal{B}}^{\mathcal{B}},$$

where $\mathcal{B}$ is **any** basis of V . (For $V = \{0\}$ we set $\operatorname{tr} T = 0$.)

In words: to find $\operatorname{tr} T$, choose a basis, write down the matrix of T with that **same** basis on input and output, and add the diagonal entries.

Well-definedness. The definition names a basis but claims the answer does not depend on it. Let $\mathcal{B}$ and $\mathcal{C}$ be bases of V . By Theorem 4.2.3, $[T]_{\mathcal{C}}^{\mathcal{C}} = P^{-1}[T]_{\mathcal{B}}^{\mathcal{B}}P$ with $P = [\operatorname{id}]_{\mathcal{C}}^{\mathcal{B}}$ invertible, so the two matrices are similar, and by Theorem 4.2.5 they have the same trace. Hence $\operatorname{tr} T$ depends only on T .

Examples.

- **Identity.** In every basis, $[\operatorname{id}_V]_{\mathcal{B}}^{\mathcal{B}} = \mathbf{I}_n$, so $\operatorname{tr} \operatorname{id}_V = n \cdot 1$, where $n = \dim V$. Over $\mathbb{R}$ this is n ; over $\mathbb{F}_2$ it is 0 when n is even.
- **Differentiation.** On $F[x]_{\leq n}$ with the basis $(1, x, \dots, x^n)$, $D(x^k) = kx^{k-1}$ has no x^k -component, so every diagonal entry of the matrix is 0 and $\operatorname{tr} D = 0$.
- **A shift of the variable.** On $\mathbb{R}[x]_{\leq 2}$, let $T(p) = p(x+1)$. In $(1, x, x^2)$, $T(1) = 1$, $T(x) = 1 + x$, $T(x^2) = 1 + 2x + x^2$, so the matrix is $\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$ and $\operatorname{tr} T = 3$.
- **The reflection.** $R(x, y) = (y, x)$ has trace 0, whether computed from $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ or from $\operatorname{diag}(1, -1)$ (Example (The reflection in a mirror-adapted basis)).

Warning. **The trace needs the same basis on both sides.** For an operator $T \in \mathcal{L}(V)$, the matrix $[T]_{\mathcal{B}}^{\mathcal{C}}$ with $\mathcal{B} \neq \mathcal{C}$ is square, but its trace depends on the bases. For id on $\mathbb{R}^2$, with $\mathcal{E}$ the standard basis and $\mathcal{C} = ((2, 0), (0, 1))$, we get $[\operatorname{id}]_{\mathcal{C}}^{\mathcal{C}} = \operatorname{diag}(\frac{1}{2}, 1)$, of trace $\frac{3}{2}$, while $\operatorname{tr} \operatorname{id} = 2$. Different bases on input and output give equivalent matrices, not similar ones, and equivalence does not preserve the trace. For the same reason there is no trace of a map $V \rightarrow W$ between different spaces.

The rules for the trace of a matrix carry over at once.

Theorem 4.5.3. Properties of the Trace of an Operator

Let V be finite-dimensional with $n = \dim V$, let $S, T \in \mathcal{L}(V)$ and $c \in F$. Then

$$\begin{aligned} \operatorname{tr}(S + T) &= \operatorname{tr} S + \operatorname{tr} T, \\ \operatorname{tr}(cT) &= c \operatorname{tr} T, \\ \operatorname{tr}(ST) &= \operatorname{tr}(TS), \\ \operatorname{tr} \operatorname{id}_V &= n \cdot 1. \end{aligned}$$

Proof. For $n = 0$ all traces are 0. Otherwise fix a basis $\mathcal{B}$ of V . The map $T \mapsto [T]_{\mathcal{B}}^{\mathcal{B}}$ is linear by Theorem 4.1.4, so $[S + T]_{\mathcal{B}}^{\mathcal{B}} = [S]_{\mathcal{B}}^{\mathcal{B}} + [T]_{\mathcal{B}}^{\mathcal{B}}$ and $[cT]_{\mathcal{B}}^{\mathcal{B}} = c[T]_{\mathcal{B}}^{\mathcal{B}}$; and $[ST]_{\mathcal{B}}^{\mathcal{B}} = [S]_{\mathcal{B}}^{\mathcal{B}}[T]_{\mathcal{B}}^{\mathcal{B}}$ by Theorem 4.1.2, and likewise for TS . The first three identities now follow from Theorem 0.9.6, since $\text{tr } T$ may be computed in $\mathcal{B}$. The last was computed in the examples. ■

A first consequence: on a finite-dimensional space V of dimension $n \geq 1$ over a field in which $n \cdot 1 \neq 0$, such as $\mathbb{R}$, there are no operators with $ST - TS = \text{id}_V$, since the left side has trace 0 and the right side has trace $n \cdot 1$.

Now the trace of a projection. Over $\mathbb{R}$ it is simply the rank, but the precise statement is an equation in the field, and reading it as an equation of integers needs a hypothesis.

Theorem 4.5.4. Trace of a Projection

Let V be a finite-dimensional vector space over F and $P \in \mathcal{L}(V)$ a projection. Then

$$\text{tr } P = (\text{rank } P) \cdot 1 \quad \text{in } F.$$

If F has characteristic 0, or characteristic $p > \dim V$, then $\text{rank } P$ is determined by $\text{tr } P$: it is the unique integer r with $0 \leq r \leq \dim V$ and $r \cdot 1 = \text{tr } P$. In particular $\text{rank } P = \text{tr } P$ for projections on real or complex spaces.

Proof. Let $r = \text{rank } P$. If $V = \{0\}$, both sides are 0. Otherwise, by Corollary 4.5.2 there is a basis $\mathcal{B}$ with $[P]_{\mathcal{B}}^{\mathcal{B}} = \text{diag}(1, \dots, 1, 0, \dots, 0)$ with r ones, so $\text{tr } P = r \cdot 1$ by Definition 4.5.2.

For the second statement, let $n = \dim V$. Since $0 \leq r \leq n$ and $r \cdot 1 = \text{tr } P$, it remains to show uniqueness. Suppose $k \cdot 1 = l \cdot 1$ for integers $0 \leq l \leq k \leq n$. By Lemma 0.7.4, $(k - l) \cdot 1 + l \cdot 1 = k \cdot 1 = l \cdot 1$, and adding $-(l \cdot 1)$ to both sides gives $(k - l) \cdot 1 = 0$. If $k - l > 0$, then by Definition 0.7.2 the characteristic of F is nonzero and at most $k - l \leq n$, which contradicts the hypothesis. Hence $k = l$. In $\mathbb{R}$ and $\mathbb{C}$, $r \cdot 1$ is the number r . ■

The hypothesis cannot be dropped. Over $\mathbb{F}_2$, the identity of $\mathbb{F}_2^2$ is a projection of rank 2, and its trace is $1 + 1 = 0$, the same as the trace of the zero projection, of rank 0. The equation $\text{tr } P = (\text{rank } P) \cdot 1$ still holds there; what fails is recovering the integer $\text{rank } P$ from the field element $\text{tr } P$.

4.5.3 Involutions and reflections

A projection satisfies $P^2 = P$. The other simple polynomial relation is $T^2 = \text{id}_V$: doing T twice changes nothing. Such an operator is called an **involution**. Examples are the reflection $R(x, y) = (y, x)$, the transpose $\mathbf{A} \mapsto \mathbf{A}^T$ on $M_n(F)$, and $f(x) \mapsto f(-x)$ on functions $\mathbb{R} \rightarrow \mathbb{R}$. In Chapter 1 the last two produced direct sums: symmetric plus skew-symmetric matrices (Example (Symmetric and skew-symmetric matrices)), and even plus odd functions (Example (Even and odd functions)). The next theorem explains both at once. For an involution T , write

$$E_+ := \{v \in V : Tv = v\}, \quad E_- := \{v \in V : Tv = -v\}.$$

These are subspaces, since $E_+ = \ker(T - \text{id}_V)$ and $E_- = \ker(T + \text{id}_V)$ (Theorem 2.2.1).

Theorem 4.5.5. Involutions Split the Space

Let F be a field of characteristic not 2, V a vector space over F , and $T \in \mathcal{L}(V)$ with $T^2 = \text{id}_V$. Then

$$V = E_+ \oplus E_-,$$

and the projection onto E_+ along E_- is $\frac{1}{2}(\text{id}_V + T)$.

Idea. Copy the even/odd split: $v = \frac{1}{2}(v + Tv) + \frac{1}{2}(v - Tv)$. Applying T swaps v and Tv , so the first piece is fixed and the second changes sign. Both halves of the argument divide by 2, which is where the characteristic enters.

Proof. Since the characteristic is not 2, $2 = 1 + 1 \neq 0$ in F (Definition 0.7.2), so $\frac{1}{2}$ exists. Let $v \in V$, and put $v_+ = \frac{1}{2}(v + Tv)$ and $v_- = \frac{1}{2}(v - Tv)$. Then $v = v_+ + v_-$, and by linearity and $T^2 = \text{id}_V$,

$$Tv_+ = \frac{1}{2}(Tv + v) = v_+, \quad Tv_- = \frac{1}{2}(Tv - v) = -v_-.$$

So $v_+ \in E_+$, $v_- \in E_-$, and $V = E_+ + E_-$. If $v \in E_+ \cap E_-$, then $v = Tv = -v$, so $2v = 0$ and $v = \frac{1}{2}(2v) = 0$. By Theorem 1.7.4, $V = E_+ \oplus E_-$.

The operator $\frac{1}{2}(\text{id}_V + T)$ is linear by Theorem 2.4.1, sends v to v_+ , fixes E_+ (where $Tv = v$), and kills E_- (where $Tv = -v$). By the uniqueness in Theorem 4.5.1 (b), it is the projection onto E_+ along E_- . ■

For the transpose, E_+ is the symmetric and E_- the skew-symmetric matrices; for $f(x) \mapsto f(-x)$, they are the even and odd functions; for the reflection $R(x, y) = (y, x)$, they are the mirror line $y = x$ and the perpendicular line $y = -x$, and the adapted basis $((1, 1), (1, -1))$ is the one of Example (The reflection in a mirror-adapted basis). In finite dimension, a basis of E_+ followed by a basis of E_- gives $[T]_{\mathcal{B}}^{\mathcal{B}} = \text{diag}(1, \dots, 1, -1, \dots, -1)$, and so $\text{tr } T = (\dim E_+) \cdot 1 - (\dim E_-) \cdot 1$. In Chapter 9 the nonzero vectors of E_+ and E_- will be called eigenvectors with eigenvalues 1 and -1 .

Warning. In characteristic 2 an involution need not split the space. Over $\mathbb{F}_2$, let $T\mathbf{x} = \mathbf{J}\mathbf{x}$ with $\mathbf{J} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Then $\mathbf{J}^2 = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} = \mathbf{I}_2$, since $2 = 0$. But $-1 = 1$, so $E_- = E_+ = \ker(\mathbf{J} - \mathbf{I}_2) = \text{span}(\mathbf{e}_1)$, and $E_+ + E_- = \text{span}(\mathbf{e}_1) \neq \mathbb{F}_2^2$.

4.5.4 Exercises

4.5.4.1 A. Check your understanding

Exercise. A1

- (a) Define a projection $P \in \mathcal{L}(V)$, and state the correspondence between projections and direct sum decompositions.
- (b) True or false: a projection is determined by its image. Justify your answer.
- (c) True or false: an operator T on $\mathbb{R}^2$ with $\text{im } T$ the x -axis and $\ker T$ the y -axis is a projection. Justify your answer.
- (d) Explain why $\text{tr } T$ does not depend on the basis used to compute it.
- (e) True or false: for every field F and every projection P on a finite-dimensional space over F , the integer $\text{rank } P$ can be recovered from $\text{tr } P \in F$. Justify your answer.
- (f) Where does the hypothesis “characteristic not 2” enter the proof of [Theorem 4.5.5](#)?

Solution.

- (a) P is a projection if $P^2 = P$ ([Definition 4.5.1](#)). By [Theorem 4.5.1](#), a projection gives $V = \text{im } P \oplus \ker P$, and each decomposition $V = U \oplus W$ gives exactly one projection with image U and kernel W .
- (b) False. On $\mathbb{R}^2$, $P(x, y) = (x, 0)$ and $P'(x, y) = (x - y, 0)$ are different projections with the same image, the x -axis. A projection is determined by its image and its kernel together.
- (c) False. $T(x, y) = (2x, 0)$ has this image and kernel, but $T^2 \neq T$.
- (d) Matrices of T in two bases are similar by [Theorem 4.2.3](#), and similar matrices have equal traces by [Theorem 4.2.5](#).
- (e) False. Over $\mathbb{F}_2$, id on $\mathbb{F}_2^2$ has rank 2 and the zero operator has rank 0, but both have trace 0. It is true when the characteristic is 0 or larger than $\dim V$ ([Theorem 4.5.4](#)).
- (f) Twice: to define $v_{\pm} = \frac{1}{2}(v \pm Tv)$, and to pass from $2v = 0$ to $v = 0$. Both need $2 \neq 0$ in F .

4.5.4.2 B. Practice

Exercise. B1: A projection in $\mathbb{R}^3$

Let $U = \{(x, y, z) \in \mathbb{R}^3 : x + y - z = 0\}$ and $L = \text{span}((1, 1, 1))$.

- (a) Show that $\mathbb{R}^3 = U \oplus L$.
- (b) Find the standard matrix $\mathbf{A}$ of the projection P onto U along L .
- (c) Verify that $\mathbf{A}^2 = \mathbf{A}$, and that $\text{tr } \mathbf{A} = \text{rank } \mathbf{A}$.

Solution.

- (a) U is the kernel of $(x, y, z) \mapsto x + y - z$, which is nonzero; its image is then a nonzero subspace of $\mathbb{R}$, hence all of $\mathbb{R}$ ([Theorem 1.6.8](#)), so $\dim U = 3 - 1 = 2$ by [Theorem 2.3.1](#). If $t(1, 1, 1) \in U$, then $t + t - t = t = 0$, so $U \cap L = \{\mathbf{0}\}$. By [Theorem 1.7.4](#), $\dim(U + L) = 2 + 1 = 3$, so $U + L = \mathbb{R}^3$ by [Theorem 1.6.8](#), and $\mathbb{R}^3 = U \oplus L$.
- (b) Split $(a, b, c) = \mathbf{u} + t(1, 1, 1)$ with $\mathbf{u} = (a - t, b - t, c - t) \in U$: the condition $(a - t) +$

$(b - t) - (c - t) = 0$ gives $t = a + b - c$. So

$$P(a, b, c) = (-b + c, -a + c, -a - b + 2c), \quad \mathbf{A} = \begin{pmatrix} 0 & -1 & 1 \\ -1 & 0 & 1 \\ -1 & -1 & 2 \end{pmatrix}.$$

Check: $\mathbf{A}(1, 1, 1) = \mathbf{0}$, and for the vectors $(1, 0, 1), (0, 1, 1) \in U$, $\mathbf{A}(1, 0, 1) = (1, 0, 1)$ and $\mathbf{A}(0, 1, 1) = (0, 1, 1)$.

(c) Row by row,

$$\mathbf{A}^2 = \begin{pmatrix} 0+1-1 & 0+0-1 & 0-1+2 \\ 0+0-1 & 1+0-1 & -1+0+2 \\ 0+1-2 & 1+0-2 & -1-1+4 \end{pmatrix} = \begin{pmatrix} 0 & -1 & 1 \\ -1 & 0 & 1 \\ -1 & -1 & 2 \end{pmatrix} = \mathbf{A}.$$

$\text{tr } \mathbf{A} = 0 + 0 + 2 = 2$, and $\text{rank } \mathbf{A} = \text{rank } P = \dim U = 2$ by [Theorem 4.5.1](#) (b), in agreement with [Theorem 4.5.4](#).

Exercise. B2: The complementary projection

Let $P \in \mathcal{L}(V)$ be a projection.

- (a) Prove that $\text{id}_V - P$ is a projection.
- (b) Prove that $\text{im}(\text{id}_V - P) = \ker P$ and $\ker(\text{id}_V - P) = \text{im } P$.
- (c) Deduce that if P is the projection onto U along W , then $\text{id}_V - P$ is the projection onto W along U .

Solution.

- (a) By [Theorem 2.4.2](#) (composition distributes over sums of maps), $(\text{id}_V - P)^2 = \text{id}_V - 2P + P^2 = \text{id}_V - 2P + P = \text{id}_V - P$.
- (b) $(\subseteq)$ $P((\text{id}_V - P)(v)) = Pv - P^2(v) = 0$, so $\text{im}(\text{id}_V - P) \subseteq \ker P$. $(\supseteq)$ If $Pv = 0$, then $v = (\text{id}_V - P)(v) \in \text{im}(\text{id}_V - P)$. For the kernel: $(\text{id}_V - P)(v) = 0$ means $v = Pv$. Such v lies in $\text{im } P$, and conversely every $x \in \text{im } P$ satisfies $Px = x$, as shown after [Definition 4.5.1](#). So $\ker(\text{id}_V - P) = \text{im } P$.
- (c) By [Theorem 4.5.1](#) (b), $\text{im } P = U$ and $\ker P = W$. By (a) and (b), $\text{id}_V - P$ is a projection with image W and kernel U . By the uniqueness in [Theorem 4.5.1](#) (b), applied to $V = W \oplus U$, it is the projection onto W along U .

Exercise. B3: Traces of operators on matrices

- (a) Let $\tau: M_n(F) \rightarrow M_n(F)$, $\tau(\mathbf{A}) = \mathbf{A}^\top$. Show that $\text{tr } \tau = n \cdot 1$.
- (b) Fix $\mathbf{B} \in M_2(F)$, and let $L_{\mathbf{B}}: M_2(F) \rightarrow M_2(F)$, $L_{\mathbf{B}}(\mathbf{X}) = \mathbf{B}\mathbf{X}$. Show that $\text{tr } L_{\mathbf{B}} = 2 \text{tr } \mathbf{B}$.

Hint: use the basis of matrix units $\mathbf{E}_{ij}$.

Solution.

- (a) Order the basis $(\mathbf{E}_{ij})$ of $M_n(F)$ in any way. Since $\tau(\mathbf{E}_{ij}) = \mathbf{E}_{ji}$, the coordinate column of $\tau(\mathbf{E}_{ij})$ has a 1 in the position of $\mathbf{E}_{ji}$ and 0 elsewhere. The diagonal entry of the matrix in the column of $\mathbf{E}_{ij}$ is therefore 1 if $\mathbf{E}_{ji} = \mathbf{E}_{ij}$, that is $i = j$, and 0 otherwise. There are n basis vectors $\mathbf{E}_{ii}$, so $\text{tr } \tau = n \cdot 1$ by [Definition 4.5.2](#). (Consistently with [Theorem 4.5.5](#) over $\mathbb{R}$: symmetric matrices have dimension $\frac{n(n+1)}{2}$, skew-symmetric ones $\frac{n(n-1)}{2}$, and the difference is n .)
- (b) Use the basis $(\mathbf{E}_{11}, \mathbf{E}_{12}, \mathbf{E}_{21}, \mathbf{E}_{22})$ and write $\mathbf{B} = (b_{ij})$. The product $\mathbf{B}\mathbf{E}_{ij}$ has the i -th column of $\mathbf{B}$ in column j and zeros elsewhere, so $\mathbf{B}\mathbf{E}_{ij} = b_{1i}\mathbf{E}_{1j} + b_{2i}\mathbf{E}_{2j}$. The diagonal entry in the column of $\mathbf{E}_{ij}$ is the coefficient of $\mathbf{E}_{ij}$ itself, which is b_{ii} . Adding over the four basis vectors, $\text{tr } L_{\mathbf{B}} = b_{11} + b_{11} + b_{22} + b_{22} = 2 \text{tr } \mathbf{B}$.

4.5.4.3 C. Going deeper

Exercise. C1: When is a sum of projections a projection?

Let F have characteristic not 2, and let $P, Q \in \mathcal{L}(V)$ be projections.

- (a) Prove that $P + Q$ is a projection if and only if $PQ = QP = 0$.
- (b) Show that (a) fails over $\mathbb{F}_2$.
- (c) Suppose V is finite-dimensional over $\mathbb{R}$ and $P + Q$ is a projection. Deduce that $\text{rank}(P + Q) = \text{rank } P + \text{rank } Q$.

Hint: for $(\Rightarrow)$ in (a), use the relation $PQ + QP = 0$ together with $P^2 = P$.

Solution.

- (a) By Theorem 2.4.2, $(P + Q)^2 = P^2 + PQ + QP + Q^2 = P + Q + PQ + QP$.
 $(\Leftarrow)$ If $PQ = QP = 0$, this is $P + Q$.
 $(\Rightarrow)$ If $(P + Q)^2 = P + Q$, then $PQ + QP = 0$. Multiplying on the left by P and using $P^2 = P$ gives $PQ + PQP = 0$; multiplying on the right by P gives $PQP + QP = 0$. Subtracting, $PQ - QP = 0$, so $PQ = QP$. Then $2PQ = PQ + QP = 0$, and since $2 \neq 0$ in F , $PQ = 0$. Hence also $QP = 0$.
- (b) Over $\mathbb{F}_2$, let $P = Q = \text{id}_V$ with $V \neq \{0\}$. Then $P + Q = 2\text{id}_V = 0$, which is a projection, but $PQ = \text{id}_V \neq 0$. The step from $2PQ = 0$ to $PQ = 0$ is the one that fails.
- (c) By [Theorem 4.5.4](#) over $\mathbb{R}$, and [Theorem 4.5.3](#),

$$\text{rank}(P + Q) = \text{tr}(P + Q) = \text{tr } P + \text{tr } Q = \text{rank } P + \text{rank } Q.$$

Exercise. C2: The trace is the only trace-like functional

Let $n \geq 1$, and let $\varphi: M_n(F) \rightarrow F$ be linear with $\varphi(\mathbf{AB}) = \varphi(\mathbf{BA})$ for all $\mathbf{A}, \mathbf{B} \in M_n(F)$.

- (a) Using $\mathbf{E}_{ij}\mathbf{E}_{jk} = \mathbf{E}_{ik}$ and $\mathbf{E}_{jk}\mathbf{E}_{ij} = \mathbf{O}$ for $k \neq i$, show that $\varphi(\mathbf{E}_{ik}) = 0$ whenever $i \neq k$.
- (b) Show that $\varphi(\mathbf{E}_{ii}) = \varphi(\mathbf{E}_{jj})$ for all i, j .

(c) Deduce that $\varphi = c \operatorname{tr}$ for some $c \in F$, and conversely that every $c \operatorname{tr}$ has the property.

Solution. Recall that $\mathbf{E}_{ij}\mathbf{E}_{kl}$ is $\mathbf{E}_{il}$ if $j = k$ and $\mathbf{O}$ otherwise, by Definition 0.9.9.

- (a) Let $i \neq k$. Then $\mathbf{E}_{ik} = \mathbf{E}_{ii}\mathbf{E}_{ik}$ and $\mathbf{E}_{ik}\mathbf{E}_{ii} = \mathbf{O}$, because $k \neq i$. By the hypothesis and linearity, $\varphi(\mathbf{E}_{ik}) = \varphi(\mathbf{E}_{ii}\mathbf{E}_{ik}) = \varphi(\mathbf{E}_{ik}\mathbf{E}_{ii}) = \varphi(\mathbf{O}) = 0$.
- (b) $\mathbf{E}_{ii} = \mathbf{E}_{ij}\mathbf{E}_{ji}$ and $\mathbf{E}_{jj} = \mathbf{E}_{ji}\mathbf{E}_{ij}$. Hence $\varphi(\mathbf{E}_{ii}) = \varphi(\mathbf{E}_{ij}\mathbf{E}_{ji}) = \varphi(\mathbf{E}_{ji}\mathbf{E}_{ij}) = \varphi(\mathbf{E}_{jj})$.
- (c) Let $c = \varphi(\mathbf{E}_{11})$. For $\mathbf{A} = (a_{ij}) = \sum_{i,j} a_{ij}\mathbf{E}_{ij}$, linearity with (a) and (b) gives

$$\varphi(\mathbf{A}) = \sum_{i,j} a_{ij}\varphi(\mathbf{E}_{ij}) = \sum_i a_{ii}\varphi(\mathbf{E}_{ii}) = c \sum_i a_{ii} = c \operatorname{tr} \mathbf{A}.$$

Conversely, $c \operatorname{tr}$ is linear and $c \operatorname{tr}(\mathbf{AB}) = c \operatorname{tr}(\mathbf{BA})$ by Theorem 0.9.6. So the linear functionals with $\varphi(\mathbf{AB}) = \varphi(\mathbf{BA})$ are exactly the multiples of the trace.