Linear Algebra Done Right Solutions

Axler. S. (2024) *Linear Algebra Done Right.* 4th edn.

August 05, 2024

Contents

Chapter 1. Vector Spaces

1.A. $\mathbf{R}^n$ and $\mathbf{C}^n$

Exercise 1.A.1. Show that $\alpha + \beta = \beta + \alpha$ for all $\alpha, \beta \in \mathbf{C}$.

Solution. If $\alpha = x + yi$ and $\beta = u + vi$, then

$$\alpha + \beta = (x+u) + (y+v)i = (u+x) + (v+y)i = \beta + \alpha,$$

where we have used the commutativity of addition in $\mathbf{R}$.

Exercise 1.A.2. Show that $(\alpha + \beta) + \lambda = \alpha + (\beta + \lambda)$ for all $\alpha, \beta \in \mathbf{C}$.

Solution. If $\alpha = x + yi, \beta = u + vi$, and $\lambda = s + ti$, then

$$\begin{aligned}(\alpha + \beta) + \lambda &= ((x+u) + (y+v))i + \lambda = ((x+u)+s) + ((y+v)+t)i \\ &= (x + (u+s)) + (y + (v+t))i = \alpha + ((u+s) + (v+t)i) = \alpha + (\beta + \lambda),\end{aligned}$$

where we have used the associativity of addition in $\mathbf{R}$.

Exercise 1.A.3. Show that $(\alpha\beta)\lambda = \alpha(\beta\lambda)$ for all $\alpha, \beta, \lambda \in \mathbf{C}$.

Solution. If $\alpha = x + yi, \beta = u + vi$, and $\lambda = s + ti$, then

$$\begin{aligned}(\alpha\beta)\lambda &= [(xu - yv) + (xv + yu)i]\lambda \\ &= [(xu - yv)s - (xv + yu)t] + [(xu - yv)t + (xv + yu)s]i \\ &= [(xu)s - (yv)s - (xv)t - (yu)t] + [(xu)t - (yv)t + (xv)s + (yu)s]i \\ &= [x(us) - x(vt) - y(ut) - y(vs)] + [x(ut) + x(vs) + y(us) - y(vt)]i \\ &= [x(us - vt) - y(ut + vs)] + [x(ut + vs) + y(us - vt)]i \\ &= \alpha[(us - vt) + (ut + vs)i] \\ &= \alpha(\beta\lambda),\end{aligned}$$

where we have used several algebraic properties of $\mathbf{R}$.

Exercise 1.A.4. Show that $\lambda(\alpha + \beta) = \lambda\alpha + \lambda\beta$ for all $\lambda, \alpha, \beta \in \mathbf{C}$.

Solution. If $\alpha = x + yi, \beta = u + vi$, and $\lambda = s + ti$, then

$$
\begin{aligned}
\lambda(\alpha+\beta) &= [s(x+u)-t(y+v)]+[s(y+v)+t(x+u)i] \\
&= (sx+su-ty-tv)+(sy+sv+tx+tu)i \\
&= [(sx-ty)+(sy+tx)i]+[(su-tv)+(sv+tu)i] \\
&= \lambda\alpha+\lambda\beta,
\end{aligned}
$$

where we have used distributivity in $\mathbf{R}$.

Exercise 1.A.5. Show that for every $\alpha \in \mathbf{C}$, there exists a unique $\beta \in \mathbf{C}$ such that $\alpha+\beta=0$.

Solution. Suppose that $\alpha = x+yi$. Let $\beta=-x-yi$ and observe that

$$\alpha+\beta=(x-x)+(y-y)i=0+0i=0.$$

To see that β is unique, suppose that β' also satisfies $\alpha+\beta'=0$ and notice that

$$\beta=\beta=0=\beta+(\alpha+\beta')=(\alpha+\beta)+\beta'=0+\beta'=\beta',$$

where we have used the associativity of addition in $\mathbf{C}$ (Exercise 1.A.2) and the commutativity of addition in $\mathbf{C}$ (Exercise 1.A.1).

Exercise 1.A.6. Show that for every $\alpha \in \mathbf{C}$ with $\alpha \neq 0$, there exists a unique $\beta \in \mathbf{C}$ such that $\alpha\beta = 1$.

Solution. Suppose that $\alpha = x+yi$. Since $\alpha \neq 0$, it must be the case that x and y are not both zero, so that $x^2+y^2 \neq 0$. Let $\beta = \frac{x}{x^2+y^2} - \frac{y}{x^2+y^2}i$ and observe that

$$\alpha\beta=(x+yi)\left(\frac{x}{x^2+y^2}-\frac{y}{x^2+y^2}i\right)=\frac{x^2+y^2}{x^2+y^2}+\frac{xy-xy}{x^2+y^2}i=1+0i=1.$$

To see that β is unique, suppose β' also satisfies $\alpha\beta'=1$ and notice that

$$\beta=\beta 1=\beta(\alpha\beta')=(\alpha\beta)\beta'=1\beta'=\beta',$$

where we have used the associativity of multiplication in $\mathbf{C}$ (Exercise 1.A.3) and the commutativity of multiplication in $\mathbf{C}$ (1.4).

Exercise 1.A.7. Show that

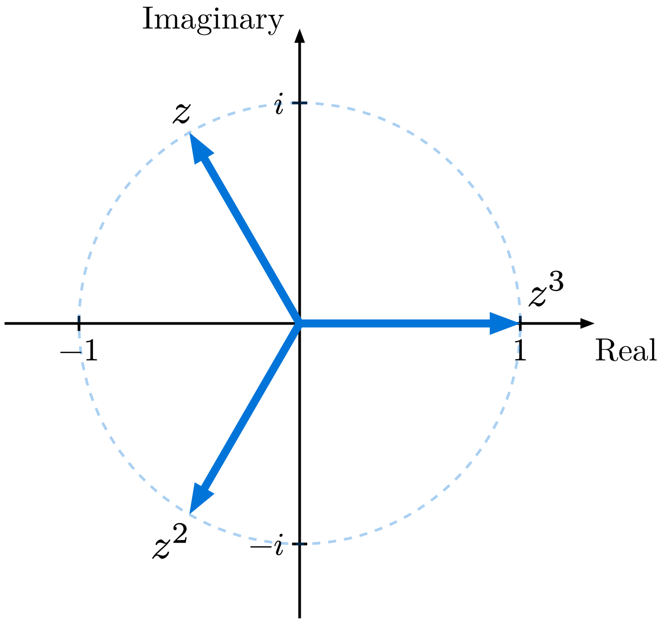
$$\frac{-1+\sqrt{3}i}{2}$$

is a cube root of 1 (meaning that its cube equals 1).

Solution. Let $z=\frac{-1+\sqrt{3}i}{2}$, so that $2z=-1+\sqrt{3}i$. Observe that

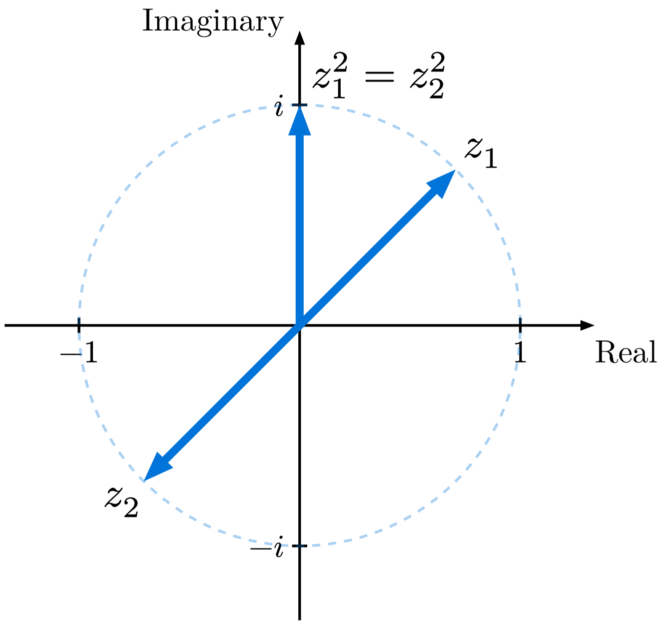
$$(2z)^2 = 4z^2 = \left(-1+\sqrt{3}i\right)^2 = 1 - 2\sqrt{3}i - 3 = -2 - 2\sqrt{3}i$$
$$\Rightarrow (2z)^3 = (4z^2)(2z) = \left(-2-2\sqrt{3}i\right)\left(-1+\sqrt{3}i\right) = 2 - 2\sqrt{3}i + 2\sqrt{3}i + 6 = 8,$$

i.e. $8z^3 = 8$. It follows that $z^3 = 1$.



Exercise 1.A.8. Find two distinct square roots of i.

Solution. Let $z_1 = \frac{1+i}{\sqrt{2}}$ and $z_2 = -z_1$ (z_1 and z_2 are distinct since $z_1 \neq 0$) and observe that
$$2z_1^2 = (1+i)^2 = 2i \quad \Rightarrow \quad z_1^2 = i,$$
i.e. z_1 is a square root of i. Furthermore, $z_2^2 = (-z_1)^2 = z_1^2 = i$, so that z_2 is a square root of i also.



Exercise 1.A.9. Find $x \in \mathbf{R}^4$ such that

$$(4,-3,1,7)+2x=(5,9,-6,8).$$

Solution. The unique solution is $x=\left(\frac{1}{2},6,-\frac{7}{2},\frac{1}{2}\right)$.

Exercise 1.A.10. Explain why there does not exist $\lambda \in \mathbf{C}$ such that

$$\lambda(2-3i,5+4i,-6+7i)=(12-5i,7+22i,-32-9i).$$

Solution. If there was such a λ, then

$$\lambda(2-3i)=12-5i \quad \Rightarrow \quad \lambda=\frac{12-5i}{2-3i}=3+2i.$$

However,

$$(3+2i)(-6+7i)=-32+9i \neq -32-9i.$$

Exercise 1.A.11. Show that $(x+y)+z=x+(y+z)$ for all $x,y,z \in \mathbf{F}^n$.

Solution. If $x=(x_1,...,x_n), y=(y_1,...,y_n)$, and $z=(z_1,...,z_n)$, then

$$\begin{aligned}(x+y)+z&=(x_1+y_1,...,x_n+y_n)+z=((x_1+y_1)+z_1,...,(x_n+y_n)+z_n)\\&=(x_1+(y_1+z_1),...,x_n+(y_n+z_n))=x+(y_1+z_1,...,y_n+z_n)=x+(y+z),\end{aligned}$$

where we have used the associativity of addition in $\mathbf{F}$ (we proved this for $\mathbf{C}$ in Exercise 1.A.2).

Exercise 1.A.12. Show that $(ab)x=a(bx)$ for all $x \in \mathbf{F}^n$ and all $a,b \in \mathbf{F}$.

Solution. If $x=(x_1,...,x_n)$, then

$$(ab)x=((ab)x_1,...,(ab)x_n)=(a(bx_1),...,a(bx_n))=a(bx_1,...,bx_n)=a(bx),$$

where we have used the associativity of multiplication in $\mathbf{F}$ (we proved this for $\mathbf{C}$ in Exercise 1.A.3).

Exercise 1.A.13. Show that $1x=x$ for all $x \in \mathbf{F}^n$.

Solution. If $x=(x_1,...,x_n)$, then

$$1x=(1x_1,...,1x_n)=(x_1,...,x_n)=x,$$

where we have used that $1x_j=x_j$ for any $x_j \in \mathbf{F}$.

Exercise 1.A.14. Show that $\lambda(x+y)=\lambda x+\lambda y$ for all $\lambda \in \mathbf{F}$ and all $x,y \in \mathbf{F}^n$.

Solution. If $x = (x_1, ..., x_n)$ and $y = (y_1, ..., y_n)$, then

$$\begin{aligned}
\lambda(x+y) &= \lambda(x_1 + y_1, ..., x_n + y_n) \\
&= (\lambda(x_1 + y_1), ..., \lambda(x_n + y_n)) \\
&= (\lambda x_1 + \lambda y_1, ..., \lambda x_n + \lambda y_n) \\
&= (\lambda x_1, ..., \lambda x_n) + (\lambda y_1, ..., \lambda y_n) \\
&= \lambda(x_1, ..., x_n) + \lambda(y_1, ..., y_n) \\
&= \lambda x + \lambda y,
\end{aligned}$$

where we have used distributivity in $\mathbf{F}$ (we proved this for $\mathbf{C}$ in Exercise 1.A.4).

Exercise 1.A.15. Show that $(a+b)x = ax + bx$ for all $a, b \in \mathbf{F}$ and all $x \in \mathbf{F}^n$.

Solution. If $x = (x_1, ..., x_n)$, then

$$\begin{aligned}
(a+b)x &= (a+b)(x_1, ..., x_n) \\
&= ((a+b)x_1, ..., (a+b)x_n) \\
&= (ax_1 + bx_1, ..., ax_n + bx_n) \\
&= (ax_1, ..., ax_n) + (bx_1, ..., bx_n) \\
&= a(x_1, ..., x_n) + b(x_1, ..., x_n) \\
&= ax + bx,
\end{aligned}$$

where we have used distributivity in $\mathbf{F}$ (we proved this for $\mathbf{C}$ in Exercise 1.A.4).

1.B. Definition of Vector Space

Exercise 1.B.1. Show that $-(-v) = v$ for every $v \in V$.

Solution. Since $v + (-v) = 0$ and the additive inverse of a vector is unique (1.27), it must be the case that $-(-v) = v$.

Exercise 1.B.2. Suppose $a \in \mathbf{F}, v \in V$, and $av = 0$. Prove that $a = 0$ or $v = 0$.

Solution. It will suffice to show that if $av = 0$ and $a \neq 0$, so that a^{-1} exists, then $v = 0$. Indeed,

$$av = 0 \quad \Rightarrow \quad a^{-1}(av) = 0 \quad \Rightarrow \quad (a^{-1}a)v = 0 \quad \Rightarrow \quad 1v = 0 \quad \Rightarrow \quad v = 0.$$

Exercise 1.B.3. Suppose $v, w \in V$. Explain why there exists a unique $x \in V$ such that $v + 3x = w$.

Solution. For $v, w, x \in V$, notice that

$$v + 3x = w \quad \Leftrightarrow \quad 3x = w - v \quad \Leftrightarrow \quad x = \tfrac{1}{3}(w - v).$$

Exercise 1.B.4. The empty set is not a vector space. The empty set fails to satisfy only one of the requirements listed in the definition of a vector space (1.20). Which one?

Solution. The empty set does not contain an additive identity.

Exercise 1.B.5. Show that in the definition of a vector space (1.20), the additive inverse condition can be replaced with the condition that

$$0v = 0 \text{ for all } v \in V.$$

Here the 0 on the left side is the number 0, and the 0 on the right side is the additive identity of V.

The phrase a "condition can be replaced" in a definition means that the collection of objects satisfying the definition is unchanged if the original condition is replaced with the new condition.

Solution. If V satisfies all of the conditions in 1.20, then as shown in 1.30 we have $0v = 0$ for all $v \in V$. Suppose that V satisfies all of the conditions in 1.20, except we have replaced the additive inverse condition with the condition that $0v = 0$ for all $v \in V$. We want to show that for each $v \in V$, there exists an element $w \in V$ such that $v + w = 0$. Indeed, for $v \in V$, let $w = (-1)v$ and observe that

$$v + w = 1v + (-1)v = (1-1)v = 0v = 0.$$

Exercise 1.B.6. Let ∞ and $-\infty$ denote two distinct objects, neither of which is in $\mathbf{R}$. Define an addition and scalar multiplication on $\mathbf{R} \cup \{\infty, -\infty\}$ as you could guess from the notation. Specifically, the sum and product of two real numbers is as usual, and for $t \in \mathbf{R}$ define

$$t\infty = \begin{cases} -\infty & \text{if } t < 0, \\ 0 & \text{if } t = 0, \\ \infty & \text{if } t > 0, \end{cases} \qquad t(-\infty) = \begin{cases} \infty & \text{if } t < 0, \\ 0 & \text{if } t = 0, \\ -\infty & \text{if } t > 0, \end{cases}$$

and

$$t + \infty = \infty + t = \infty + \infty = \infty,$$
$$t + (-\infty) = (-\infty) + t = (-\infty) + (-\infty) = -\infty,$$
$$\infty + (-\infty) = (-\infty) + \infty = 0.$$

With these operations of addition and scalar multiplication, is $\mathbf{R} \cup \{\infty, -\infty\}$ a vector space over $\mathbf{R}$? Explain.

Solution. This is not a vector space over $\mathbf{R}$, since addition is not associative:

$$(1 + \infty) + (-\infty) = \infty + (-\infty) = 0 \neq 1 = 1 + 0 = 1 + (\infty + (-\infty)).$$

Exercise 1.B.7. Suppose S is a nonempty set. Let V^S denote the set of functions from S to V. Define a natural addition and scalar multiplication on V^S, and show that V^S is a vector space with these definitions.

Solution. We define addition and scalar multiplication on V^S as in 1.24, i.e. for $f, g \in V^S$ the sum $f + g \in V^S$ is the function

$$\begin{aligned} f + g : S &\to V \\ x &\mapsto f(x) + g(x); \end{aligned}$$

the addition $f(x) + g(x)$ is vector addition in V. Similarly, for $\lambda \in \mathbf{F}$ and $f \in V^S$, the product $\lambda f \in V^S$ is the function

$$\begin{aligned} \lambda f : S &\to V \\ x &\mapsto \lambda f(x); \end{aligned}$$

the product $\lambda f(x)$ is scalar multiplication in V. We now show that V^S with these definitions satisfies each condition in definition 1.20.

Commutativity. Let $f, g \in V^S$ and $x \in S$ be given. Observe that

$$(f + g)(x) = f(x) + g(x) = g(x) + f(x) = (g + f)(x),$$

where we have used the commutativity of addition in V for the second equality. It follows that $f + g = g + f$.

Associativity. Let $f, g, h \in V^S$ and $x \in S$ be given. Observe that

$$\begin{aligned}((f+g)+h)(x) &= (f+g)(x) + h(x) = (f(x)+g(x)) + h(x) \\ &= f(x) + (g(x)+h(x)) = f(x) + (g+h)(x) = (f+(g+h))(x),\end{aligned}$$

where we have used the associativity of addition in V for the third equality. It follows that $(f + g) + h = f + (g + h)$. Similarly, let $f \in V^S$ and $a, b \in \mathbf{F}$ be given. Observe that, for any $x \in S$,

$$((ab)f)(x) = (ab)f(x) = a(bf(x)) = a((bf)(x)) = (a(bf))(x),$$

where we have used the associativity of scalar multiplication in V for the second equality. It follows that $(ab)f = a(bf)$.

Additive identity. We claim that the additive identity in V^S is the function $0 : S \to V$ given by $0(x) = 0$ for any $x \in S$; the 0 on the right-hand side is the additive identity in V. Indeed, for any $f \in V^S$ and $x \in S$ we have

$$(f + 0)(x) = f(x) + 0(x) = f(x) + 0 = f(x).$$

It follows that $f + 0 = f$.

Additive inverse. For $f \in V^S$, define $g : S \to V$ by $g(x) = -f(x)$ for $x \in S$, where $-f(x)$ is the additive inverse in V of $f(x)$. We claim that g is the additive inverse of f. To see this, let $x \in S$ be given and observe that

$$(f + g)(x) = f(x) + g(x) = f(x) + (-f(x)) = 0 = 0(x);$$

it follows that $f + g = 0$.

Multiplicative identity. Let $f \in V^S$ and $x \in S$ be given. Observe that

$$(1f)(x) = 1f(x) = f(x),$$

where we have used that $1v = v$ for any $v \in V$. It follows that $1f = f$.

Distributive properties. Let $a \in \mathbf{F}$ and $f, g \in V^S$ be given. Observe that, for any $x \in S$,

$$\begin{aligned}(a(f+g))(x) &= a(f+g)(x) = a((f(x)+g(x)) \\ &= af(x) + ag(x) = (af)(x) + (ag)(x) = (af+ag)(x),\end{aligned}$$

where we have used the first distributive property in V for the third equality. It follows that $a(f + g) = af + ag$. Similarly, let $a, b \in \mathbf{F}$ and $f \in V^S$ be given. For any $x \in S$, observe that

$$((a+b)f)(x) = (a+b)f(x) = af(x) + bf(x) = (af)(x) + (bf)(x) = (af+bf)(x),$$

where we have used the second distributive property in V for the second equality. It follows that $(a + b)f = af + bf$.

We may conclude that V^S is a vector space over $\mathbf{F}$.

Exercise 1.B.8. Suppose V is a real vector space.

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

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

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 $\mathbf{R}^n$.

Solution. We need to verify each condition in definition 1.20. The algebraic manipulations required to show that commutativity, associativity, and the first distributive property hold for $V_{\mathbf{C}}$ are essentially the same algebraic manipulations we performed in Exercise 1.A.1, Exercise 1.A.2, Exercise 1.A.3, and Exercise 1.A.4, except instead of using the algebraic properties of $\mathbf{R}$, we use the algebraic properties of V (i.e. the properties listed in 1.20); we will avoid repeating ourselves and instead verify the remaining conditions.

Additive identity. We claim that the additive identity in $V_{\mathbf{C}}$ is $0 + i0$, where 0 is the additive identity in V. Indeed, for any $u + iv \in V_{\mathbf{C}}$ we have
$$(u + iv) + (0 + i0) = (u + 0) + i(v + 0) = u + iv.$$

Additive inverse. We claim that the additive inverse of an element $u + iv \in V_{\mathbf{C}}$ is the element $(-u) + i(-v)$, where $-u$ is the additive inverse of u in V. Indeed,
$$(u + iv) + ((-u) + i(-v)) = (u + (-u)) + i(v + (-v)) = 0 + i0.$$

Multiplicative identity. For any $u + iv \in V_{\mathbf{C}}$, we have
$$(1 + 0i)(u + iv) = (1u - 0v) + i(1v + 0u) = u + iv.$$

Distributive properties. For the second distributive property, let $a + bi, c + di \in \mathbf{C}$ and $u + iv \in V_{\mathbf{C}}$ be given. Observe that
$$\begin{aligned}((a + bi) + (c + di))(u + iv) &= ((a + c) + (b + d)i)(u + iv)\\ &= ((a + c)u - (b + d)v) + i((a + c)v + (b + d)u)\\ &= (au + cu - bv - dv) + i(av + cv + bu + du)\end{aligned}$$

$$
\begin{aligned}
&= ((au - bv) + i(av + bu)) + ((cu - dv) + i(cv + du)) \\
&= (a + bi)(u + iv) + (c + di)(u + iv),
\end{aligned}
$$

where we have used the second distributive property for V for the third equality.

1.C. Subspaces

Exercise 1.C.1. For each of the following subsets of $\mathbf{F}^3$, determine whether it is a subspace of $\mathbf{F}^3$.

(a) $\{(x_1, x_2, x_3) \in \mathbf{F}^3 : x_1 + 2x_2 + 3x_3 = 0\}$

(b) $\{(x_1, x_2, x_3) \in \mathbf{F}^3 : x_1 + 2x_2 + 3x_3 = 4\}$

(c) $\{(x_1, x_2, x_3) \in \mathbf{F}^3 : x_1 x_2 x_3 = 0\}$

(d) $\{(x_1, x_2, x_3) \in \mathbf{F}^3 : x_1 = 5x_3\}$

Solution. Let U denote the set in each part of this question.

(a) This is a subspace of $\mathbf{F}^3$. Certainly the zero vector belongs to U. Suppose that $x = (x_1, x_2, x_3), y = (y_1, y_2, y_3) \in U$ and $\alpha \in \mathbf{F}$ and observe that

$$(x_1 + y_1) + 2(x_2 + y_2) + 3(x_3 + y_3) = (x_1 + 2x_2 + 3x_3) + (y_1 + 2y_2 + 3y_3) = 0 + 0 = 0,$$

$$\alpha x_1 + 2(\alpha x_2) + 3(\alpha x_3) = \alpha(x_1 + 2x_2 + 3x_3) = \alpha 0 = 0.$$

Thus $x + y$ and αx also belong to U. It follows from 1.34 that U is a subspace of V.

(b) This is not a subspace of $\mathbf{F}^3$ because it does not contain the zero vector.

(c) This is not a subspace of $\mathbf{F}^3$ because it is not closed under addition: $(1, 1, 0)$ and $(0, 0, 1)$ belong to U, but $(1, 1, 0) + (0, 0, 1) = (1, 1, 1)$ does not belong to U.

(d) This is a subspace of $\mathbf{F}^3$. Note that $U = \{(x_1, x_2, x_3) \in \mathbf{F}^3 : x_1 - 5x_3 = 0\}$. Certainly the zero vector belongs to U. Suppose that $x = (x_1, x_2, x_3), y = (y_1, y_2, y_3) \in U$ and $\alpha \in \mathbf{F}$ and observe that

$$(x_1 + y_1) - 5(x_3 + y_3) = (x_1 - 5x_3) + (y_1 - 5y_3) = 0 + 0 = 0,$$

$$\alpha x_1 - 5(\alpha x_3) = \alpha(x_1 - 5x_3) = \alpha 0 = 0.$$

Thus $x + y$ and αx also belong to U. It follows from 1.34 that U is a subspace of V.

Exercise 1.C.2. Verify all assertions about subspaces in Example 1.35.

Solution.

(a) The assertion is that if $b \in \mathbf{F}$, then

$$U = \{(x_1, x_2, x_3, x_4) \in \mathbf{F}^4 : x_3 = 5x_4 + b\} = \{(x_1, x_2, x_3, x_4) \in \mathbf{F}^4 : x_3 - 5x_4 = b\}$$

is a subspace of $\mathbf{F}^4$ if and only if $b = 0$. Indeed, if $b \neq 0$ then U is not a subspace of $\mathbf{F}^4$ because the zero vector does not belong to U, and if $b = 0$ then we may argue as in Exercise 1.C.1 (d) to see that U is a subspace of $\mathbf{F}^4$.

(b) The assertion is that the set of continuous real-valued functions on the interval $[0, 1]$ is a subspace of $\mathbf{R}^{[0,1]}$, i.e.

$$U = \{f : [0, 1] \to \mathbf{R}, f \text{ continuous}\}$$

is a subspace of $\mathbf{R}^{[0,1]}$. Certainly the zero function $x \mapsto 0$ on $[0, 1]$ is continuous and hence belongs to U, and it is well-known from elementary real analysis that sums and constant multiples of continuous functions are again continuous. It follows from 1.34 that U is a subspace of $\mathbf{R}^{[0,1]}$.

(c) The assertion is that the set of differentiable real-valued functions on $\mathbf{R}$ is a subspace of $\mathbf{R}^{\mathbf{R}}$, i.e.

$$U = \{f : \mathbf{R} \to \mathbf{R}, f \text{ differentiable}\}$$

is a subspace of $\mathbf{R}^{\mathbf{R}}$. Certainly the zero function $x \mapsto 0$ on $\mathbf{R}$ is differentiable and hence belongs to U, and it is well-known from elementary real analysis that sums and constant multiples of differentiable functions are again differentiable. It follows from 1.34 that U is a subspace of $\mathbf{R}^{\mathbf{R}}$.

(d) The assertion is that the set U of differentiable real-valued functions f on the interval $(0, 3)$ such that $f'(2) = b$ is a subspace of $\mathbf{R}^{(0,3)}$ if and only if $b = 0$. If $b \neq 0$, then the zero function $x \mapsto 0$ on $(0, 3)$, which has derivative $0 \neq b$ at $x = 2$, does not belong to U and thus U is not a subspace of $\mathbf{R}^{(0,3)}$.

Suppose that $b = 0$ and note that the zero function now belongs to U. If $f, g \in U$ and $\alpha \in \mathbf{R}$, then

$$(f + g)'(2) = f'(2) + g'(2) = 0 + 0 = 0 \quad \text{and} \quad (\alpha f)'(2) = \alpha f'(2) = \alpha 0 = 0.$$

Thus $f + g$ and αf belong to U. It follows from 1.34 that U is a subspace of $\mathbf{R}^{(0,3)}$.

(e) The assertion is that the set U of all sequences of complex numbers with limit 0 is a subspace of $\mathbf{C}^\infty$. Certainly the zero sequence $(0, 0, 0, ...)$ has limit 0 and hence belongs to U. Suppose that $x = (x_n)_{n=1}^\infty$ and $y = (y_n)_{n=1}^\infty$ belong to U and $\alpha \in \mathbf{C}$. Using basic results about limits, observe that

$$\lim_{n\to\infty} (x_n + y_n) = \lim_{n\to\infty} x_n + \lim_{n\to\infty} y_n = 0 + 0 = 0$$
$$\text{and} \quad \lim_{n\to\infty} (\alpha x_n) = \alpha \lim_{n\to\infty} x_n = \alpha 0 = 0.$$

Thus $x + y$ and αx belong to U. It follows from 1.34 that U is a subspace of $\mathbf{C}^{(0,3)}$.

Exercise 1.C.3. Show that the set of differentiable real-valued functions f on the interval $(-4, 4)$ such that $f'(-1) = 3f(2)$ is a subspace of $\mathbf{R}^{(-4,4)}$.

Solution. Let U be the set in question; it is straightforward to verify that the zero function belongs to U. Suppose that $f, g \in U$ and $\alpha \in \mathbf{R}$. Observe that

$$(f + g)'(-1) = f'(-1) + g'(-1) = 3f(2) + 3g(2) = 3(f(2) + g(2)) = 3(f + g)(2)$$

$$\text{and} \quad (\alpha f)'(-1) = \alpha f'(-1) = \alpha(3f(2)) = 3(\alpha f(2)) = 3(\alpha f)(2).$$

Thus $f+g$ and αf belong to U. It follows from 1.34 that U is a subspace of $\mathbf{R}^{(-4,4)}$.

Exercise 1.C.4. Suppose $b \in \mathbf{R}$. Show that the set of continuous real-valued functions f on the interval $[0,1]$ such that $\int_0^1 f = b$ is a subspace of $\mathbf{R}^{[0,1]}$ if and only if $b = 0$.

Solution. Let U be the set in question. If $b \neq 0$ then the zero function $x \mapsto 0$ on $[0,1]$, which has integral $0 \neq b$ over $[0,1]$, does not belong to U and thus U is not a subspace of $\mathbf{R}^{[0,1]}$.

Suppose that $b = 0$ and note that the zero function now belongs to U. If $f, g \in U$ and $\alpha \in \mathbf{R}$, then using basic properties of integration we have

$$\int_0^1 (f+g) = \int_0^1 f + \int_0^1 g = 0 + 0 = 0 \quad \text{and} \quad \int_0^1 (\alpha f) = \alpha \int_0^1 f = \alpha 0 = 0.$$

Thus $f+g$ and αf belong to U. It follows from 1.34 that U is a subspace of $\mathbf{R}^{[0,1]}$.

Exercise 1.C.5. Is $\mathbf{R}^2$ a subspace of the complex vector space $\mathbf{C}^2$?

Solution. The question is whether the subset

$$\mathbf{R}^2 = \{(x,y) : x, y \in \mathbf{R}\} \subseteq \{(z,w) : z, w \in \mathbf{C}\} = \mathbf{C}^2$$

is a subspace, where we are taking complex scalars in $\mathbf{C}^2$. This is not a subspace because it is not closed under scalar multiplication: $(1,0) \in \mathbf{R}^2$ but $i(1,0) = (i,0) \notin \mathbf{R}^2$.

Exercise 1.C.6.

(a) Is $\{(a,b,c) \in \mathbf{R}^3 : a^3 = b^3\}$ a subspace of $\mathbf{R}^3$?

(b) Is $\{(a,b,c) \in \mathbf{C}^3 : a^3 = b^3\}$ a subspace of $\mathbf{C}^3$?

Solution.

(a) Let U be the set in question. For $a, b \in \mathbf{R}$ we have $a^3 = b^3$ if and only if $a = b$ and thus the set U can be expressed as

$$U = \{(a,a,c) \in \mathbf{R}^3 : a, c \in \mathbf{R}\}.$$

Certainly $(0,0,0) \in U$. If $(a,a,c), (x,x,y) \in U$ and $\lambda \in \mathbf{R}$, then observe that

$$(a,a,c) + (x,x,y) = (a+x, a+x, c+y) \in U \quad \text{and} \quad \lambda(a,a,c) = (\lambda a, \lambda a, \lambda c) \in U.$$

It follows from 1.34 that U is a subspace of $\mathbf{R}^3$.

(b) Let U be the set in question. Observe that

$$\left(\frac{-1+\sqrt{3}i}{2}\right)^3 = \left(\frac{-1-\sqrt{3}i}{2}\right)^3 = 1.$$

It follows that $u := \left(\frac{-1+\sqrt{3}i}{2}, 1, 0\right)$ and $v := \left(\frac{-1-\sqrt{3}i}{2}, 1, 0\right)$ belong to U. However,

$$u + v = (-1, 2, 0) \notin U.$$

Thus U is not a subspace of $\mathbf{C}^3$ because it is not closed under addition.

Exercise 1.C.7. Prove or give a counterexample: If U is a nonempty subset of $\mathbf{R}^2$ such that U is closed under addition and under taking additive inverses (meaning $-u \in U$ whenever $u \in U$), then U is a subspace of $\mathbf{R}^2$.

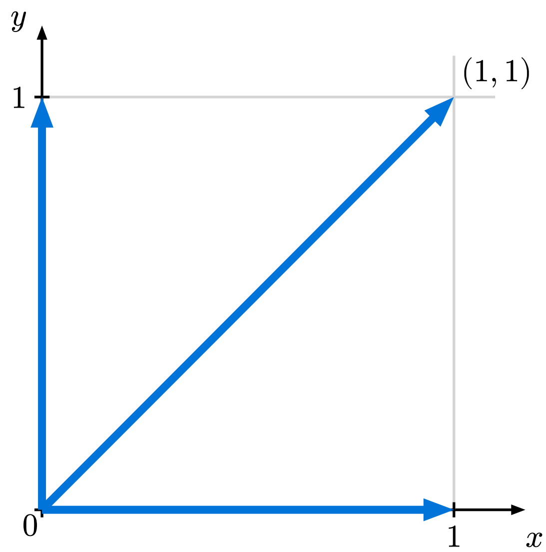
Solution. For a counterexample, consider $U = \{(a, b) : a, b \in \mathbf{Q}\} \subseteq \mathbf{R}^2$, which satisfies the required conditions since the sum of two rational numbers is a rational number and the additive inverse of a rational number is a rational number. However, U is not a subspace of $\mathbf{R}^2$ because it is not closed under scalar multiplication: $(1, 0) \in U$ but $\sqrt{2}(1, 0) = \left(\sqrt{2}, 0\right) \notin U$.

Exercise 1.C.8. Give an example of a nonempty subset U of $\mathbf{R}^2$ such that U is closed under scalar multiplication, but U is not a subspace of $\mathbf{R}^2$.

Solution. Let U be the union of the x- and y-axes, i.e.

$$U = \{(x, 0) : x \in \mathbf{R}\} \cup \{(0, y) : y \in \mathbf{R}\}.$$

It is straightforward to verify that U is closed under scalar multiplication. However, U is not a subspace of $\mathbf{R}^2$ because it is not closed under addition: $(1, 0)$ and $(0, 1)$ belong to U, but $(1, 0) + (0, 1) = (1, 1)$ does not.



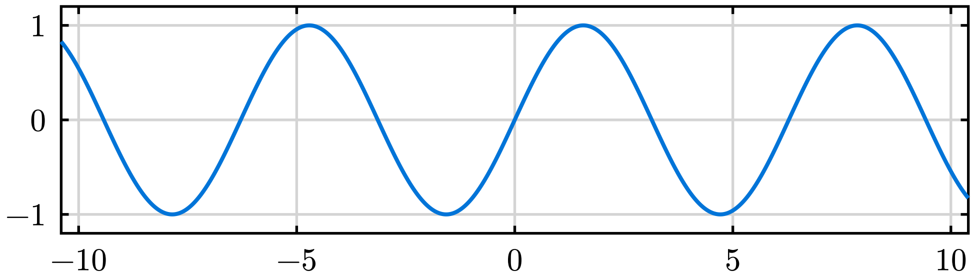
Exercise 1.C.9. A function $f : \mathbf{R} \to \mathbf{R}$ is called *periodic* if there exists a positive number p such that $f(x) = f(x + p)$ for all $x \in \mathbf{R}$. Is the set of periodic functions from $\mathbf{R}$ to $\mathbf{R}$ a subspace of $\mathbf{R}^{\mathbf{R}}$? Explain.

Solution. Consider the periodic functions $\sin(x)$ and $\sin\left(\sqrt{2}x\right)$ and let

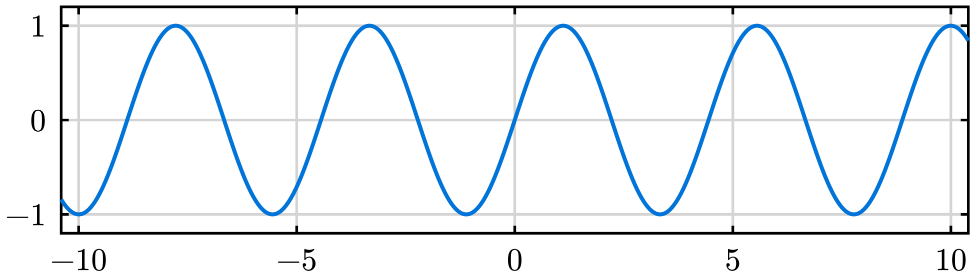
$$f(x) = \sin(x) + \sin\left(\sqrt{2}x\right).$$

We will show that f is not periodic.

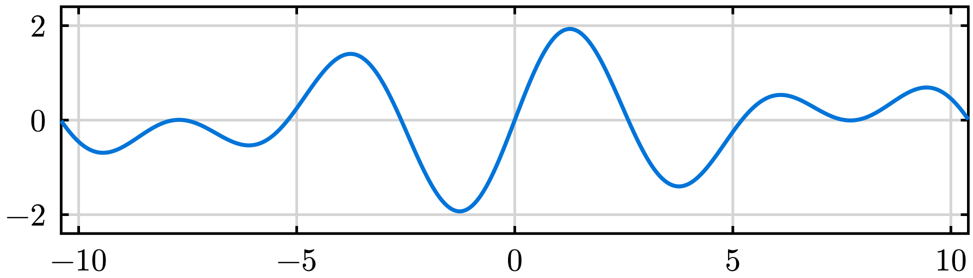
$\sin(x)$



$\sin\left(\sqrt{2}x\right)$



$\sin(x) + \sin\left(\sqrt{2}x\right)$



Suppose there was a positive real number p such that $f(x) = f(x+p)$ for all $x \in \mathbf{R}$, i.e.

$$\sin(x) + \sin\left(\sqrt{2}x\right) = \sin(x+p) + \sin\left(\sqrt{2}x + \sqrt{2}p\right) \text{ for all } x \in \mathbf{R}. \tag{1}$$

By differentiating this equation twice, we see that

$$\sin(x) + 2\sin\left(\sqrt{2}x\right) = \sin(x+p) + 2\sin\left(\sqrt{2}x + \sqrt{2}p\right) \text{ for all } x \in \mathbf{R}. \tag{2}$$

Subtracting equation (1) from equation (2) gives us

$$\sin\left(\sqrt{2}x\right) = \sin\left(\sqrt{2}x + \sqrt{2}p\right) \text{ for all } x \in \mathbf{R}, \tag{3}$$

which together with equation (1) implies that

$$\sin(x) = \sin(x+p) \text{ for all } x \in \mathbf{R}. \tag{4}$$

By taking $x = 0$ in equation (4) we see that $0 = \sin(p)$, which is the case if and only if $p = n\pi$ for some positive integer n (p was assumed to be positive). Substituting this value of p and $x = 0$ into equation (3) gives $0 = \sin\left(n\sqrt{2}\pi\right)$, which is the case if and only if $n\sqrt{2}\pi = m\pi$ for some integer m, which must be positive since n is positive. It follows that $\sqrt{2} = \frac{m}{n}$, contradicting the irrationality of $\sqrt{2}$.

Thus f is not periodic and we may conclude that the set of periodic functions from $\mathbf{R}$ to $\mathbf{R}$ is not a subspace of $\mathbf{R}^{\mathbf{R}}$ because it is not closed under addition.

Exercise 1.C.10. Suppose V_1 and V_2 are subspaces of V. Prove that the intersection $V_1 \cap V_2$ is a subspace of V.

Solution. Because V_1 and V_2 are subspaces of V, we have $0 \in V_1$ and $0 \in V_2$ and thus $0 \in V_1 \cap V_2$. Suppose $u, v \in V_1 \cap V_2$ and $\lambda \in \mathbf{F}$. Since $u, v \in V_1$ and V_1 is a subspace of V, we have $u + v \in V_1$ and $\lambda u \in V_1$. Similarly, $u + v \in V_2$ and $\lambda u \in V_2$. Thus $u + v \in V_1 \cap V_2$ and $\lambda u \in V_1 \cap V_2$. We may use 1.34 to conclude that $V_1 \cap V_2$ is a subspace of V.

Exercise 1.C.11. Prove that the intersection of every collection of subspaces of V is a subspace of V.

Solution. Let $\mathcal{U}$ be an arbitrary collection of subspaces of V. We will show that $\bigcap \mathcal{U}$ is a subspace of V. The zero vector belongs to $\bigcap \mathcal{U}$ because each $U \in \mathcal{U}$ is a subspace of V and hence contains the zero vector. If $u, v \in \bigcap \mathcal{U}$, $\lambda \in \mathbf{F}$, and $U \in \mathcal{U}$, then $u, v \in U$ and thus $u + v \in U$ and $\lambda u \in U$ since U is a subspace of V. Because $U \in \mathcal{U}$ was arbitrary, it follows that $u + v \in \bigcap \mathcal{U}$ and $\lambda u \in \bigcap \mathcal{U}$. We may use 1.34 to conclude that $\bigcap \mathcal{U}$ is a subspace of V.

Exercise 1.C.12. Prove that the union of two subspaces of V is a subspace of V if and only if one of the subspaces is contained in the other.

Solution. Suppose that U and W are subspaces of V. We want to show that $U \cup W$ is a subspace of V if and only if $U \subseteq W$ or $W \subseteq U$. If one of U or W is contained in the other then either $U \cup W = U$ or $U \cup W = W$; in either case, $U \cup W$ is then a subspace of V by assumption.

For the converse, it will suffice to show that if $U \cup W$ is a subspace of V and $U \not\subseteq W$, then $W \subseteq U$. Since $U \not\subseteq W$, there is some $u \in U$ such that $u \notin W$. Let $w \in W$ be given and note that, because $U \cup W$ is a subspace of V and $u, w \in U \cup W$, we must have $u + w \in U \cup W$. It cannot be the case that $u + w \in W$, since then $u + w - w = u \in W$, so it must be the case that $u + w \in U$. It follows that $u + w - u = w \in U$ and hence that $W \subseteq U$, as desired.

Exercise 1.C.13. Prove that the union of three subspaces of V is a subspace of V if and only if one of the subspaces contains the other two.

This exercise is surprisingly harder than Exercise 12, possibly because this exercise is not true if we replace $\mathbf{F}$ *with a field containing only two elements.*

Solution. Let U_1, U_2, and U_3 be subspaces of V. We want to show that $U = U_1 \cup U_2 \cup U_3$ is a subspace of V if and only if some U_j contains the other two. If some U_j contains the other two, then $U = U_j$ is a subspace of V by assumption.

Suppose that U is a subspace of V. If any U_j is contained in the union of the other two, say $U_1 \subseteq U_2 \cup U_3$, then $U = U_2 \cup U_3$ and we may apply Exercise 1.C.12 to see that either $U_2 \subseteq U_3$ or $U_3 \subseteq U_2$; in either case, one U_j contains the other two. Suppose therefore that no U_j is contained in the union of the other two. Seeking a contradiction, suppose further that no U_j contains the other two, so that

$$U_1 \nsubseteq (U_2 \cup U_3) \quad \text{and} \quad (U_2 \cup U_3) \nsubseteq U_1.$$

It follows that there exists some $u \in U_1$ such that $u \notin U_2 \cup U_3$ and some $v \in U_2 \cup U_3$ such that $v \notin U_1$. Let $W = \{v + \lambda u : \lambda \in \mathbf{F}\} \subseteq U$ and observe that no element of W belongs to U_1, for if $v + \lambda u \in U_1$ then $v + \lambda u - \lambda u = v \in U_1$—but $v \notin U_1$. Thus

$$W \cap U_1 = \emptyset \quad \text{and} \quad W \subseteq (U_1 \cup U_2 \cup U_3) \quad \Rightarrow \quad W \subseteq (U_2 \cup U_3).$$

Because W contains infinitely many elements, there must be some $i \in \{2, 3\}$ such that U_i contains infinitely many elements of W. There then exist $\lambda, \mu \in \mathbf{F}$ such that $\lambda \neq \mu$ and such that $v + \lambda u$ and $v + \mu u$ both belong to U_i, which implies that $(\lambda - \mu)u \in U_i$ since U_i is a subspace of V. This gives $u \in U_i$ since $\lambda - \mu \neq 0$, contradicting that $u \notin U_2 \cup U_3$. We may conclude that some U_j contains the other two.

Exercise 1.C.14. Suppose

$$U = \{(x, -x, 2x) \in \mathbf{F}^3 : x \in \mathbf{F}\} \quad \text{and} \quad W = \{(x, x, 2x) \in \mathbf{F}^3 : x \in \mathbf{F}\}.$$

Describe $U + W$ using symbols, and also give a description of $U + W$ that uses no symbols.

Solution. We claim that $U + W$ is the subspace

$$E = \{(x, y, 2x) \in \mathbf{F}^3 : x, y \in \mathbf{F}\}.$$

To see this, let $(x, -x, 2x) \in U$ and $(y, y, 2y) \in W$ be given and notice that

$$(x, -x, 2x) + (y, y, 2y) = (x + y, -x + y, 2(x + y)) \in E.$$

Thus $U + W \subseteq E$. For the reverse inclusion, let $(x, y, 2x) \in E$ be given and observe that

$$(x, y, 2x) = \left(\frac{x - y}{2}, \frac{y - x}{2}, x - y\right) + \left(\frac{x + y}{2}, \frac{x + y}{2}, x + y\right) \in U + W.$$

Thus $U + W = E$, as claimed. In words, $U + W$ is the subspace of $\mathbf{F}^3$ consisting of those vectors whose third coordinate is twice their first coordinate.

Exercise 1.C.15. Suppose U is a subspace of V. What is $U + U$?

Solution. For $u + v \in U + U$ we have $u + v \in U$ since U is a subspace of V, and for $u \in U$ we have $u = u + 0 \in U + U$. Thus $U + U = U$.

Exercise 1.C.16. Is the operation of addition on the subspaces of V commutative? In other words, if U and W are subspaces of V, is $U + W = W + U$?

Solution. The operation is commutative, since addition of vectors in V is commutative. If $u + w \in U + W$, then $u + w = w + u \in W + U$, so that $U + W \subseteq W + U$. Similarly, $W + U \subseteq U + W$.

Exercise 1.C.17. Is the operation of addition on the subspaces of V associative? In other words, if V_1, V_2, V_3 are subspaces of V, is

$$(V_1 + V_2) + V_3 = V_1 + (V_2 + V_3)?$$

Solution. The operation is associative, since addition of vectors in V is associative. If $(u_1 + u_2) + u_3 \in (U_1 + U_2) + U_3$, then

$$(u_1 + u_2) + u_3 = u_1 + (u_2 + u_3) \in U_1 + (U_2 + U_3),$$

so that $(U_1 + U_2) + U_3 \subseteq U_1 + (U_2 + U_3)$. Similarly, $U_1 + (U_2 + U_3) \subseteq (U_1 + U_2) + U_3$.

Exercise 1.C.18. Does the operation of addition on the subspaces of V have an additive identity? Which subspaces have additive inverses?

Solution. The subspace $\{0\}$ is the additive identity for the operation. If U is a subspace of V then $u + 0 = u$ for any $u \in U$; it follows that $U + \{0\} = U$.

Since $\{0\} + \{0\} = \{0\}$, the subspace $\{0\}$ is its own additive inverse. We claim that no other subspace of V has an additive inverse, i.e. if U is a subspace of V with $U \neq \{0\}$, then there does not exist a subspace W satisfying $U + W = \{0\}$. Indeed, simply observe that $U \subseteq U + W$ for any subspace W, so that $U + W \neq \{0\}$.

Exercise 1.C.19. Prove or give a counterexample: If V_1, V_2, U are subspaces of V such that

$$V_1 + U = V_2 + U,$$

then $V_1 = V_2$.

Solution. This is false. For a counterexample, consider the real vector space $\mathbf{R}$ and observe that

$$\{0\} + \mathbf{R} = \mathbf{R} + \mathbf{R} = \mathbf{R},$$

but $\{0\} \neq \mathbf{R}$.

Exercise 1.C.20. Suppose

$$U = \{(x, x, y, y) \in \mathbf{F}^4 : x, y \in \mathbf{F}\}.$$

Find a subspace W of $\mathbf{F}^4$ such that $\mathbf{F}^4 = U \oplus W$.

Solution. Let

$$W = \{(0, a, 0, b) \in \mathbf{F}^4 : a, b \in \mathbf{F}\};$$

it is straightforward to verify that W is a subspace of $\mathbf{F}^4$. If $v \in U \cap W$, then

$$v \in W \quad \Rightarrow \quad v = (0, a, 0, b) \text{ for some } a, b \in \mathbf{F},$$
$$v \in U \quad \Rightarrow \quad a = b = 0 \quad \Rightarrow \quad v = 0.$$

Thus $U \cap W = \{0\}$ and it follows from 1.46 that the sum $U + W$ is direct.

Let $(v_1, v_2, v_3, v_4) \in \mathbf{F}^4$ be given and observe that

$$(v_1, v_2, v_3, v_4) = (v_1, v_1, v_3, v_3) + (0, v_2 - v_1, 0, v_4 - v_3) \in U \oplus W.$$

Thus $\mathbf{F}^4 = U \oplus W$.

Exercise 1.C.21. Suppose

$$U = \{(x, y, x + y, x - y, 2x) \in \mathbf{F}^5 : x, y \in \mathbf{F}\}.$$

Find a subspace W of $\mathbf{F}^5$ such that $\mathbf{F}^5 = U \oplus W$.

Solution. Let

$$W = \{(0, 0, a, b, c) \in \mathbf{F}^5 : a, b, c \in \mathbf{F}\};$$

it is straightforward to verify that W is a subspace of $\mathbf{F}^5$. If $v \in U \cap W$, then

$$v \in U \quad \Rightarrow \quad v = (x, y, x + y, x - y, 2x) \text{ for some } x, y \in \mathbf{F},$$
$$v \in W \quad \Rightarrow \quad x = y = 0 \quad \Rightarrow \quad v = 0.$$

Thus $U \cap W = \{0\}$ and it follows from 1.46 that the sum $U + W$ is direct.

Let $v = (v_1, v_2, v_3, v_4, v_5) \in \mathbf{F}^5$ be given and observe that

$$\begin{aligned}(v_1, v_2, v_3, v_4, v_5) &= (v_1, v_2, v_1 + v_2, v_1 - v_2, 2v_1) \\ &\qquad + (0, 0, v_3 - (v_1 + v_2), v_4 - (v_1 - v_2), v_5 - 2v_1) \in U \oplus W.\end{aligned}$$

Thus $\mathbf{F}^5 = U \oplus W$.

Exercise 1.C.22. Suppose

$$U = \{(x, y, x+y, x-y, 2x) \in \mathbf{F}^5 : x, y \in \mathbf{F}\}.$$

Find three subspaces W_1, W_2, W_3 of $\mathbf{F}^5$, none of which equals $\{0\}$, such that $\mathbf{F}^5 = U \oplus W_1 \oplus W_2 \oplus W_3$.

Solution. Let

$$W_1 = \{(0,0,a,0,0) \in \mathbf{F}^5 : a \in \mathbf{F}\}, \quad W_2 = \{(0,0,0,b,0) \in \mathbf{F}^5 : b \in \mathbf{F}\},$$
$$W_3 = \{(0,0,0,0,c) \in \mathbf{F}^5 : c \in \mathbf{F}\};$$

it is straightforward to verify that W_1, W_2, and W_3 are subspaces of $\mathbf{F}^5$. Suppose that

$$u = (x, y, x+y, x-y, 2x) \in U, \qquad w_1 = (0,0,a,0,0) \in W_1,$$
$$w_2 = (0,0,0,b,0) \in W_2, \quad \text{and} \quad w_3 = (0,0,0,0,c) \in W_3$$

are such that $u + w_1 + w_2 + w_3 = 0$. That is,

$$(x, y, x+y+a, x-y+b, 2x+c) = (0,0,0,0,0),$$

from which it follows that $x = y = a = b = c = 0$. Thus the only way to express the zero vector as a sum $u + w_1 + w_2 + w_3 \in U + W_1 + W_2 + W_3$ is to take $u = w_1 = w_2 = w_3 = 0$ and so it follows from 1.45 that the sum $U + W_1 + W_2 + W_3$ is direct.

Let $(v_1, v_2, v_3, v_4, v_5) \in \mathbf{F}^5$ be given and observe that

$$\begin{aligned}(v_1, v_2, v_3, v_4, v_5) &= (v_1, v_2, v_1+v_2, v_1-v_2, 2v_1) + (0,0,v_3-(v_1+v_2),0,0) \\ &\quad + (0,0,0,v_4-(v_1-v_2),0) + (0,0,0,0,v_5-2v_1) \in U \oplus W_1 \oplus W_2 \oplus W_3.\end{aligned}$$

Thus $\mathbf{F}^5 = U \oplus W_1 \oplus W_2 \oplus W_3$.

Exercise 1.C.23. Prove or give a counterexample: If V_1, V_2, U are subspaces of V such that

$$V = V_1 \oplus U \quad \text{and} \quad V = V_2 \oplus U,$$

then $V_1 = V_2$.

Hint: When trying to discover whether a conjecture in linear algebra is true or false, it is often useful to start by experimenting in $\mathbf{F}^2$.

Solution. This is false. For a counterexample, consider $V = \mathbf{R}^2$,

$$U = \{(x,0) \in \mathbf{R}^2 : x \in \mathbf{R}\}, \quad V_1 = \{(0,y) \in \mathbf{R}^2 : y \in \mathbf{R}\}, \quad V_2 = \{(y,y) \in \mathbf{R}^2 : y \in \mathbf{R}\}.$$

It is straightforward to verify that $U \cap V_1 = U \cap V_2 = \{0\}$, so that $U + V_1$ and $U + V_2$ are both direct sums (1.46), and that $U \oplus V_1 = U \oplus V_2 = \mathbf{R}^2$. However, $V_1 \neq V_2$ since $(1,1) \in V_2$ but $(1,1) \notin V_1$.

Exercise 1.C.24. A function $f : \mathbf{R} \to \mathbf{R}$ is called *even* if

$$f(-x) = f(x)$$

for all $x \in \mathbf{R}$. A function $f : \mathbf{R} \to \mathbf{R}$ is called *odd* if

$$f(-x) = -f(x)$$

for all $x \in \mathbf{R}$. Let V_e denote the set of real-valued even functions on $\mathbf{R}$ and let V_o denote the set of real-valued odd functions on $\mathbf{R}$. Show that $\mathbf{R}^\mathbf{R} = V_\mathrm{e} \oplus V_\mathrm{o}$.

Solution. Suppose that $f \in V_\mathrm{e} \cap V_\mathrm{o}$, so that $f(x) = -f(x)$ for all $x \in \mathbf{R}$. This implies that $f(x) = 0$ for all $x \in \mathbf{R}$, i.e. $f = 0$. Thus $V_\mathrm{e} \cap V_\mathrm{o} = \{0\}$ and it follows from 1.46 that the sum $V_\mathrm{e} + V_\mathrm{o}$ is direct. For $f : \mathbf{R} \to \mathbf{R}$, define $f_\mathrm{e} : \mathbf{R} \to \mathbf{R}$ and $f_\mathrm{o} : \mathbf{R} \to \mathbf{R}$ by

$$f_\mathrm{e}(x) = \frac{f(x) + f(-x)}{2} \quad \text{and} \quad f_\mathrm{o}(x) = \frac{f(x) - f(-x)}{2}.$$

It is straightforward to verify that f_e is an even function, f_o is an odd function, and $f = f_\mathrm{e} + f_\mathrm{o}$. We may conclude that $\mathbf{R}^\mathbf{R} = V_\mathrm{e} \oplus V_\mathrm{o}$.

Chapter 2. Finite-Dimensional Vector Spaces

2.A. Span and Linear Independence

Exercise 2.A.1. Find a list of four distinct vectors in $\mathbf{F}^3$ whose span equals

$$\{(x, y, z) \in \mathbf{F}^3 : x + y + z = 0\}.$$

Solution. Let W be the subspace in question and consider the list

$$v_1 = (1, 0, -1), \quad v_2 = (0, 1, -1), \quad v_3 = (1, 1, -2), \quad v_4 = (1, -1, 0).$$

We claim that $\operatorname{span}(v_1, v_2, v_3, v_4) = W$. If $a_1, a_2, a_3, a_4 \in \mathbf{F}$, then

$$\begin{aligned} a_1 v_1 + a_2 v_2 + a_3 v_3 + a_4 v_4 = (a_1 + a_3 + a_4, a_2 + a_3 - a_4, -a_1 - a_2 - 2a_3) \in W \\ \text{since} \quad (a_1 + a_3 + a_4) + (a_2 + a_3 - a_4) + (-a_1 - a_2 - 2a_3) = 0. \end{aligned}$$

Thus $\operatorname{span}(v_1, v_2, v_3, v_4) \subseteq W$. Now suppose that $(x, y, z) \in W$ and observe that $z = -x - y$. It follows that

$$(x, y, z) = (x, y, -x - y) = x v_1 + y v_2 \in \operatorname{span}(v_1, v_2, v_3, v_4).$$

Thus $W \subseteq \operatorname{span}(v_1, v_2, v_3, v_4)$ and we may conclude that $\operatorname{span}(v_1, v_2, v_3, v_4) = W$, as claimed.

Exercise 2.A.2. Prove or give a counterexample: If v_1, v_2, v_3, v_4 spans V, then the list

$$v_1 - v_2, v_2 - v_3, v_3 - v_4, v_4$$

also spans V.

Solution. This is true. Let $v \in V$ be given. Since $V = \operatorname{span}(v_1, v_2, v_3, v_4)$, there are scalars a_1, a_2, a_3, a_4 such that $v = a_1 v_1 + a_2 v_2 + a_3 v_3 + a_4 v_4$. Observe that

$$\begin{aligned} a_1(v_1 - v_2) + (a_1 + a_2)(v_2 - v_3) + (a_1 + a_2 + a_3)(v_3 - v_4) + (a_1 + a_2 + a_3 + a_4) v_4 \\ = a_1 v_1 + a_2 v_2 + a_3 v_3 + a_4 v_4 = v. \end{aligned}$$

Thus $v \in \operatorname{span}(v_1 - v_2, v_2 - v_3, v_3 - v_4, v_4)$. It follows that

$$V = \operatorname{span}(v_1 - v_2, v_2 - v_3, v_3 - v_4, v_4).$$

Exercise 2.A.3. Suppose $v_1, ..., v_m$ is a list of vectors in V. For $k \in \{1, ..., m\}$, let

$$w_k = v_1 + \cdots + v_k.$$

Show that $\operatorname{span}(v_1, ..., v_m) = \operatorname{span}(w_1, ..., w_m)$.

Solution. For any scalars $a_1, \ldots, a_m \in \mathbf{F}$, observe that

$$a_1 v_1 + a_2(v_1 + v_2) + \cdots + a_m(v_1 + \cdots + v_m)$$
$$= (a_1 + \cdots + a_m)v_1 + (a_2 + \cdots + a_m)v_2 + \cdots + a_m v_m.$$

It follows that $\operatorname{span}(w_1, \ldots, w_m) \subseteq \operatorname{span}(v_1, \ldots, v_m)$. Similarly, for any scalars $a_1, \ldots, a_m \in \mathbf{F}$, notice that

$$a_1 v_1 + a_2 v_2 + \cdots + a_m v_m = (a_1 - a_2)v_1 + (a_2 - a_3)(v_1 + v_2)$$
$$+ \cdots + (a_{m-1} - a_m)(v_1 + \cdots + v_{m-1}) + a_m(v_1 + \cdots + v_m).$$

Thus $\operatorname{span}(v_1, \ldots, v_m) \subseteq \operatorname{span}(w_1, \ldots, w_m)$.

Exercise 2.A.4.

(a) Show that a list of length one in a vector space is linearly independent if and only if the vector in the list is not 0.

(b) Show that a list of length two in a vector space is linearly independent if and only if neither of the two vectors in the list is a scalar multiple of the other.

Solution.

(a) Suppose the list consists of the single vector $v \in V$. If $v \neq 0$ and $a \in \mathbf{F}$ is such that $av = 0$, then Exercise 1.B.2 shows that we must have $a = 0$; it follows that the list v is linearly independent. If $v = 0$ then simply observe that $1v = 0$, demonstrating that the list v is linearly dependent.

(b) Suppose that the list consists of the vectors $u, v \in V$. If one of these vectors is a scalar multiple of the other, say $v = \lambda u$ for some $\lambda \in \mathbf{F}$, then observe that $v - \lambda u = 0$. Because the coefficient of v in this linear combination is non-zero, we see that the list u, v is linearly dependent.

Conversely, suppose that the list u, v is linearly dependent, so that $\mu v + \lambda u = 0$ with at least one of the coefficients μ, λ non-zero, say $\mu \neq 0$; it follows that $v = -\frac{\lambda}{\mu} u$.

Exercise 2.A.5. Find a number t such that

$$(3, 1, 4), (2, -3, 5), (5, 9, t)$$

is not linearly independent in $\mathbf{R}^3$.

Solution. Let $t = 2$ and observe that

$$3(3, 1, 4) - 2(2, -3, 5) - (5, 9, 2) = (0, 0, 0).$$

Exercise 2.A.6. Show that the list $(2, 3, 1), (1, -1, 2), (7, 3, c)$ is linearly dependent in $\mathbf{F}^3$ if and only if $c = 8$.

Solution. That the list is linearly dependent if $c = 8$ was shown in the first bullet point of (2.20). Conversely, suppose that the list is linearly dependent. Since $(1, -1, 2)$ is evidently not a scalar multiple of $(2, 3, 1)$, the linear dependence lemma (2.19) implies that $(7, 3, c)$ lies in the span of $(2, 3, 1)$ and $(1, -1, 2)$, i.e. there are scalars x and y such that

$$x(2, 3, 1) + y(1, -1, 2) = (7, 3, c).$$

Solving the equations $2x + y = 7$ and $3x - y = 3$ gives $x = 2$ and $y = 3$, whence $c = x + 2y = 8$.

Exercise 2.A.7.

(a) Show that if we think of $\mathbf{C}$ as a vector space over $\mathbf{R}$, then the list $1 + i, 1 - i$ is linearly independent.

(b) Show that if we think of $\mathbf{C}$ as a vector space over $\mathbf{C}$, then the list $1 + i, 1 - i$ is linearly dependent.

Solution.

(a) Suppose that x and y are real numbers such that

$$x(1 + i) + y(1 - i) = (x + y) + (x - y)i = 0.$$

Since a complex number is zero if and only if its real and imaginary parts are zero, we must have

$$x + y = 0 \text{ and } x - y = 0 \quad \Leftrightarrow \quad x = y = 0.$$

Thus the list $1 + i, 1 - i$ is linearly independent.

(b) Observe that $i(1 - i) = 1 + i$, so that $1 + i$ is a scalar multiple of $1 - i$. It follows from Exercise 2.A.4 (b) that the list $1 + i, 1 - i$ is linearly dependent.

Exercise 2.A.8. Suppose v_1, v_2, v_3, v_4 is linearly independent in V. Prove that the list

$$v_1 - v_2, v_2 - v_3, v_3 - v_4, v_4$$

is also linearly independent.

Solution. Suppose that a_1, a_2, a_3, a_4 are scalars such that

$$\begin{aligned} &a_1(v_1 - v_2) + a_2(v_2 - v_3) + a_3(v_3 - v_4) + a_4 v_4 = 0 \\ &\qquad \Leftrightarrow \quad a_1 v_1 + (a_2 - a_1)v_2 + (a_3 - a_2)v_3 + (a_4 - a_3)v_4 = 0. \end{aligned}$$

Since the list v_1, v_2, v_3, v_4 is linearly independent, we must have

$$a_1 = a_2 - a_1 = a_3 - a_2 = a_4 - a_3 = 0,$$

which implies that $a_1 = a_2 = a_3 = a_4 = 0$. It follows that the list $v_1 - v_2, v_2 - v_3, v_3 - v_4, v_4$ is linearly independent.

Exercise 2.A.9. Prove or give a counterexample: If $v_1, v_2, ..., v_m$ is a linearly independent list of vectors in V, then

$$5v_1 - 4v_2, v_2, v_3, ..., v_m$$

is linearly independent.

Solution. Suppose that $a_1, a_2, ..., a_m$ are scalars such that

$$\begin{aligned} &a_1(5v_1 - 4v_2) + a_2v_2 + a_3v_3 + \cdots + a_mv_m = 0 \\ &\qquad\qquad \Leftrightarrow \; 5a_1v_1 + (a_2 - 4a_1)v_2 + a_3v_3 + \cdots + a_mv_m = 0. \end{aligned}$$

Since the list $v_1, v_2, ..., v_m$ is linearly independent, we must have

$$5a_1 = a_2 - 4a_1 = a_3 = \cdots = a_m = 0,$$

which implies that $a_1 = a_2 = a_3 = \cdots = a_m = 0$. It follows that the list $5v_1 - 4v_2, v_2, v_3, ..., v_m$ is linearly independent.

Exercise 2.A.10. Prove or give a counterexample: If $v_1, v_2, ..., v_m$ is a linearly independent list of vectors in V and $\lambda \in \mathbf{F}$ with $\lambda \neq 0$, then $\lambda v_1, \lambda v_2, ..., \lambda v_m$ is linearly independent.

Solution. Suppose that $a_1, a_2, ..., a_m$ are scalars such that

$$a_1\lambda v_1 + a_2\lambda v_2 + \cdots + a_m\lambda v_m = 0.$$

Since $\lambda \neq 0$, we may multiply both sides of this equation by λ^{-1} to obtain the equation

$$a_1v_1 + a_2v_2 + \cdots + a_mv_m = 0.$$

Since the list $v_1, v_2, ..., v_m$ is linearly independent, this implies that $a_1 = a_2 = \cdots = a_m = 0$. It follows that the list $\lambda v_1, \lambda v_2, ..., \lambda v_m$ is linearly independent.

Exercise 2.A.11. Prove or give a counterexample: If $v_1, ..., v_m$ and $w_1, ..., w_m$ are linearly independent lists of vectors in V, then the list $v_1 + w_1, ..., v_m + w_m$ is linearly independent.

Solution. This is false. Consider $\mathbf{R}$ as a vector space over itself. We have two linearly independent lists 1 and -1, but the list $1 + (-1) = 0$ is linearly dependent.

Exercise 2.A.12. Suppose $v_1, ..., v_m$ is linearly independent in V and $w \in V$. Prove that if $v_1 + w, ..., v_m + w$ is linearly dependent, then $w \in \operatorname{span}(v_1, ..., v_m)$.

Solution. By the linear dependence lemma (2.19), there is a $j \in \{1, 2, ..., m\}$ such that $v_j + w \in \operatorname{span}(v_1 + w, ..., v_{j-1} + w)$. If $j = 1$ then $v_1 + w = 0$, i.e. $w = -v_1$. It follows that $w \in \operatorname{span}(v_1, ..., v_m)$.

If $j \geq 2$, then there are scalars $a_1, ..., a_{j-1}$ such that

$$v_j + w = a_1(v_1 + w) + \cdots + a_{j-1}(v_{j-1} + w) \quad \Leftrightarrow \quad v_j + \lambda w = a_1 v_1 + \cdots + a_{j-1} v_{j-1},$$

where $\lambda = 1 - (a_1 + \cdots + a_{j-1})$. Note that λ must be non-zero: if this were not the case, then v_j would lie in the span of $v_1, ..., v_{j-1}$, which cannot happen since the list $v_1, ..., v_j$ is linearly independent. It follows that

$$w = \lambda^{-1}(a_1 v_1 + \cdots + a_{j-1} v_{j-1} - v_j),$$

so that $w \in \operatorname{span}(v_1, ..., v_m)$.

Exercise 2.A.13. Suppose $v_1, ..., v_m$ is linearly independent in V and $w \in V$. Show that

$$v_1, ..., v_m, w \text{ is linearly independent} \quad \Leftrightarrow \quad w \notin \operatorname{span}(v_1, ..., v_m).$$

Solution. If $w \in \operatorname{span}(v_1, ..., v_m)$ then the list $v_1, ..., v_m, w$ is linearly dependent by the third bullet point of 2.18. Conversely, suppose that the list $v_1, ..., v_m, w$ is linearly dependent. By the linear dependence lemma (2.19), one of the vectors in the list must be in the span of the previous vectors. It cannot be the case that some v_j belongs to $\operatorname{span}(v_1, ..., v_{j-1})$ since this would contradict the linear independence of the list $v_1, ..., v_m$, so it must be the case that $w \in \operatorname{span}(v_1, ..., v_m)$.

Exercise 2.A.14. Suppose $v_1, ..., v_m$ is a list of vectors in V. For $k \in \{1, ..., m\}$, let

$$w_k = v_1 + \cdots + v_k.$$

Show that the list $v_1, ..., v_m$ is linearly independent if and only if the list $w_1, ..., w_m$ is linearly independent.

Solution. Let $W = \operatorname{span}(w_1, ..., w_m)$; by Exercise 2.A.3 we also have $W = \operatorname{span}(v_1, ..., v_m)$. If the list $w_1, ..., w_m$ is linearly dependent, then using the linear dependence lemma (2.19) we may remove some w_j from the list $w_1, ..., w_m$ to obtain a spanning list for W of length $m - 1$. It follows from 2.22 that the list $v_1, ..., v_m$, which spans W, must be linearly dependent. A similar argument shows that the list $w_1, ..., w_m$ must be linearly dependent if the list $v_1, ..., v_m$ is linearly dependent.

Exercise 2.A.15. Explain why there does not exist a list of six polynomials that is linearly independent in $\mathcal{P}_4(\mathbf{F})$.

Solution. As noted in the textbook, $\mathcal{P}_4(\mathbf{F})$ is spanned by the list $1, z, z^2, z^3, z^4$, which has length 5. It follows from 2.22 that any linearly independent list in $\mathcal{P}_4(\mathbf{F})$ can have length at most 5.

Exercise 2.A.16. Explain why no list of four polynomials spans $\mathcal{P}_4(\mathbf{F})$.

Solution. As shown in (2.16) (b), the list $1, z, z^2, z^3, z^4$, which has length 5, is linearly independent in $\mathcal{P}_4(\mathbf{F})$. It follows from 2.22 that any spanning list for $\mathcal{P}_4(\mathbf{F})$ must have length at least 5.

Exercise 2.A.17. Prove that V is infinite-dimensional if and only if there is a sequence $v_1, v_2, \ldots$ of vectors in V such that $v_1, \ldots, v_m$ is linearly independent for every positive integer m.

Solution. First suppose that V is finite-dimensional, so that it is spanned by some list $w_1, \ldots, w_m$, and let $v_1, v_2, \ldots$ be any sequence of vectors in V; by 2.22, the list $v_1, v_2, \ldots, v_{m+1}$ must be linearly dependent.

Now suppose that V is infinite-dimensional, so that no list of vectors in V is a spanning list. Certainly $V \neq \{0\}$, so pick any $v_1 \neq 0$ in V and note that the list v_1 is linearly independent. Suppose that after m steps we have chosen a linearly independent list $v_1, \ldots, v_m$. By assumption $V \neq \operatorname{span}(v_1, \ldots, v_m)$, so pick any $v_{m+1} \notin \operatorname{span}(v_1, \ldots, v_m)$ and note that, by Exercise 2.A.13, the list $v_1, \ldots, v_m, v_{m+1}$ is linearly independent. This process recursively defines a sequence of vectors $v_1, v_2, \ldots$ such that $v_1, \ldots, v_m$ is linearly independent for each positive integer m.

Exercise 2.A.18. Prove that $\mathbf{F}^\infty$ is infinite-dimensional.

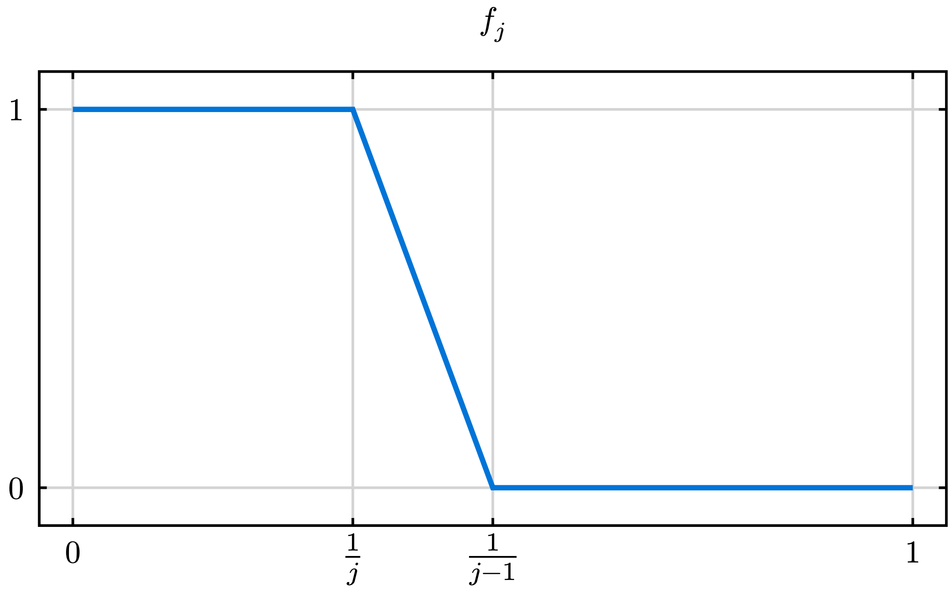
Solution. Consider the sequence of vectors $v_1, v_2, \ldots$, where $v_j \in \mathbf{F}^\infty$ is the sequence with a 1 in the j^{th} position and 0′s elsewhere. For each positive integer m, it is straightforward to verify that the list $v_1, \ldots, v_m$ is linearly independent; it follows from Exercise 2.A.17 that $\mathbf{F}^\infty$ is infinite-dimensional.

Exercise 2.A.19. Prove that the real vector space of all continuous real-valued functions on the interval $[0, 1]$ is infinite-dimensional.

Solution. Consider the sequence of continuous functions $f_1, f_2, \ldots$ on the interval $[0, 1]$, where $f_1(x) = 1$ for all $x \in [0, 1]$ and, for $j \geq 2$,

$$f_j(x) = \begin{cases} 1 & \text{if } 0 \leq x \leq \frac{1}{j}, \\ 0 & \text{if } \frac{1}{j-1} \leq x \leq 1. \end{cases}$$

On the interval $\left(\frac{1}{j}, \frac{1}{j-1}\right)$, take f_j to be the line segment joining the points $\left(\frac{1}{j}, 1\right)$ and $\left(\frac{1}{j-1}, 0\right)$, so that f_j is a continuous function on $[0, 1]$.



Let m be a positive integer and suppose we have real numbers $a_1, \ldots, a_m$ such that

$$a_1 f_1(x) + a_2 f_2(x) + \cdots + a_m f_m(x) = 0$$

for all $x \in [0, 1]$. Taking $x = 1$ gives us

$$a_1 f_1(1) + a_2 f_2(1) + \cdots + a_m f_m(1) = a_1 = 0.$$

Similarly, taking $x = \frac{1}{2}$ gives us

$$a_2 f_2\left(\tfrac{1}{2}\right) + \cdots + a_m f_m\left(\tfrac{1}{2}\right) = a_2 = 0.$$

By continuing in this fashion, taking $x = \frac{1}{j}$ for each $j \in \{1, 2, \ldots, m\}$, we see that $a_1 = a_2 = \cdots = a_m = 0$. It follows that the list $f_1, f_2, \ldots, f_m$ is linearly independent for each positive integer m and thus by Exercise 2.A.17 the real vector space of all continuous real-valued functions on the interval $[0, 1]$ is infinite-dimensional.

Exercise 2.A.20. Suppose $p_0, p_1, \ldots, p_m$ are polynomials in $\mathcal{P}_m(\mathbf{F})$ such that $p_k(2) = 0$ for each $k \in \{0, \ldots, m\}$. Prove that $p_0, p_1, \ldots, p_m$ is not linearly independent in $\mathcal{P}_m(\mathbf{F})$.

Solution. Since $\mathcal{P}_m(\mathbf{F})$ is spanned by the list $1, x, \ldots, x^m$ of length $m + 1$, 2.22 implies that the list $p_0, p_1, \ldots, p_m, x$ of length $m + 2$ is linearly dependent. The linear dependence lemma (2.19) implies that one of the vectors from this list belongs to the span of the previous vectors. Notice that for any scalars $a_0, \ldots, a_m$,

$$x = a_0 p_0(x) + \cdots + a_m p_m(x) \text{ for all } x \in \mathbf{F} \quad \Rightarrow \quad 2 = a_0 p_0(2) + \cdots + a_m p_m(2) = 0,$$

which is a contradiction; it follows that $x \notin \operatorname{span}(p_0, p_1, \ldots, p_m)$ and thus there must be some $j \in \{0, \ldots, m\}$ such that $p_j \in \operatorname{span}(p_0, p_1, \ldots, p_{j-1})$. The third bullet point of 2.18 then implies that the list $p_0, p_1, \ldots, p_m$ is linearly dependent.

2.B. Bases

Exercise 2.B.1. Find all vector spaces that have exactly one basis.

Solution. We will consider only finite-dimensional vector spaces over $\mathbf{R}$ or $\mathbf{C}$.

First consider the trivial vector space $\{0\}$. There are two possible lists of vectors: the empty list and the list 0. Since any list containing the zero vector is linearly dependent, the list 0 cannot be a basis of $\{0\}$. By definition the empty list is linearly independent and has span $\{0\}$; it follows that the empty list is a basis of $\{0\}$. Thus the trivial vector space has exactly one basis.

Now suppose that $V \neq \{0\}$. By 2.31, V has a basis $v_1, \ldots, v_m$. Since $V \neq \{0\}$, this basis is not the empty list, so v_1 exists and is non-zero. It follows that $B = 2v_1, \ldots, 2v_m$ is distinct from $v_1, \ldots, v_m$. By Exercise 2.A.10, B is linearly independent. Furthermore, we claim that $\operatorname{span} B = V$. Let $v \in V$ be given. Since $v_1, \ldots, v_m$ is a basis, there are scalars $a_1, \ldots, a_m$ such that $v = \sum_{j=1}^{m} a_j v_j$. This is equivalent to

$$v = \sum_{j=1}^{m} \left(\tfrac{1}{2} a_j\right)(2v_j);$$

it follows that $v \in \operatorname{span} B$ and hence that $\operatorname{span} B = V$, as claimed. Thus B is a basis of V, distinct from the original basis $v_1, \ldots, v_m$. We may conclude that the trivial vector space is the only vector space which has exactly one basis.

Exercise 2.B.2. Verify all assertions in Example 2.27.

Solution.

(a) The assertion is that the list $B = (1,0,\ldots,0), (0,1,0,\ldots,0), \ldots, (0,\ldots,0,1)$ is a basis of $\mathbf{F}^n$. Since any $(x_1, x_2, \ldots, x_n) \in \mathbf{F}^n$ can be expressed as

$$x_1(1,0,\ldots,0) + x_2(0,1,0,\ldots,0) + \cdots + x_n(0,\ldots,0,1),$$

we see that $\operatorname{span} B = \mathbf{F}^n$. Setting the above expression equal to $(0,0,0,\ldots,0)$ immediately gives us $x_1 = x_2 = \cdots = x_n = 0$, so that the list B is linearly independent. Thus B is a basis of $\mathbf{F}^n$.

(b) The assertion is that the list $B = (1,2), (3,5)$ is a basis of $\mathbf{F}^2$. Since neither of these vectors is a scalar multiple of the other, Exercise 2.A.4 (b) shows that B is linearly independent. If $(a,b) \in \mathbf{F}^2$, then observe that

$$(-5a+3b)(1,2) + (2a-b)(3,5) = (a,b).$$

Thus $\operatorname{span} B = V$ and we may conclude that B is a basis of $\mathbf{F}^2$.

(c) The assertion is that the list $B = (1,2,-4),(7,-5,6)$ is linearly independent in $\mathbf{F}^3$ but is not a basis of $\mathbf{F}^3$ because it does not span $\mathbf{F}^3$. Since neither of these vectors is a scalar multiple of the other, Exercise 2.A.4 (b) shows that B is linearly independent. However, since the list $(1,0,0),(0,1,0),(0,0,1)$ of length 3 is linearly independent in $\mathbf{F}^3$ (see (a)), 2.22 implies that B cannot span $\mathbf{F}^3$.

(d) The assertion is that the list $B = (1,2),(3,5),(4,13)$ spans $\mathbf{F}^2$ but is not a basis of $\mathbf{F}^2$ because it is not linearly independent. Indeed, part (b) shows that B spans $\mathbf{F}^2$ and that $(4,13)$ lies in the span of $(1,2)$ and $(3,5)$, so that B is linearly dependent.

(e) The assertion is that the list $B = (1,1,0),(0,0,1)$ is a basis of

$$U = \{(x,x,y) \in \mathbf{F}^3 : x,y \in \mathbf{F}\}.$$

Indeed, $\operatorname{span} B = U$ since $x(1,1,0) + y(0,0,1) = (x,x,y)$ for any scalars x,y, and B is linearly independent since $(x,x,y) = (0,0,0)$ forces $x = y = 0$.

(f) The assertion is that the list $B = (1,-1,0),(1,0,-1)$ is a basis of

$$U = \{(x,y,z) \in \mathbf{F}^3 : x+y+z = 0\}.$$

Observe that B is linearly independent since

$$x(1,-1,0) + y(1,0,-1) = (x+y,-x,-y) = (0,0,0)$$

gives us $x = y = 0$, and B spans U by Exercise 2.A.1 (using the notation of that exercise, we have $B = v_4, v_1$).

(g) The assertion is that the list $B = 1, z, \ldots, z^m$ is a basis of $\mathcal{P}_m(\mathbf{F})$. The fact that $\operatorname{span} B = \mathcal{P}_m(\mathbf{F})$ was noted on p. 31 of the textbook, and the linear independence of B was shown in 2.16(b).

Exercise 2.B.3.

(a) Let U be the subspace of $\mathbf{R}^5$ defined by

$$U = \{(x_1,x_2,x_3,x_4,x_5) \in \mathbf{R}^5 : x_1 = 3x_2 \text{ and } x_3 = 7x_4\}.$$

Find a basis of U.

(b) Extend the basis in (a) to a basis of $\mathbf{R}^5$.

(c) Find a subspace W of $\mathbf{R}^5$ such that $\mathbf{R}^5 = U \oplus W$.

Solution.

(a) Note that

$$U = \{(3x_1,x_1,7x_2,x_2,x_3) \in \mathbf{R}^5 : x_1,x_2,x_3 \in \mathbf{R}\}.$$

Let $u_1 = (3,1,0,0,0), u_2 = (0,0,7,1,0), u_3 = (0,0,0,0,1)$ and $B = u_1,u_2,u_3$. Since

$$x_1u_1 + x_2u_2 + x_3u_3 = (3x_1,x_1,7x_2,x_2,x_3)$$

for scalars x_1, x_2, x_3, we see that $\operatorname{span} B = U$. Setting the above expression equal to $(0,0,0,0,0)$, it is immediate that $x_1 = x_2 = x_3 = 0$, so that B is linearly independent. Thus B is a basis of U.

(b) Denote the j^{th} standard basis vector of $\mathbf{R}^5$ by e_j. Following the procedure outlined in 2.30 and 2.32, we adjoin the five standard basis vectors to B to obtain the spanning list

$$u_1, u_2, u_3, e_1, e_2, e_3, e_4, e_5.$$

- e_1 does not belong to $\operatorname{span}(u_1, u_2, u_3)$, so we do not delete it.
- Note that $e_2 = u_1 - 3e_1$, so we delete e_2 from the list.
- e_3 does not belong to $\operatorname{span}(u_1, u_2, u_3, e_1)$, so we do not delete it.
- Note that $e_4 = u_2 - 7e_3$, so we delete e_4 from the list.
- Since $e_5 = u_3$, we delete e_5 from the list.

We are left with the list u_1, u_2, u_3, e_1, e_3; as the proof of (2.32) shows, this must be a basis of $\mathbf{R}^5$.

(c) As shown in the proof of (2.33), if we let

$$W = \operatorname{span}(e_1, e_3) = \{(x_1, 0, x_3, 0, 0) \in \mathbf{R}^5 : x_1, x_3 \in \mathbf{R}\},$$

then $\mathbf{R}^5 = U \oplus W$.

Exercise 2.B.4.

(a) Let U be the subspace of $\mathbf{C}^5$ defined by

$$U = \{(z_1, z_2, z_3, z_4, z_5) \in \mathbf{C}^5 : 6z_1 = z_2 \text{ and } z_3 + 2z_4 + 3z_5 = 0\}.$$

Find a basis of U.

(b) Extend the basis in (a) to a basis of $\mathbf{C}^5$.

(c) Find a subspace W of $\mathbf{C}^5$ such that $\mathbf{C}^5 = U \oplus W$.

Solution.

(a) Note that

$$U = \{(z_1, 6z_1, -2z_2 - 3z_3, z_2, z_3) \in \mathbf{C}^5 : z_1, z_2, z_3 \in \mathbf{C}\}.$$

Let $u_1 = (1,6,0,0,0), u_2 = (0,0,-2,1,0), u_3 = (0,0,-3,0,1)$, and $B = u_1, u_2, u_3$. Since

$$z_1u_1 + z_2u_2 + z_3u_3 = (z_1, 6z_1, -2z_2 - 3z_3, z_2, z_3)$$

for scalars z_1, z_2, z_3, we see that $\operatorname{span} B = U$. Setting the above expression equal to $(0,0,0,0,0)$, it is immediate that $z_1 = z_2 = z_3 = 0$, so that B is linearly independent. Thus B is a basis of U.

(b) Denote the j^{th} standard basis vector of $\mathbf{C}^5$ by e_j. Following the procedure outlined in 2.30 and 2.32, we adjoin the five standard basis vectors to B to obtain the spanning list

$$u_1, u_2, u_3, e_1, e_2, e_3, e_4, e_5.$$

- e_1 does not belong to $\operatorname{span}(u_1, u_2, u_3)$, so we do not delete it.
- Note that $e_2 = \frac{1}{6}(u_1 - e_1)$, so we delete e_2 from the list.
- e_3 does not belong to $\operatorname{span}(u_1, u_2, u_3, e_1)$, so we do not delete it.
- Note that $e_4 = u_2 + 2e_3$, so we delete e_4 from the list.
- Since $e_5 = u_3 + 3e_3$, we delete e_5 from the list.

We are left with the list u_1, u_2, u_3, e_1, e_3; as the proof of 2.32 shows, this must be a basis of $\mathbf{C}^5$.

(c) As shown in the proof of (2.33), if we let

$$W = \operatorname{span}(e_1, e_3) = \{(z_1, 0, z_3, 0, 0) \in \mathbf{C}^5 : x_1, x_3 \in \mathbf{C}\},$$

then $\mathbf{C}^5 = U \oplus W$.

Exercise 2.B.5. Suppose V is finite-dimensional and U, W are subspaces of V such that $V = U + W$. Prove that there exists a basis of V consisting of vectors in $U \cup W$.

Solution. Let $u_1, ..., u_m$ be a basis of U and let $w_1, ..., w_n$ be a basis of W; these bases exist by 2.25 and 2.31. Since

$$U = \operatorname{span}(u_1, ..., u_m), \quad W = \operatorname{span}(w_1, ..., w_n), \quad \text{and} \quad V = U + W,$$

we see that $V = \operatorname{span}(u_1, ..., u_m, w_1, ..., w_n)$. Thus, using the procedure of 2.30, we can reduce the list $u_1, ..., u_m, w_1, ..., w_n$ to a basis of V consisting of vectors in $U \cup W$.

Exercise 2.B.6. Prove or give a counterexample: If p_0, p_1, p_2, p_3 is a list in $\mathcal{P}_3(\mathbf{F})$ such that none of the polynomials p_0, p_1, p_2, p_3 has degree 2, then p_0, p_1, p_2, p_3 is not a basis of $\mathcal{P}_3(\mathbf{F})$.

Solution. For a counterexample, consider $B = 1, x, x^2 + x^3, x^3$; none of the polynomials in this list has degree 2. Suppose a_0, a_1, a_2, a_3 are scalars such that

$$a_0 + a_1 x + a_2(x^2 + x^3) + a_3 x^3 = a_0 + a_1 x + a_2 x^2 + (a_2 + a_3)x^3 = 0$$

for all $x \in \mathbf{F}$. This implies that $a_0 = a_1 = a_2 = a_2 + a_3 = 0$ (we will prove this in 4.8), which in turn gives $a_3 = 0$. It follows that B is linearly independent. Now suppose that $p = a_0 + a_1 x + a_2 x^2 + a_3 x^3 \in \mathcal{P}_3(\mathbf{F})$ is given and observe that

$$a_0 + a_1 x + a_2(x^2 + x^3) + (a_3 - a_2)x^3 = p,$$

so that $p \in \operatorname{span} B$. It follows that $\mathcal{P}_3(\mathbf{F}) = \operatorname{span} B$ and hence that B is a basis of $\mathcal{P}_3(\mathbf{F})$.

Exercise 2.B.7. Suppose v_1, v_2, v_3, v_4 is a basis of V. Prove that

$$v_1 + v_2, v_2 + v_3, v_3 + v_4, v_4$$

is also a basis of V.

Solution. Let $B = v_1 + v_2, v_2 + v_3, v_3 + v_4, v_4$. Suppose there are scalars a_1, a_2, a_3, a_4 such that

$$\begin{aligned} &a_1(v_1 + v_2) + a_2(v_2 + v_3) + a_3(v_3 + v_4) + a_4 v_4 \\ &\qquad\qquad = a_1 v_1 + (a_1 + a_2)v_2 + (a_2 + a_3)v_3 + (a_3 + a_4)v_4 = 0. \end{aligned}$$

Since v_1, v_2, v_3, v_4 is a basis, this implies that

$$a_1 = a_1 + a_2 = a_2 + a_3 = a_3 + a_4 = 0 \quad \Rightarrow \quad a_1 = a_2 = a_3 = a_4 = 0.$$

Thus the list B is linearly independent. Let $v \in V$ be given. Since v_1, v_2, v_3, v_4 is a basis of V, there are scalars a_1, a_2, a_3, a_4 such that $v = a_1 v_1 + a_2 v_2 + a_3 v_3 + a_4 v_4$. Observe that

$$\begin{aligned} &a_1(v_1 + v_2) + (a_2 - a_1)(v_2 + v_3) + (a_3 - a_2 + a_1)(v_3 + v_4) \\ &\qquad\qquad + (a_4 - a_3 + a_2 - a_1)v_4 = a_1 v_1 + a_2 v_2 + a_3 v_3 + a_4 v_4 = v. \end{aligned}$$

It follows that $\operatorname{span} B = V$ and hence that B is a basis of V.

Exercise 2.B.8. Prove or give a counterexample: If v_1, v_2, v_3, v_4 is a basis of V and U is a subspace of V such that $v_1, v_2 \in U$ and $v_3 \notin U$ and $v_4 \notin U$, then v_1, v_2 is a basis of U.

Solution. For a counterexample, consider $V = \mathbf{R}^4$ and let e_j be the j^{th} standard basis vector of $\mathbf{R}^4$. It is straightforward to verify that the list

$$v_1 = e_1, \quad v_2 = e_2, \quad v_3 = e_3 + e_4, \quad v_4 = e_1 + e_4$$

is a basis of $\mathbf{R}^4$. Let $U = \operatorname{span}(e_1, e_2, e_3)$ and note that $v_1, v_2 \in U$. Note further that, since each vector $(a_1, a_2, a_3, a_4) \in U$ must satisfy $a_4 = 0$, we have $v_3, v_4 \notin U$. However, v_1, v_2 is not a basis for U: since e_1, e_2, e_3 is linearly independent, any spanning list for U must contain at least three vectors.

Exercise 2.B.9. Suppose $v_1, \ldots, v_m$ is a list of vectors in V. For $k \in \{1, \ldots, m\}$, let

$$w_k = v_1 + \cdots + v_k.$$

Show that $v_1, \ldots, v_m$ is a basis of V if and only if $w_1, \ldots, w_m$ is a basis of V.

Solution. This is immediate from Exercise 2.A.3 and Exercise 2.A.14.

Exercise 2.B.10. Suppose U and W are subspaces of V such that $V = U \oplus W$. Suppose also that $u_1, \dots, u_m$ is a basis of U and $w_1, \dots, w_n$ is a basis of W. Prove that

$$u_1, \dots, u_m, w_1, \dots, w_n$$

is a basis of V.

Solution. Let $v \in V$ be given. Since the sum $V = U \oplus W$ is direct, there are unique vectors $u \in U$ and $w \in W$ such that $v = u + w$. Because $u_1, \dots, u_m$ is a basis of U, 2.28 implies that there are unique scalars $a_1, \dots, a_m$ such that $u = a_1 u_1 + \cdots + a_m u_m$. Similarly, there are unique scalars $b_1, \dots, b_n$ such that $w = b_1 w_1 + \cdots + b_n w_n$. It follows that v can be uniquely represented as

$$v = a_1 u_1 + \cdots + a_m u_m + b_1 w_1 + \cdots + b_n w_n.$$

Thus, by 2.28, $u_1, \dots, u_m, w_1, \dots, w_n$ is a basis of V.

Exercise 2.B.11. Suppose V is a real vector space. Show that if $v_1, \dots, v_n$ is a basis of V (as a real vector space), then $v_1, \dots, v_n$ is also a basis of the complexification $V_{\mathbf{C}}$ (as a complex vector space).

See Exercise 8 in Section 1B for the definition of the complexification $V_{\mathbf{C}}$.

Solution. Let $u + iv \in V_{\mathbf{C}}$ be given. By 2.28, there are unique real scalars $a_1, \dots, a_n, b_1, \dots, b_n$ such that

$$u = a_1 v_1 + \cdots + a_n v_n \quad \text{and} \quad v = b_1 v_1 + \cdots + b_n v_n.$$

Using the definitions of vector addition and complex scalar multiplication in $V_{\mathbf{C}}$ given in Exercise 1.B.8, observe that

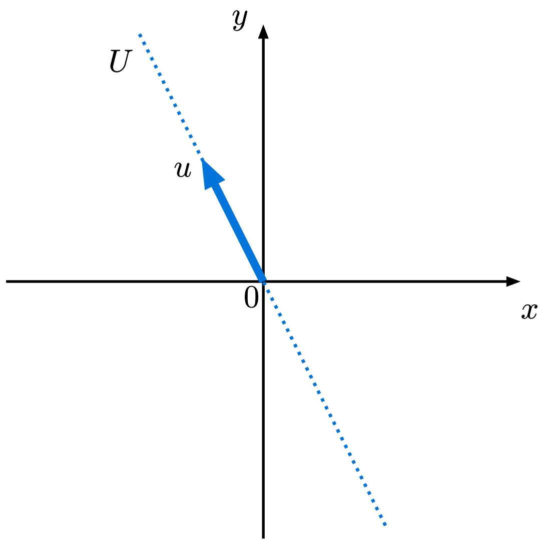
$$\sum_{k=1}^{n} (a_k + b_k i) v_k = \left(\sum_{k=1}^{n} a_k v_k \right) + i \left(\sum_{k=1}^{n} b_k v_k \right) = u + iv. \tag{1}$$

Because two ordered pairs (u, v) and (w, x) are equal if and only if $u = w$ and $v = x$, and the real scalars $a_1, \dots, a_n, b_1, \dots, b_n$ are unique, we see that the complex scalars $a_1 + b_1 i, \dots, a_n + b_n i$ in the linear combination on the left-hand side of equation (1) are unique. It follows from 2.28 that $v_1, \dots, v_n$ is a basis of $V_{\mathbf{C}}$.

2.C. Dimension

Exercise 2.C.1. Show that the subspaces of $\mathbf{R}^2$ are precisely $\{0\}$, all lines in $\mathbf{R}^2$ containing the origin, and $\mathbf{R}^2$.

Solution. It is easily verified that $\{0\}$, all lines in $\mathbf{R}^2$ through the origin, and $\mathbf{R}^2$ are indeed subspaces of $\mathbf{R}^2$. To see that these are the only subspaces of $\mathbf{R}^2$, suppose that U is a subspace of $\mathbf{R}^2$ and note that by 2.37 we must have $\dim U \in \{0, 1, 2\}$. If $\dim U = 0$ then $U = \{0\}$, if $\dim U = 2$ then $U = \mathbf{R}^2$ by 2.39, and if $\dim U = 1$ then there exists a basis $u \neq 0$ of U, so that $U = \operatorname{span}(u) = \{\lambda u : \lambda \in \mathbf{R}\}$, i.e. U is a line through the origin with direction vector u.



Exercise 2.C.2. Show that the subspaces of $\mathbf{R}^3$ are precisely $\{0\}$, all lines in $\mathbf{R}^3$ containing the origin, all planes in $\mathbf{R}^3$ containing the origin, and $\mathbf{R}^3$.

Solution. It is easily verified that $\{0\}$, all lines in $\mathbf{R}^3$ through the origin, all planes in $\mathbf{R}^3$ through the origin, and $\mathbf{R}^3$ are indeed subspaces of $\mathbf{R}^3$. To see that these are the only subspaces of $\mathbf{R}^3$, suppose that U is a subspace of $\mathbf{R}^3$ and note that by 2.37 we must have $\dim U \in \{0, 1, 2, 3\}$. If $\dim U = 0$ then $U = \{0\}$, if $\dim U = 3$ then $U = \mathbf{R}^3$ by 2.39, and if $\dim U = 1$ then there exists a basis $u \neq 0$ of U, so that $U = \operatorname{span}(u) = \{\lambda u : \lambda \in \mathbf{R}\}$, i.e. U is a line through the origin with direction vector u. If $\dim U = 2$ then there is a basis u_1, u_2 of U, so that $U = \operatorname{span}(u_1, u_2) = \{\lambda_1 u_1 + \lambda_2 u_2 : \lambda_1, \lambda_2 \in \mathbf{R}\}$. Because neither of u_1, u_2 is a scalar multiple of the other, i.e. u_1 and u_2 are not collinear, this describes a plane through the origin in $\mathbf{R}^3$.

Exercise 2.C.3.

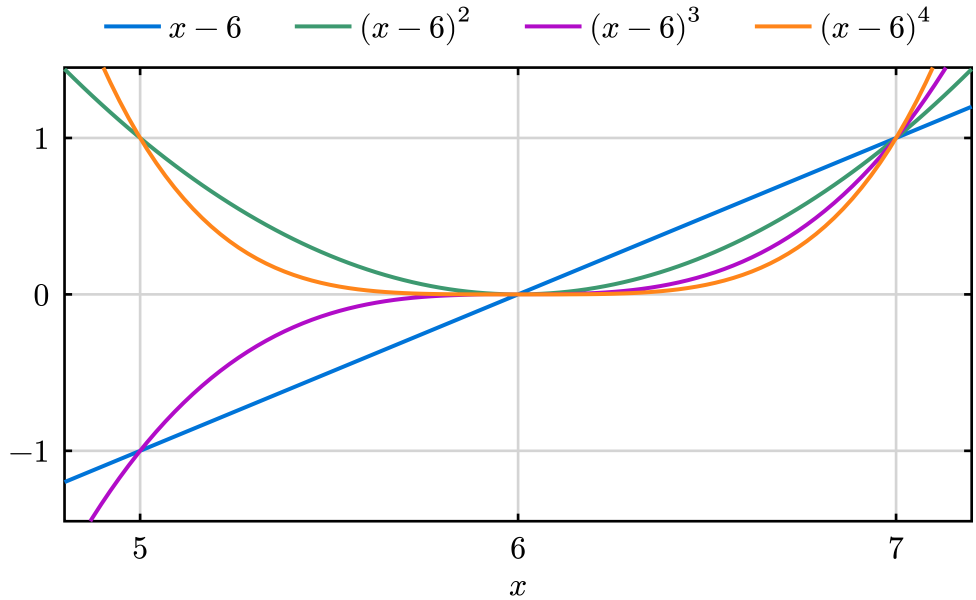
(a) Let $U = \{p \in \mathcal{P}_4(\mathbf{F}) : p(6) = 0\}$. Find a basis of U.

(b) Extend the basis in (a) to a basis of $\mathcal{P}_4(\mathbf{F})$.

(c) Find a subspace W of $\mathcal{P}_4(\mathbf{F})$ such that $\mathcal{P}_4(\mathbf{F}) = U \oplus W$.

Solution.

(a) Let $B = x - 6, (x-6)^2, (x-6)^3, (x-6)^4$; certainly each of these polynomials belongs to U.



Suppose we have scalars a_1, a_2, a_3, a_4 such that

$$a_1(x-6) + a_2(x-6)^2 + a_3(x-6)^3 + a_4(x-6)^4 = 0$$

for all $x \in \mathbf{F}$. Using the reasoning of 2.41, we see that this equation implies that $a_1 = a_2 = a_3 = a_4 = 0$. It follows that B is linearly independent and thus by 2.22 we have $\dim U \geq 4$. Using 2.37, we also find that $\dim U \leq \dim \mathcal{P}_4(\mathbf{F}) = 5$. However, notice that $U \neq \mathcal{P}_4(\mathbf{F})$ because the non-zero constant polynomials do not belong to U; it follows from 2.39 that $\dim U$ must be strictly less than $\dim \mathcal{P}_4(\mathbf{F}) = 5$. Thus $\dim U = 4$ and using 2.38 we may conclude that B is a basis of U.

(b) Certainly the constant polynomial 1 does not belong to $U = \operatorname{span} B$. It then follows from Exercise 2.A.13 that the list $B' = 1, x-6, (x-6)^2, (x-6)^3, (x-6)^4$ is linearly independent. Since $\dim \mathcal{P}_4(\mathbf{F}) = 5$, 2.38 allows us to conclude that B' is a basis of $\mathcal{P}_4(\mathbf{F})$.

(c) Let $W = \operatorname{span}(1)$, i.e. the subspace of all constant polynomials. As the proof of 2.33 shows, we then have $\mathcal{P}_4(\mathbf{F}) = U \oplus W$.

Exercise 2.C.4.

(a) Let $U = \{p \in \mathcal{P}_4(\mathbf{R}) : p''(6) = 0\}$. Find a basis of U.

(b) Extend the basis in (a) to a basis of $\mathcal{P}_4(\mathbf{R})$.

(c) Find a subspace W of $\mathcal{P}_4(\mathbf{R})$ such that $\mathcal{P}_4(\mathbf{R}) = U \oplus W$.

Solution.

(a) Let $B = 1, x, (x-6)^3, (x-6)^4$; it is straightforward to verify that each of these polynomials belongs to U. Suppose we have scalars a_0, a_1, a_3, a_4 such that

$$a_0 + a_1 x + a_3(x-6)^3 + a_4(x-6)^4 = 0$$

for all $x \in \mathbf{R}$. Using the reasoning of 2.41, we see that this equation implies that $a_0 = a_1 = a_3 = a_4 = 0$. It follows that B is linearly independent and thus by 2.22 we have $\dim U \geq 4$. Using 2.37, we also find that $\dim U \leq \dim \mathcal{P}_4(\mathbf{R}) = 5$. However, notice that $U \neq \mathcal{P}_4(\mathbf{R})$ because $x^2 \notin U$; it follows from 2.39 that $\dim U$ must be strictly less than $\dim \mathcal{P}_4(\mathbf{R}) = 5$. Thus $\dim U = 4$ and using 2.38 we may conclude that B is a basis of U.

(b) As noted in part (a), $x^2 \notin U = \operatorname{span} B$. It then follows from Exercise 2.A.13 that the list $B' = 1, x, x^2, (x-6)^3, (x-6)^4$ is linearly independent. Since $\dim \mathcal{P}_4(\mathbf{R}) = 5$, (2.38) allows us to conclude that B' is a basis of $\mathcal{P}_4(\mathbf{R})$.

(c) Let $W = \operatorname{span}(x^2)$. As the proof of 2.33 shows, we then have $\mathcal{P}_4(\mathbf{R}) = U \oplus W$.

Exercise 2.C.5.

(a) Let $U = \{p \in \mathcal{P}_4(\mathbf{F}) : p(2) = p(5)\}$. Find a basis of U.

(b) Extend the basis in (a) to a basis of $\mathcal{P}_4(\mathbf{F})$.

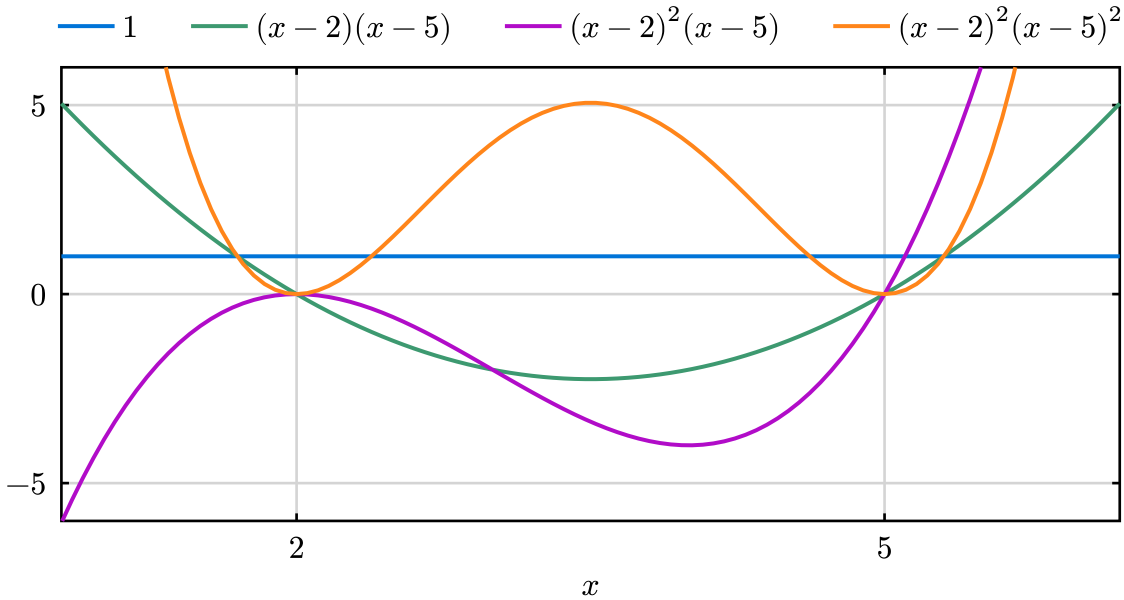
(c) Find a subspace W of $\mathcal{P}_4(\mathbf{F})$ such that $\mathcal{P}_4(\mathbf{F}) = U \oplus W$.

Solution.

(a) Let $B = 1, (x-2)(x-5), (x-2)^2(x-5), (x-2)^2(x-5)^2$; it is straightforward to verify that each of these polynomials belongs to U. Suppose we have scalars a_0, a_2, a_3, a_4 such that

$$a_0 + a_2(x-2)(x-5) + a_3(x-2)^2(x-5) + a_4(x-2)^2(x-5)^2 = 0$$

for all $x \in \mathbf{F}$. Using the reasoning of 2.41, we see that this equation implies that $a_0 = a_2 = a_3 = a_4 = 0$. It follows that B is linearly independent and thus by 2.22 we have $\dim U \geq 4$. Using 2.37, we also find that $\dim U \leq \dim \mathcal{P}_4(\mathbf{F}) = 5$. However, notice that $U \neq \mathcal{P}_4(\mathbf{F})$ because $x \notin U$; it follows from 2.39 that $\dim U$ must be strictly less than $\dim \mathcal{P}_4(\mathbf{F}) = 5$. Thus $\dim U = 4$ and using 2.38 we may conclude that B is a basis of U.



(b) As noted in part (a), $x \notin U = \operatorname{span} B$. It then follows from Exercise 2.A.13 that the list $B' = 1, x, (x-2)(x-5), (x-2)^2(x-5), (x-2)^2(x-5)^2$ is linearly independent. Since $\dim \mathcal{P}_4(\mathbf{F}) = 5$, 2.38 allows us to conclude that B' is a basis of $\mathcal{P}_4(\mathbf{F})$.

(c) Let $W = \operatorname{span}(x)$. As the proof of 2.33 shows, we then have $\mathcal{P}_4(\mathbf{F}) = U \oplus W$.

Exercise 2.C.6.

(a) Let $U = \{p \in \mathcal{P}_4(\mathbf{F}) : p(2) = p(5) = p(6)\}$. Find a basis of U.

(b) Extend the basis in (a) to a basis of $\mathcal{P}_4(\mathbf{F})$.

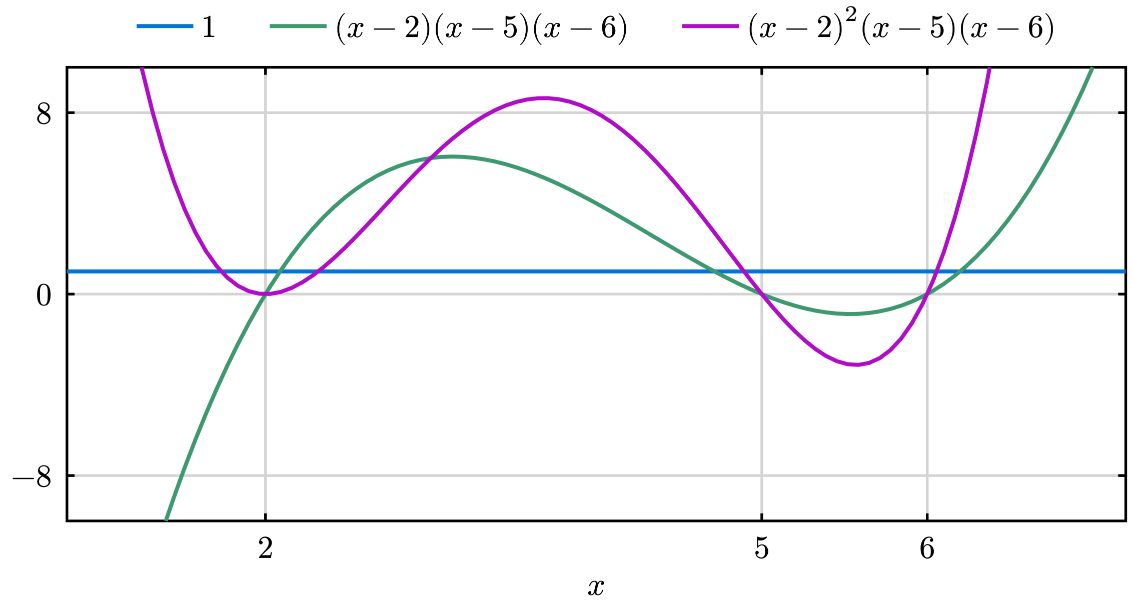
(c) Find a subspace W of $\mathcal{P}_4(\mathbf{F})$ such that $\mathcal{P}_4(\mathbf{F}) = U \oplus W$.

Solution.

(a) Let $B = 1, (x-2)(x-5)(x-6), (x-2)^2(x-5)(x-6)$; it is straightforward to verify that each of these polynomials belongs to U. Suppose we have scalars a_0, a_3, a_4 such that

$$a_0 + a_3(x-2)(x-5)(x-6) + a_4(x-2)^2(x-5)(x-6) = 0$$

for all $x \in \mathbf{F}$. Using the reasoning of 2.41, we see that this equation implies that $a_0 = a_3 = a_4 = 0$. It follows that B is linearly independent and thus by 2.22 we have $\dim U \geq 3$. Let Y denote the subspace from Exercise 2.C.5 and notice that U is a subspace of Y. Using 2.37, we then find that $\dim U \leq \dim Y = 4$. However, notice that $U \neq Y$ because $(x-2)(x-5) \in Y$ but $(x-2)(x-5) \notin U$; it follows from 2.39 that $\dim U$ must be strictly less than $\dim Y = 4$. Thus $\dim U = 3$ and using 2.38 we may conclude that B is a basis of U.



(b) As noted in part (a), $(x-2)(x-5) \in Y$ but $(x-2)(x-5) \notin U = \operatorname{span} B$. It then follows from Exercise 2.A.13 that the list

$$B' = 1, (x-2)(x-5), (x-2)(x-5)(x-6), (x-2)^2(x-5)(x-6)$$

is linearly independent. Since $\dim Y = 4$, 2.38 shows that B' is a basis of Y. We can now argue as in Exercise 2.C.5 (b) to conclude that the list

$$1, x, (x-2)(x-5), (x-2)(x-5)(x-6), (x-2)^2(x-5)(x-6)$$

is a basis of $\mathcal{P}_4(\mathbf{F})$.

(c) Let $W = \operatorname{span}(x, (x-2)(x-5))$. As the proof of 2.33 shows, we then have $\mathcal{P}_4(\mathbf{F}) = U \oplus W$.

Exercise 2.C.7.

(a) Let $U = \left\{p \in \mathcal{P}_4(\mathbf{R}) : \int_{-1}^{1} p = 0\right\}$. Find a basis of U.

(b) Extend the basis in (a) to a basis of $\mathcal{P}_4(\mathbf{R})$.

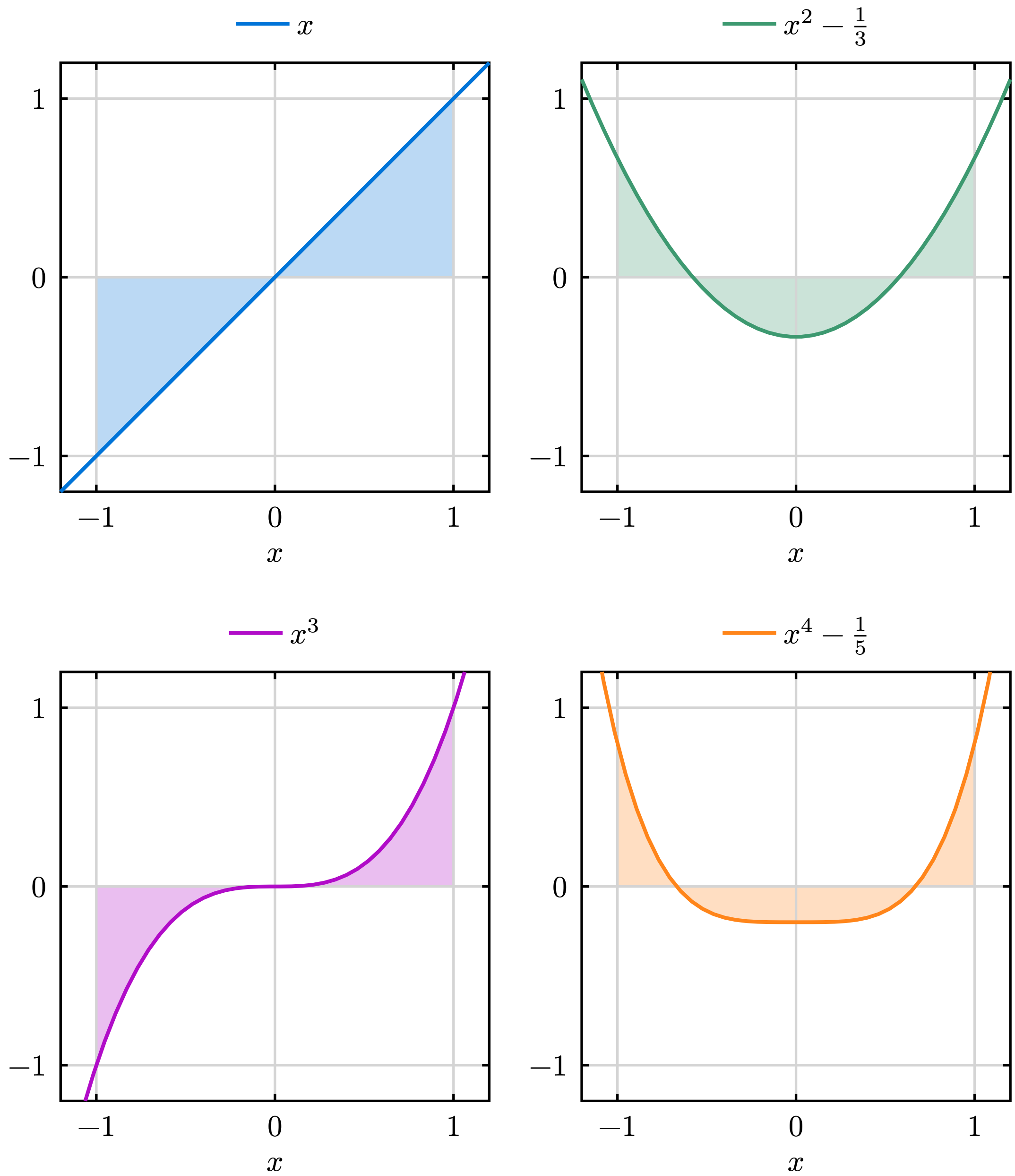
(c) Find a subspace W of $\mathcal{P}_4(\mathbf{R})$ such that $\mathcal{P}_4(\mathbf{R}) = U \oplus W$.

Solution.

(a) Let $B = x, x^2 - \frac{1}{3}, x^3, x^4 - \frac{1}{5}$; it is straightforward to verify that each of these polynomials belongs to U. Suppose we have scalars a_1, a_2, a_3, a_4 such that

$$a_1 x + a_2\left(x^2 - \tfrac{1}{3}\right) + a_3 x^3 + a_4\left(x^4 - \tfrac{1}{5}\right) = 0$$

for all $x \in \mathbf{R}$. Using the reasoning of 2.41, we see that this equation implies that $a_1 = a_2 = a_3 = a_4 = 0$. It follows that B is linearly independent and thus by 2.22 we have $\dim U \geq 4$. Using 2.37, we also find that $\dim U \leq \dim \mathcal{P}_4(\mathbf{R}) = 5$. However, notice that $U \neq \mathcal{P}_4(\mathbf{R})$ because $1 \notin U$; it follows from 2.39 that $\dim U$ must be strictly less than $\dim \mathcal{P}_4(\mathbf{R}) = 5$. Thus $\dim U = 4$ and using 2.38 we may conclude that B is a basis of U.



(b) As noted in part (a), $1 \notin U = \operatorname{span} B$. It then follows from Exercise 2.A.13 that the list $B' = 1, x, x^2 - \frac{1}{3}, x^3, x^4 - \frac{1}{5}$ is linearly independent. Since $\dim \mathcal{P}_4(\mathbf{R}) = 5$, 2.38 allows us to conclude that B' is a basis of $\mathcal{P}_4(\mathbf{R})$.

(c) Let $W = \operatorname{span}(1)$. As the proof of 2.33 shows, we then have $\mathcal{P}_4(\mathbf{R}) = U \oplus W$.

Exercise 2.C.8. Suppose $v_1, ..., v_m$ is linearly independent in V and $w \in V$. Prove that

$$\dim \operatorname{span}(v_1 + w, ..., v_m + w) \geq m - 1.$$

Solution. If $m = 1$ then certainly $\dim \operatorname{span}(v_1 + w) \geq 0$, so suppose that $m \geq 2$. Because $v_1, ..., v_m$ is linearly independent, notice that:

- $-v_1 \notin \operatorname{span}(v_2, \ldots, v_m)$;
- $v_2, \ldots, v_m$ is linearly independent.

It then follows from the contrapositive of Exercise 2.A.12 that the list $B = v_2 - v_1, \ldots, v_m - v_1$ is linearly independent. Now observe that

$$v_j - v_1 = (v_j + w) - (v_1 + w) \in \operatorname{span}(v_1 + w, \ldots, v_m + w)$$

for any $2 \leq j \leq m$. Thus B is a linearly independent list of length $m - 1$ contained in $\operatorname{span}(v_1 + w, \ldots, v_m + w)$ and we may use 2.22 to conclude that

$$\dim \operatorname{span}(v_1 + w, \ldots, v_m + w) \geq m - 1.$$

Exercise 2.C.9. Suppose m is a positive integer and $p_0, p_1, \ldots, p_m \in \mathcal{P}(\mathbf{F})$ are such that each p_k has degree k. Prove that $p_0, p_1, \ldots, p_m$ is a basis of $\mathcal{P}_m(\mathbf{F})$.

Solution. Suppose we have scalars $a_0, a_1, \ldots, a_m$ such that

$$a_0 p_0(x) + a_1 p_1(x) + \cdots + a_m p_m(x) = 0 \tag{$*$}$$

for all $x \in \mathbf{F}$. Let c be the coefficient of x^m in the polynomial p_m and note that $c \neq 0$ since p_m has degree m. Because each p_k has degree k, the coefficient of x^m in the polynomial p_k must be zero for $k < m$. Thus the left-hand side of $(*)$ has an $a_m c x^m$ term whereas the right-hand side has no x^m term. It follows that $a_m c = 0$ and hence that $a_m = 0$, since $c \neq 0$. Repeating this argument for the lower degree terms, we find that $a_0 = a_1 = \cdots = a_m = 0$. Thus $p_0, p_1, \ldots, p_m$ is a linearly independent list of length $m + 1$ contained in $\mathcal{P}_m(\mathbf{F})$. Since $\dim \mathcal{P}_m(\mathbf{F}) = m + 1$, 2.38 allows us to conclude that $p_0, p_1, \ldots, p_m$ is a basis of $\mathcal{P}_m(\mathbf{F})$.

Exercise 2.C.10. Suppose m is a positive integer. For $0 \leq k \leq m$, let

$$p_k(x) = x^k (1 - x)^{m-k}.$$

Show that $p_0, \ldots, p_m$ is a basis of $\mathcal{P}_m(\mathbf{F})$.

> *The basis in this exercise leads to what are called* ***Bernstein polynomials****. You can do a web search to learn how Bernstein polynomials are used to approximate continuous functions on* $[0, 1]$.

Solution. To remind us that these polynomials depend on m, let us use the notation

$$p_{k,m}(x) = x^k (1 - x)^{m-k}.$$

For a positive integer m, let $S(m)$ be the statement that the list $p_{0,m}, \ldots, p_{m,m}$ is linearly independent. We will use induction (twice) to show that $S(m)$ holds for all positive integers. First, we show the truth of $S(1)$ and $S(2)$.

For $S(1)$, suppose that a_0, a_1 are scalars such that

$$a_0 p_{0,1}(x) + a_1 p_{1,1}(x) = a_0(1 - x) + a_1 x = 0$$

for all $x \in \mathbf{F}$. Taking $x = 0$ and $x = 1$ immediately gives us $a_0 = a_1 = 0$.

For $S(2)$, suppose that a_0, a_1, a_2 are scalars such that

$$a_0 p_{0,2}(x) + a_1 p_{1,2}(x) + a_2 p_{2,2}(x) = a_0(1-x)^2 + a_1 x(1-x) + a_2 x^2 = 0$$

for all $x \in \mathbf{F}$. Taking $x = 0$ and $x = 1$ immediately gives us $a_0 = a_2 = 0$, and then taking any $x \notin \{0, 1\}$ gives us $a_2 = 0$.

Now suppose that $S(m)$ holds for some positive integer m and let $a_0, ..., a_{m+2}$ be scalars such that

$$\sum_{k=0}^{m+2} a_k p_{k,m+2}(x) = \sum_{k=0}^{m+2} a_k x^k (1-x)^{m+2-k} = 0$$

for all $x \in \mathbf{F}$. Taking $x = 0$ and $x = 1$ immediately gives us $a_0 = a_{m+2} = 0$, so that we now have the equation

$$\sum_{k=1}^{m+1} a_k x^k (1-x)^{m+2-k} = 0$$

for all $x \in \mathbf{F}$. Observe that

$$\begin{aligned}
\sum_{k=1}^{m+1} a_k x^k (1-x)^{m+2-k} &= x(1-x) \sum_{k=1}^{m+1} a_k x^{k-1} (1-x)^{m+1-k} \\
&= x(1-x) \sum_{k=0}^{m} a_{k+1} x^k (1-x)^{m-k} \\
&= x(1-x) \sum_{k=0}^{m} a_{k+1} p_{k,m}(x).
\end{aligned}$$

Thus $x(1-x) \sum_{k=0}^{m} a_{k+1} p_{k,m}(x) = 0$ for all $x \in \mathbf{F}$, which implies that $\sum_{k=0}^{m} a_{k+1} p_{k,m}(x) = 0$ for all $x \neq 0, 1$. Because the only polynomial with infinitely many roots is the zero polynomial (we will prove this in 4.8), in fact we must have

$$\sum_{k=0}^{m} a_{k+1} p_{k,m}(x) = 0$$

for all $x \in \mathbf{F}$. The induction hypothesis now implies that $a_1 = \cdots = a_m = 0$ and thus the list $p_0, ..., p_{m+2}$ is linearly independent, i.e. $S(m+2)$ holds. This completes the induction step.

We have now shown that $S(m) \Rightarrow S(m+2)$ for a positive integer m. Since $S(1)$ holds, an application of induction shows that $S(1), S(3), S(5), ...$ all hold. Similarly, since $S(2)$ holds, another application of induction shows that $S(2), S(4), S(6), ...$ all hold. Thus $S(m)$ holds for all positive integers m.

To complete the exercise, let m be a positive integer. As we just showed, $p_0, ..., p_m$ is a linearly independent list of length $m+1$ contained in $\mathcal{P}_m(\mathbf{F})$. Since $\dim \mathcal{P}_m(\mathbf{F}) = m+1$, 2.38 allows us to conclude that $p_0, ..., p_m$ is a basis of $\mathcal{P}_m(\mathbf{F})$.

Exercise 2.C.11. Suppose U and W are both four-dimensional subspaces of $\mathbf{C}^6$. Prove that there exist two vectors in $U \cap W$ such that neither of these vectors is a scalar multiple of the other.

Solution. Notice that

$$6 = \dim \mathbf{C}^6 \geq \dim(U + W) = \dim U + \dim W - \dim(U \cap W) = 8 - \dim(U \cap W),$$

where we have used 2.37 and 2.43. It follows that $\dim(U \cap W) \geq 2$ and thus we can find a linearly independent list v_1, v_2 in $U \cap W$; by Exercise 2.A.4 (b), neither of these vectors is a scalar multiple of the other.

Exercise 2.C.12. Suppose that U and W are subspaces of $\mathbf{R}^8$ such that $\dim U = 3$, $\dim W = 5$, and $U + W = \mathbf{R}^8$. Prove that $\mathbf{R}^8 = U \oplus W$.

Solution. By 2.43 we have

$$8 = \dim \mathbf{R}^8 = \dim(U + W) = \dim U + \dim W - \dim(U \cap W) = 8 - \dim(U \cap W).$$

It follows that $\dim(U \cap W) = 0$ and hence that $U \cap W = \{0\}$. Thus, by 1.46, the sum $\mathbf{R}^8 = U \oplus W$ is direct.

Exercise 2.C.13. Suppose U and W are both five-dimensional subspaces of $\mathbf{R}^9$. Prove that $U \cap W \neq \{0\}$.

Solution. By 2.43 we have

$$9 = \dim \mathbf{R}^9 \geq \dim(U + W) = \dim U + \dim W - \dim(U \cap W) = 10 - \dim(U \cap W).$$

It follows that $\dim(U \cap W) \geq 1$ and hence that $U \cap W \neq \{0\}$.

Exercise 2.C.14. Suppose V is a ten-dimensional vector space and V_1, V_2, V_3 are subspaces of V with $\dim V_1 = \dim V_2 = \dim V_3 = 7$. Prove that $V_1 \cap V_2 \cap V_3 \neq \{0\}$.

Solution. By 2.43 we have

$$\dim((V_1 \cap V_2) + V_3) = \dim(V_1 \cap V_2) + 7 - \dim(V_1 \cap V_2 \cap V_3), \tag{1}$$

$$\dim(V_1 + V_2) = 14 - \dim(V_1 \cap V_2). \tag{2}$$

Combining equations (1) and (2) gives us

$$\dim(V_1 \cap V_2 \cap V_3) = 21 - \dim(V_1 + V_2) - \dim((V_1 \cap V_2) + V_3).$$

Now we use the above equation and 2.37:

$$\dim(V_1 \cap V_2 \cap V_3) \geq 21 - 2 \dim V = 1.$$

Thus $V_1 \cap V_2 \cap V_3 \neq \{0\}$.

Exercise 2.C.15. Suppose V is finite-dimensional and V_1, V_2, V_3 are subspaces of V with $\dim V_1 + \dim V_2 + \dim V_3 > 2 \dim V$. Prove that $V_1 \cap V_2 \cap V_3 \neq \{0\}$.

Solution. By 2.43 we have

$$\dim((V_1 \cap V_2) + V_3) = \dim(V_1 \cap V_2) + \dim V_3 - \dim(V_1 \cap V_2 \cap V_3), \tag{1}$$

$$\dim(V_1 + V_2) = \dim V_1 + \dim V_2 - \dim(V_1 \cap V_2). \tag{2}$$

Combining equations (1) and (2) gives us

$$\dim(V_1 \cap V_2 \cap V_3) = \dim V_1 + \dim V_2 + \dim V_3 - \dim(V_1 + V_2) - \dim((V_1 \cap V_2) + V_3).$$

Now we use the above equation and 2.37:

$$\dim(V_1 \cap V_2 \cap V_3) \geq \dim V_1 + \dim V_2 + \dim V_3 - 2 \dim V > 0.$$

Thus $V_1 \cap V_2 \cap V_3 \neq \{0\}$.

Exercise 2.C.16. Suppose V is finite-dimensional and U is a subspace of V with $U \neq V$. Let $n = \dim V$ and $m = \dim U$. Prove that there exist $n - m$ subspaces of V, each of dimension $n - 1$, whose intersection equals U.

Solution. Let $u_1, ..., u_m$ be a basis of U and, using 2.32, extend this to a basis $B = u_1, ..., u_m, v_1, ..., v_k$ of V; note that $k \geq 1$ since $U \neq V$ and that $n - m = k$. For each $j \in \{1, ..., k\}$, let B_j be the list of vectors obtained by removing v_j from B, i.e.

$$B_j = u_1, ..., u_m, v_1, ..., v_{j-1}, v_{j+1}, ..., v_k.$$

Now let $U_j = \operatorname{span} B_j$. Observe that B_j is linearly independent since B is linearly independent and thus B_j is a basis of U_j, so that $\dim U_j = m + k - 1 = n - 1$. Furthermore, for $i \neq j$ we have $v_j \in U_i$ but $v_j \notin U_j$ by the linear independence of B; it follows that $U_i \neq U_j$. Thus the collection $\{U_j : 1 \leq j \leq k\}$ consists of $k = n - m$ distinct subspaces of V each of dimension $n - 1$.

We now show that $U = U_1 \cap \cdots \cap U_k$. If $k = 1$ then $U_1 = U$ and the equality is clear, so suppose that $k \geq 2$. Certainly $U \subseteq U_j$ for each $j \in \{1, ..., k\}$ and thus $U \subseteq U_1 \cap \cdots \cap U_k$. Let $u \in U_1 \cap \cdots \cap U_k$ be given. In particular $u \in U_1$, so there are scalars $a_1, ..., a_m, c_2, ..., c_k$ such that

$$u = a_1 u_1 + \cdots + a_m u_m + c_2 v_2 + \cdots + c_k v_k. \tag{$*$}$$

For each $j \in \{2, ..., k\}$ we have $u \in U_j$ and thus u can also be expressed as a linear combination of the list B_j. The coefficient of v_j in this linear combination is zero and, because B is a basis of V, it then follows from unique representation 2.28 that the coefficient c_j in the linear combination $(*)$ is also zero. Thus $u = a_1 u_1 + \cdots + a_m u_m \in U$, so that $U_1 \cap \cdots \cap U_m \subseteq U$. We may conclude that $U = U_1 \cap \cdots \cap U_k$.

Exercise 2.C.17. Suppose that $V_1, \ldots, V_m$ are finite-dimensional subspaces of V. Prove that $V_1 + \cdots + V_m$ is finite-dimensional and

$$\dim(V_1 + \cdots + V_m) \leq \dim V_1 + \cdots + \dim V_m.$$

The inequality above is an equality if and only if $V_1 + \cdots + V_m$ is a direct sum, as will be shown in 3.94.

Solution. Each V_j has a basis B_j by 2.31, so that $V_j = \operatorname{span} B_j$. Let B be the list $B_1, B_2, \ldots, B_m$ (removing duplicate vectors if necessary) and, letting $|B|$ denote the length of the list B, notice that $|B| \leq |B_1| + \cdots + |B_m|$. Notice further that $V_1 + \cdots + V_m = \operatorname{span} B$. It follows that $V_1 + \cdots + V_m$ is finite-dimensional and furthermore, by 2.22,

$$\dim(V_1 + \cdots + V_m) \leq |B| \leq |B_1| + \cdots + |B_m| = \dim V_1 + \cdots + \dim V_m.$$

Exercise 2.C.18. Suppose V is finite-dimensional, with $\dim V = n \geq 1$. Prove that there exist one-dimensional subspaces $V_1, \ldots, V_n$ of V such that

$$V = V_1 \oplus \cdots \oplus V_n.$$

Solution. V has a non-empty (since $n \geq 1$) basis $v_1, \ldots, v_n$. For each $j \in \{1, \ldots, n\}$ let $V_j = \operatorname{span}(v_j)$ and note that $\dim V_j = 1$ because $v_j \neq 0$. By 2.28 each vector in V is a linear combination of the basis vectors $v_1, \ldots, v_n$, so that $V = V_1 + \cdots + V_n$, and furthermore this linear combination is unique, so that the sum $V = V_1 \oplus \cdots + \oplus V_n$ is direct.

Exercise 2.C.19. Explain why you might guess, motivated by analogy with the formula for the number of elements in the union of three finite sets, that if V_1, V_2, V_3 are subspaces of a finite-dimensional vector space, then

$$\begin{aligned}
&\dim(V_1 + V_2 + V_3) \\
&\qquad = \dim V_1 + \dim V_2 + \dim V_3 \\
&\qquad\quad - \dim(V_1 \cap V_2) - \dim(V_1 \cap V_3) - \dim(V_2 \cap V_3) \\
&\qquad\quad + \dim(V_1 \cap V_2 \cap V_3).
\end{aligned}$$

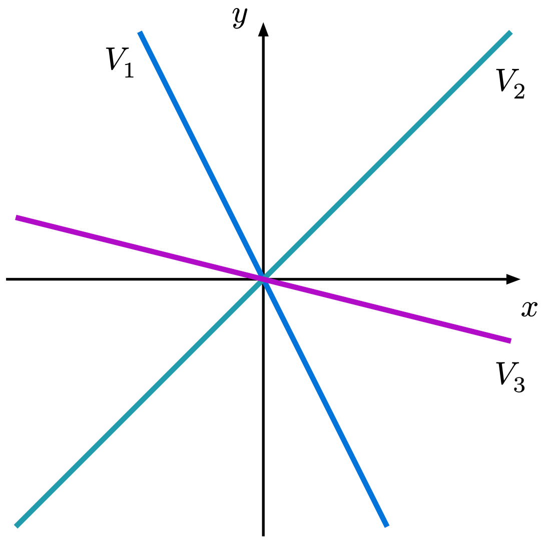
Then either prove the formula above or give a counterexample.

Solution. If S_1, S_2, S_3 are finite sets and $|S|$ denotes the number of elements in a finite set S, then the inclusion-exclusion principle gives us the formula

$$|S_1 \cup S_2 \cup S_3| = |S_1| + |S_2| + |S_3| - |S_1 \cap S_2| - |S_1 \cap S_3| - |S_2 \cap S_3| + |S_1 \cap S_2 \cap S_3|.$$

However, the analogous formula for the dimensions of finite-dimensional subspaces does not hold, as the following counterexample shows. Consider $\mathbf{R}^2$ and let V_1, V_2, V_3 be three distinct lines through the origin, say

$$V_1 = \operatorname{span}((-1,2)), \quad V_2 = \operatorname{span}((1,1)), \quad \text{and} \quad V_3 = \operatorname{span}((4,-1)).$$



It is straightforward to verify that $V_1 + V_2 + V_3 = \mathbf{R}^2$ and that

$$V_1 \cap V_2 = V_1 \cap V_3 = V_2 \cap V_3 = V_1 \cap V_2 \cap V_3 = \{0\}.$$

Thus $\dim(V_1 + V_2 + V_3) = \dim \mathbf{R}^2 = 2$, whereas the right-hand side of the proposed formula is

$$1 + 1 + 1 - 0 - 0 - 0 + 0 = 3.$$

Exercise 2.C.20. Prove that if V_1, V_2, and V_3 are subspaces of a finite-dimensional vector space, then

$$\begin{aligned}
&\dim(V_1 + V_2 + V_3) \\
&\quad = \dim V_1 + \dim V_2 + \dim V_3 \\
&\qquad - \frac{\dim(V_1 \cap V_2) + \dim(V_1 \cap V_3) + \dim(V_2 \cap V_3)}{3} \\
&\qquad - \frac{\dim((V_1 + V_2) \cap V_3) + \dim((V_1 + V_3) \cap V_2) + \dim((V_2 + V_3) \cap V_1)}{3}.
\end{aligned}$$

The formula above may seem strange because the right side does not look like an integer.

Solution. Using 2.43 twice, observe that

$$\begin{aligned}
\dim(V_1 + V_2 + V_3) &= \dim(V_1 + V_2) + \dim V_3 - \dim((V_1 + V_2) \cap V_3) \\
&= \dim V_1 + \dim V_2 + \dim V_3 - \dim(V_1 \cap V_2) - \dim((V_1 + V_2) \cap V_3).
\end{aligned}$$

Similarly, we find that

$$\dim(V_1 + V_2 + V_3) = \dim V_1 + \dim V_2 + \dim V_3 - \dim(V_1 \cap V_3) - \dim((V_1 + V_3) \cap V_2),$$

$$\dim(V_1 + V_2 + V_3) = \dim V_1 + \dim V_2 + \dim V_3 - \dim(V_2 \cap V_3) - \dim((V_2 + V_3) \cap V_1).$$

Adding these three formulas together and then dividing through by 3 gives us the desired formula.

Chapter 3. Linear Maps

3.A. Vector Space of Linear Maps

Exercise 3.A.1. Suppose $b, c \in \mathbf{R}$. Define $T : \mathbf{R}^3 \to \mathbf{R}^2$ by

$$T(x, y, z) = (2x - 4y + 3z + b, 6x + cxyz).$$

Show that T is linear if and only if $b = c = 0$.

Solution. First suppose that $b = c = 0$, so that T is the map

$$T(x, y, z) = (2x - 4y + 3z, 6x).$$

Let $(x_1, y_1, z_1), (x_2, y_2, z_2) \in \mathbf{R}^3$ and $\lambda \in \mathbf{R}$ be given. Observe that

$$\begin{aligned}
T(x_1 + x_2, y_1 + y_2, z_1 + z_2) &= (2(x_1 + x_2) - 4(y_1 + y_2) + 3(z_1 + z_2), 6(x_1 + x_2)) \\
&= (2x_1 + 2x_2 - 4y_1 - 4y_2 + 3z_1 + 3z_2, 6x_1 + 6x_2) \\
&= (2x_1 - 4y_1 + 3z_1, 6x_1) + (2x_2 - 4y_2 + 3z_2, 6x_2) \\
&= T(x_1, y_1, z_1) + T(x_2, y_2, z_2)
\end{aligned}$$

$$\begin{aligned}
T(\lambda x_1, \lambda y_1, \lambda z_1) &= (2\lambda x_1 - 4\lambda y_1 + 3\lambda z_1, 6\lambda x_1) \\
&= (\lambda(2x_1 - 4y_1 + 3z_1), \lambda(6x_1)) \\
&= \lambda(2x_1 - 4y_1 + 3z_1, 6x_1) \\
&= \lambda T(x_1, y_1, z_1).
\end{aligned}$$

Thus T is linear.

Now suppose that $b \neq 0$ and notice that $T(0, 0, 0) = (b, 0) \neq (0, 0)$; it follows from 3.10 that T is not linear. If $c \neq 0$ then note that

$$T(1, 1, 1) = (1 + b, 6 + c) \quad \text{and} \quad T(2, 2, 2) = (2 + b, 12 + 8c).$$

Since $2(6 + c) = 12 + 2c \neq 12 + 8c$ for $c \neq 0$, we see that $2T(1, 1, 1) \neq T(2, 2, 2)$ and thus T is not linear.

Exercise 3.A.2. Suppose $b, c \in \mathbf{R}$. Define $T : \mathcal{P}(\mathbf{R}) \to \mathbf{R}^2$ by

$$Tp = \left(3p(4) + 5p'(6) + bp(1)p(2), \int_{-1}^{2} x^3 p(x)\,\mathrm{d}x + c\sin p(0)\right).$$

Show that T is linear if and only if $b = c = 0$.

Solution. First suppose that $b = c = 0$, so that T is the map

$$Tp = \left(3p(4) + 5p'(6), \int_{-1}^{2} x^3 p(x)\,\mathrm{d}x\right).$$

Let $p, q \in \mathcal{P}(\mathbf{R})$ and $\lambda \in \mathbf{R}$ be given. Observe that

$$\begin{aligned}
T(p+q) &= \left(3(p+q)(4) + 5(p+q)'(6), \int_{-1}^{2} x^3 (p+q)(x)\,\mathrm{d}x\right) \\
&= \left(3(p(4)+q(4)) + 5(p'(6)+q'(6)), \int_{-1}^{2} x^3 (p(x)+q(x))\,\mathrm{d}x\right) \\
&= \left(3p(4) + 3q(4) + 5p'(6) + 5q'(6), \int_{-1}^{2} x^3 p(x)\,\mathrm{d}x + \int_{-1}^{2} x^3 q(x)\,\mathrm{d}x\right) \\
&= \left(3p(4) + 5p'(6), \int_{-1}^{2} x^3 p(x)\,\mathrm{d}x\right) + \left(3q(4) + 5q'(6), \int_{-1}^{2} x^3 q(x)\,\mathrm{d}x\right) \\
&= Tp + Tq.
\end{aligned}$$

$$\begin{aligned}
T(\lambda p) &= \left(3(\lambda p)(4) + 5(\lambda p)'(6), \int_{-1}^{2} x^3 (\lambda p)(x)\,\mathrm{d}x\right) \\
&= \left(3(\lambda p(4)) + 5(\lambda p'(6)), \int_{-1}^{2} x^3 (\lambda p(x))\,\mathrm{d}x\right) \\
&= \left(\lambda(3p(4) + 5p'(6)), \lambda\int_{-1}^{2} x^3 p(x)\,\mathrm{d}x\right) \\
&= \lambda\left(3p(4) + 5p'(6), \int_{-1}^{2} x^3 p(x)\,\mathrm{d}x\right) \\
&= \lambda Tp.
\end{aligned}$$

Thus T is linear.

Now suppose that T is linear and observe that

$$2T(\pi) = \left(6\pi + 2b\pi^2, \tfrac{15}{2}\pi + 2c\right) \quad \text{and} \quad T(2\pi) = \left(6\pi + 4b\pi^2, \tfrac{15}{2}\pi\right).$$

Since T is linear we must have $2T(\pi) = T(2\pi)$:

$$\left(6\pi + 2b\pi^2, \tfrac{15}{2}\pi + 2c\right) = \left(6\pi + 4b\pi^2, \tfrac{15}{2}\pi\right) \quad \Leftrightarrow \quad \left(2b\pi^2, 2c\right) = \left(4b\pi^2, 0\right) \quad \Leftrightarrow \quad b = c = 0.$$

Exercise 3.A.3. Suppose that $T \in \mathcal{L}(\mathbf{F}^n, \mathbf{F}^m)$. Show that there exist scalars $A_{j,k} \in \mathbf{F}$ for $j = 1, ..., m$ and $k = 1, ..., n$ such that

$$T(x_1, ..., x_n) = \left(A_{1,1}x_1 + \cdots + A_{1,n}x_n, ..., A_{m,1}x_1 + \cdots + A_{m,n}x_n\right)$$

for every $(x_1, ..., x_n) \in \mathbf{F}^n$.

This exercise shows that the linear map T has the form promised in the second to last item of Example 3.3.

Solution. Let $e_1, ..., e_n$ be the standard basis of $\mathbf{F}^n$ and let $f_1, ..., f_m$ be the standard basis of $\mathbf{F}^m$. For any $k \in \{1, ..., n\}$, there are scalars $A_{1,k}, ..., A_{m,k}$ such that

$$Te_k = \sum_{j=1}^{m} A_{j,k} f_j.$$

Let $x = (x_1, ..., x_n) = \sum_{k=1}^{n} x_k e_k$ be given and observe that by linearity,

$$\begin{aligned} Tx = T\left(\sum_{k=1}^{n} x_k e_k\right) &= \sum_{k=1}^{n} x_k T e_k = \sum_{k=1}^{n} x_k \sum_{j=1}^{m} A_{j,k} f_j = \sum_{j=1}^{m} \left(\sum_{k=1}^{n} A_{j,k} x_k\right) f_j \\ &= \left(\sum_{k=1}^{n} A_{1,k} x_k, ..., \sum_{k=1}^{n} A_{m,k} x_k\right) = \left(A_{1,1}x_1 + \cdots + A_{1,n}x_n, ..., A_{m,1}x_1 + \cdots + A_{m,n}x_n\right). \end{aligned}$$

Exercise 3.A.4. Suppose $T \in \mathcal{L}(V, W)$ and $v_1, ..., v_m$ is a list of vectors in V such that $Tv_1, ..., Tv_m$ is a linearly independent list in W. Prove that $v_1, ..., v_m$ is linearly independent.

Solution. Suppose we have scalars $a_1, ..., a_m$ such that $a_1 v_1 + \cdots + a_m v_m = 0$. Applying T to both sides of this equation and using linearity and 3.10, we obtain

$$T(a_1 v_1 + \cdots + a_m v_m) = T(0) \quad \Leftrightarrow \quad a_1 T v_1 + \cdots + a_m T v_m = 0.$$

Since the list $Tv_1, ..., Tv_m$ is linearly independent, this equation implies that $a_1 = \cdots = a_m = 0$. Thus the list $v_1, ..., v_m$ is linearly independent.

Exercise 3.A.5. Prove that $\mathcal{L}(V, W)$ is a vector space, as was asserted in 3.6.

Solution. First let us show that $\mathcal{L}(V, W)$ is closed under addition and scalar multiplication. Let $S, T \in \mathcal{L}(V, W)$, $u, v \in V$, and $\lambda, \alpha \in \mathbf{F}$ be given. Observe that

$$\begin{aligned} (S+T)(u+v) &= S(u+v) + T(u+v) = Su + Sv + Tu + Tv \\ &= Su + Tu + Sv + Tv = (S+T)(u) + (S+T)(v), \end{aligned}$$

$$(S+T)(\alpha v) = S(\alpha v) + T(\alpha v) = \alpha Sv + \alpha Tv = \alpha(Sv + Tv) + \alpha(S+T)(v).$$

Thus $S + T \in \mathcal{L}(V, W)$. Similarly,

$$(\lambda S)(u+v) = \lambda S(u+v) = \lambda(Su + Sv) = \lambda Su + \lambda Sv = (\lambda S)(u) + (\lambda S)(v),$$

$$(\lambda S)(\alpha v) = \lambda S(\alpha v) = \lambda(\alpha Sv) = \alpha(\lambda Sv) = \alpha(\lambda S)(v).$$

Thus $\lambda S \in \mathcal{L}(V, W)$. We now verify each requirement in definition 1.20.

Commutativity. Suppose $S, T \in \mathcal{L}(V, W)$ and $v \in V$. Observe that

$$(S+T)(v) = Sv + Tv = Tv + Sv = (T+S)(v).$$

Thus $S + T = T + S$.

Associativity. Suppose $R, S, T \in \mathcal{L}(V, W), a, b \in \mathbf{F}$, and $v \in V$. Observe that

$$\begin{aligned}((R+S)+T)(v) = (R+S)(v) = Tv = (Rv + Sv) = Tv \\ = Rv + (Sv + Tv) = Rv + (S+T)(v) = (R + (S+T))(v),\end{aligned}$$

$$((ab)R)(v) = (ab)Rv = a(bRv) = a((bR)(v)) = (a(bR))(v).$$

Thus $(R+S)+T = R+(S+T)$ and $(ab)R = a(bR)$.

Additive identity. Certainly the map $0 : V \to W$ given by $v \mapsto 0$ belongs to $\mathcal{L}(V, W)$; we claim that this map is the additive identity in $\mathcal{L}(V, W)$. Indeed, let $S \in \mathcal{L}(V, W)$ and $v \in V$ be given and observe that

$$(S+0)(v) = Sv + 0v = Sv + 0 = Sv.$$

Thus $S + 0 = S$.

Additive inverse. For $S \in \mathcal{L}(V, W)$, define $T : V \to W$ by $Tv = -Sv$; it is straightforward to verify that T is linear. We claim that T is the additive inverse to S. Indeed, for any $v \in V$,

$$(S+T)(v) = Sv + Tv = Sv + (-Sv) = 0.$$

Thus $S + T = 0$.

Multiplicative identity. Let $S \in \mathcal{L}(V, W)$ and $v \in V$ be given and observe that

$$(1S)(v) = 1Sv = Sv.$$

Thus $1S = S$.

Distributive properties. Let $S, T \in \mathcal{L}(V, W), a, b \in \mathbf{F}$, and $v \in V$ be given. Observe that

$$(a(S+T))(v) = a(S+T)(v) = a(Sv + Tv) = aSv + aTv = (aS)(v) + (aT)(v),$$

$$((a+b)S)(v) = (a+b)Sv = aSv + bSv = (aS)(v) + (bS)(v).$$

Thus $a(S+T) = aS + aT$ and $(a+b)S = aS + bS$.

Exercise 3.A.6. Prove that multiplication of linear maps has the associative, identity, and distributive properties asserted in 3.8.

Solution. The associative property is immediate from the associativity of composition of functions.

For the identity property, let I_V be the identity map on V and let I_W be the identity map on W. For $T \in \mathcal{L}(V, W)$ and $v \in V$, observe that

$$(TI_V)(v) = T(I_V v) = Tv \quad \text{and} \quad (I_W T)(v) = I_W(Tv) = Tv.$$

Thus $TI_V = I_W T = T$.

For the distributive properties, let $T, T_1, T_2 \in \mathcal{L}(U, V)$ and $S, S_1, S_2 \in \mathcal{L}(V, W)$ be given. For any $u \in U$, observe that

$$((S_1 + S_2)T)(u) = (S_1 + S_2)(Tu) = S_1(Tu) + S_2(Tu) = (S_1T)(u) = (S_2T)(u),$$

$$\begin{aligned}(S(T_1 + T_2))(u) = S((T_1 + T_2)(u)) = S(T_1u + T_2u)& \\ = S(T_1u) + S(T_2u) = (ST_1)(u) + (ST_2)(u).&\end{aligned}$$

Thus $(S_1 + S_2)T = S_1T + S_2T$ and $S(T_1 + T_2) = ST_1 + ST_2$.

Exercise 3.A.7. Show that every linear map from a one-dimensional vector space to itself is multiplication by some scalar. More precisely, prove that if $\dim V = 1$ and $T \in \mathcal{L}(V)$, then there exists $\lambda \in \mathbf{F}$ such that $Tv = \lambda v$ for all $v \in V$.

Solution. Since $\dim V = 1$ there is a basis u of V and thus $Tu = \lambda u$ for some $\lambda \in \mathbf{F}$. Let $v = \alpha u \in V$ be given and observe that

$$Tv = T(\alpha u) = \alpha Tu = \alpha(\lambda u) = \lambda(\alpha u) = \lambda v.$$

Exercise 3.A.8. Give an example of a function $\varphi : \mathbf{R}^2 \to \mathbf{R}$ such that

$$\varphi(av) = a\varphi(v)$$

for all $a \in \mathbf{R}$ and all $v \in \mathbf{R}^2$ but φ is not linear.

This exercise and the next exercise show that neither homogeneity nor additivity alone is enough to imply that a function is a linear map.

Solution. Let $\varphi : \mathbf{R}^2 \to \mathbf{R}$ be given by $\varphi(x, y) = \left(x^3 + y^3\right)^{1/3}$ and observe that for any $a \in \mathbf{R}$ and $(x, y) \in \mathbf{R}^2$,

$$\varphi(ax, ay) = \left((ax)^3 + (ay)^3\right)^{1/3} = \left(a^3\right)^{1/3}\left(x^3 + y^3\right)^{1/3} = a\left(x^3 + y^3\right)^{1/3} = a\varphi(x, y).$$

However, notice that

$$\varphi(1,0)+\varphi(0,1)=1+1=2\neq 2^{1/3}=\varphi(1,1).$$

Thus φ is not linear.

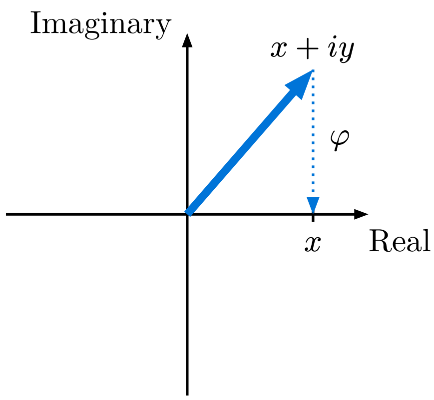
Exercise 3.A.9. Give an example of a function $\varphi:\mathbf{C}\to\mathbf{C}$ such that

$$\varphi(w+z)=\varphi(w)+\varphi(z)$$

for all $w,z\in\mathbf{C}$ but φ is not linear. (Here $\mathbf{C}$ is thought of as a complex vector space.)

There also exists a function $\varphi:\mathbf{R}\to\mathbf{R}$ such that φ satisfies the additivity condition above but φ is not linear. However, showing the existence of such a function involves considerably more advanced tools.

Solution. Let $\varphi:\mathbf{C}\to\mathbf{C}$ be given by $\varphi(x+iy)=x$, i.e. φ takes a complex number to its real part.



Observe that

$$\varphi((x+iy)+(u+iv))=\varphi((x+u)+i(y+v))=x+u=\varphi(x+iy)+\varphi(u+iv).$$

However, $\varphi(i)=0$ but $\varphi(i^2)=\varphi(-1)=-1\neq i\varphi(i)$. Thus φ is not linear.

Exercise 3.A.10. Prove or give a counterexample: If $q\in\mathcal{P}(\mathbf{R})$ and $T:\mathcal{P}(\mathbf{R})\to\mathcal{P}(\mathbf{R})$ is defined by $Tp=q\circ p$, then T is linear map.

The function T defined here differs from the function T defined in the last bullet point of 3.3 by the order of the functions in the compositions.

Solution. This does not necessarily define a linear map. For example, consider $q(x)=1$ and observe that $(T0)(x)=q(0(x))=q(0)=1\neq 0$. It follows from 3.10 that T is not linear.

Exercise 3.A.11. Suppose V is finite-dimensional and $T\in\mathcal{L}(V)$. Prove that T is a scalar multiple of the identity if and only if $ST=TS$ for every $S\in\mathcal{L}(V)$.

Solution. Suppose that T is a scalar multiple of the identity, say $T=\lambda I$ for some $\lambda\in\mathbf{F}$, and let $S\in\mathcal{L}(V)$ and $v\in V$ be given. Observe that

$$S(Tv) = S((\lambda I)(v)) = S(\lambda v) = \lambda(Sv) = (\lambda I)(Sv) = T(Sv).$$

Thus $ST = TS$.

Now suppose that $ST = TS$ for every $S \in \mathcal{L}(V)$. If $V = \{0\}$ then $T = 0I$ and we are done. Otherwise, let $v_1, ..., v_m$ be a basis of V and, using the linear map lemma (3.4), define a linear map $\varphi : V \to \mathbf{F}$ satisfying $\varphi(v_1) = \cdots = \varphi(v_m) = 1$. Let $\lambda = \varphi(Tv_1) \in \mathbf{F}$. For a fixed $v \in V$, define $S_v \in \mathcal{L}(V)$ by $S_v u = \varphi(u)v$; the linearity of S follows from the linearity of φ. Now observe that

$$Tv = T(\varphi(v_1)v) = T(S_v v_1) = S_v(Tv_1) = \varphi(Tv_1)v = \lambda v,$$

where we have used that T commutes with every $S \in \mathcal{L}(V)$ for the third equality. Because $v \in V$ was arbitrary, we may conclude that $T = \lambda I$.

Exercise 3.A.12. Suppose U is a subspace of V with $U \neq V$. Suppose $S \in \mathcal{L}(U, W)$ and $S \neq 0$ (which means that $Su \neq 0$ for some $u \in U$). Define $T : V \to W$ by

$$Tv = \begin{cases} Sv & \text{if } v \in U, \\ 0 & \text{if } v \in V \text{ and } v \notin U. \end{cases}$$

Prove that T is not a linear map on V.

Solution. There is some $u \in U$ such that $Su \neq 0$ and since $U \neq V$ there is some $v \in V$ such that $v \notin U$. This implies that $u - v \notin U$, otherwise $v = -(u - v) + u \in U$. Observe that

$$Tv + T(u - v) = 0 + 0 = 0 \neq Su = Tu = T(v + u - v).$$

Thus T is not linear.

Exercise 3.A.13. Suppose V is finite-dimensional. Prove that every linear map on a subspace of V can be extended to a linear map on V. In other words, show that if U is a subspace of V and $S \in \mathcal{L}(U, W)$, then there exists $T \in \mathcal{L}(V, W)$ such that $Tu = Su$ for all $u \in U$.

The result in this exercise is used in the proof of 3.125.

Solution. Let $u_1, ..., u_m$ be a basis of U and extend this to a basis $u_1, ..., u_m, v_1, ..., v_n$ of V. If we use the linear map lemma (3.4) to define $T \in \mathcal{L}(V, W)$ by $Tu_k = Su_k$ and $Tv_k = 0$, then for any $u = a_1u_1 + \cdots + a_mu_m \in U$ we have

$$Tu = a_1Tu_1 + \cdots + a_mTu_m = a_1Su_1 + \cdots + a_mSu_m = Su.$$

Thus T extends S.

Exercise 3.A.14. Suppose V is finite-dimensional with $\dim V > 0$, and suppose W is infinite-dimensional. Prove that $\mathcal{L}(V, W)$ is infinite-dimensional.

Solution. By Exercise 2.A.17 there is a sequence $w_1, w_2, w_3, \ldots$ in W such that $w_1, \ldots, w_m$ is linearly independent for each positive integer m. Let $v_1, \ldots, v_n$ be a basis of V and note that $n \geq 1$. For each positive integer k, use the linear map lemma (3.4) to define a linear map $T_k \in \mathcal{L}(V, W)$ satisfying

$$T_k v_j = \begin{cases} w_k & \text{if } j = 1, \\ 0 & \text{otherwise.} \end{cases}$$

Let m be a positive integer and suppose we have scalars $a_1, \ldots, a_m$ such that $a_1 T_1 + \cdots + a_m T_m = 0$. In particular,

$$0 = (a_1 T_1 + \cdots + a_m T_m)(v_1) = a_1 T_1 v_1 + \cdots + a_m T_m v_1 = a_1 w_1 + \cdots + a_m w_m.$$

This implies that $a_1 = \cdots = a_m = 0$ since the list $w_1, \ldots, w_m$ is linearly independent and it follows that the list $T_1, \ldots, T_m$ is linearly independent. We may use Exercise 2.A.17 to conclude that $\mathcal{L}(V, W)$ is infinite-dimensional.

Exercise 3.A.15. Suppose $v_1, \ldots, v_m$ is a linearly dependent list of vectors in V. Suppose also that $W \neq \{0\}$. Prove that there exist $w_1, \ldots, w_m \in W$ such that no $T \in \mathcal{L}(V, W)$ satisfies $T v_k = w_k$ for each $k = 1, \ldots, m$.

Solution. (We will use complex conjugation for this solution; complex conjugation is defined and its properties are studied in Chapter 4 of the textbook.)

We will prove the contrapositive statement. That is, assuming that for all lists $w_1, \ldots, w_m \in W$ there is a linear map $T \in \mathcal{L}(V, W)$ such that $T v_k = w_k$ for each $k \in \{1, \ldots, m\}$, we will prove that the list $v_1, \ldots, v_m$ is linearly independent. Indeed, suppose we have scalars $a_1, \ldots, a_m$ such that $a_1 v_1 + \cdots + a_m v_m = 0$. There is some non-zero $w \in W$; by assumption there is a linear map $T \in \mathcal{L}(V, W)$ such that $T v_k = \overline{a_k} w$ for each $k \in \{1, \ldots, m\}$. It follows that

$$0 = T\left(\sum_{k=1}^{m} a_k v_k\right) = \sum_{k=1}^{m} a_k T v_k = \sum_{k=1}^{m} a_k \overline{a_k} w = \left(\sum_{k=1}^{m} |a_k|^2\right) w.$$

Since $w \neq 0$, Exercise 1.B.2 implies that $\sum_{k=1}^{m} |a_k|^2 = 0$, which is the case if and only if $a_k = 0$ for all $k \in \{1, \ldots, m\}$. Thus the list $v_1, \ldots, v_m$ is linearly independent.

Exercise 3.A.16. Suppose V is finite-dimensional with $\dim V > 1$. Prove that there exist $S, T \in \mathcal{L}(V)$ such that $ST \neq TS$.

Solution. There is a basis $v_1, v_2, \ldots, v_n$ for V with $n \geq 2$. Using the linear map lemma (3.4), define linear maps $S, T \in \mathcal{L}(V)$ satisfying

$$S v_k = \begin{cases} v_2 & \text{if } k = 1, \\ v_1 & \text{if } k = 2, \\ 0 & \text{otherwise,} \end{cases} \qquad T v_k = \begin{cases} 2 v_2 & \text{if } k = 1, \\ v_1 & \text{if } k = 2, \\ 0 & \text{otherwise.} \end{cases}$$

Observe that

$$(ST - TS)(v_1) = S(Tv_1) - T(Sv_1) = S(2v_2) - Tv_2 = 2v_1 - v_1 = v_1 \neq 0.$$

Thus $ST \neq TS$.

Exercise 3.A.17. Suppose V is finite-dimensional. Show that the only two-sided ideals of $\mathcal{L}(V)$ are $\{0\}$ and $\mathcal{L}(V)$.

*A subspace $\mathcal{E}$ of $\mathcal{L}(V)$ is called a **two-sided ideal** of $\mathcal{L}(V)$ if $TE \in \mathcal{E}$ and $ET \in \mathcal{E}$ for all $E \in \mathcal{E}$ and all $T \in \mathcal{L}(V)$.*

Solution. Certainly $\{0\}$ is a two-sided ideal of $\mathcal{L}(V)$, so suppose that $\mathcal{E}$ is a two-sided ideal of $\mathcal{L}(V)$ containing some non-zero linear map $T \in \mathcal{L}(V)$; we must show that $\mathcal{E} = \mathcal{L}(V)$.

Let $v_1, ..., v_m$ be a basis of V. For each pair (i, j), use the linear map lemma (3.4) to define a linear map $E_{i,j} \in \mathcal{L}(V)$ which sends v_i to v_j and each other basis vector to 0. Since $T \neq 0$, there must be some $\ell \in \{1, ..., m\}$ such that

$$Tv_\ell = a_1 v_1 + \cdots + a_m v_m \neq 0;$$

there must then be some non-zero coefficient in this linear combination, say a_n. For each $i \in \{1, ..., m\}$, observe that

$$E_{n,i} T E_{i,\ell} v_i = E_{n,i} T v_\ell = E_{n,i}(a_1 v_1 + \cdots + a_m v_m) = a_n v_i,$$

$$\text{and for } k \neq i, \qquad E_{n,i} T E_{i,\ell} v_k = E_{n,i} T(0) = 0.$$

Thus, letting L_i be the linear map $E_{n,i} T E_{i,\ell}$, we have $L_i v_i = a_n v_i$ and $L_i v_k = 0$ for $k \neq i$; it follows that $L_1 + \cdots + L_m = a_n I$, where I is the identity map on V. Now observe that

$$\begin{aligned} T \in \mathcal{E} \quad \Rightarrow \quad & E_{n,i} T E_{i,\ell} = L_i \in \mathcal{E} \text{ for each } i \in \{1, ..., m\} \\ & \Rightarrow \quad L_1 + \cdots + L_m = a_n I \in \mathcal{E} \quad \Rightarrow \quad a_n^{-1} a_n I = I \in \mathcal{E}; \end{aligned}$$

each of these implications is justified because $\mathcal{E}$ is a two-sided ideal of $\mathcal{L}(V)$. It follows that $SI = S \in \mathcal{E}$ for any $S \in \mathcal{L}(V)$ and we may conclude that $\mathcal{E} = \mathcal{L}(V)$.

3.B. Null Spaces and Ranges

Exercise 3.B.1. Give an example of a linear map T with $\dim \operatorname{null} T = 3$ and $\dim \operatorname{range} T = 2$.

Solution. Let $T : \mathbf{R}^5 \to \mathbf{R}^2$ be given by

$$T(x_1, x_2, x_3, x_4, x_5) = (x_1, x_2)$$

and observe that

$$\operatorname{null} T = \{(0, 0, x_3, x_4, x_5) \in \mathbf{R}^5 : x_3, x_4, x_5 \in \mathbf{R}\} \quad \text{and} \quad \operatorname{range} T = \mathbf{R}^2.$$

Thus $\dim \operatorname{null} T = 3$ and $\dim \operatorname{range} T = 2$.

Exercise 3.B.2. Suppose $S, T \in \mathcal{L}(V)$ are such that $\operatorname{range} S \subseteq \operatorname{null} T$. Prove that $(ST)^2 = 0$.

Solution. For any $v \in V$ we have $S(Tv) \in \operatorname{range} S \subseteq \operatorname{null} T$, so that $T(S(Tv)) = 0$. It follows that

$$(ST)^2(v) = S(T(S(Tv))) = S(0) = 0.$$

Thus $(ST)^2 = 0$.

Exercise 3.B.3. Suppose $v_1, \dots, v_m$ is a list of vectors in V. Define $T \in \mathcal{L}(\mathbf{F}^m, V)$ by

$$T(z_1, \dots, z_m) = z_1 v_1 + \cdots + z_m v_m.$$

(a) What property of T corresponds to $v_1, \dots, v_m$ spanning V?

(b) What property of T corresponds to the list $v_1, \dots, v_m$ being linearly independent?

Solution.

(a) The surjectivity of T corresponds to $v_1, \dots, v_m$ spanning V. Indeed, observe that T is surjective if and only if for every $v \in V$ there exists $(z_1, \dots, z_m) \in \mathbf{F}^m$ such that $T(z_1, \dots, z_m) = z_1 v_1 + \cdots + z_m v_m = v$, which is the case if and only if $V = \operatorname{span}(v_1, \dots, v_m)$.

(b) The injectivity of T corresponds to $v_1, \dots, v_m$ being linearly independent. By 3.15, T is injective if and only if $\operatorname{null} T = \{0\}$, i.e. if and only if the only choice of $(z_1, \dots, z_m) \in \mathbf{F}^m$ which gives $z_1 v_1 + \cdots + z_m v_m = 0$ is $(0, \dots, 0)$; this is the definition of linear independence.

Exercise 3.B.4. Show that $\{T \in \mathcal{L}(\mathbf{R}^5, \mathbf{R}^4) : \dim \operatorname{null} T > 2\}$ is not a subspace of $\mathcal{L}(\mathbf{R}^5, \mathbf{R}^4)$.

Solution. Let $W = \{T \in \mathcal{L}(\mathbf{R}^5, \mathbf{R}^4) : \dim \operatorname{null} T > 2\}$. Define $S, T \in \mathcal{L}(\mathbf{R}^5, \mathbf{R}^4)$ by

$$S(x_1, x_2, x_3, x_4, x_5) = (x_1, x_2, 0, 0) \quad \text{and} \quad T(x_1, x_2, x_3, x_4, x_5) = (0, 0, x_3, x_4)$$

and observe that

$$\begin{aligned} &\operatorname{null} S = \{(0, 0, x_3, x_4, x_5) \in \mathbf{R}^5 : x_3, x_4, x_5 \in \mathbf{R}\} \\ &\qquad\qquad \text{and} \quad \operatorname{null} T = \{(x_1, x_2, 0, 0, x_5) \in \mathbf{R}^5 : x_1, x_2, x_5 \in \mathbf{R}\}, \end{aligned}$$

so that $\dim \operatorname{null} S = \dim \operatorname{null} T = 3$. Thus $S, T \in W$. However, note that

$$\begin{aligned} &(S + T)(x_1, x_2, x_3, x_4, x_5) = (x_1, x_2, x_3, x_4) \\ &\qquad \Rightarrow \quad \operatorname{null}(S + T) = \{(0, 0, 0, 0, x_5) \in \mathbf{R}^5 : x_5 \in \mathbf{R}\} \quad \Rightarrow \quad \dim \operatorname{null}(S + T) = 1. \end{aligned}$$

Thus $S + T \notin W$. It follows that W is not closed under addition and hence that W is not a subspace of $\mathcal{L}(\mathbf{R}^5, \mathbf{R}^4)$.

Exercise 3.B.5. Give an example of $T \in \mathcal{L}(\mathbf{R}^4)$ such that $\operatorname{range} T = \operatorname{null} T$.

Solution. Let $T \in \mathcal{L}(\mathbf{R}^4)$ be given by

$$T(x_1, x_2, x_3, x_4) = (x_3, x_4, 0, 0),$$

which satisfies $\operatorname{range} T = \operatorname{null} T = \{(a, b, 0, 0) \in \mathbf{R}^4 : a, b \in \mathbf{R}\}$.

Exercise 3.B.6. Prove that there does not exist $T \in \mathcal{L}(\mathbf{R}^5)$ such that $\operatorname{range} T = \operatorname{null} T$.

Solution. If $T \in \mathcal{L}(V)$ for some finite-dimensional vector space V and $\operatorname{range} T = \operatorname{null} T$, then the fundamental theorem of linear maps (3.21) implies that

$$\dim V = \dim \operatorname{null} T + \dim \operatorname{range} T = 2 \dim \operatorname{null} T.$$

Thus the dimension of V must be a non-negative even integer, which 5 is not.

Exercise 3.B.7. Suppose V and W are finite-dimensional with $2 \leq \dim V \leq \dim W$. Show that $\{T \in \mathcal{L}(V, W) : T \text{ is not injective}\}$ is not a subspace of $\mathcal{L}(V, W)$.

Solution. Let $X = \{T \in \mathcal{L}(V, W) : T \text{ is not injective}\}$ and note that by 3.15 we have

$$X = \{T \in \mathcal{L}(V, W) : \operatorname{null} T \neq \{0\}\}.$$

Let $v_1, \ldots, v_m$ be a basis of V and let $w_1, \ldots, w_n$ be a basis of W; by assumption we have $2 \leq m \leq n$. Define $S, T \in \mathcal{L}(V, W)$ by

$$Sv_k = \begin{cases} 0 & \text{if } k = 1, \\ w_2 & \text{if } k = 2, \\ \frac{1}{2}w_k & \text{otherwise,} \end{cases} \qquad Tv_k = \begin{cases} w_1 & \text{if } k = 1, \\ 0 & \text{if } k = 2, \\ \frac{1}{2}w_k & \text{otherwise.} \end{cases}$$

Notice that $S, T \in X$ since v_1, v_2 are non-zero, $v_1 \in \operatorname{null} S$, and $v_2 \in \operatorname{null} T$. Notice further that

$$(S+T)(v_k) = w_k \text{ for all } k \in \{1, ..., m\} \quad \Rightarrow \quad \operatorname{range}(S+T) = \operatorname{span}(w_1, ..., w_m).$$

The linear independence of the list $w_1, ..., w_n$ then implies that $w_1, ..., w_m$ is a basis of $\operatorname{range}(S+T)$ and thus $\dim \operatorname{range}(S+T) = \dim \operatorname{span}(w_1, ..., w_m) = m$. Because $\dim V = m$, it follows from the fundamental theorem of linear maps (3.21) that $\dim \operatorname{null}(S+T) = 0$, hence $\operatorname{null}(S+T) = \{0\}$, hence $S+T \notin X$. We may conclude that X is not closed under addition and hence is not a subspace of $\mathcal{L}(V, W)$.

Exercise 3.B.8. Suppose V and W are finite-dimensional with $\dim V \geq \dim W \geq 2$. Show that $\{T \in \mathcal{L}(V, W) : T \text{ is not surjective}\}$ is not a subspace of $\mathcal{L}(V, W)$.

Solution. Let $X = \{T \in \mathcal{L}(V, W) : T \text{ is not surjective}\}$, let $v_1, ..., v_m$ be a basis of V, and let $w_1, ..., w_n$ be a basis of W; by assumption we have $m \geq n \geq 2$. Define $S, T \in \mathcal{L}(V, W)$ by

$$Sv_k = \begin{cases} 0 & \text{if } k = 1 \text{ or } k > n, \\ w_2 & \text{if } k = 2, \\ \frac{1}{2}w_k & \text{otherwise,} \end{cases} \qquad Tv_k = \begin{cases} w_1 & \text{if } k = 1, \\ 0 & \text{if } k = 2 \text{ or } k > n, \\ \frac{1}{2}w_k & \text{otherwise.} \end{cases}$$

It is straightforward to verify that $\operatorname{range} S = \operatorname{span}(w_2, ..., w_n)$. The linear independence of the list $w_1, ..., w_n$ then implies that $w_1 \notin \operatorname{range} S$ and thus S is not surjective, i.e. $S \in X$. Similarly, we find that $w_2 \notin \operatorname{range} T$ and thus $T \in X$. Now observe that

$$(S+T)(v_k) = \begin{cases} w_k & \text{if } k \leq n, \\ 0 & \text{if } k > n, \end{cases} \qquad \Rightarrow \quad \operatorname{range}(S+T) = \operatorname{span}(w_1, ..., w_n) = W.$$

Thus $S+T$ is surjective, i.e. $S+T \notin X$. It follows that X is not closed under addition and hence is not a subspace of $\mathcal{L}(V, W)$.

Exercise 3.B.9. Suppose $T \in \mathcal{L}(V, W)$ is injective and $v_1, ..., v_n$ is linearly independent in V. Prove that $Tv_1, ..., Tv_n$ is linearly independent in W.

Solution. Suppose we have scalars $a_1, ..., a_n$ such that $a_1Tv_1 + \cdots + a_nTv_n = 0$. By linearity, this is equivalent to $T(a_1v_1 + \cdots + a_nv_n) = 0$, i.e. $a_1v_1 + \cdots + a_nv_n \in \operatorname{null} T$. Because T is injective, it follows from 3.15 that $a_1v_1 + \cdots + a_nv_n = 0$. The linear independence of $v_1, ..., v_n$ then implies that $a_1 = \cdots = a_n = 0$. Thus $Tv_1, ..., Tv_n$ is linearly independent.

Exercise 3.B.10. Suppose $v_1, ..., v_n$ spans V and $T \in \mathcal{L}(V, W)$. Show that $Tv_1, ..., Tv_n$ spans $\operatorname{range} T$.

Solution. Let $w \in \operatorname{range} T$ be given, so that $w = Tv$ for some $v \in V$. Since $v_1, ..., v_n$ spans V, there are scalars $a_1, ..., a_n$ such that $v = a_1 + \cdots + a_n v_n$. It follows from the linearity of T that

$$w = Tv = T(a_1 v_1 + \cdots + a_n v_n) = a_1 Tv_1 + \cdots + a_n Tv_n \in \operatorname{span}(Tv_1, ..., Tv_n).$$

Thus $\operatorname{range} T \subseteq \operatorname{span}(Tv_1, ..., Tv_n)$. Now let $a_1 Tv_1 + \cdots + a_n Tv_n \in \operatorname{span}(Tv_1, ..., Tv_n)$ be given. The linearity of T gives us

$$a_1 Tv_1 + \cdots + a_n Tv_n = T(a_1 v_1 + \cdots + a_n v_n) \in \operatorname{range} T.$$

Thus $\operatorname{span}(Tv_1, ..., Tv_n) \subseteq \operatorname{range} T$ and we may conclude that $\operatorname{range} T = \operatorname{span}(Tv_1, ..., Tv_n)$.

Exercise 3.B.11. Suppose that V is finite-dimensional and that $T \in \mathcal{L}(V, W)$. Prove that there exists a subspace U of V such that

$$U \cap \operatorname{null} T = \{0\} \quad \text{and} \quad \operatorname{range} T = \{Tu : u \in U\}.$$

Solution. By 2.33 there is a subspace U of V such that $V = U \oplus \operatorname{null} T$ and 1.46 then gives us $U \cap \operatorname{null} T = \{0\}$. Suppose that $w \in \operatorname{range} T$, so that $w = Tv$ for some $v \in V$. Since $V = U \oplus \operatorname{null} T$, there are vectors $u \in U$ and $x \in \operatorname{null} T$ such that $v = u + x$ and it follows that

$$w = Tv = T(u + x) = Tu + Tx = Tu.$$

Thus $\operatorname{range} T \subseteq \{Tu : u \in U\}$; the reverse inclusion $\{Tu : u \in U\} \subseteq \operatorname{range} T$ is clear.

Exercise 3.B.12. Suppose T is a linear map from $\mathbf{F}^4$ to $\mathbf{F}^2$ such that

$$\operatorname{null} T = \{(x_1, x_2, x_3, x_4) \in \mathbf{F}^4 : x_1 = 5x_2 \text{ and } x_3 = 7x_4\}.$$

Prove that T is surjective.

Solution. It is straightforward to verify that $(5, 1, 0, 0), (0, 0, 7, 1)$ is a basis of $\operatorname{null} T$, so that $\dim \operatorname{null} T = 2$. The fundamental theorem of linear maps (3.21) then gives

$$\dim \operatorname{range} T = \dim \mathbf{F}^4 - \dim \operatorname{null} T = 2.$$

Thus $\dim \operatorname{range} T = 2 = \dim \mathbf{F}^2$ and it follows from 2.39 that $\operatorname{range} T = \mathbf{F}^2$, i.e. T is surjective.

Exercise 3.B.13. Suppose U is a three-dimensional subspace of $\mathbf{R}^8$ and that T is a linear map from $\mathbf{R}^8$ to $\mathbf{R}^5$ such that $\operatorname{null} T = U$. Prove that T is surjective.

Solution. The fundamental theorem of linear maps (3.21) gives us

$$\dim \operatorname{range} T = \dim \mathbf{R}^8 - \dim \operatorname{null} T = 8 - \dim U = 5.$$

Thus $\dim \operatorname{range} T = 5 = \dim \mathbf{R}^5$ and it follows from (2.39) that $\operatorname{range} T = \mathbf{R}^5$, i.e. T is surjective.

Exercise 3.B.14. Prove that there does not exist a linear map from $\mathbf{F}^5$ to $\mathbf{F}^2$ whose null space equals $\{(x_1, x_2, x_3, x_4, x_5) \in \mathbf{F}^5 : x_1 = 3x_2 \text{ and } x_3 = x_4 = x_5\}$.

Solution. Let $U = \{(x_1, x_2, x_3, x_4, x_5) \in \mathbf{F}^5 : x_1 = 3x_2 \text{ and } x_3 = x_4 = x_5\}$. It is straightforward to verify that $(3,1,0,0,0), (0,0,1,1,1)$ is a basis of U, so that $\dim U = 2$. For any $T \in \mathcal{L}(\mathbf{F}^5, \mathbf{F}^2)$, the fundamental theorem of linear maps (3.21) implies that

$$\dim \operatorname{null} T = \dim \mathbf{F}^5 - \dim \operatorname{range} T \geq 5 - \dim \mathbf{F}^2 = 3.$$

It follows that U cannot be the null space of T.

Exercise 3.B.15. Suppose there exists a linear map on V whose null space and range are both finite-dimensional. Prove that V is finite-dimensional.

Solution. Let T be the linear map in question, let $Tv_1, \dots, Tv_m$ be a basis of $\operatorname{range} T$, and let $w_1, \dots, w_n$ be a basis of $\operatorname{null} T$. For any $v \in V$ there are scalars $a_1, \dots, a_m$ such that

$$\begin{aligned} Tv &= a_1 Tv_1 + \cdots + a_m Tv_m = T(a_1 v_1 + \cdots + a_m v_m) \\ &\Rightarrow \quad T(v - (a_1 v_1 + \cdots + a_m v_m)) = 0 \quad \Rightarrow \quad v - (a_1 v_1 + \cdots + a_m v_m) \in \operatorname{null} T. \end{aligned}$$

It follows that there are scalars $b_1, \dots, b_n$ such that

$$\begin{aligned} v - (a_1 v_1 + \cdots + a_m v_m) &= b_1 w_1 + \cdots + b_n w_n \\ &\Rightarrow \quad v = a_1 v_1 + \cdots + a_m v_m + b_1 w_1 + \cdots + b_n w_n. \end{aligned}$$

Thus the list $v_1, \dots, v_m, w_1, \dots, w_n$ spans V. We may conclude that V is finite-dimensional.

Exercise 3.B.16. Suppose V and W are finite-dimensional. Prove that there exists an injective linear map from V to W if and only if $\dim V \leq \dim W$.

Solution. If $\dim V > \dim W$ then 3.22 guarantees that no linear map from V to W is injective. Suppose therefore that $\dim V = m \leq n = \dim W$, let $v_1, \dots, v_m$ be a basis of V, and let $w_1, \dots, w_n$ be a basis of W. Define $T \in \mathcal{L}(V, W)$ by $Tv_k = w_k$. The linear independence of the list $w_1, \dots, w_m$ and Exercise 3.B.10 imply that $w_1, \dots, w_m$ is a basis of $\operatorname{range} T$, so that $\dim \operatorname{range} T = m = \dim V$. It then follows from the fundamental theorem of linear maps (3.21) that $\dim \operatorname{null} T = 0$, i.e. T is injective.

Exercise 3.B.17. Suppose V and W are both finite-dimensional. Prove that there exists a surjective linear map from V onto W if and only if $\dim V \geq \dim W$.

Solution. If $\dim V < \dim W$ then 3.24 guarantees that no linear map from V to W is surjective. Suppose therefore that $\dim V = m \geq n = \dim W$, let $v_1, ..., v_m$ be a basis of V, and let $w_1, ..., w_n$ be a basis of W. Define $T \in \mathcal{L}(V, W)$ by

$$Tv_k = \begin{cases} w_k & \text{if } k \leq n, \\ 0 & \text{otherwise.} \end{cases}$$

It follows from Exercise 3.B.10 that

$$\operatorname{range} T = \operatorname{span}(Tv_1, ..., Tv_m) = \operatorname{span}(w_1, ..., w_n) = W.$$

Thus T is surjective.

Exercise 3.B.18. Suppose V and W are finite-dimensional and that U is a subspace of V. Prove that there exists $T \in \mathcal{L}(V, W)$ such that $\operatorname{null} T = U$ if and only if $\dim U \geq \dim V - \dim W$.

Solution. If there exists such a T then the fundamental theorem of linear maps (3.21) and 2.37 give us

$$\dim U = \dim \operatorname{null} T = \dim V - \dim \operatorname{range} T \geq \dim V - \dim W.$$

Conversely, suppose that $\dim U \geq \dim V - \dim W$. Let $u_1, ..., u_m$ be a basis of U, extend this to a basis $u_1, ..., u_m, v_1, ..., v_n$ of V, and let $X = \operatorname{span}(v_1, ..., v_n)$, so that $V = U \oplus X$ and $\dim V = \dim U + \dim X$. Combining this with our hypothesis $\dim U \geq \dim V - \dim W$ gives us $\dim X \leq \dim W$ and so we may invoke Exercise 3.B.16 to obtain an injective linear map $S : X \to W$. Extend this to a linear map $T : V \to W$ satisfying $T(u + x) = Sx$ as in Exercise 3.A.13 (X is playing the role of U here). Now observe that

$$T(u + x) = 0 \quad \Leftrightarrow \quad Sx = 0 \quad \Leftrightarrow \quad x = 0 \quad \Leftrightarrow \quad u + x \in U,$$

where we have used the injectivity of S for the second equivalence. It follows that $\operatorname{null} T = U$.

Exercise 3.B.19. Suppose W is finite-dimensional and $T \in \mathcal{L}(V, W)$. Prove that T is injective if and only if there exists $S \in \mathcal{L}(W, V)$ such that ST is the identity operator on V.

Solution. Suppose there exists such an S and let $v \in \operatorname{null} T$ be given. Observe that

$$v = (ST)(v) = S(0) = 0.$$

Thus $\operatorname{null} T = \{0\}$, which implies that T is injective (by 3.15).

Now suppose that T is injective. Let $Tv_1, ..., Tv_m$ be a basis of range T and extend this to a basis $Tv_1, ..., Tv_m, w_1, ..., w_n$ of W. Define $S \in \mathcal{L}(W, V)$ by $S(Tv_k) = v_k$ and $Sw_k = 0$. For any $v \in V$ we have $Tv = a_1Tv_1 + \cdots + a_mTv_m$ for some scalars $a_1, ..., a_m$, which implies

$$T(v - (a_1v_1 + \cdots + a_mv_m)) = 0.$$

Because T is injective it must then be the case that $v = a_1v_1 + \cdots + a_mv_m$. Now observe that

$$S(Tv) = a_1S(Tv_1) + \cdots + a_mS(Tv_m) = a_1v_1 + \cdots + a_mv_m = v.$$

Thus ST is the identity map on V.

Exercise 3.B.20. Suppose W is finite-dimensional and $T \in \mathcal{L}(V, W)$. Prove that T is surjective if and only if there exists $S \in \mathcal{L}(W, V)$ such that TS is the identity operator on W.

Solution. Suppose there exists such a map S and let $w \in W$ be given. Observe that

$$w = T(Sw) \in \operatorname{range} T.$$

It follows that $W = \operatorname{range} T$, i.e. T is surjective.

Now suppose that T is surjective, i.e. $W = \operatorname{range} T$. Let $Tv_1, ..., Tv_m$ be a basis of $\operatorname{range} T = W$ and define $S \in \mathcal{L}(W, V)$ by $S(Tv_k) = v_k$ for $k \in \{1, ..., m\}$. For any $w = a_1Tv_1 + \cdots + a_mTv_m \in W$, observe that

$$\begin{aligned} Sw = a_1S(Tv_1) + \cdots + a_mS(Tv_m) = a_1v_1 + \cdots + a_mv_m& \\ \Rightarrow \quad T(Sw) = a_1Tv_1 + \cdots + a_mTv_m = w.& \end{aligned}$$

Thus TS is the identity map on W.

Exercise 3.B.21. Suppose V is finite-dimensional, $T \in \mathcal{L}(V, W)$, and U is a subspace of W. Prove that $\{v \in V : Tv \in U\}$ is a subspace of V and

$$\dim\{v \in V : Tv \in U\} = \dim \operatorname{null} T + \dim(U \cap \operatorname{range} T).$$

Solution. Let $X = \{v \in V : Tv \in U\}$ and note that $0 \in X$ since $T(0) = 0 \in U$. Suppose $v_1, v_2 \in X$, so that $Tv_1, Tv_2 \in U$, and $\lambda \in \mathbf{F}$. Because T is linear and U is a subspace, we then have

$$Tv_1, Tv_2 \in U \quad \Rightarrow \quad T(v_1) + T(v_2) = T(v_1 + v_2) \in U \quad \text{and} \quad \lambda Tv_1 = T(\lambda v_1) \in U.$$

Thus $v_1 + v_2$ and λv_1 belong to X. It follows from 1.34 that X is a subspace of V.

Let S be the restriction of T to X, i.e. $S : X \to W$ is given by $Sv = Tv$, and notice that S is linear because T is linear. Notice further that

$$v \in \operatorname{null} T \quad \Rightarrow \quad Tv = 0 \in U \quad \Rightarrow \quad v \in X \quad \Rightarrow \quad Sv = Tv = 0 \quad \Rightarrow \quad v \in \operatorname{null} S,$$

$$v \in \operatorname{null} S \quad \Rightarrow \quad Sv = Tv = 0 \quad \Rightarrow \quad v \in \operatorname{null} T.$$

Thus $\operatorname{null} S = \operatorname{null} T$. Similarly,

$$Tv \in U \cap \operatorname{range} T \text{ for some } v \in V \quad \Rightarrow \quad v \in X \quad \Rightarrow \quad Tv = Sv \in \operatorname{range} S,$$

$$Sv \in \operatorname{range} S \text{ for some } v \in X \quad \Rightarrow \quad Sv = Tv \in U \cap \operatorname{range} T.$$

Thus $\operatorname{range} S = U \cap \operatorname{range} T$. The fundamental theorem of linear maps (3.21) now implies that

$$\dim X = \dim \operatorname{null} S + \dim \operatorname{range} S = \dim \operatorname{null} T + \dim(U \cap \operatorname{range} T).$$

Exercise 3.B.22. Suppose U and V are finite-dimensional vector spaces and $S \in \mathcal{L}(V, W)$ and $T \in \mathcal{L}(U, V)$. Prove that

$$\dim \operatorname{null} ST \leq \dim \operatorname{null} S + \dim \operatorname{null} T.$$

Solution. For any $u \in U$ observe that

$$u \in \operatorname{null} ST \quad \Leftrightarrow \quad S(Tu) = 0 \quad \Leftrightarrow \quad Tu \in \operatorname{null} S.$$

Thus $\operatorname{null} ST = \{u \in U : Tu \in \operatorname{null} S\}$. It follows from Exercise 3.B.21 that

$$\dim \operatorname{null} ST = \dim \operatorname{null} T + \dim(\operatorname{null} S \cap \operatorname{range} T) \leq \dim \operatorname{null} T + \dim \operatorname{null} S.$$

Exercise 3.B.23. Suppose U and V are finite-dimensional vector spaces and $S \in \mathcal{L}(V, W)$ and $T \in \mathcal{L}(U, V)$. Prove that

$$\dim \operatorname{range} ST \leq \min\{\dim \operatorname{range} S, \dim \operatorname{range} T\}.$$

Solution. Certainly $\operatorname{range} ST \subseteq \operatorname{range} S$ and thus $\dim \operatorname{range} ST \leq \dim \operatorname{range} S$. Let $R \in \mathcal{L}(\operatorname{range} T, W)$ be the restriction of S to $\operatorname{range} T$, so that $\operatorname{range} R = \operatorname{range} ST$. The fundamental theorem of linear maps (3.21) then implies that

$$\dim \operatorname{range} ST = \dim \operatorname{range} R = \dim \operatorname{range} T - \dim \operatorname{null} R \leq \dim \operatorname{range} T.$$

Thus $\dim \operatorname{range} ST \leq \dim \operatorname{range} S$ and $\dim \operatorname{range} ST \leq \dim \operatorname{range} T$; it follows that

$$\dim \operatorname{range} ST \leq \min\{\dim \operatorname{range} S, \dim \operatorname{range} T\}.$$

Exercise 3.B.24.

(a) Suppose $\dim V = 5$ and $S, T \in \mathcal{L}(V)$ are such that $ST = 0$. Prove that $\dim \operatorname{range} TS \leq 2$.

(b) Give an example of $S, T \in \mathcal{L}(\mathbf{F}^5)$ with $ST = 0$ and $\dim \operatorname{range} TS = 2$.

Solution.

(a) The fundamental theorem of linear maps (3.21) and Exercise 3.B.23 give us

$$\dim \operatorname{null} S = 5 - \dim \operatorname{range} S \leq 5 - \dim \operatorname{range} TS. \tag{1}$$

Note that $ST = 0$ implies $\operatorname{range} T \subseteq \operatorname{null} S$, so that $\dim \operatorname{range} T \leq \dim \operatorname{null} S$. It then follows from Exercise 3.B.23 and equation (1) that

$$\dim \operatorname{range} TS \leq \dim \operatorname{range} T \leq \dim \operatorname{null} S \leq 5 - \dim \operatorname{range} TS$$
$$\Rightarrow \quad 2 \dim \operatorname{range} TS \leq 5 \quad \Rightarrow \quad \dim \operatorname{range} TS \leq \tfrac{5}{2}.$$

Thus $\dim \operatorname{range} TS \leq 2$, since the dimension of a vector space must be an integer.

(b) Let $S, T \in \mathcal{L}(\mathbf{F}^5)$ be given by

$$S(x_1, x_2, x_3, x_4, x_5) = (0, 0, 0, x_4, x_5) \quad \text{and} \quad T(x_1, x_2, x_3, x_4, x_5) = (x_3, x_4, x_5, 0, 0).$$

Observe that

$$(ST)(x_1, x_2, x_3, x_4, x_5) = (0, 0, 0, 0, 0)$$
$$\text{and} \quad (TS)(x_1, x_2, x_3, x_4, x_5) = (0, x_4, x_5, 0, 0).$$

Thus $ST = 0$ and $\dim \operatorname{range} TS = 2$.

Exercise 3.B.25. Suppose that W is finite-dimensional and $S, T \in \mathcal{L}(V, W)$. Prove that $\operatorname{null} S \subseteq \operatorname{null} T$ if and only if there exists $E \in \mathcal{L}(W)$ such that $T = ES$.

Solution. If there exists such a map E then for any $v \in \operatorname{null} S$ we have

$$Tv = E(Sv) = E(0) = 0 \quad \Rightarrow \quad v \in \operatorname{null} T.$$

Thus $\operatorname{null} S \subseteq \operatorname{null} T$.

Now suppose that $\operatorname{null} S \subseteq \operatorname{null} T$. Let $Sv_1, \ldots, Sv_m$ be a basis of $\operatorname{range} S$ and extend this to a basis $Sv_1, \ldots, Sv_m, w_1, \ldots, w_n$ of W. Define $E \in \mathcal{L}(W)$ by $E(Sv_k) = Tv_k$ and $Ew_k = 0$. For any $v \in V$ we have $Sv = a_1 Sv_1 + \cdots + a_m Sv_m$ for some scalars $a_1, \ldots, a_m$, which implies

$$S(v - (a_1 v_1 + \cdots + a_m v_m)) = 0 \quad \Rightarrow \quad v - (a_1 v_1 + \cdots + a_m v_m) \in \operatorname{null} S$$
$$\Rightarrow \quad v - (a_1 v_1 + \cdots + a_m v_m) \in \operatorname{null} T \quad \Rightarrow \quad T(v - (a_1 v_1 + \cdots + a_m v_m)) = 0.$$

Thus $Tv = a_1 Tv_1 + \cdots + a_m Tv_m$. It follows that

$$E(Sv) = a_1 E(Sv_1) + \cdots + a_m E(Sv_m) = a_1 Tv_1 + \cdots + a_m Tv_m = Tv.$$

Hence $T = ES$.

Exercise 3.B.26. Suppose that V is finite-dimensional and $S, T \in \mathcal{L}(V, W)$. Prove that $\operatorname{range} S \subseteq \operatorname{range} T$ if and only if there exists $E \in \mathcal{L}(V)$ such that $S = TE$.

Solution. If there exists such a map E then for any $Sv \in \operatorname{range} S$ we have $Sv = T(Ev) \in \operatorname{range} T$ also. Thus $\operatorname{range} S \subseteq \operatorname{range} T$.

Now suppose that $\operatorname{range} S \subseteq \operatorname{range} T$ and let $v_1, \ldots, v_m$ be a basis of V. Since $\operatorname{range} S \subseteq \operatorname{range} T$, for each $k \in \{1, \ldots, m\}$ we have $Sv_k = Tu_k$ for some $u_k \in V$. Define $E \in \mathcal{L}(V)$ by $Ev_k = u_k$ and observe that for any $v = a_1 v_1 + \cdots + a_m v_m \in V$ we have

$$\begin{aligned}(TE)(v) &= (TE)(a_1v_1 + \cdots + a_mv_m) = a_1T(E(v_1)) + \cdots + a_mT(E(v_m)) \\ &= a_1Tu_1 + \cdots + a_mTu_m = a_1Sv_1 + \cdots + a_mSv_m = S(a_1v_1 + \cdots + a_mv_m) = Sv.\end{aligned}$$

Thus $S = TE$.

Exercise 3.B.27. Suppose $P \in \mathcal{L}(V)$ and $P^2 = P$. Prove that $V = \operatorname{null} P \oplus \operatorname{range} P$.

Solution. For any $v \in V$ note that

$$\begin{aligned}&P^2v = Pv \quad\Rightarrow\quad P(Pv) - Pv = 0 \quad\Rightarrow\quad P(Pv - v) = 0 \\ &\qquad\Rightarrow\quad Pv - v = u \in \operatorname{null} P \quad\Rightarrow\quad v = u + Pv \in \operatorname{null} P + \operatorname{range} P.\end{aligned}$$

Thus $V = \operatorname{null} P + \operatorname{range} P$. Suppose that $v = Pu \in \operatorname{null} P \cap \operatorname{range} P$. It follows that

$$0 = Pv = P^2u = Pu = v.$$

Thus $\operatorname{null} P \cap \operatorname{range} P = \{0\}$ and it follows from 1.46 that the sum $V = \operatorname{null} P \oplus \operatorname{range} P$ is direct.

Exercise 3.B.28. Suppose $D \in \mathcal{L}(\mathcal{P}(\mathbf{R}))$ is such that $\deg Dp = (\deg p) - 1$ for every non-constant polynomial $p \in \mathcal{P}(\mathbf{R})$. Prove that D is surjective.

The notation D is used above to remind you of the differentiation map that sends a polynomial p to p'.

Solution. First, we will recursively define a sequence of polynomials $(p_k)_{k=0}^{\infty}$ such that $Dp_k = x^k$. By assumption $\deg Dx = (\deg x) - 1 = 0$, so that $Dx = b$ for some non-zero $b \in \mathbf{F}$. Define $p_0 = b^{-1}x$ and, using the linearity of D, observe that

$$Dp_0 = D(b^{-1}x) = b^{-1}Dx = b^{-1}b = 1.$$

Now suppose that we have defined polynomials $p_0, ..., p_n$ such that $Dp_k = x^k$ for each $k \in \{0, ..., n\}$. By assumption $D(x^{n+2})$ must have degree $n + 1$, i.e. must be of the form

$$D(x^{n+2}) = b_{n+1}x^{n+1} + b_nx^n + \cdots + b_1x + b_0$$

where $b_{n+1} \neq 0$. Because $Dp_k = x^k$ and D is linear, it follows that

$$\begin{aligned}&b_{n+1}^{-1}D(x^{n+2}) = x^{n+1} + b_{n+1}^{-1}(b_nDp_n + \cdots + b_1Dp_1 + b_0Dp_0) \\ &\qquad\Rightarrow\quad x^{n+1} = D(b_{n+1}^{-1}(x^{n+2} - b_np_n - \cdots - b_1p_1 - b_0p_0)).\end{aligned}$$

Thus, defining $p_{n+1} = b_{n+1}^{-1}(x^{n+2} - b_np_n - \cdots - b_1p_1 - b_0p_0)$, we have $Dp_{n+1} = x^{n+1}$. We now obtain the desired sequence $(p_k)_{k=0}^{\infty}$ of polynomials by recursion.

We can now show that D is surjective. Let $p = \sum_{k=0}^{\deg p} a_kx^k \in \mathcal{P}(\mathbf{R})$ be given. Because D is linear, it follows that

$$D\left(\sum_{k=0}^{\deg p} a_k p_k\right) = \sum_{k=0}^{\deg p} a_k D p_k = \sum_{k=0}^{\deg p} a_k x^k = p.$$

Thus D is surjective.

Exercise 3.B.29. Suppose $p \in \mathcal{P}(\mathbf{R})$. Prove that there exists a polynomial $q \in \mathcal{P}(\mathbf{R})$ such that $5q'' + 3q' = p$.

> *This exercise can be done without linear algebra, but it's more fun to do it using linear algebra.*

Solution. Define a map $D : \mathcal{P}(\mathbf{R}) \to \mathcal{P}(\mathbf{R})$ by $Dq = 5q'' + 3q'$; it will suffice to show that D is surjective. The linearity of D follows from the linearity of differentiation. Suppose $q \in \mathcal{P}(\mathbf{R})$ is a non-constant polynomial of degree $n \geq 1$, so that $q = \sum_{k=0}^{n} a_k x^k$ with $a_n \neq 0$. Some calculations reveals that

$$Dq = \begin{cases} 3a_n & \text{if } n = 1, \\ 3na_n x^{n-1} + \sum_{k=0}^{n-2}(k+1)[3a_{k+1} + 5(k+2)a_{k+2}]x^k & \text{if } n \geq 2. \end{cases}$$

In either case, because $a_n \neq 0$, the polynomial Dq has degree $n-1$. Thus D satisfies the hypotheses of Exercise 3.B.26 and hence must be surjective.

Exercise 3.B.30. Suppose $\varphi \in \mathcal{L}(V, \mathbf{F})$ and $\varphi \neq 0$. Suppose $u \in V$ is not in $\operatorname{null} \varphi$. Prove that

$$V = \operatorname{null} \varphi \oplus \{au : a \in \mathbf{F}\}.$$

Solution. For any $v \in V$, the linearity of φ gives us

$$1 = \varphi\left(\frac{1}{\varphi(u)}u\right) \quad \Rightarrow \quad \varphi(v) = \varphi\left(\frac{\varphi(v)}{\varphi(u)}u\right) \quad \Rightarrow \quad \varphi\left(v - \frac{\varphi(v)}{\varphi(u)}u\right) = 0$$

$$\Rightarrow \quad v - \frac{\varphi(v)}{\varphi(u)}u = w \text{ for some } w \in \operatorname{null}\varphi \quad \Rightarrow \quad v = w + \frac{\varphi(v)}{\varphi(u)}u \in \operatorname{null}\varphi + \{au : a \in \mathbf{F}\}.$$

Thus $V = \operatorname{null}\varphi + \{au : a \in \mathbf{F}\}$. Suppose that $v \in \operatorname{null}\varphi \cap \{au : a \in \mathbf{F}\}$, so that $v = au$ for some $a \in \mathbf{F}$. Observe that

$$0 = \varphi(v) = \varphi(au) = a\varphi(u).$$

Since $\varphi(u) \neq 0$, Exercise 1.B.2 implies that $a = 0$ and hence that $v = 0$. Thus

$$\operatorname{null}\varphi \cap \{au : a \in \mathbf{F}\} = \{0\}$$

and it follows from 1.46 that the sum $V = \operatorname{null}\varphi \oplus \{au : a \in \mathbf{F}\}$ is direct.

Exercise 3.B.31. Suppose V is finite-dimensional, X is a subspace of V, and Y is a finite-dimensional subspace of W. Prove that there exists $T \in \mathcal{L}(V, W)$ such that $\operatorname{null} T = X$ and $\operatorname{range} T = Y$ if and only if $\dim X + \dim Y = \dim V$.

Solution. If there exists such a map T then the equality $\dim X + \dim Y = \dim V$ is immediate from the fundamental theorem of linear maps (3.21).

Suppose that $\dim X + \dim Y = \dim V$. Let $x_1, \ldots, x_\ell$ be a basis of X, which we extend to a basis $x_1, \ldots, x_\ell, v_1, \ldots, v_m$, and let $y_1, \ldots, y_n$ be a basis of Y. By assumption we have

$$\dim X + \dim Y = \dim V \quad \Leftrightarrow \quad \ell + n = \ell + m \quad \Leftrightarrow \quad n = m,$$

so the map $T \in \mathcal{L}(V, W)$ given by $Tx_k = 0$ and $Tv_k = y_k$ is well-defined (i.e. there are enough y_k's to define this map). Suppose $v = a_1 x_1 + \cdots + a_\ell x_\ell + b_1 v_1 + \cdots + b_m v_m \in V$ is such that $Tv = 0$. Observe that

$$0 = Tv = a_1 T x_1 + \cdots + a_\ell T x_\ell + b_1 T v_1 + \cdots + b_m T v_m = b_1 y_1 + \cdots + b_m y_m.$$

The linear independence of $y_1, \ldots, y_n$ then implies that $b_1 = \cdots = b_m = 0$ and thus $v \in X$, so that $\operatorname{null} T \subseteq X$. Certainly $X \subseteq \operatorname{null} T$ and it follows that $\operatorname{null} T = X$. Furthermore, Exercise 3.B.10 implies that

$$\operatorname{range} T = \operatorname{span}(Tx_1, \ldots, Tx_\ell, Tv_1, \ldots, Tv_m) = \operatorname{span}(y_1, \ldots, y_m) = \operatorname{span}(y_1, \ldots, y_n) = Y,$$

where we have used $m = n$ for the third equality.

Exercise 3.B.32. Suppose V is finite-dimensional with $\dim V > 1$. Show that if $\varphi : \mathcal{L}(V) \to \mathbf{F}$ is a linear map such that $\varphi(ST) = \varphi(S)\varphi(T)$ for all $S, T \in \mathcal{L}(V)$, then $\varphi = 0$.

Hint: The description of the two-sided ideals of $\mathcal{L}(V)$ given by Exercise 17 in Section 3A might be useful.

Solution. First note that by Exercise 3.A.16 there exist $S, T \in \mathcal{L}(V)$ such that $ST - TS \neq 0$. Note further that

$$\varphi(ST - TS) = \varphi(ST) - \varphi(TS) = \varphi(S)\varphi(T) - \varphi(T)\varphi(S) = 0,$$

where we have used that multiplication in $\mathbf{F}$ is commutative. It follows that $ST - TS \in \operatorname{null} \varphi$ and hence that $\operatorname{null} \varphi \neq \{0\}$.

Now suppose that $E \in \operatorname{null} \varphi$ and $T \in \mathcal{L}(V)$. Observe that

$$\varphi(ET) = \varphi(E)\varphi(T) = 0 \cdot \varphi(T) = 0 \quad \text{and} \quad \varphi(TE) = \varphi(T)\varphi(E) = \varphi(T) \cdot 0 = 0.$$

Thus ET and TE also belong to $\operatorname{null} \varphi$, so that $\operatorname{null} \varphi$ is a two-sided ideal of $\mathcal{L}(V)$. As we showed in Exercise 3.A.17, any non-zero two-sided ideal of $\mathcal{L}(V)$ must be $\mathcal{L}(V)$ itself, i.e. $\operatorname{null} \varphi = \mathcal{L}(V)$. It follows that $\varphi = 0$.

Exercise 3.B.33. Suppose that V and W are real vector spaces and $T \in \mathcal{L}(V, W)$. Define $T_{\mathbf{C}} : V_{\mathbf{C}} \to W_{\mathbf{C}}$ by

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

for all $u, v \in V$.

(a) Show that $T_{\mathbf{C}}$ is a (complex) linear map from $V_{\mathbf{C}}$ to $W_{\mathbf{C}}$.

(b) Show that $T_{\mathbf{C}}$ is injective if and only if T is injective.

(c) Show that $\operatorname{range} T_{\mathbf{C}} = W_{\mathbf{C}}$ if and only if $\operatorname{range} T = W$.

*See Exercise 8 in Section 1B for the definition of the complexification $V_{\mathbf{C}}$. The linear map $T_{\mathbf{C}}$ is called the **complexification** of the linear map T.*

Solution.

(a) Let $u_1 + iv_1$ and $u_2 + iv_2 \in V_{\mathbf{C}}$ be given. Using the linearity of T, observe that

$$\begin{aligned} T_{\mathbf{C}}((u_1 + iv_1) + (u_2 + iv_2)) &= T_{\mathbf{C}}((u_1 + u_2) + i(v_1 + v_2)) \\ &= T(u_1 + u_2) + iT(v_1 + v_2) \\ &= Tu_1 + Tu_2 + iTv_1 + iTv_2 \\ &= Tu_1 + iTv_1 + Tu_2 + iTv_2 \\ &= T_{\mathbf{C}}(u_1 + iv_1) + T_{\mathbf{C}}(u_2 + iv_2). \end{aligned}$$

Similarly, let $u + iv \in V_{\mathbf{C}}$ and $a + bi \in \mathbf{C}$ be given. The linearity of T gives us

$$\begin{aligned} T_{\mathbf{C}}((a + bi)(u + iv)) &= T_{\mathbf{C}}((au - bv) + i(av + bu)) \\ &= T(au - bv) + iT(av + bu) \\ &= (aTu - bTv) + i(aTv + bTu) \\ &= (a + bi)(Tu + iTv) \\ &= (a + bi)T_{\mathbf{C}}(u + iv). \end{aligned}$$

Thus $T_{\mathbf{C}}$ is a $\mathbf{C}$-linear map from $V_{\mathbf{C}}$ to $W_{\mathbf{C}}$.

(b) Suppose that T is injective and observe that for any $u + iv \in V_{\mathbf{C}}$,

$$\begin{aligned} T_{\mathbf{C}}(u + iv) = 0 \quad &\Leftrightarrow \quad Tu + iTv = 0 \quad \Leftrightarrow \quad Tu = 0 \text{ and } Tv = 0 \\ &\Leftrightarrow \quad u = 0 \text{ and } v = 0 \quad \Leftrightarrow \quad u + iv = 0, \end{aligned}$$

where we have used the injectivity of T for the third equivalence. Thus $\operatorname{null} T_{\mathbf{C}} = \{0\}$, i.e. $T_{\mathbf{C}}$ is injective.

Now suppose that $T_{\mathbf{C}}$ is injective and observe that for any $u \in V$,

$$Tu = 0 \quad \Leftrightarrow \quad T_{\mathbf{C}}(u + i0) = 0 \quad \Leftrightarrow \quad u + i0 = 0 \quad \Leftrightarrow \quad u = 0,$$

where we have used the injectivity of $T_{\mathbf{C}}$ for the second equivalence. It follows that $\operatorname{null} T = \{0\}$, i.e. T is injective.

(c) Suppose that $\operatorname{range} T = W$ and let $w + ix \in W_{\mathbf{C}}$ be given. There exist $u, v \in V$ such that $Tu = w$ and $Tv = x$. It follows that

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

Thus $\operatorname{range} T_{\mathbf{C}} = W_{\mathbf{C}}$.

Now suppose that $\operatorname{range} T_{\mathbf{C}} = W_{\mathbf{C}}$ and let $w \in W$ be given. There exists $u + iv \in V_{\mathbf{C}}$ such that $T_{\mathbf{C}}(u + iv) = w + i0$. It follows that

$$T_{\mathbf{C}}(u + iv) = Tu + iTv = w + i0 \quad \Rightarrow \quad Tu = w.$$

Thus $\operatorname{range} T = W$.

3.C. Matrices

Exercise 3.C.1. Suppose $T \in \mathcal{L}(V, W)$. Show that with respect to each choice of bases of V and W, the matrix of T has at least $\dim \operatorname{range} T$ nonzero entries.

Solution. Let $v_1, \ldots, v_n$ be a basis of V and $w_1, \ldots, w_m$ be a basis of W, so that the matrix of T with respect to these bases is the m-by-n matrix $\mathcal{M}(T)$ whose entries $A_{j,k}$ are defined by

$$Tv_k = A_{1,k}w_1 + \cdots + A_{m,k}w_m.$$

Let $p = \dim \operatorname{null} T$ and $q = \dim \operatorname{range} T$, so that $p + q = n$. Because the list $v_1, \ldots, v_n$ is linearly independent, at most p of the v_k's can belong to $\operatorname{null} T$. Equivalently, at least $n - p = q$ of the v_k's do not belong to $\operatorname{null} T$. Letting v_k be such a vector, we have

$$Tv_k = A_{1,k}w_1 + \cdots + A_{m,k}w_m \neq 0.$$

This implies that at least one of the scalars $A_{j,k}$ is non-zero, i.e. column k of $\mathcal{M}(T)$ has at least one non-zero entry. Since there are at least q choices of k resulting in a non-zero column k of $\mathcal{M}(T)$, we see that $\mathcal{M}(T)$ has at least $q = \dim \operatorname{range} T$ non-zero entries.

Exercise 3.C.2. Suppose V and W are finite-dimensional and $T \in \mathcal{L}(V, W)$. Prove that $\dim \operatorname{range} T = 1$ if and only if there exist a basis of V and a basis of W such that with respect to these bases, all entries of $\mathcal{M}(T)$ equal 1.

Solution. Suppose there exists a basis $v_1, \ldots, v_n$ of V and a basis $w_1, \ldots, w_m$ of W such that with respect to these bases all entries of $\mathcal{M}(T)$ equal 1. That is,

$$Tv_1 = \cdots = Tv_n = w_1 + \cdots + w_m.$$

It follows from Exercise 3.B.10 that

$$\operatorname{range} T = \operatorname{span}(Tv_1, \ldots, Tv_n) = \operatorname{span}(w_1 + \cdots + w_m).$$

Because $w_1, \ldots, w_m$ is linearly independent, we must have $w_1 + \cdots + w_m \neq 0$ and thus $w_1, \ldots, w_m$ is a basis of $\operatorname{range} T$. It follows that $\dim \operatorname{range} T = 1$.

To prove the converse, let us first prove the following lemmas.

Lemma L.1. If W is finite-dimensional with $\dim W = m$ and $w \in W$ is non-zero, then there exists a basis $w_1, \ldots, w_m$ of W such that $w = w_1 + \cdots + w_m$.

Proof. Note that w being non-zero implies $m \geq 1$. If $m = 1$ then take $w_1 = w$; otherwise, extend the list w to a basis $w, w_1, \ldots, w_{m-1}$ of W and define

$$w_m = w - w_1 - \cdots - w_{m-1},$$

so that $w = w_1 + \cdots + w_m$. Observe that each vector in the basis $w, w_1, \ldots, w_{m-1}$ can be expressed as a linear combination of vectors from the list $w_1, \ldots, w_m$. It follows that $W = \operatorname{span}(w_1 \ldots, w_m)$ and hence, by 2.42, $w_1, \ldots, w_m$ is a basis of W. $\square$

Lemma L.2. If V is finite-dimensional with $\dim V = n$ and U is a subspace of V with $U \neq V$, then there exists a basis $v_1, \ldots, v_n$ of V such that $v_k \notin U$ for each $k \in \{1, \ldots, n\}$.

Proof. Note that $U \neq V$ implies $n \geq 1$. We will construct the required basis $v_1, \ldots, v_n$ via the following process.

Step 1. Because $U \neq V$, there exists some $v_1 \in V$ such that $v_1 \notin U$, which implies that $v_1 \neq 0$. If $\operatorname{span}(v_1) = V$ then the process stops and v_1 is the required basis. Otherwise, move to step 2.

Step k. Suppose we have chosen linearly independent vectors $v_1, \ldots, v_{k-1}$, none of which belong to U, such that $\operatorname{span}(v_1, \ldots, v_{k-1}) \neq V$. Observe that

$$\begin{aligned} \operatorname{span}(v_1, \ldots, v_{k-1}) \cup U = V \quad &\Rightarrow \quad U \subseteq \operatorname{span}(v_1, \ldots, v_{k-1}) \\ &\Rightarrow \quad \operatorname{span}(v_1, \ldots, v_{k-1}) = V, \end{aligned}$$

where we have used Exercise 1.C.12 and the fact that $\operatorname{span}(v_1, \ldots, v_{k-1})$ is not contained in U (since $v_1 \notin U$) for the first implication. Given that $\operatorname{span}(v_1, \ldots, v_{k-1}) \neq V$, it must be the case that $\operatorname{span}(v_1, \ldots, v_{k-1}) \cup U \neq V$ and thus there exists some $v_k \in V$ such that

$$v_k \notin \operatorname{span}(v_1, \ldots, v_{k-1}) \quad \text{and} \quad v_k \notin U.$$

It follows from Exercise 2.A.13 that the list $v_1, \ldots, v_k$ is linearly independent. If $\operatorname{span}(v_1, \ldots, v_k) = V$ then the process stops and $v_1, \ldots, v_k$ is the required basis. Otherwise, move to step $k + 1$.

Because V is finite-dimensional, this process must stop after a finite number of steps (indeed, it stops after n steps). $\square$

Returning to the exercise, suppose that $\dim \operatorname{range} T = 1$, so that $\operatorname{range} T$ has a basis w. By Lemma L.1 there is a basis $w_1, \ldots, w_m$ of W such that $w = w_1 + \cdots + w_m$, and by Lemma L.2 there is a basis $u_1, \ldots, u_n$ of V such that each $u_k \notin \operatorname{null} T$. For any $k \in \{1, \ldots, n\}$ we have $Tu_k \in \operatorname{range} T = \operatorname{span}(w)$ and thus $Tu_k = \lambda_k w$ for some scalar λ_k; this scalar must be non-zero since $Tu_k \neq 0$. Let $v_k = \lambda_k^{-1} u_k$ and observe that, because each λ_k^{-1} is non-zero, $v_1, ..., v_n$ is a basis of V. It follows that

$$Tv_k = w = w_1 + \cdots + w_m$$

for each $k \in \{1, ..., n\}$. Thus with respect to the bases $v_1, ..., v_n$ and $w_1, ..., w_m$ all entries of $\mathcal{M}(T)$ equal 1.

Exercise 3.C.3. Suppose $v_1, ..., v_n$ is a basis of V and $w_1, ..., w_m$ is a basis of W.

(a) Show that if $S, T \in \mathcal{L}(V, W)$, then $\mathcal{M}(S+T) = \mathcal{M}(S) + \mathcal{M}(T)$.

(b) Show that if $\lambda \in \mathbf{F}$ and $T \in \mathcal{L}(V, W)$, then $\mathcal{M}(\lambda T) = \lambda \mathcal{M}(T)$.

This exercise asks you to verify 3.35 and 3.38.

Solution.

(a) Suppose $\mathcal{M}(S)$ has entries $A_{j,k}$ and $\mathcal{M}(T)$ has entries $B_{j,k}$, i.e.

$$Sv_k = A_{1,k} w_1 + \cdots + A_{m,k} w_m \quad \text{and} \quad Tv_k = B_{1,k} w_1 + \cdots + B_{m,k} w_m.$$

It follows that

$$(S+T)(v_k) = Sv_k + Tv_k = \left(A_{1,k} + B_{1,k}\right) w_1 + \cdots + \left(A_{m,k} + B_{m,k}\right) w_m.$$

Thus $\mathcal{M}(S+T)$ has entries $A_{j,k} + B_{j,k}$. That is, $\mathcal{M}(S+T) = \mathcal{M}(S) + \mathcal{M}(T)$.

(b) Suppose $\mathcal{M}(T)$ has entries $A_{j,k}$, i.e.

$$Tv_k = A_{1,k} w_1 + \cdots + A_{m,k} w_m.$$

It follows that

$$(\lambda T)(v_k) = \lambda T v_k = \left(\lambda A_{1,k}\right) w_1 + \cdots + \left(\lambda A_{m,k}\right) w_m.$$

Thus $\mathcal{M}(S+T)$ has entries $\lambda A_{j,k}$. That is, $\mathcal{M}(\lambda T) = \lambda \mathcal{M}(T)$.

Exercise 3.C.4. Suppose that $D \in \mathcal{L}(\mathcal{P}_3(\mathbf{R}), \mathcal{P}_2(\mathbf{R}))$ is the differentiation map defined by $Dp = p'$. Find a basis of $\mathcal{P}_3(\mathbf{R})$ and a basis of $\mathcal{P}_2(\mathbf{R})$ such that the matrix of D with respect to these bases is

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Compare with Example 3.33. The next exercise generalizes this exercise.

Solution. Take $x^3, x^2, x, 1$ as a basis of $\mathcal{P}_3(\mathbf{R})$ and $3x^2, 2x, 1$ as a basis of $\mathcal{P}_2(\mathbf{R})$ and observe that

$$D(x^3) = 3x^2, \quad D(x^2) = 2x, \quad D(x) = 1, \quad \text{and} \quad D(1) = 0.$$

Thus the matrix of D with respect to these bases is

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Exercise 3.C.5. Suppose V and W are finite-dimensional and $T \in \mathcal{L}(V, W)$. Prove that there exist a basis of V and a basis of W such that with respect to these bases, all entries of $\mathcal{M}(T)$ are 0 except that the entries in row k, column k, equal 1 if $1 \le k \le \dim \operatorname{range} T$.

Solution. Let $v_1, \dots, v_m$ be a basis of $\operatorname{null} T$ and extend this to a basis $u_1, \dots, u_\ell, v_1, \dots, v_m$ of V (note the ordering). As the proof of 2.33 shows, we then have $U \cap \operatorname{null} T = \{0\}$, where $U = \operatorname{span}(u_1, \dots, u_\ell)$. The restriction of T to U is a linear map in its own right. Moreover, this restriction is injective:

$$u \in U \text{ and } Tu = 0 \quad \Rightarrow \quad u \in U \cap \operatorname{null} T = \{0\} \quad \Rightarrow \quad u = 0.$$

It then follows from Exercise 3.B.9, Exercise 3.B.10, and Exercise 3.B.11 that the list $Tu_1, \dots, Tu_\ell$ is a basis of $\{Tu : u \in U\} = \operatorname{range} T$. Thus $\dim \operatorname{range} T = \ell$. Extend the list $Tu_1, \dots, Tu_\ell$ to a basis $Tu_1, \dots, Tu_\ell, w_1, \dots, w_n$ of W and let $\mathcal{M}(T)$ be the matrix of T with respect to the bases $u_1, \dots, u_\ell, v_1, \dots, v_m$ and $Tu_1, \dots, Tu_\ell, w_1, \dots, w_n$. If $\mathcal{M}(T)$ has entries $A_{j,k}$, then notice that $A_{j,k} = 1$ if $j = k$ and $1 \le k \le \ell = \dim \operatorname{range} T$, and $A_{j,k} = 0$ otherwise.

Exercise 3.C.6. Suppose $v_1, \dots, v_m$ is a basis of V and W is finite-dimensional. Suppose $T \in \mathcal{L}(V, W)$. Prove that there exists a basis $w_1, \dots, w_n$ of W such that all entries in the first column of $\mathcal{M}(T)$ [with respect to the bases $v_1, \dots, v_m$ and $w_1, \dots, w_n$] are 0 except for possibly a 1 in the first row, first column.

In this exercise, unlike Exercise 5, you are given the basis of V instead of being able to choose a basis of V.

Solution. If $v_1 \in \operatorname{null} T$ then let $w_1, \dots, w_n$ be any basis of W. Since $Tv_1 = 0$, it follows that the first column of $\mathcal{M}(T)$ is zero.

Suppose that $v_1 \notin \operatorname{null} T$, so that $Tv_1 \neq 0$. Let $w_1 = Tv_1$ and extend this to a basis $w_1, \dots, w_n$ of W. It follows that the entries in the first column of $\mathcal{M}(T)$ are all 0 except for a 1 in the first row.

Exercise 3.C.7. Suppose $w_1, \dots, w_n$ is a basis of W and V is finite-dimensional. Suppose $T \in \mathcal{L}(V, W)$. Prove that there exists a basis $v_1, \dots, v_m$ of V such that all entries in the first row of $\mathcal{M}(T)$ [with respect to the bases $v_1, \dots, v_m$ and $w_1, \dots, w_n$] are 0 except for possibly a 1 in the first row, first column.

In this exercise, unlike Exercise 5, you are given the basis of W instead of being able to choose a basis of W.

Solution. Let $u_1, ..., u_m$ be any basis of V and suppose $\mathcal{M}(T, (u_1, ..., u_m), (w_1, ..., w_n))$ has entries $A_{j,k}$. If the first row of this matrix is zero then $u_1, ..., u_m$ is the desired basis. Otherwise, there exists some $i \in \{1, ..., m\}$ such that $A_{1,i}$ is non-zero. Let $\lambda = A_{1,i}^{-1}$ and define

$$v_1 = \lambda u_i, \quad v_i = u_1 - \lambda A_{1,1} u_i, \quad \text{and} \quad v_k = u_k - \lambda A_{1,k} u_i \text{ for } 2 \leq k \leq m \text{ and } k \neq i.$$

Because each u_k belongs to $\operatorname{span}(v_1, ..., v_m)$, the list $v_1, ..., v_m$ spans V. It follows from 2.42 that $v_1, ..., v_m$ is a basis of V. Now observe that

$$Tv_1 = \lambda T u_i = \lambda\big(A_{1,i} w_1 + \cdots + A_{n,i} w_n\big) = 1 w_1 + \cdots + \lambda A_{n,i} w_n,$$

$$\begin{aligned} Tv_i = Tu_1 - A_{1,1}(\lambda T u_i) = A_{1,1} w_1 + \cdots + A_{n,1} w_n - A_{1,1}\big(w_1 + \cdots + \lambda A_{n,i} w_n\big) \\ = 0 w_1 + \cdots + \big(A_{n,1} - \lambda A_{1,1} A_{n,i}\big) w_n. \end{aligned}$$

For $2 \leq k \leq m$ and $k \neq i$,

$$\begin{aligned} Tv_k = Tu_k - A_{1,k}(\lambda T u_i) = A_{1,k} w_1 + \cdots + A_{n,k} w_n - A_{1,k}\big(w_1 + \cdots + \lambda A_{n,i} w_n\big) \\ = 0 w_1 + \cdots + \big(A_{n,k} - \lambda A_{1,k} A_{n,i}\big) w_n. \end{aligned}$$

Thus the entries in the first row of the matrix of T with respect to the bases $v_1, ..., v_m$ and $w_1, ..., w_n$ are 0, except for a 1 in the first column.

Exercise 3.C.8. Suppose A is an m-by-n matrix and B is an n-by-p matrix. Prove that

$$(AB)_{j,\cdot} = A_{j,\cdot} B$$

for each $1 \leq j \leq m$. In other words, show that row j of AB equals (row j of A) times B.

This exercise gives the row version of 3.48.

Solution. $(AB)_{j,\cdot}$ is a 1-by-p matrix whose entry in the k^{th} column is

$$\big((AB)_{j,\cdot}\big)_{1,k} = (AB)_{j,k} = \sum_{r=1}^{n} A_{j,r} B_{r,k}.$$

$A_{j,\cdot}$ is a 1-by-n matrix and so $A_{j,\cdot} B$ is a 1-by-p matrix whose entry in the k^{th} column is

$$\big(A_{j,\cdot} B\big)_{1,k} = \sum_{r=1}^{n} \big(A_{j,\cdot}\big)_{1,r} B_{r,k} = \sum_{r=1}^{n} A_{j,r} B_{r,k}.$$

Thus $(AB)_{j,\cdot} = A_{j,\cdot} B$.

Exercise 3.C.9. Suppose $a = \begin{pmatrix} a_1 & \cdots & a_n \end{pmatrix}$ is a 1-by-n matrix and B is an n-by-p matrix. Prove that

$$aB = a_1 B_{1,\cdot} + \cdots + a_n B_{n,\cdot}\,.$$

In other words, show that aB is a linear combination of the rows of B, with the scalars that multiply the rows coming from a.

This exercise gives the row version of 3.50.

Solution. aB is a 1-by-p matrix whose entry in the k^{th} column is

$$(aB)_{1,k} = \sum_{r=1}^{n} a_r B_{r,k} = a_1 B_{1,k} + \cdots + a_n B_{n,k}.$$

Thus

$$aB = a_1 B_{1,\cdot} + \cdots + a_n B_{n,\cdot}\,.$$

Exercise 3.C.10. Give an example of 2-by-2 matrices A and B such that $AB \neq BA$.

Solution. Let

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$$

and observe that

$$AB = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = BA.$$

Exercise 3.C.11. Prove that the distributive property holds for matrix addition and matrix multiplication. In other words, suppose A, B, C, D, E, and F are matrices whose sizes are such that $A(B + C)$ and $(D + E)F$ make sense. Explain why $AB + AC$ and $DF + EF$ both make sense and prove that

$$A(B + C) = AB + AC \quad \text{and} \quad (D + E)F = DF + EF.$$

Solution. For $B + C$ to make sense, B and C must have the same sizes; suppose they are both n-by-p matrices. For $A(B + C)$ to make sense, A must then be an m-by-n matrix for some m. Given this, both AB and AC are m-by-p matrices and thus $AB + AC$ makes sense.

Similarly, suppose D and E are both m-by-n matrices. For $(D + E)F$ to make sense, F must be an n-by-p matrix for some p. Given this, both DF and EF are m-by-p matrices and thus $DF + EF$ makes sense.

The entry of $A(B + C)$ in row j, column k is given by

$$\sum_{r=1}^{n} A_{j,r}(B+C)_{r,k} = \sum_{r=1}^{n} A_{j,r}\big(B_{r,k}+C_{r,k}\big) = \sum_{r=1}^{n} A_{j,r}B_{r,k} + \sum_{r=1}^{n} A_{j,r}C_{r,k},$$

where we have used distributivity in $\mathbf{F}$. The expression on the right-hand side gives the entry of $AB + AC$ in row j, column k. Thus $A(B + C) = AB + AC$. A similar argument shows that $(D + E)F = DF + EF$.

Exercise 3.C.12. Prove that matrix multiplication is associative. In other words, suppose A, B, and C are matrices whose sizes are such that $(AB)C$ makes sense. Explain why $A(BC)$ makes sense and prove that

$$(AB)C = A(BC).$$

Try to find a clean proof that illustrates the following quote from Emil Artin: "It is my experience that proofs involving matrices can be shortened by 50% if one throws the matrices out."

Solution. If A is an m-by-n matrix then for AB to make sense, B must be an n-by-p matrix for some p, so that AB is an m-by-p matrix. For $(AB)C$ to make sense, C must then be a p -by-q matrix for some q, so that $(AB)C$ is an m-by-q matrix. Thus BC is an n-by-q matrix and $A(BC)$ is an m-by-q matrix.

Let $e_1, ..., e_n$ be the standard basis of $\mathbf{F}^n$, let $f_1, ..., f_m$ be the standard basis of $\mathbf{F}^m$, and define $R \in \mathcal{L}(\mathbf{F}^n, \mathbf{F}^m)$ by $Re_k = A_{1,k}f_1 + \cdots + A_{m,k}f_m$. Thus, with respect to the standard bases, $\mathcal{M}(R) = A$. Define $S \in \mathcal{L}(\mathbf{F}^p, \mathbf{F}^n)$ and $T \in \mathcal{L}(\mathbf{F}^q, \mathbf{F}^p)$ similarly, so that $\mathcal{M}(S) = B$ and $\mathcal{M}(T) = C$, with respect to the standard bases. Now observe that

$$\begin{aligned}(AB)C = (\mathcal{M}(R)\mathcal{M}(S))\mathcal{M}(T) &\overset{(3.43)}{=} \mathcal{M}(RS)\mathcal{M}(T) \overset{(3.43)}{=} \mathcal{M}((RS)T) \\ &\overset{(3.8)}{=} \mathcal{M}(R(ST)) \overset{(3.43)}{=} \mathcal{M}(R)\mathcal{M}(ST) \overset{(3.43)}{=} \mathcal{M}(R)(\mathcal{M}(S)\mathcal{M}(T)) = A(BC);\end{aligned}$$

the number above the equals sign is the textbook reference justifying the equality.

Exercise 3.C.13. Suppose A is an n-by-n matrix and $1 \leq j, k \leq n$. Show that the entry in row j, column k, of A^3 (which is defined to mean AAA) is

$$\sum_{p=1}^{n}\sum_{r=1}^{n} A_{j,p}A_{p,r}A_{r,k}.$$

Solution. By the definition of matrix multiplication, we have

$$(A^3)_{j,k} = (A^2A)_{j,k} = \sum_{r=1}^{n} (A^2)_{j,r}A_{r,k} = \sum_{r=1}^{n}\left(\sum_{p=1}^{n} A_{j,p}A_{p,r}\right)A_{r,k} = \sum_{p=1}^{n}\sum_{r=1}^{n} A_{j,p}A_{p,r}A_{r,k}.$$

Exercise 3.C.14. Suppose m and n are positive integers. Prove that the function $A \mapsto A^{\mathrm{t}}$ is a linear map from $\mathbf{F}^{m,n}$ to $\mathbf{F}^{n,m}$.

Solution. Let $A, B \in \mathbf{F}^{m,n}$ and $\lambda \in \mathbf{F}$ be given. Observe that

$$\left((A+B)^{\mathrm{t}}\right)_{k,j} = (A+B)_{j,k} = A_{j,k} + B_{j,k} = (A^{\mathrm{t}})_{k,j} + (B^{\mathrm{t}})_{k,j},$$

$$\left((\lambda A)^{\mathrm{t}}\right)_{k,j} = (\lambda A)_{j,k} = \lambda A_{j,k} = \lambda (A^{\mathrm{t}})_{k,j}.$$

Thus $(A+B)^{\mathrm{t}} = A^{\mathrm{t}} + B^{\mathrm{t}}$ and $(\lambda A)^{\mathrm{t}} = \lambda A^{\mathrm{t}}$.

Exercise 3.C.15. Prove that if A is an m-by-n matrix and C is an n-by-p matrix, then

$$(AC)^{\mathrm{t}} = C^{\mathrm{t}} A^{\mathrm{t}}.$$

This exercise shows that the transpose of the product of two matrices is the product of the transposes in the opposite order.

Solution. Observe that

$$\left((AC)^{\mathrm{t}}\right)_{k,j} = (AC)_{j,k} = \sum_{r=1}^{n} A_{j,r} C_{r,k} = \sum_{r=1}^{n} (C^{\mathrm{t}})_{k,r} (A^{\mathrm{t}})_{r,j} = (C^{\mathrm{t}} A^{\mathrm{t}})_{k,j}.$$

Thus $(AC)^{\mathrm{t}} = C^{\mathrm{t}} A^{\mathrm{t}}$.

Exercise 3.C.16. Suppose A is an m-by-n matrix with $A \neq 0$. Prove that the rank of A is 1 if and only if there exist $(c_1, \ldots, c_m) \in \mathbf{F}^m$ and $(d_1, \ldots, d_n) \in \mathbf{F}^n$ such that $A_{j,k} = c_j d_k$ for every $j = 1, \ldots, m$ and every $k = 1, \ldots, n$.

Solution. If the rank of A is 1 then by 3.56 there is an m-by-1 matrix C and a 1-by-n matrix R such that $A = CR$; take c_j to be the j^{th} entry of C and take d_k to be the k^{th} entry of R.

Now suppose there exist $(c_1, \ldots, c_m) \in \mathbf{F}^m$ and $(d_1, \ldots, d_n) \in \mathbf{F}^n$ such that $A_{j,k} = c_j d_k$ for every $j = 1, \ldots, m$ and every $k = 1, \ldots, n$. If we define

$$C = \begin{pmatrix} c_1 \\ \vdots \\ c_m \end{pmatrix} \in \mathbf{F}^{m,1} \quad \text{and} \quad D = \begin{pmatrix} d_1 & \cdots & d_n \end{pmatrix} \in \mathbf{F}^{1,n},$$

then $A = CD$. By (3.51) (a), every column of A is a scalar multiple of C. It follows that $\operatorname{rank} A \leq 1$, and since $A \neq 0$ we must have $\operatorname{rank} A \geq 1$. Thus $\operatorname{rank} A = 1$.

Exercise 3.C.17. Suppose $T \in \mathcal{L}(V)$, and $u_1, ..., u_n$ and $v_1, ..., v_n$ are bases of V. Prove that the following are equivalent.

(a) T is injective.

(b) The columns of $\mathcal{M}(T)$ are linearly independent in $\mathbf{F}^{n,1}$.

(c) The columns of $\mathcal{M}(T)$ span $\mathbf{F}^{n,1}$.

(d) The rows of $\mathcal{M}(T)$ span $\mathbf{F}^{1,n}$.

(e) The rows of $\mathcal{M}(T)$ are linearly independent in $\mathbf{F}^{1,n}$.

Here $\mathcal{M}(T)$ means $\mathcal{M}(T, (u_1, ..., u_n), (v_1, ..., v_n))$.

Solution. Suppose that $\mathcal{M}(T)$ has entries $A_{j,k}$ and suppose that (a) holds. Let $b_1, ..., b_n$ be scalars such that

$$b_1 \begin{pmatrix} A_{1,1} \\ \vdots \\ A_{n,1} \end{pmatrix} + \cdots + b_n \begin{pmatrix} A_{1,n} \\ \vdots \\ A_{n,n} \end{pmatrix} = \begin{pmatrix} b_1 A_{1,1} + \cdots + b_n A_{1,n} \\ \vdots \\ b_1 A_{n,1} + \cdots + b_n A_{n,n} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Let $u = b_1 u_1 + \cdots + b_n u_n$ and observe that

$$\begin{aligned} Tu &= b_1 T u_1 + \cdots + b_n T u_n \\ &= b_1 (A_{1,1} v_1 + \cdots + A_{n,1} v_n) + \cdots + b_n (A_{1,n} v_1 + \cdots + A_{n,n} v_n) \\ &= (b_1 A_{1,1} + \cdots + b_n A_{1,n}) v_1 + \cdots + (b_1 A_{n,1} + \cdots + b_n A_{n,n}) v_n \\ &= 0 v_1 + \cdots + 0 v_n = 0. \end{aligned}$$

Thus $u \in \operatorname{null} T$. Since T is injective, this implies that $u = 0$. The linear independence of the basis $u_1, ..., u_n$ then gives us $b_1 = \cdots = b_n = 0$. Thus the columns of $\mathcal{M}(T)$ are linearly independent, i.e. (b) holds.

Now suppose that (b) holds and let $u = b_1 u_1 + \cdots + b_n u_n$ be such that $Tu = 0$. As in the previous paragraph, this is equivalent to

$$(b_1 A_{1,1} + \cdots + b_n A_{1,n}) v_1 + \cdots + (b_1 A_{n,1} + \cdots + b_n A_{n,n}) v_n = 0.$$

The linear independence of the basis $v_1, ..., v_n$ then implies that

$$b_1 A_{1,1} + \cdots + b_n A_{1,n} = \cdots = b_1 A_{n,1} + \cdots + b_n A_{n,n} = 0,$$

which in turn gives us

$$\begin{pmatrix} b_1 A_{1,1} + \cdots + b_n A_{1,n} \\ \vdots \\ b_1 A_{n,1} + \cdots + b_n A_{n,n} \end{pmatrix} = b_1 \begin{pmatrix} A_{1,1} \\ \vdots \\ A_{n,1} \end{pmatrix} + \cdots + b_n \begin{pmatrix} A_{1,n} \\ \vdots \\ A_{n,n} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

It follows from the linear independence of the columns of $\mathcal{M}(T)$ that $b_1 = \cdots = b_n = 0$, so that $u = 0$. Thus T is injective, i.e. (a) holds. This gives us the equivalence of (a) and (b).

For the equivalence of (b) and (c), note that $\dim \mathbf{F}^{n,1} = n$ by 3.40 and thus, by 2.38 and 2.40, a list of n vectors in $\mathbf{F}^{n,1}$ is linearly independent if and only if it spans $\mathbf{F}^{n,1}$. The same argument gives us the equivalence of (d) and (e), since we also have $\dim \mathbf{F}^{1,n} = n$ by 3.40.

Finally, the equivalence of (c) and (d) is given by 3.57.

3.D. Invertibility and Isomorphisms

Exercise 3.D.1. Suppose $T \in \mathcal{L}(V, W)$ is invertible. Show that T^{-1} is invertible and

$$\left(T^{-1}\right)^{-1} = T.$$

Solution. By the definition of invertibility of T we have $TT^{-1} = I$ and $T^{-1}T = I$. These equations show that T^{-1} is invertible and its inverse is T.

Exercise 3.D.2. Suppose $T \in \mathcal{L}(U, V)$ and $S \in \mathcal{L}(V, W)$ are both invertible linear maps. Prove that $ST \in \mathcal{L}(U, W)$ is invertible and that $(ST)^{-1} = T^{-1}S^{-1}$.

Solution. Using the algebraic properties of 3.8, observe that

$$T^{-1}S^{-1}ST = T^{-1}IT = T^{-1}T = I \quad \text{and} \quad STT^{-1}S^{-1} = SIS^{-1} = SS^{-1} = I.$$

Thus $T^{-1}S^{-1}ST$ is the identity map on U and $STT^{-1}S^{-1}$ is the identity map on W. It follows that ST is invertible and that $(ST)^{-1} = T^{-1}S^{-1}$.

Exercise 3.D.3. Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Prove that the following are equivalent.

(a) T is invertible.

(b) $Tv_1, \ldots, Tv_n$ is a basis of V for every basis $v_1, \ldots, v_n$ of V.

(c) $Tv_1, \ldots, Tv_n$ is a basis of V for some basis $v_1, \ldots, v_n$ of V.

Solution. Suppose that (a) holds and let $v_1, \ldots, v_n$ be a basis of V. Since T is invertible, it is injective. It follows from Exercise 3.B.9 that $Tv_1, \ldots, Tv_n$ is linearly independent. Thus by 2.38, $Tv_1, \ldots, Tv_n$ is a basis of V. Hence (b) holds.

Suppose that (b) holds. By 2.31 there is a basis $v_1, \ldots, v_n$ of V. By assumption $Tv_1, \ldots, Tv_n$ is a basis of V. Thus (c) holds.

Suppose that (c) holds, so that there is a basis $v_1, \ldots, v_n$ of V such that $Tv_1, \ldots, Tv_n$ is a basis of V. Let $v = a_1v_1 + \cdots + a_nv_n$ be such that $Tv = 0$ and observe that

$$0 = Tv = T(a_1v_1 + \cdots + a_nv_n) = a_1Tv_1 + \cdots + a_nTv_n.$$

The linear independence of $Tv_1, \ldots, Tv_n$ then implies that $a_1 = \cdots = a_n = 0$ and thus $v = 0$. It follows that $\operatorname{null} T = \{0\}$ and hence that T is injective. Thus by 3.65, T is invertible.

Exercise 3.D.4. Suppose V is finite-dimensional and $\dim V > 1$. Prove that the set of noninvertible linear maps from V to itself is not a subspace of $\mathcal{L}(V)$.

Solution. By 3.65 this is equivalent to showing that the set of linear maps from V to itself which are not injective is not a subspace of $\mathcal{L}(V)$. We showed this in Exercise 3.B.7.

Exercise 3.D.5. Suppose V is finite-dimensional, U is a subspace of V, and $S \in \mathcal{L}(U, V)$. Prove that there exists an invertible linear map T from V to itself such that $Tu = Su$ for every $u \in U$ if and only if S is injective.

Solution. If there is such a map T, then

$$\operatorname{null} S = \operatorname{null} T \cap U = \{0\} \cap U = \{0\}.$$

Thus S is injective.

Suppose that S is injective. Let $u_1, \ldots, u_m$ be a basis of U, which we extend to a basis $u_1, \ldots, u_m, x_1, \ldots, x_n$ of V. The injectivity of S and Exercise 3.B.9 imply that $Su_1, \ldots, Su_m$ is linearly independent and thus can be extended to a basis $Su_1, \ldots, Su_m, y_1, \ldots, y_n$ of V. Define $T \in \mathcal{L}(V)$ by $Tu_k = Su_k$ and $Tx_k = y_k$. Certainly T extends S. Furthermore, since T maps a basis of V to a basis of V, T must be invertible by Exercise 3.B.3.

Exercise 3.D.6. Suppose that W is finite-dimensional and $S, T \in \mathcal{L}(V, W)$. Prove that $\operatorname{null} S = \operatorname{null} T$ if and only if there exists an invertible $E \in \mathcal{L}(W)$ such that $S = ET$.

Solution. Suppose there exists an invertible $E \in \mathcal{L}(W)$ such that $S = ET$, which implies $T = E^{-1}S$. It follows from Exercise 3.B.25 that $\operatorname{null} T \subseteq \operatorname{null} S$ and $\operatorname{null} S \subseteq \operatorname{null} T$. Thus $\operatorname{null} S = \operatorname{null} T$.

Now suppose that $\operatorname{null} S = \operatorname{null} T$. It follows from Exercise 3.B.25 that there are maps $R, R' \in \mathcal{L}(W)$ such that $T = RS$ and $S = R'T$ and thus by Exercise 3.B.23 we have $\dim \operatorname{range} T = \dim \operatorname{range} S$. Let $Tv_1, \ldots, Tv_m$ be a basis of $\operatorname{range} T$ and notice that this linearly independent list is equal to $R(Sv_1), \ldots, R(Sv_m)$. It follows from Exercise 3.A.4 that the list $Sv_1, \ldots, Sv_m$ is linearly independent and hence is a basis of $\operatorname{range} S$, since $\dim \operatorname{range} S = m$. Extend these lists to bases

$$Tv_1, \ldots, Tv_m, x_1, \ldots, x_n \quad \text{and} \quad Sv_1, \ldots, Sv_m, y_1, \ldots, y_n$$

of W, and define $E \in \mathcal{L}(W)$ by $E(Tv_k) = Sv_k$ and $Ex_k = y_k$. Notice that E maps a basis of W to a basis of W; it follows from Exercise 3.D.3 that E is invertible. For any $v \in V$ we have $Tv = a_1 Tv_1 + \cdots + a_m Tv_m$ for some scalars $a_1, \ldots, a_m$. As we showed in Exercise 3.B.25, this implies $Sv = a_1 Sv_1 + \cdots + a_m Sv_m$ since $\operatorname{null} T = \operatorname{null} S$. It follows that

$$E(Tv) = a_1 E(Tv_1) + \cdots + a_m E(Tv_m) = a_1 Sv_1 + \cdots + a_m Sv_m = Sv.$$

Thus $S = ET$.

Exercise 3.D.7. Suppose that V is finite-dimensional and $S, T \in \mathcal{L}(V, W)$. Prove that $\operatorname{range} S = \operatorname{range} T$ if and only if there exists an invertible $E \in \mathcal{L}(V)$ such that $S = TE$.

Solution. Suppose there exists an invertible $E \in \mathcal{L}(V)$ such that $S = TE$, which implies $T = SE^{-1}$. It follows from Exercise 3.B.26 that $\operatorname{range} T \subseteq \operatorname{range} S$ and $\operatorname{range} S \subseteq \operatorname{range} T$. Thus $\operatorname{range} S = \operatorname{range} T$.

Now suppose that $\operatorname{range} S = \operatorname{range} T$. Let $u_1, \ldots, u_m$ be a basis of $\operatorname{null} S$ and extend this to a basis of $u_1, \ldots, u_m, x_1, \ldots, x_n$ of V; as the proof of the fundamental theorem of linear maps (3.21) shows, $Sx_1, \ldots, Sx_n$ is a basis of $\operatorname{range} S$. Our assumption $\operatorname{range} S = \operatorname{range} T$ implies that $Sx_k = Ty_k$ for some $y_1, \ldots, y_n \in V$, and also that $\dim \operatorname{null} T = \dim \operatorname{null} S = m$. Let $v_1, \ldots, v_m$ be a basis of $\operatorname{null} T$ and suppose we have scalars $a_1, \ldots, a_m, b_1, \ldots, b_n$ such that

$$a_1 v_1 + \cdots + a_m v_m + b_1 y_1 + \cdots + b_n y_n = 0.$$

Applying T to both sides of this equation gives us

$$b_1 T y_1 + \cdots + b_n T y_n = b_1 S x_1 + \cdots + b_n S x_n = 0.$$

The linear independence of $Sx_1, \ldots, Sx_n$ implies that $b_1 = \cdots = b_n = 0$ and the linear independence of $v_1, \ldots, v_m$ then gives us $a_1 = \cdots = a_m = 0$. Thus the list $v_1, \ldots, v_m, y_1, \ldots, y_n$ is linearly independent and hence is a basis of V by 2.38. Define $E \in \mathcal{L}(V)$ by $Eu_k = v_k$ and $Ex_k = y_k$. Because E maps a basis of V to a basis of V, Exercise 3.D.3 shows that E is invertible. For any $v = a_1 u_1 + \cdots + a_m u_m + b_1 x_1 + \cdots + b_n x_n \in V$, observe that

$$\begin{aligned} T(Ev) &= T(a_1 v_1 + \cdots + a_m v_m + b_1 y_1 + \cdots + b_n y_n) \\ &= b_1 T y_1 + \cdots + b_n T y_n = b_1 S x_1 + \cdots + b_n S x_n = Sv. \end{aligned}$$

Thus $S = TE$.

Exercise 3.D.8. Suppose V and W are finite-dimensional and $S, T \in \mathcal{L}(V, W)$. Prove that there exist invertible $E_1 \in \mathcal{L}(V)$ and $E_2 \in \mathcal{L}(W)$ such that $S = E_2 T E_1$ if and only if $\dim \operatorname{null} S = \dim \operatorname{null} T$.

Solution. Suppose there exist such maps E_1, E_2. It follows from Exercise 3.B.22 that

$$\dim \operatorname{null} S \leq \dim \operatorname{null} E_2 + \dim \operatorname{null} T + \dim \operatorname{null} E_1 = \dim \operatorname{null} T.$$

Notice that $T = E_2^{-1} S E_1^{-1}$; repeating the previous argument gives us $\dim \operatorname{null} T \leq \dim \operatorname{null} S$ and thus $\dim \operatorname{null} S = \dim \operatorname{null} T$.

Now suppose that $\dim \operatorname{null} S = \dim \operatorname{null} T$. Let $u_1, \ldots, u_m$ be a basis of $\operatorname{null} S$ and let $v_1, \ldots, v_m$ be a basis of $\operatorname{null} T$. Extend these to bases

$$u_1, \ldots, u_m, x_1, \ldots, x_n \quad \text{and} \quad v_1, \ldots, v_m, y_1, \ldots, y_n$$

of V, and define $E_1 \in \mathcal{L}(V)$ by $E_1 u_k = v_k$ and $E_1 x_k = y_k$. Because E_1 maps a basis to a basis, Exercise 3.D.3 shows that E_1 is invertible. It is straightforward to verify that $\operatorname{null} S = \operatorname{null} TE_1$ and thus by Exercise 3.D.6 there is an invertible $E_2 \in \mathcal{L}(W)$ such that $S = E_2 T E_1$.

Exercise 3.D.9. Suppose V is finite-dimensional and $T : V \to W$ is a surjective linear map of V onto W. Prove that there is a subspace U of V such that $T|_U$ is an isomorphism of U onto W.

Here $T|_U$ means the function T restricted to U. Thus $T|_U$ is the function whose domain is U, with $T|_U$ defined by $T|_U\ (u) = Tu$ for every $u \in U$.

Solution. By Exercise 3.B.11 there is a subspace U of V such that

$$U \cap \operatorname{null} T = \{0\} \quad \text{and} \quad W = \operatorname{range} T = \{Tu : u \in U\}.$$

The equation $U \cap \operatorname{null} T = \{0\}$ shows that $T|_U$ is injective and the equation $W = \{Tu : u \in U\}$ shows that $T|_U$ is surjective. Thus $T|_U$ is an isomorphism of U onto W.

Exercise 3.D.10. Suppose V and W are finite-dimensional and U is a subspace of V. Let

$$\mathcal{E} = \{T \in \mathcal{L}(V, W) : U \subseteq \operatorname{null} T\}.$$

(a) Show that $\mathcal{E}$ is a subspace of $\mathcal{L}(V, W)$.

(b) Find a formula for $\dim \mathcal{E}$ in terms of $\dim V, \dim W$, and $\dim U$.

Hint: Define $\Phi : \mathcal{L}(V, W) \to \mathcal{L}(U, W)$ by $\Phi(T) = T|_U$. What is $\operatorname{null} \Phi$? What is $\operatorname{range} \Phi$?

Solution.

(a) Because the null space of the zero map is all of V, we certainly have $0 \in \mathcal{E}$. Suppose that $S, T \in \mathcal{E}$ and $\lambda \in \mathbf{F}$. For any $u \in U$, observe that

$$(S + T)(u) = Su + Tu = 0 \quad \text{and} \quad (\lambda T)(u) = \lambda Tu = 0,$$

where we have used that $u \in \operatorname{null} S$ and $u \in \operatorname{null} T$. It follows that $U \subseteq \operatorname{null}(S + T)$ and $U \subseteq \operatorname{null}(\lambda T)$, so that $S + T \in \mathcal{E}$ and $\lambda T \in \mathcal{E}$. Thus $\mathcal{E}$ is a subspace of $\mathcal{L}(V, W)$.

(b) Following the hint, define $\Phi : \mathcal{L}(V, W) \to \mathcal{L}(U, W)$ by $\Phi(T) = T|_U$; it is straightforward to verify that Φ is linear. Note that

$$\Phi(T) = 0 \quad \Leftrightarrow \quad T|_U = 0 \quad \Leftrightarrow \quad Tu = 0 \text{ for all } u \in U \quad \Leftrightarrow \quad U \subseteq \operatorname{null} T.$$

Thus $\operatorname{null} \Phi = \mathcal{E}$. For any $S \in \mathcal{L}(U, W)$ we can use Exercise 3.B.13 to extend S to a linear map $T \in \mathcal{L}(V, W)$; it follows that $\Phi(T) = S$. Thus Φ is surjective, i.e. $\operatorname{range} \Phi = \mathcal{L}(U, W)$. It now follows from the fundamental theorem of linear maps (3.21) and 3.72 that

$$\dim \mathcal{L}(V, W) = \dim \operatorname{null} \Phi + \dim \operatorname{range} \Phi \quad \Rightarrow \quad \dim \mathcal{E} = \dim W(\dim V - \dim U).$$

Exercise 3.D.11. Suppose V is finite-dimensional and $S, T \in \mathcal{L}(V)$. Prove that

$$ST \text{ is invertible} \quad \Leftrightarrow \quad S \text{ and } T \text{ are invertible.}$$

Solution. If S and T are invertible then Exercise 3.D.2 shows that ST is invertible. If S is not invertible then S is not surjective by 3.65, so that $\dim \operatorname{range} S < \dim V$. It follows from Exercise 3.B.23 that

$$\dim \operatorname{range} ST \leq \dim \operatorname{range} S < \dim V.$$

Thus ST is not surjective and hence not invertible. A similar argument shows that if T is not invertible then ST is not invertible.

Exercise 3.D.12. Suppose V is finite-dimensional and $S, T, U \in \mathcal{L}(V)$ and $STU = I$. Show that T is invertible and that $T^{-1} = US$.

Solution. It follows from 3.68 that S commutes with TU and that ST commutes with U. Thus

$$STU = I \quad \Rightarrow \quad TUS = I \text{ and } UST = I.$$

Hence T is invertible and $T^{-1} = US$.

Exercise 3.D.13. Show that the result in Exercise 12 can fail without the hypothesis that V is finite-dimensional.

Solution. Consider $V = \mathbf{F}^\infty$. Let S be the backward shift operator, let T be the forward shift operator, and let U be the identity on $\mathbf{F}^\infty$. For any $(x_1, x_2, x_3, \ldots) \in \mathbf{F}^\infty$, observe that

$$(STU)(x_1, x_2, x_3, \ldots) = S(T(x_1, x_2, x_3, \ldots)) = S(0, x_1, x_2, x_3, \ldots) = (x_1, x_2, x_3, \ldots).$$

Thus $STU = I$. However, T is not invertible because T is not surjective: $(1, 0, 0, \ldots) \notin \operatorname{range} T$.

Exercise 3.D.14. Prove or give a counterexample: If V is a finite-dimensional vector space and $R, S, T \in \mathcal{L}(V)$ are such that RST is surjective, then S is injective.

Solution. This is true. RST must be invertible by 3.65 and thus by Exercise 3.D.11 S is invertible and hence injective.

Exercise 3.D.15. Suppose $T \in \mathcal{L}(V)$ and $v_1, \ldots, v_m$ is a list in V such that $Tv_1, \ldots, Tv_m$ spans V. Prove that $v_1, \ldots, v_m$ spans V.

Solution. Using 2.30 we can reduce the list $Tv_1, \ldots, Tv_m$ to a basis $Tv_{k_1}, \ldots, Tv_{k_n}$ for some indices $1 \leq k_1 < \cdots < k_n \leq m$. It follows from Exercise 3.A.4 and 2.38 that $v_{k_1}, \ldots, v_{k_n}$ is a basis of V and thus

$$V = \operatorname{span}\left(v_{k_1}, \ldots, v_{k_n}\right) \subseteq \operatorname{span}(v_1, \ldots, v_m) \quad \Rightarrow \quad V = \operatorname{span}(v_1, \ldots, v_m).$$

Exercise 3.D.16. Prove that every linear map from $\mathbf{F}^{n,1}$ to $\mathbf{F}^{m,1}$ is given by a matrix multiplication. In other words, prove that if $T \in \mathcal{L}(\mathbf{F}^{n,1}, \mathbf{F}^{m,1})$, then there exists an m-by-n matrix A such that $Tx = Ax$ for every $x \in \mathbf{F}^{n,1}$.

Solution. Let A be the matrix of T with respect to the standard bases of $\mathbf{F}^{n,1}$ and $\mathbf{F}^{m,1}$. With respect to these standard bases we have $\mathcal{M}(x) = x$ and $\mathcal{M}(y) = y$ for any $x \in \mathbf{F}^{n,1}$ and $y \in \mathbf{F}^{m,1}$. It then follows from 3.76 that

$$Tx = \mathcal{M}(Tx) = \mathcal{M}(T)\mathcal{M}(x) = Ax$$

for every $x \in \mathbf{F}^{n,1}$.

Exercise 3.D.17. Suppose V is finite-dimensional and $S \in \mathcal{L}(V)$. Define $\mathcal{A} \in \mathcal{L}(\mathcal{L}(V))$ by

$$\mathcal{A}(T) = ST$$

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

(a) Show that $\dim \operatorname{null} \mathcal{A} = (\dim V)(\dim \operatorname{null} S)$.

(b) Show that $\dim \operatorname{range} \mathcal{A} = (\dim V)(\dim \operatorname{range} S)$.

Solution.

(a) For $T \in \mathcal{L}(V)$, note that $ST = 0$ if and only if $\operatorname{range} T \subseteq \operatorname{null} S$. Thus we can identify $\operatorname{null} \mathcal{A}$ with $\mathcal{L}(V, \operatorname{null} S)$ and it follows from 3.72 that

$$\dim \operatorname{null} \mathcal{A} = (\dim V)(\dim \operatorname{null} S).$$

(b) For $R \in \mathcal{L}(V)$, Exercise 3.B.26 implies that $R = ST$ for some $T \in \mathcal{L}(V)$ if and only if $\operatorname{range} R \subseteq \operatorname{range} S$. Thus we can identify $\operatorname{range} \mathcal{A}$ with $\mathcal{L}(V, \operatorname{range} S)$ and it follows from 3.72 that

$$\dim \operatorname{range} \mathcal{A} = (\dim V)(\dim \operatorname{range} S).$$

Exercise 3.D.18. Show that V and $\mathcal{L}(\mathbf{F}, V)$ are isomorphic vector spaces.

Solution. Define a map $\Phi : \mathcal{L}(\mathbf{F}, V) \to V$ by

$$\Phi(T) = T(1).$$

It is straightforward to verify that Φ is linear. Now define a map $\Psi : V \to \mathcal{L}(\mathbf{F}, V)$ by

$$[\Psi(v)](x) = xv.$$

It is straightforward to check that $\Psi(v)$ indeed belongs to $\mathcal{L}(\mathbf{F}, V)$ for any $v \in V$. For any $T \in \mathcal{L}(\mathbf{F}, V)$, observe that

$$[\Psi(\Phi(T))](x) = [\Psi(T(1))](x) = xT(1) = T(x).$$

Thus $\Psi(\Phi(T)) = T$, i.e. $\Psi \circ \Phi$ is the identity map on $\mathcal{L}(\mathbf{F}, V)$. Now let $v \in V$ be given and observe that

$$\Phi(\Psi(v)) = [\Psi(v)](1) = v.$$

Thus $\Phi \circ \Psi$ is the identity map on V. As the proof of 3.63 shows, it now follows that Ψ is a linear map. Thus Φ is an isomorphism from $\mathcal{L}(\mathbf{F}, V)$ to V and its inverse is Ψ.

Exercise 3.D.19. Suppose V is a finite-dimensional and $T \in \mathcal{L}(V)$. Prove that T has the same matrix with respect to every basis of V if and only if T is a scalar multiple of the identity operator.

Solution. If $T = \lambda I$ for some $\lambda \in \mathbf{F}$ then the matrix of T with respect to any basis of V must be λI (here I is the identity matrix).

Suppose that T has the same matrix with respect to every basis of V, i.e. there is some matrix A with entries $A_{j,k}$ such that

$$Tu_k = A_{1,k}u_1 + \cdots + A_{m,k}u_m$$

for any basis $u_1, ..., u_m$ of V. Let $v_1, ..., v_m$ be a fixed basis of V and let $k \in \{1, ..., m\}$ be given; it is straightforward to verify that $v_1, ..., \frac{1}{2}v_k, ..., v_m$ is also a basis of V. By assumption we must then have

$$\begin{aligned}
T\left(\tfrac{1}{2}v_k\right) = A_{1,k}v_1 + \cdots + A_{k,k}\left(\tfrac{1}{2}v_k\right) + \cdots + A_{m,k}v_m& \\
\Rightarrow \quad Tv_k = 2A_{1,k}v_1 + \cdots + A_{k,k}v_k + \cdots + 2A_{m,k}v_m&.
\end{aligned}$$

On the other hand we must have

$$Tv_k = A_{1,k}v_1 + \cdots + A_{k,k}v_k + \cdots + A_{m,k}v_m.$$

Hence by unique representation we must have $A_{j,k} = 2A_{j,k}$, so that $A_{j,k} = 0$, for all $j \neq k$; it follows that $Tu_k = A_{k,k}u_k$ for any basis $u_1, ..., u_m$ of V. Let $1 \leq j < k \leq m$ be given and consider the basis $v_1, ..., v_k, ..., v_j, ..., v_m$ of V, i.e. the basis obtained by swapping the basis vectors v_j and v_k. This gives us the two equations

$$Tv_k = A_{k,k}v_k \quad \text{and} \quad Tv_k = A_{j,j}v_k.$$

It follows from unique representation that $A_{j,j} = A_{k,k}$ and thus, letting $\lambda = A_{1,1}$, we have $Tv_k = \lambda v_k$ for all $k \in \{1, ..., m\}$. Thus $T = \lambda I$.

Exercise 3.D.20. Suppose $q \in \mathcal{P}(\mathbf{R})$. Prove that there exists a polynomial $p \in \mathcal{P}(\mathbf{R})$ such that

$$q(x) = (x^2 + x)p''(x) + 2xp'(x) + p(3)$$

for all $x \in \mathbf{R}$.

Solution. There is a non-negative integer m such that $q \in \mathcal{P}_m(\mathbf{R})$; either $m = \deg q$ if $q \neq 0$ or $m = 0$ if $q = 0$. Define $T : \mathcal{P}_m(\mathbf{R}) \to \mathcal{P}_m(\mathbf{R})$ by

$$Tp = (x^2 + x)p''(x) + 2xp'(x) + p(3).$$

It is straightforward to check that T is linear and some calculations reveal that $\deg Tp = \deg p$ for any $p \in \mathcal{P}_m(\mathbf{R})$. It follows that if $p \in \mathcal{P}_m(\mathbf{R})$ is such that $Tp = 0$, so that $\deg Tp = -\infty$, then $\deg p = -\infty$, i.e. $p = 0$. Thus T is injective. By 3.65 this implies that T is surjective and so there must exist some $p \in \mathcal{P}_m(\mathbf{R})$ such that $Tp = q$.

Exercise 3.D.21. Suppose n is a positive integer and $A_{j,k} \in \mathbf{F}$ for all $j, k = 1, \ldots, n$. Prove that the following are equivalent (note that in both parts below, the number of equations equals the number of variables).

(a) The trivial solution $x_1 = \cdots = x_n = 0$ is the only solution to the homogeneous system of equations

$$\sum_{k=1}^{n} A_{1,k} x_k = 0$$

$$\vdots$$

$$\sum_{k=1}^{n} A_{n,k} x_k = 0.$$

(b) For every $c_1, \ldots, c_n \in \mathbf{F}$, there exists a solution to the system of equations

$$\sum_{k=1}^{n} A_{1,k} x_k = c_1$$

$$\vdots$$

$$\sum_{k=1}^{n} A_{n,k} x_k = c_n.$$

Solution. Define $T \in \mathcal{L}(\mathbf{F}^n)$ by

$$T(x_1, \ldots, x_n) = \left(\sum_{k=1}^{n} A_{1,k} x_k, \ldots, \sum_{k=1}^{n} A_{n,k} x_k \right)$$

and notice that (a) is equivalent to the injectivity of T and (b) is equivalent to the surjectivity of (b). It then follows from 3.65 that (a) and (b) are equivalent.

Exercise 3.D.22. Suppose $T \in \mathcal{L}(V)$ and $v_1, ..., v_n$ is a basis of V. Prove that

$$\mathcal{M}(T, (v_1, ..., v_n)) \text{ is invertible} \quad \Leftrightarrow \quad T \text{ is invertible.}$$

Solution. In what follows, all matrices of linear maps and vectors are understood to be with respect to the basis $v_1, ..., v_n$ of V.

Suppose that $\mathcal{M}(T)$ is invertible and let B be its inverse. Define $S \in \mathcal{L}(V)$ by

$$Sv_k = B_{1,k}v_1 + \cdots + B_{n,k}v_n,$$

so that $\mathcal{M}(S) = B$. Note that for any $u, v \in V$ we have $u = v$ if and only if $\mathcal{M}(u) = \mathcal{M}(v)$, by unique representation in the basis $v_1, ..., v_n$. Let $v \in V$ be given and observe that

$$\mathcal{M}((ST)(v)) \overset{(3.76)}{=} \mathcal{M}(ST)\mathcal{M}(v) \overset{(3.43)}{=} \mathcal{M}(S)\mathcal{M}(T)\mathcal{M}(v) = I\mathcal{M}(v) = \mathcal{M}(v).$$

The number above the equals sign is the textbook reference justifying the equality; the third equality is justified as $\mathcal{M}(S) = B$ is the inverse of $\mathcal{M}(T)$. Thus $(ST)(v) = v$ for all $v \in V$, so that ST is the identity map on V. It follows from 3.68 that TS is also the identity map on V and we may conclude that T is invertible with inverse S.

The converse statement is the content of 3.86, which we now prove. Suppose that T is invertible. Using 3.43, observe that

$$\mathcal{M}(T)\mathcal{M}(T^{-1}) = \mathcal{M}(TT^{-1}) = \mathcal{M}(I) = I = \mathcal{M}(I) = \mathcal{M}(T^{-1}T) = \mathcal{M}(T^{-1})\mathcal{M}(T).$$

Thus $(\mathcal{M}(T))^{-1} = \mathcal{M}(T^{-1})$.

Exercise 3.D.23. Suppose that $u_1, ..., u_n$ and $v_1, ..., v_n$ are bases of V. Let $T \in \mathcal{L}(V)$ be such that $Tv_k = u_k$ for each $k = 1, ..., n$. Prove that

$$\mathcal{M}(T, (v_1, ..., v_n)) = \mathcal{M}(I, (u_1, ..., u_n), (v_1, ..., v_n)).$$

Solution. For ease of notation, let us write

$$\mathcal{M}(T, u, v) = \mathcal{M}(T, (u_1, ..., u_n), (v_1, ..., v_n)) \quad \text{and} \quad \mathcal{M}(T, v) = \mathcal{M}(T, (v_1, ..., v_n)).$$

Note that, by 3.81,

$$\mathcal{M}(T, u)\mathcal{M}(I, v, u) = \mathcal{M}(TI, v, u) = \mathcal{M}(T, v, u) = I, \tag{$*$}$$

where the last equality follows from the definition of T. Using 3.84, 3.82, and $(*)$, observe that

$$\mathcal{M}(T, v) = (\mathcal{M}(I, v, u))^{-1}\mathcal{M}(T, u)\mathcal{M}(I, v, u) = \mathcal{M}(I, u, v)\mathcal{M}(T, u)\mathcal{M}(I, v, u) = \mathcal{M}(I, u, v).$$

Exercise 3.D.24. Suppose A and B are square matrices of the same size and $AB = I$. Prove that $BA = I$.

Solution. Suppose that A and B are n-by-n matrices and let $e_1, ..., e_n$ be the standard basis of $\mathbf{F}^n$. In what follows, all matrices of linear maps are understood to be with respect to this standard basis.

Let $S, T \in \mathcal{L}(\mathbf{F}^n)$ be given by

$$Se_k = \sum_{j=1}^{n} A_{j,k} e_k \quad \text{and} \quad Te_k = \sum_{j=1}^{n} B_{j,k} e_k,$$

so that $\mathcal{M}(S) = A$ and $\mathcal{M}(T) = B$. Using 3.43, we then have

$$\mathcal{M}(ST) = \mathcal{M}(S)\mathcal{M}(T) = AB = I = \mathcal{M}(I).$$

Thus, by the uniqueness part of the linear map lemma (3.4), we have $ST = I$, which implies $TS = I$ by 3.68. It follows from 3.43 that

$$BA = \mathcal{M}(T)\mathcal{M}(S) = \mathcal{M}(TS) = \mathcal{M}(I) = I.$$

3.E. Products and Quotients of Vector Spaces

Exercise 3.E.1. Suppose T is a function from V to W. The *graph* of T is the subset of $V \times W$ defined by

$$\text{graph of } T = \{(v, Tv) \in V \times W : v \in V\}.$$

Prove that T is a linear map if and only if the graph of T is a subspace of $V \times W$.

Formally, a function T from V to W is a subset T of $V \times W$ such that for each $v \in V$, there exists exactly one element $(v, w) \in T$. In other words, formally a function is what is called above its graph. We do not usually think of functions in this formal manner. However, if we do become formal, then this exercise could be rephrased as follows: Prove that a function T from V to W is a linear map if and only if T is a subspace of $V \times W$.

Solution. Let $\mathsf{graph}(T)$ be the graph of T. First suppose that T is linear, so that $T(0) = 0$; it follows that $(0, 0) = (0, T(0)) \in \mathsf{graph}(T)$. For any $(u, Tu), (v, Tv) \in \mathsf{graph}(T)$ and $\lambda \in \mathbf{F}$, the linearity of T implies that

$$(u, Tu) + (v, Tv) = (u + v, Tu + Tv) = (u + v, T(u + v)) \in \mathsf{graph}(T),$$

$$\lambda(v, Tv) = (\lambda v, \lambda Tv) = (\lambda v, T(\lambda v)) \in \mathsf{graph}(T).$$

Thus $\mathsf{graph}(T)$ is a subspace of $V \times W$.

Now suppose that $\mathsf{graph}(T)$ is a subspace of $V \times W$. Let $u, v \in V$ and $\lambda \in \mathbf{F}$ be given. Observe that

$$(u, Tu), (v, Tv) \in \mathsf{graph}(T) \quad \Rightarrow \quad (u, Tu) + (v, Tv) = (u + v, Tu + Tv) \in \mathsf{graph}(T).$$

Because the second component of an element of $\mathsf{graph}(T)$ must be T applied to the first component, and $(u + v, Tu + Tv)$ belongs to $\mathsf{graph}(T)$, it must be that $T(u + v) = Tu + Tv$. Similarly,

$$(v, Tv) \in \mathsf{graph}(T) \quad \Rightarrow \quad \lambda(v, Tv) = (\lambda v, \lambda Tv) \in \mathsf{graph}(T) \quad \Rightarrow \quad T(\lambda v) = \lambda Tv.$$

Thus T is linear.

Exercise 3.E.2. Suppose that $V_1, \dots, V_m$ are vector spaces such that $V_1 \times \cdots \times V_m$ is finite-dimensional. Prove that V_k is finite-dimensional for each $k = 1, \dots, m$.

Solution. For any $k \in \{1, \dots, m\}$, let $p_k : V_1 \times \cdots \times V_m \to V_k$ be given by $p_k(v_1, \dots, v_m) = v_k$. It is straightforward to verify that p_k is a surjective linear map and thus, by the fundamental theorem of linear maps (3.21), $\operatorname{range} p_k = V_k$ is finite-dimensional.

Exercise 3.E.3. Suppose $V_1, ..., V_m$ are vector spaces. Prove that $\mathcal{L}(V_1 \times \cdots \times V_m, W)$ and $\mathcal{L}(V_1, W) \times \cdots \times \mathcal{L}(V_m, W)$ are isomorphic vector spaces.

Solution. Define a map

$$\Phi : \mathcal{L}(V_1, W) \times \cdots \times \mathcal{L}(V_m, W) \to \mathcal{L}(V_1 \times \cdots \times V_m, W),$$

where $\Phi(T_1, ..., T_m)$ is the map $V_1 \times \cdots \times V_m \to W$ given by

$$(v_1, ..., v_m) \mapsto T_1 v_1 + \cdots + T_m v_m.$$

It is straightforward to verify that $\Phi(T_1, ..., T_m)$ is indeed a linear map for any $(T_1, ..., T_m)$, and that Φ itself is linear.

For $k \in \{1, ..., m\}$, define $\iota_k : V_k \to V_1 \times \cdots \times V_m$ by $\iota_k(v) = (0, ..., v, ..., 0)$, where the v is in the k^{th} position; it is straightforward to check that each ι_k is a linear map. Define a map

$$\Psi : \mathcal{L}(V_1 \times \cdots \times V_m, W) \to \mathcal{L}(V_1, W) \times \cdots \times \mathcal{L}(V_m, W),$$

where $\Psi(T)$ is given by $(T \circ \iota_1, ..., T \circ \iota_m)$. The linearity of each $T \circ \iota_k$ follows from the linearity of T and the linearity of ι_k. Let $(T_1, ..., T_m) \in \mathcal{L}(V_1, W) \times \cdots \times \mathcal{L}(V_m, W)$ be given and observe that

$$\Psi(\Phi(T_1, ..., T_m)) = (\Phi(T_1, ..., T_m) \circ \iota_1, ..., \Phi(T_1, ..., T_m) \circ \iota_m).$$

For any $k \in \{1, ..., m\}$ and $v \in V_k$ we have

$$[\Phi(T_1, ..., T_m)](\iota_k(v)) = [\Phi(T_1, ..., T_m)](0, ..., v, ..., 0) = T_1(0) + \cdots + T_k v + \cdots + T_m(0) = T_k v.$$

Thus $\Phi(T_1, ..., T_m) \circ \iota_k = T_k$ and it follows that $\Psi(\Phi(T_1, ..., T_m)) = (T_1, ..., T_m)$, i.e. $\Psi \circ \Phi$ is the identity map on $\mathcal{L}(V_1, W) \times \cdots \times \mathcal{L}(V_m, W)$. Now let $T \in \mathcal{L}(V_1 \times \cdots \times V_m, W)$ be given and observe that

$$\begin{aligned}
[\Phi(\Psi(T))](v_1, ..., v_m) &= [\Phi(T \circ \iota_1, ..., T \circ \iota_m)](v_1, ..., v_m) \\
&= (T \circ \iota_1)(v_1) + \cdots + (T \circ \iota_m)(v_m) \\
&= T(v_1, ..., 0) + \cdots + T(0, ..., v_m) \\
&= T(v_1, ..., v_m).
\end{aligned}$$

Thus $\Phi(\Psi(T)) = T$, i.e. $\Phi \circ \Psi$ is the identity map on $\mathcal{L}(V_1 \times \cdots \times V_m, W)$. As the proof of 3.63 shows, it now follows that Ψ is a linear map. Thus Φ is an isomorphism from $\mathcal{L}(V_1, W) \times \cdots \times \mathcal{L}(V_m, W)$ to $\mathcal{L}(V_1 \times \cdots \times V_m, W)$ and its inverse is Ψ.

Exercise 3.E.4. Suppose $W_1, ..., W_m$ are vector spaces. Prove that $\mathcal{L}(V, W_1 \times \cdots \times W_m)$ and $\mathcal{L}(V, W_1) \times \cdots \times \mathcal{L}(V, W_m)$ are isomorphic vector spaces.

Solution. Define a map

$$\Phi : \mathcal{L}(V, W_1) \times \cdots \times \mathcal{L}(V, W_m) \to \mathcal{L}(V, W_1 \times \cdots \times W_m),$$

where $\Phi(T_1, ..., T_m)$ is the map $V \to W_1 \times \cdots \times W_m$ given by

$$v \mapsto (T_1 v, ..., T_m v).$$

It is straightforward to verify that $\Phi(T_1, ..., T_m)$ is indeed a linear map for any $(T_1, ..., T_m)$, and that Φ itself is linear.

For each $k \in \{1, ..., m\}$, define $p_k : W_1 \times \cdots \times W_m \to W_k$ by $p_k(w_1, ..., w_m) = w_k$; it is straightforward to check that each p_k is a linear map. Define a map

$$\Psi : \mathcal{L}(V, W_1 \times \cdots \times W_m) \to \mathcal{L}(V, W_1) \times \cdots \times \mathcal{L}(V, W_m),$$

where $\Psi(T)$ is given by $(p_1 \circ T, ..., p_m \circ T)$. The linearity of each $p_k \circ T$ is given by the linearity of p_k and the linearity of T. Let $(T_1, ..., T_m) \in \mathcal{L}(V, W_1) \times \cdots \times \mathcal{L}(V, W_m)$ be given and observe that

$$\Psi(\Phi(T_1, ..., T_m)) = (p_1 \circ \Phi(T_1, ..., T_m), ..., p_m \circ \Phi(T_1, ..., T_m)).$$

For any $k \in \{1, ..., m\}$ and $v \in V$ we have

$$p_k([\Phi(T_1, ..., T_m)](v)) = p_k(T_1 v, ..., T_m v) = T_k v.$$

Thus $p_k \circ \Phi(T_1, ..., T_m) = T_k$ and it follows that $\Psi(\Phi(T_1, ..., T_m)) = (T_1, ..., T_m)$, i.e. $\Psi \circ \Phi$ is the identity map on $\mathcal{L}(V, W_1) \times \cdots \times \mathcal{L}(V, W_m)$. Now let $T \in \mathcal{L}(V, W_1 \times \cdots \times W_m)$ be given and observe that

$$[\Phi(\Psi(T))](v) = [\Phi(p_1 \circ T, ..., p_m \circ T)](v) = (p_1(Tv), ..., p_m(Tv)) = Tv.$$

Thus $\Phi(\Psi(T)) = T$, i.e. $\Phi \circ \Psi$ is the identity map on $\mathcal{L}(V, W_1 \times \cdots \times W_m)$. As the proof of 3.63 shows, it now follows that Ψ is a linear map. Thus Φ is an isomorphism from $\mathcal{L}(V, W_1) \times \cdots \times \mathcal{L}(V, W_m)$ to $\mathcal{L}(V, W_1 \times \cdots \times W_m)$ and its inverse is Ψ.

Exercise 3.E.5. For m a positive integer, define V^m by

$$V^m = \underbrace{V \times \cdots \times V}_{m \text{ times}}.$$

Prove that V^m and $\mathcal{L}(\mathbf{F}^m, V)$ are isomorphic vector spaces.

Solution. Define a map $\Phi : \mathcal{L}(\mathbf{F}^m, V) \to V^m$ by

$$\Phi(T) = (Te_1, ..., Te_m),$$

where $e_1, ..., e_m$ is the standard basis of $\mathbf{F}^m$. It is straightforward to verify that Φ is linear. Now define a map $\Psi : V^m \to \mathcal{L}(\mathbf{F}^m, V)$ by

$$[\Psi(v_1, ..., v_m)](x_1, ..., x_m) = x_1 v_1 + \cdots + x_m v_m.$$

It is straightforward to check that $\Psi(v_1, ..., v_m)$ indeed belongs to $\mathcal{L}(\mathbf{F}^m, V)$ for any $(v_1, ..., v_m) \in V^m$. For any $T \in \mathcal{L}(\mathbf{F}^m, V)$, observe that

$$
\begin{aligned}
[\Psi(\Phi(T))](x_1, ..., x_m) &= [\Psi(Te_1, ..., Te_m)](x_1, ..., x_m) \\
&= x_1 Te_1 + \cdots + x_m Te_m \\
&= T(x_1 e_1 + \cdots + x_m e_m) \\
&= T(x_1, ..., x_m).
\end{aligned}
$$

Thus $\Psi(\Phi(T)) = T$, i.e. $\Psi \circ \Phi$ is the identity map on $\mathcal{L}(\mathbf{F}^m, V)$. Now let $(v_1, ..., v_m) \in V^m$ be given and observe that

$$\Phi(\Psi(v_1, ..., v_m)) = ([\Psi(v_1, ..., v_m)](e_1), ..., [\Psi(v_1, ..., v_m)](e_m)) = (v_1, ..., v_m).$$

Thus $\Phi \circ \Psi$ is the identity map on V^m. As the proof of 3.63 shows, it now follows that Ψ is a linear map. Thus Φ is an isomorphism from $\mathcal{L}(\mathbf{F}^m, V)$ to V^m and its inverse is Ψ.

Exercise 3.E.6. Suppose that v, x are vectors in V and that U, W are subspaces of V such that $v + U = x + W$. Prove that $U = W$.

Solution. Since $v \in v + U = x + W$, there is some $w \in W$ such that $v = x + w$, which implies that $x - v = -w \in W$. For any $u \in U$ we have $v + u = x + w$ for some $w \in W$, so that $u = x - v + w \in W$. Thus $U \subseteq W$. A similar argument shows that $W \subseteq U$ and it follows that $U = W$.

Exercise 3.E.7. Let $U = \{(x, y, z) \in \mathbf{R}^3 : 2x + 3y + 5z = 0\}$. Suppose $A \subseteq \mathbf{R}^3$. Prove that A is a translate of U if and only if there exists $c \in \mathbf{R}$ such that

$$A = \{(x, y, z) \in \mathbf{R}^3 : 2x + 3y + 5z = c\}.$$

Solution. Suppose that A is a translate of U, i.e. there is some $(a_1, a_2, a_3) \in \mathbf{R}^3$ such that $A = (a_1, a_2, a_3) + U$. Let $c = 2a_1 + 3a_2 + 5a_3$ and let

$$W = \{(x_1, x_2, x_3) \in \mathbf{R}^3 : 2x_1 + 3x_2 + 5x_3 = c\}.$$

We need to show that $A = W$. Suppose that $(x_1, x_2, x_3) \in A$, so that

$$(x_1, x_2, x_3) = (a_1, a_2, a_3) + (u_1, u_2, u_3)$$

for some $(u_1, u_2, u_3) \in U$. It follows that

$$2x_1 + 3x_2 + 5x_3 = 2u_1 + 3u_2 + 5u_3 + 2a_1 + 3a_2 + 5a_3 = 0 + c = c.$$

Thus $A \subseteq W$. If $(x_1, x_2, x_3) \in W$ then note that

$$2x_1 + 3x_2 + 5x_3 = c = 2a_1 + 3a_2 + 5a_3 \quad \Rightarrow \quad (x_1 - a_1, x_2 - a_2, x_3 - a_3) \in U.$$

Thus $(x_1, x_2, x_3) = (a_1, a_2, a_3) + (x_1 - a_1, x_2 - a_2, x_3 - a_3) \in (a_1, a_2, a_3) + U = A$ and it follows that $W \subseteq A$. Hence $A = W$.

Now suppose that there is some $c \in \mathbf{R}$ such that

$$A = \big\{(x_1, x_2, x_3) \in \mathbf{R}^3 : 2x_1 + 3x_2 + 5x_3 = c\big\}.$$

We claim that $A = \left(\frac{c}{6}, \frac{c}{9}, \frac{c}{15}\right) + U$. If $(x_1, x_2, x_3) \in A$ then observe that

$$(x_1, x_2, x_3) = \left(\frac{c}{6}, \frac{c}{9}, \frac{c}{15}\right) + \left(x_1 - \frac{c}{6}, x_2 - \frac{c}{9}, x_3 - \frac{c}{15}\right) \in \left(\frac{c}{6}, \frac{c}{9}, \frac{c}{15}\right) + U,$$

where $\left(x_1 - \frac{c}{6}, x_2 - \frac{c}{9}, x_3 - \frac{c}{15}\right) \in U$ since

$$2\Big(x_1 - \frac{c}{6}\Big) + 3\Big(x_2 - \frac{c}{9}\Big) + 5\Big(x_3 - \frac{c}{15}\Big) = 2x_1 + 3x_2 + 5x_3 - \left(\frac{c}{3} + \frac{c}{3} + \frac{c}{3}\right) = c - c = 0.$$

Thus $A \subseteq \left(\frac{c}{6}, \frac{c}{9}, \frac{c}{15}\right) + U$. If $\left(\frac{c}{6}, \frac{c}{9}, \frac{c}{15}\right) + (u_1, u_2, u_3) \in \left(\frac{c}{6}, \frac{c}{9}, \frac{c}{15}\right) + U$ then

$$2\Big(\frac{c}{6} + u_1\Big) + 3\Big(\frac{c}{9} + u_2\Big) + 5\Big(\frac{c}{15} + u_3\Big) = \left(\frac{c}{3} + \frac{c}{3} + \frac{c}{3}\right) + 2u_1 + 3u_2 + 5u_3 = c + 0 = c.$$

Thus $\left(\frac{c}{6}, \frac{c}{9}, \frac{c}{15}\right) + (u_1, u_2, u_3) \in A$ and we may conclude that $A = \left(\frac{c}{6}, \frac{c}{9}, \frac{c}{15}\right) + U$.

Exercise 3.E.8.

(a) Suppose $T \in \mathcal{L}(V, W)$ and $c \in W$. Prove that $\{x \in V : Tx = c\}$ is either the empty set or is a translate of $\operatorname{null} T$.

(b) Explain why the set of solutions to a system of linear equations such as 3.27 is either the empty set or is a translate of some subspace of $\mathbf{F}^n$.

Solution.

(a) If $c \notin \operatorname{range} T$ then $\{x \in V : Tx = c\}$ must be empty. Suppose that $c \in \operatorname{range} T$, so that $c = Tu$ for some $u \in V$. We claim that $\{x \in V : Tx = c\} = u + \operatorname{null} T$. Indeed, for $x = u + v \in u + \operatorname{null} T$ we have

$$Tx = Tu + Tv = c \quad \Rightarrow \quad x \in \{x \in V : Tx = c\}.$$

Thus $u + \operatorname{null} T \subseteq \{x \in V : Tx = c\}$. If $x \in V$ is such that $Tx = c$ then observe that

$$\begin{aligned} Tx = c = Tu \quad \Rightarrow \quad T(x-u) = 0 \quad \Rightarrow \quad & x - u \in \operatorname{null} T \\ \Rightarrow \quad & x = u + v \text{ for some } v \in \operatorname{null} T. \end{aligned}$$

Thus $\{x \in V : Tx = c\} \subseteq u + \operatorname{null} T$ and our claim follows.

(b) Consider a system of linear equations

$$\begin{gathered} \sum_{k=1}^{n} A_{1,k} x_k = c_1 \\ \vdots \\ \sum_{k=1}^{n} A_{m,k} x_k = c_m. \end{gathered}$$

Define $T \in \mathcal{L}(\mathbf{F}^n, \mathbf{F}^m)$ by

$$T(x_1, \dots, x_n) = \left(\sum_{k=1}^{n} A_{1,k} x_k, \dots, \sum_{k=1}^{n} A_{m,k} x_k \right)$$

and let $c = (c_1, \dots, c_m) \in \mathbf{F}^m$. The solution set of the system of linear equations is then precisely the set

$$\{(x_1, \dots, x_n) \in \mathbf{F}^n : T(x_1, \dots, x_n) = c\}.$$

As we showed in part (a), this set is either empty or is a translate of the subspace $\operatorname{null} T$.

Exercise 3.E.9. Prove that a nonempty subset A of V is a translate of some subspace of V if and only if $\lambda v + (1 - \lambda)w \in A$ for all $v, w \in A$ and all $\lambda \in \mathbf{F}$.

Solution. Suppose that $A = x + U$ for some $x \in V$ and some subspace U of V. Let $v, w \in A$ and $\lambda \in \mathbf{F}$ be given and note that $v = x + t$ and $w = x + u$ for some $t, u \in U$. It follows that

$$\lambda v + (1 - \lambda)w = \lambda(x + t) + (1 - \lambda)(x + u) = x + (\lambda t + (1 - \lambda)u) \in x + U = A.$$

Now suppose that $\lambda v + (1 - \lambda)w \in A$ for all $v, w \in A$ and all $\lambda \in \mathbf{F}$. Because A is non-empty, there is some $x \in A$. We claim that $-x + A$ is a subspace of V. Certainly $0 = -x + x \in -x + A$. Suppose that $-x + v, -x + w \in -x + A$ and $\lambda \in \mathbf{F}$, and observe that

$$x, v \in A \quad \Rightarrow \quad -x + 2v \in A \quad \text{and} \quad x, w \in A \quad \Rightarrow \quad -x + 2w \in A.$$

It follows that

$$\left(-\tfrac{1}{2}x + v\right) + \left(-\tfrac{1}{2}x + w\right) = -x + v + w \in A \quad \Rightarrow \quad -2x + v + w \in -x + A.$$

Thus $(-x + v) + (-x + w) \in -x + A$, so that $-x + A$ is closed under vector addition. Furthermore, $-x + A$ is closed under scalar multiplication:

$$x, v \in A \quad \Rightarrow \quad \lambda(-x + v) = -x + (\lambda v + (1 - \lambda)x) \in -x + A.$$

Thus $-x + A$ is a subspace of V. It follows that $A = x + (-x + A)$ is a translate of the subspace $-x + A$.

Exercise 3.E.10. Suppose $A_1 = v + U_1$ and $A_2 = w + U_2$ for some $v, w \in V$ and some subspaces U_1, U_2 of V. Prove that the intersection $A_1 \cap A_2$ is either a translate of some subspace of V or is the empty set.

Solution. If $A_1 \cap A_2$ is non-empty, so that there is some $x \in A_1 \cap A_2$, then by 3.101 we have $A_1 = x + U_1$ and $A_2 = x + U_2$. We claim that $A_1 \cap A_2 = x + (U_1 \cap U_2)$. If $y \in A_1 \cap A_2$ then $y = x + u_1 = x + u_2$ for some $u_1 \in U_1$ and some $u_2 \in U_2$. It follows that $u_1 = u_2 \in U_1 \cap U_2$ and thus $y \in x + (U_1 \cap U_2)$. This gives us the inclusion $A_1 \cap A_2 \subseteq x + (U_1 \cap U_2)$. If $y = x + u \in x + (U_1 \cap U_2)$ then $y = x + u \in x + U_1 = A_1$ and $y = x + u \in x + U_2 = A_2$, so that $y \in A_1 \cap A_2$. Thus $x + (U_1 \cap U_2) \subseteq A_1 \cap A_2$ and our claim follows.

Exercise 3.E.11. Suppose $U = \{(x_1, x_2, ...) \in \mathbf{F}^\infty : x_k \neq 0 \text{ for only finitely many } k\}$.

(a) Show that U is a subspace of $\mathbf{F}^\infty$.

(b) Prove that $\mathbf{F}^\infty/U$ is infinite-dimensional.

Solution.

(a) Notice that

$$U = \{(x_1, x_2, ...) \in \mathbf{F}^\infty : \text{there exists } K \text{ such that } x_k = 0 \text{ for all } k \geq K\}.$$

Certainly $(0, 0, 0, ...) \in U$. Suppose that $(x_1, x_2, ...), (y_1, y_2, ...) \in U$ and $\lambda \in \mathbf{F}$. There are positive integers K, L such that $x_k = 0$ for all $k \geq K$ and $y_k = 0$ for all $k \geq L$. It follows that $x_k + y_k = 0$ for all $k \geq \max\{K, L\}$ and $\lambda x_k = 0$ for all $k \geq K$. Thus

$$(x_1, x_2, ...) + (y_1, y_2, ...) \in U \quad \text{and} \quad \lambda(x_1, x_2, ...) \in U.$$

It follows that U is a subspace of $\mathbf{F}^\infty$.

(b) For $x \in \mathbf{F}^\infty$ we will use the notation $x(k)$ to denote the k^{th} term of x. Let $e_n \in \mathbf{F}^\infty$ be the sequence given by

$$e_n(k) = \begin{cases} 1 & \text{if } k \text{ is divisible by } 2^n, \\ 0 & \text{otherwise.} \end{cases}$$

$$\begin{aligned} e_1 &= (0, 1, 0, 1, 0, 1, 0, 1, ...), \\ e_2 &= (0, 0, 0, 1, 0, 0, 0, 1, ...), \\ e_3 &= (0, 0, 0, 0, 0, 0, 0, 1, ...), \text{ etc.} \end{aligned}$$

Let $m \in \mathbf{N}$ be given. We claim that the list $e_1 + U, ..., e_m + U$ is linearly independent. Suppose $a_1, ..., a_m$ are such that

$$a_1(e_1 + U) + \cdots + a_m(e_m + U) = (a_1 e_1 + \cdots + a_m e_m) + U = 0.$$

This is the case if and only if $e \in U$, where $e = a_1 e_1 + \cdots + a_m e_m$, which implies that there is a positive integer K such that $e(k) = 0$ for all $k \geq K$. Let $N \in \mathbf{N}$ be such that $2^{mN} \geq K$ and note that $2^{mN} + 2$ is divisible by 2 but not by $2^2, ..., 2^m$. It follows that

$$0 = e(2^{mN} + 2) = a_1 e_1(2^{mN} + 2) + a_2 e_2(2^{mN} + 2) + \cdots + a_m e_m(2^{mN} + 2) = a_1.$$

Similarly, $2^{mN} + 2^2$ is divisible by 2^2 but not by $2^3, ..., 2^m$ and thus

$$\begin{aligned} 0 = e(2^{mN} + 2^2) = a_2 e_2(2^{mN} + 2^2) &+ a_3 e_3(2^{mN} + 2^2) \\ &+ \cdots + a_m e_m(2^{mN} + 2^2) = a_2. \end{aligned}$$

Continuing in this manner, we find that $a_1 = \cdots = a_m = 0$ and our claim follows.

We can now use Exercise 2.A.17 to conclude that $\mathbf{F}^\infty/U$ is infinite-dimensional.

Exercise 3.E.12. Suppose $v_1, \ldots, v_m \in V$. Let

$$A = \{\lambda_1 v_1 + \cdots + \lambda_m v_m : \lambda_1, \ldots, \lambda_m \in \mathbf{F} \text{ and } \lambda_1 + \cdots + \lambda_m = 1\}.$$

(a) Prove that A is a translate of some subspace of V.

(b) Prove that if B is a translate of some subspace of V and $\{v_1, \ldots, v_m\} \subseteq B$, then $A \subseteq B$.

(c) Prove that A is a translate of some subspace of V of dimension less than m.

Solution.

(a) Suppose $v = \sum_{k=1}^m \lambda_k v_k$ and $w = \sum_{k=1}^m \mu_k v_k$ belong to A and $\gamma \in \mathbf{F}$. Observe that

$$\sum_{k=1}^m [\gamma\lambda_k + (1-\gamma)\mu_k] = \gamma \sum_{k=1}^m \lambda_k + (1-\gamma) \sum_{k=1}^m \mu_k = \gamma + (1-\gamma) = 1.$$

It follows that

$$\gamma v + (1-\gamma) w = \gamma \sum_{k=1}^m \lambda_k v_k + (1-\gamma) \sum_{k=1}^m \mu_k v_k = \sum_{k=1}^m [\gamma\lambda_k + (1-\gamma)\mu_k] v_k$$

belongs to A. Thus A is a translate of some subspace of V by Exercise 3.E.9.

(b) Suppose that $B = v + U$ for some $v \in U$ and some subspace U of V, and suppose that $\{v_1, \ldots, v_m\} \subseteq B$, so that each $v_k = v + u_k$ for some $u_k \in U$. Let $\sum_{k=1}^m \lambda_k v_k \in A$ be given and observe that

$$\sum_{k=1}^m \lambda_k v_k = \sum_{k=1}^m \lambda_k (v + u_k) = \left(\sum_{k=1}^m \lambda_k\right) v + \sum_{k=1}^m \lambda_k u_k = v + \sum_{k=1}^m \lambda_k u_k \in v + U = B.$$

Thus $A \subseteq B$.

(c) If $m = 1$ then $A = \{v_1\} = v_1 + \{0\}$. If $m \geq 2$ then let $U = \operatorname{span}(v_2 - v_1, \ldots, v_m - v_1)$ and note that $\dim U \leq m - 1$. Note further that $v_1 + U$ is a translate of U containing $\{v_1, \ldots, v_m\}$; it follows from part (b) that $A \subseteq v_1 + U$. Let $v_1 + \sum_{k=2}^m a_k (v_k - v_1)$ in $v_1 + U$ be given and observe that

$$v_1 + \sum_{k=2}^m a_k (v_k - v_1) = \left(1 - \sum_{k=2}^m a_k\right) v_1 + \sum_{k=2}^m a_k v_k \in A.$$

Thus $v_1 + U \subseteq A$ and we may conclude that $A = v_1 + U$, where $\dim U \leq m - 1$.

Exercise 3.E.13. Suppose U is a subspace of V such that V/U is finite-dimensional. Prove that V is isomorphic to $U \times (V/U)$.

Solution. Let $v_1 + U, \ldots, v_m + U$ be a basis of V/U. Define $T \in \mathcal{L}(V/U, V)$ by $T(v_k + U) = v_k$ and notice that $\pi \circ T$ is the identity map on V/U. Now define

$S : U \times (V/U) \to V$ by $S(u, v+U) = u + T(v+U)$; the linearity of S follows from the linearity of T.

Suppose $(u, v+U)$ is such that

$$S(u, v+U) = u + T(v+U) = 0.$$

Using $\operatorname{null} \pi = U$ and $\pi \circ T = I$, applying π to the equation above shows that $v + U = 0$. It follows that $T(v+U) = 0$ and hence that $u = 0$. Thus S is injective.

For any $v \in V$ there are scalars $a_1, ..., a_m$ such that $v + U = a_1 v_1 + \cdots + a_m v_m + U$ and thus by 3.101 we have $v = u + a_1 v_1 + \cdots + a_m v_m$ for some $u \in U$. Observe that

$$\begin{aligned} S(u, a_1 v_1 + \cdots + a_m v_m + U) &= u + T(a_1 v_1 + \cdots + a_m v_m + U) \\ &= u + a_1 v_1 + \cdots + a_m v_m = v. \end{aligned}$$

Thus S is surjective and we may conclude that S is an isomorphism from $U \times (V/U)$ to V.

Exercise 3.E.14. Suppose U and W are subspaces of V and $V = U \oplus W$. Suppose $w_1, ..., w_m$ is a basis of W. Prove that $w_1 + U, ..., w_m + U$ is a basis of V/U.

Solution. Suppose $a_1, ..., a_m$ are scalars such that $a_1 w_1 + \cdots + a_m w_m + U = 0$, which by 3.101 is the case if and only if $a_1 w_1 + \cdots + a_m w_m \in U$. Because the sum $U \oplus W$ is direct, 1.46 shows that $a_1 w_1 + \cdots + a_m w_m = 0$ and it follows from the linear independence of $w_1, ..., w_m$ that $a_1 = \cdots = a_m = 0$. Thus $w_1 + U, ..., w_m + U$ is linearly independent.

For any $v + U \in V/U$, we have $v = u + a_1 w_1 + \cdots + a_m w_m$ for some $u \in U$ and some scalars $a_1, ..., a_m$. It follows that

$$v + U = \pi(v) = a_1 w_1 + \cdots + a_m w_m + U = a_1(w_1 + U) + \cdots + a_m(w_m + U).$$

Thus $w_1 + U, ..., w_m + U$ spans V/U and we may conclude that $w_1 + U, ..., w_m + U$ is a basis of V/U.

Exercise 3.E.15. Suppose U is a subspace of V and $v_1 + U, ..., v_m + U$ is a basis of V/U and $u_1, ..., u_n$ is a basis of U. Prove that $v_1, ..., v_m, u_1, ..., u_n$ is a basis of V.

Solution. Suppose there are scalars $a_1, ..., a_m, b_1, ..., b_n$ such that

$$a_1 v_1 + \cdots + a_m v_m + b_1 u_1 + \cdots + b_n u_n = 0.$$

Applying π to both sides of this equation shows that $a_1(v_1 + U) + \cdots + a_m(v_m + U) = 0$. The linear independence of $v_1 + U, ..., v_m + U$ then implies that $a_1 = \cdots = a_m = 0$, and the linear independence of $u_1, ..., u_n$ then gives us $b_1 = \cdots = b_n = 0$. Thus $v_1, ..., v_m, u_1, ..., u_n$ is linearly independent.

Let $v \in V$ be given. There are scalars $a_1, ..., a_m$ such that $v + U = a_1 v_1 + \cdots + a_m v_m + U$. It follows from 3.101 that $v = a_1 v_1 + \cdots + a_m v_m + u$ for some $u \in U$, so that there are scalars $b_1, ..., b_n$ such that

$$v = a_1v_1 + \cdots + a_mv_m + b_1u_1 + \cdots + b_nu_n.$$

Thus $v_1, \ldots, v_m, u_1, \ldots, u_n$ spans V and we may conclude that $v_1, \ldots, v_m, u_1, \ldots, u_n$ is a basis of V.

Exercise 3.E.16. Suppose $\varphi \in \mathcal{L}(V, \mathbf{F})$ and $\varphi \neq 0$. Prove that $\dim V/(\operatorname{null}\varphi) = 1$.

Solution. There is some $u \in V$ such that $\varphi(u) \neq 0$; Exercise 3.B.30 then implies that

$$V = \operatorname{null}\varphi \oplus \{au : a \in \mathbf{F}\}.$$

Notice that u is a basis of $\{au : a \in \mathbf{F}\}$. It follows from Exercise 3.E.14 that $u + \operatorname{null}\varphi$ is a basis of $V/(\operatorname{null}\varphi)$. Thus $\dim V/(\operatorname{null}\varphi) = 1$.

Exercise 3.E.17. Suppose U is a subspace of V such that $\dim V/U = 1$. Prove that there exists $\varphi \in \mathcal{L}(V, \mathbf{F})$ such that $\operatorname{null}\varphi = U$.

Solution. Let $w + U$ be a basis of V/U. For any $v \in V$ there is a unique $a \in \mathbf{F}$ such that $v + U = aw + U$. Given this uniqueness, the map $\varphi : V \to \mathbf{F}$ defined by $\varphi(v) = a$ is well-defined. Moreover, φ is linear. For any $v_1, v_2 \in V$ there are unique scalars a_1, a_2 such that $v_1 + U = a_1w + U$ and $v_2 + U = a_2w + U$. Let $\lambda \in \mathbf{F}$ be given. Since

$$(v_1 + v_2) + U = (a_1 + a_2)w + U \quad \text{and} \quad \lambda v_1 + U = (\lambda a_1)w + U,$$

$a_1 + a_2$ must be the unique coefficient of w in the linear combination representing $(v_1 + v_2) + U$ in the basis $w + U$ of V/U; similarly, λa_1 must be the unique coefficient of w in the representation of $\lambda v_1 + U$. It follows that

$$\varphi(v_1 + v_2) = a_1 + a_2 = \varphi(v_1) + \varphi(v_2) \quad \text{and} \quad \varphi(\lambda v_1) = \lambda a_1 = \lambda\varphi(v_1).$$

Thus φ is linear.

For any $u \in U$ we have $u + U = 0w + U$ and thus $\varphi(u) = 0$, so that $U \subseteq \operatorname{null}\varphi$. Conversely, if $v \in V$ is such that $\varphi(v) = 0$ then $v + U = 0w + U$; it follows from 3.101 that $v \in U$. Thus $\operatorname{null}\varphi = U$.

Exercise 3.E.18. Suppose that U is a subspace of V such that V/U is finite-dimensional.

(a) Show that if W is a finite-dimensional subspace of V and $V = U + W$, then $\dim W \geq \dim V/U$.

(b) Prove that there exists a finite-dimensional subspace W of V such that $\dim W = \dim V/U$ and $V = U \oplus W$.

Solution.

(a) Let $w_1, \ldots, w_m$ be a basis of W. For any $v + U \in V/U$ we have

$$v = u + a_1 w_1 + \cdots + a_m w_m$$

for some $u \in U$ and some scalars $a_1, ..., a_m$, so that

$$v + U = a_1(w_1 + U) + \cdots + a_m(w_m + U).$$

Thus $w_1 + U, ..., w_m + U$ spans V/U. It follows that $\dim V/U \le m = \dim W$.

(b) Let $v_1 + U, ..., v_m + U$ be a basis of V/U and define $T \in \mathcal{L}(V/U, V)$ by $T(v_k + U) = v_k$. Notice that

$$\pi \circ T = I \quad \text{and} \quad \operatorname{range} T = \operatorname{span}(v_1, ..., v_m).$$

It follows from Exercise 3.B.19 that T is injective; letting $W = \operatorname{range} T$, we see that T is an isomorphism between V/U and W. Thus $\dim W = \dim V/U$.

For any $v \in V$ we have $v + U = a_1 v_1 + \cdots + a_m v_m + U$ for some scalars $a_1, ..., a_m$. It follows from 3.101 that $v = u + a_1 v_1 + \cdots + a_m v_m$ for some $u \in U$. Thus $V = U + W$. Suppose that $v \in U \cap W$, so that $v + U = 0$ and $v = a_1 v_1 + \cdots + a_m v_m$ for some scalars $a_1, ..., a_m$. It follows that

$$0 = v + U = a_1(v_1 + U) + \cdots + a_m(v_m + U).$$

The linear independence of $v_1 + U, ..., v_m + U$ then implies that $a_1 = \cdots = a_m = 0$, whence $v = 0$. Thus $U \cap W = \{0\}$ and we may use 1.46 to conclude that the sum $V = U \oplus W$ is direct.

Exercise 3.E.19. Suppose $T \in \mathcal{L}(V, W)$ and U is a subspace of V. Let π denote the quotient map from V onto V/U. Prove that there exists $S \in \mathcal{L}(V/U, W)$ such that $T = S \circ \pi$ if and only if $U \subseteq \operatorname{null} T$.

Solution. If there exists such a map S and $u \in U$, then

$$Tu = S(\pi(u)) = S(0) = 0.$$

Thus $U \subseteq \operatorname{null} T$.

Now suppose $U \subseteq \operatorname{null} T$. Define $S : V/U \to W$ by $S(v + U) = Tv$. This map is well-defined:

$$v_1 + U = v_2 + U \quad \Leftrightarrow \quad v_1 - v_2 \in U \quad \Rightarrow \quad v_1 - v_2 \in \operatorname{null} T \quad \Rightarrow \quad Tv_1 = Tv_2.$$

The linearity of S follows from the linearity of T and the definition of S makes it clear that $T = S \circ \pi$.

3.F. Duality

Exercise 3.F.1. Explain why each linear functional is surjective or is the zero map.

Solution. Suppose $\varphi \in V'$ is non-zero, so that there is some $v \in V$ such that $\varphi(v) \neq 0$, and notice that

$$\varphi\left(\frac{\lambda}{\varphi(v)} v\right) = \lambda$$

for any $\lambda \in \mathbf{F}$. Thus φ is surjective.

Exercise 3.F.2. Give three distinct examples of linear functionals on $\mathbf{R}^{[0,1]}$.

Solution. For $k \in \{0, 1, 2\}$, define $\varphi_k : \mathbf{R}^{[0,1]} \to \mathbf{R}$ by $\varphi_k(f) = f\left(\frac{k}{2}\right)$. It is straightforward to check that each φ_k is a linear functional on $\mathbf{R}^{[0,1]}$.

Exercise 3.F.3. Suppose V is finite-dimensional and $v \in V$ with $v \neq 0$. Prove that there exists $\varphi \in V'$ such that $\varphi(v) = 1$.

Solution. Let $v_1 = v$ and extend this to a basis $v_1, ..., v_n$ of V. Now define $\varphi \in V'$ by $\varphi(v_k) = 1$ for each $k \in \{1, ..., n\}$.

Exercise 3.F.4. Suppose V is finite-dimensional and U is a subspace of V such that $U \neq V$. Prove that there exists $\varphi \in V'$ such that $\varphi(u) = 0$ for every $u \in U$ but $\varphi \neq 0$.

Solution. Let $u_1, ..., u_m$ be a basis of U and extend this to a basis $u_1, ..., u_m, v_1, ..., v_n$ of V; there must be at least one v_k since $U \neq V$. Define $\varphi \in V'$ by

$$\varphi(u_1) = \cdots = \varphi(u_m) = 0 \quad \text{and} \quad \varphi(v_1) = \cdots = \varphi(v_n) = 1.$$

It follows that $\varphi(u) = 0$ for all $u \in U$ but $\varphi \neq 0$.

Exercise 3.F.5. Suppose $T \in \mathcal{L}(V, W)$ and $w_1, ..., w_m$ is a basis of $\operatorname{range} T$. Hence for each $v \in V$, there exist unique numbers $\varphi_1(v), ..., \varphi_m(v)$ such that

$$Tv = \varphi_1(v) w_1 + \cdots + \varphi_m(v) w_m,$$

thus defining functions $\varphi_1, ..., \varphi_m$ from V to $\mathbf{F}$. Show that each of the functions $\varphi_1, ..., \varphi_m$ is a linear functional on V.

Solution. Let us think of T as a linear map $V \to \operatorname{range} T$, so that the dual map T' is a linear map $(\operatorname{range} T)' \to V'$. Let $\psi_1, ..., \psi_m$ be the dual basis to $w_1, ..., w_m$. For any $v \in V$, the definition of each φ_k and 3.114 show that

$$\varphi_1(v)w_1 + \cdots + \varphi_m(v)w_m = Tv = \psi_1(Tv)w_1 + \cdots + \psi_m(Tv)w_m.$$

It follows from unique representation that $\varphi_k(v) = \psi_k(Tv)$ for all $v \in V$, i.e.

$$\varphi_k = \psi_k \circ T = T'(\psi_k) \in V'.$$

Exercise 3.F.6. Suppose $\varphi, \beta \in V'$. Prove that $\operatorname{null}\varphi \subseteq \operatorname{null}\beta$ if and only if there exists $c \in \mathbf{F}$ such that $\beta = c\varphi$.

Solution. By Exercise 3.B.25 we have $\operatorname{null}\varphi \subseteq \operatorname{null}\beta$ if and only if there exists $E \in \mathcal{L}(\mathbf{F})$ such that $\beta = E \circ \varphi$. If there exists such an E, then let $c = E(1)$ and observe that, for any $v \in V$,

$$\beta(v) = E(\varphi(v)) = E(1)\varphi(v) = c\varphi(v) \quad \Rightarrow \quad \beta = c\varphi.$$

Conversely, if there exists such a $c \in \mathbf{F}$ then define $E \in \mathcal{L}(\mathbf{F})$ by $E(x) = cx$ and observe that, for any $v \in V$,

$$\beta(v) = c\varphi(v) = E(\varphi(v)) \quad \Rightarrow \quad \beta = E \circ \varphi.$$

Exercise 3.F.7. Suppose that $V_1, \ldots, V_m$ are vector spaces. Prove that $(V_1 \times \cdots \times V_m)'$ and $V_1' \times \cdots \times V_m'$ are isomorphic vector spaces.

Solution. This is immediate from Exercise 3.E.3, taking $W = \mathbf{F}$.

Exercise 3.F.8. Suppose $v_1, \ldots, v_n$ is a basis of V and $\varphi_1, \ldots, \varphi_n$ is the dual basis of V'. Define $\Gamma : V \to \mathbf{F}^n$ and $\Lambda : \mathbf{F}^n \to V$ by

$$\Gamma(v) = (\varphi_1(v), \ldots, \varphi_n(v)) \quad \text{and} \quad \Lambda(a_1, \ldots, a_n) = a_1v_1 + \cdots + a_nv_n.$$

Explain why Γ and Λ are inverses of each other.

Solution. It is straightforward to verify that Γ and Λ are linear. For any $(a_1, \ldots, a_n) \in \mathbf{F}^n$, 3.114 shows that

$$\begin{aligned}\Gamma(\Lambda(a_1, \ldots, a_n)) &= \Gamma(a_1v_1 + \cdots + a_nv_n) \\ &= (\varphi_1(a_1v_1 + \cdots + a_nv_n), \ldots, \varphi_n(a_1v_1 + \cdots + a_nv_n)) = (a_1, \ldots, a_n).\end{aligned}$$

Thus $\Gamma\Lambda$ is the identity on $\mathbf{F}^n$. For any $v \in V$, 3.114 gives us

$$\Lambda(\Gamma(v)) = \Lambda(\varphi_1(v), \ldots, \varphi_n(v)) = \varphi_1(v)v_1 + \cdots + \varphi_n(v)v_n = v.$$

Thus $\Lambda\Gamma$ is the identity on V and we may conclude that Γ and Λ are inverses of each other.

Exercise 3.F.9. Suppose m is a positive integer. Show that the dual basis of the basis $1, x, ..., x^m$ of $\mathcal{P}_m(\mathbf{R})$ is $\varphi_0, \varphi_1, ..., \varphi_m$, where

$$\varphi_k(p) = \frac{p^{(k)}(0)}{k!}.$$

Here $p^{(k)}$ denotes the k^{th} derivative of p, with the understanding that the 0^{th} derivative of p is p.

Solution. Let $k \in \{0, ..., m\}$ be given and observe that, for $j \neq k$,

$$\varphi_k(x^k) = \frac{(x^k)^{(k)}(0)}{k!} = \frac{k!}{k!} = 1 \quad \text{and} \quad \varphi_k(x^j) = \frac{(x^j)^{(k)}(0)}{k!} = \frac{0}{k!} = 0.$$

Thus the linear functionals φ_k and $p \mapsto \frac{p^{(k)}(0)}{k!}$ agree on the basis $1, x, ..., x^m$; it follows that they are equal as functions.

Exercise 3.F.10. Suppose m is a positive integer.

(a) Show that $1, x-5, ..., (x-5)^m$ is a basis of $\mathcal{P}_m(\mathbf{R})$.

(b) What is the dual basis of the basis in (a)?

Solution.

(a) This is immediate from Exercise 2.C.9.

(b) We can argue as we did in Exercise 3.F.9 to see that the dual basis is $\varphi_0, ..., \varphi_m$, where

$$\varphi_k(p) = \frac{p^{(k)}(5)}{k!}.$$

Exercise 3.F.11. Suppose $v_1, ..., v_n$ is a basis of V and $\varphi_1, ..., \varphi_n$ is the corresponding dual basis of V'. Suppose $\psi \in V'$. Prove that

$$\psi = \psi(v_1)\varphi_1 + \cdots + \psi(v_n)\varphi_n.$$

Solution. Let $v = a_1v_1 + \cdots + a_nv_n \in V$ be given and note that, by 3.114, we have $\varphi_k(a_1v_1 + \cdots + a_nv_n) = a_k$. It follows that

$$\begin{aligned}
[\psi(v_1)\varphi_1 + \cdots + \psi(v_n)\varphi_n](a_1v_1 + \cdots + a_nv_n) &= \psi(v_1)\varphi_1(a_1v_1 + \cdots + a_nv_n) \\
&\quad + \cdots + \psi(v_n)\varphi_n(a_1v_1 + \cdots + a_nv_n) \\
&= \psi(v_1)a_1 + \cdots + \psi(v_n)a_n \\
&= \psi(a_1v_1 + \cdots + a_nv_n).
\end{aligned}$$

Thus $\psi(v_1)\varphi_1 + \cdots + \psi(v_n)\varphi_n = \psi$.

Exercise 3.F.12. Suppose $S, T \in \mathcal{L}(V, W)$.

(a) Prove that $(S+T)' = S' + T'$.

(b) Prove that $(\lambda T)' = \lambda T'$ for all $\lambda \in \mathbf{F}$.

This exercise asks you to verify (a) and (b) in 3.120.

Solution.

(a) For any $\psi \in W'$ and $v \in V$, observe that

$$\begin{aligned}\big[(S+T)'(\psi)\big](v) &= \psi((S+T)(v)) = \psi(Sv+Tv) = \psi(Sv)+\psi(Tv)\\ &= [S'(\psi)](v)+[T'(\psi)](v) = [S'(\psi)+T'(\psi)](v) = [(S'+T')(\psi)](v).\end{aligned}$$

Thus $(S+T)' = S' + T'$.

(b) For any $\psi \in W'$ and $v \in V$, observe that

$$\begin{aligned}\big[(\lambda T)'(\psi)\big](v) &= \psi((\lambda T)(v)) = \psi(\lambda Tv) = \lambda\psi(Tv)\\ &= \lambda[T'(\psi)](v) = [\lambda T'(\psi)](v) = [(\lambda T')(\psi)](v).\end{aligned}$$

Thus $(\lambda T)' = \lambda T'$.

Exercise 3.F.13. Show that the dual map of the identity operator on V is the identity operator on V'.

Solution. For any $\varphi \in V'$ and $v \in V$, observe that

$$[I'(\varphi)](v) = \varphi(Iv) = \varphi(v).$$

Thus $I'(\varphi) = \varphi$, i.e. I' is the identity operator on V'.

Exercise 3.F.14. Define $T : \mathbf{R}^3 \to \mathbf{R}^2$ by

$$T(x,y,z) = (4x+5y+6z, 7x+8y+9z).$$

Suppose φ_1, φ_2 denotes the dual basis of the standard basis of $\mathbf{R}^2$ and ψ_1, ψ_2, ψ_3 denotes the dual basis of the standard basis of $\mathbf{R}^3$.

(a) Describe the linear functionals $T'(\varphi_1)$ and $T'(\varphi_2)$.

(b) Write $T'(\varphi_1)$ and $T'(\varphi_2)$ as linear combinations of ψ_1, ψ_2, ψ_3.

Solution.

(a) By the definition of the dual map, we have

$$[T'(\varphi_1)](x,y,z) = \varphi_1(T(x,y,z)) = \varphi_1(4x+5y+6z, 7x+8y+9z) = 4x+5y+6z,$$
$$[T'(\varphi_2)](x,y,z) = \varphi_2(T(x,y,z)) = \varphi_2(4x+5y+6z, 7x+8y+9z) = 7x+8y+9z.$$

(b) Note that

$$\psi_1(x,y,z) = x, \quad \psi_2(x,y,z) = y, \quad \text{and} \quad \psi_3(x,y,z) = z.$$

Thus

$$T'(\varphi_1) = 4\psi_1 + 5\psi_2 + 6\psi_3 \quad \text{and} \quad T'(\varphi_2) = 6\psi_1 + 7\psi_2 + 8\psi_3.$$

Exercise 3.F.15. Define $T : \mathcal{P}(\mathbf{R}) \to \mathcal{P}(\mathbf{R})$ by

$$(Tp)(x) = x^2 p(x) + p''(x)$$

for each $x \in \mathbf{R}$.

(a) Suppose $\varphi \in \mathcal{P}(\mathbf{R})'$ is defined by $\varphi(p) = p'(4)$. Describe the linear functional $T'(\varphi)$ on $\mathcal{P}(\mathbf{R})$.

(b) Suppose $\varphi \in \mathcal{P}(\mathbf{R})'$ is defined by $\varphi(p) = \int_0^1 p$. Evaluate $(T'(\varphi))(x^3)$.

Solution.

(a) We have

$$\begin{aligned}[T'(\varphi)](p) = \varphi(Tp) = \varphi(x^2 p + p'') = (x^2 p + p'')'(4) \\ = (2xp + x^2 p' + p''')(4) = 8p(4) + 16p'(4) + p'''(4).\end{aligned}$$

(b) We have

$$[T'(\varphi)](x^3) = \varphi(Tx^3) = \varphi(x^5 + 6x) = \int_0^1 x^5 + 6x \,\mathrm{d}x = \frac{19}{6}.$$

Exercise 3.F.16. Suppose W is finite-dimensional and $T \in \mathcal{L}(V,W)$. Prove that

$$T' = 0 \quad \Leftrightarrow \quad T = 0.$$

Solution. If $T = 0$ and $\psi \in W'$, then

$$T'(\psi) = \psi \circ T = \psi \circ 0 = 0.$$

Thus $T' = 0$.

Now suppose that $T' = 0$, so that $\operatorname{null} T' = W'$. It follows from 3.128(a) that $(\operatorname{range} T)^0 = W'$, which by 3.127(b) is equivalent to $\operatorname{range} T = \{0\}$. Thus $T = 0$.

Exercise 3.F.17. Suppose V and W are finite-dimensional and $T \in \mathcal{L}(V,W)$. Prove that T is invertible if and only if $T' \in \mathcal{L}(W', V')$ is invertible.

Solution. If either of T, T' is invertible then, using 3.111, it must be the case that

$$\dim V' = \dim V = \dim W = \dim W'.$$

Thus, by 3.65 and 3.129,

$$T \text{ is invertible} \quad \Leftrightarrow \quad T \text{ is surjective} \quad \Leftrightarrow \quad T' \text{ is injective} \quad \Leftrightarrow \quad T' \text{ is invertible.}$$

Exercise 3.F.18. Suppose V and W are finite-dimensional. Prove that the map that takes $T \in \mathcal{L}(V, W)$ to $T' \in \mathcal{L}(W', V')$ is an isomorphism of $\mathcal{L}(V, W)$ onto $\mathcal{L}(W', V')$.

Solution. Let $\Phi : \mathcal{L}(V, W) \to \mathcal{L}(W', V')$ be the map in question, i.e. $\Phi(T) = T'$. Exercise 3.F.12 shows that Φ is linear and Exercise 3.F.16 shows that Φ is injective. Note that, by 3.72 and 3.111,

$$\dim \mathcal{L}(V, W) = (\dim V)(\dim W) = (\dim V')(\dim W') = \dim \mathcal{L}(W', V').$$

3.65 allows us to conclude that Φ is an isomorphism.

Exercise 3.F.19. Suppose $U \subseteq V$. Explain why

$$U^0 = \{\varphi \in V' : U \subseteq \operatorname{null} \varphi\}.$$

Solution. This is immediate from the equivalence

$$\varphi(u) = 0 \text{ for all } u \in U \quad \Leftrightarrow \quad U \subseteq \operatorname{null} \varphi.$$

Exercise 3.F.20. Suppose V is finite-dimensional and U is a subspace of V. Show that

$$U = \{v \in V : \varphi(v) = 0 \text{ for every } \varphi \in U^0\}.$$

Solution. If $v \in U$ then certainly $\varphi(v) = 0$ for every $\varphi \in U^0$. Suppose $v \notin U$ and let $u_1, \ldots, u_m$ be a basis of U. Let $v_1 = v$ and note that the list $u_1, \ldots, u_m, v_1$ is linearly independent since $v_1 \notin U = \operatorname{span}(u_1, \ldots, u_m)$. Extend this list to a basis $u_1, \ldots, u_m, v_1, \ldots, v_n$ of V and define $\varphi \in V'$ by

$$\varphi(u_1) = \cdots = \varphi(u_m) = 0 \quad \text{and} \quad \varphi(v_1) = \cdots = \varphi(v_n) = 1.$$

It follows that $\varphi \in U^0$ and $\varphi(v) \neq 0$. Thus

$$v \in U \quad \Leftrightarrow \quad v \in \{v \in V : \varphi(v) = 0 \text{ for every } \varphi \in U^0\}.$$

Exercise 3.F.21. Suppose V is finite-dimensional and U and W are subspaces of V.

(a) Prove that $W^0 \subseteq U^0$ if and only if $U \subseteq W$.

(b) Prove that $W^0 = U^0$ if and only if $U = W$.

Solution. For a subspace U of V, let

$$A_U = \{v \in V : \varphi(v) = 0 \text{ for every } \varphi \in U^0\}.$$

(a) If $W^0 \subseteq U^0$ and $v \in A_U$ then in particular $\varphi(v) = 0$ for every $\varphi \in W^0$, i.e. $v \in A_W$. Thus $A_U \subseteq A_W$, which by Exercise 3.F.21 is equivalent to $U \subseteq W$. If $U \subseteq W$ and $\varphi \in W^0$ then in particular $\varphi(v) = 0$ for all $v \in U$, i.e. $\varphi \in U^0$. Thus $W^0 \subseteq U^0$.

(b) This follows from part (a):

$$W^0 = U^0 \quad \Leftrightarrow \quad W^0 \subseteq U^0 \text{ and } U^0 \subseteq W^0 \quad \Leftrightarrow \quad U \subseteq W \text{ and } W \subseteq U \quad \Leftrightarrow \quad U = W.$$

Exercise 3.F.22. Suppose V is finite-dimensional and U and W are subspaces of V.

(a) Show that $(U + W)^0 = U^0 \cap W^0$.

(b) Show that $(U \cap W)^0 = U^0 + W^0$.

Solution.

(a) Suppose $\varphi \in (U + W)^0$. For any $u \in U$ and $w \in W$ observe that $u, w \in U + W$, so that $\varphi(u) = \varphi(w) = 0$. Thus $\varphi \in U^0 \cap W^0$ and it follows that $(U + W)^0 \subseteq U^0 \cap W^0$.

Suppose $\varphi \in U^0 \cap W^0$ and observe that, for any $u + w \in U + W$, we have

$$\varphi(u + w) = \varphi(u) + \varphi(w) = 0 \quad \Rightarrow \quad \varphi \in (U + W)^0.$$

It follows that $U^0 \cap W^0 \subseteq (U + W)^0$ and thus $(U + W)^0 = U^0 \cap W^0$.

(b) Suppose $\varphi \in (U \cap W)^0$. There are subspaces X, Y of V such that $V = U \oplus X$ and $V = W \oplus Y$. Define $\psi, \beta \in V'$ by

$$\psi(u + x) = \tfrac{1}{2}\varphi(x) \quad \text{and} \quad \beta(w + y) = \tfrac{1}{2}\varphi(y).$$

It is straightforward to verify that $\psi \in U^0$ and $\beta \in W^0$. Let $v = u + x = w + y \in V$ be given and note that, because $\varphi \in (U \cap W)^0$, $\varphi(v) = \varphi(x) = \varphi(y)$. It follows that

$$\varphi(v) = \tfrac{1}{2}\varphi(v) + \tfrac{1}{2}\varphi(v) = \tfrac{1}{2}\varphi(x) + \tfrac{1}{2}\varphi(y) = \psi(v) + \beta(v).$$

Thus $\varphi \in U^0 + W^0$, whence $(U \cap W)^0 \subseteq U^0 + W^0$.

Now suppose that $\varphi \in U^0 + W^0$, so that $\varphi = \psi + \beta$ for some $\psi \in U^0$ and some $\beta \in W^0$. Let $v \in U \cap W$ be given and observe that

$$\varphi(v) = \psi(v) + \beta(v) = 0 \quad \Rightarrow \quad \varphi \in (U \cap W)^0.$$

It follows that $U^0 + W^0 \subseteq (U \cap W)^0$ and thus $(U \cap W)^0 = U^0 + W^0$.

Exercise 3.F.23. Suppose V is finite-dimensional and $\varphi_1, \ldots, \varphi_m \in V'$. Prove that the following three sets are equal to each other.

(a) $\operatorname{span}(\varphi_1, \ldots, \varphi_m)$

(b) $((\operatorname{null} \varphi_1) \cap \cdots \cap (\operatorname{null} \varphi_m))^0$

(c) $\{\varphi \in V' : (\operatorname{null} \varphi_1) \cap \cdots \cap (\operatorname{null} \varphi_m) \subseteq \operatorname{null} \varphi\}$

Solution. (b) and (c) are equal by Exercise 3.F.19.

To show that (a) and (b) are equal, let us prove the following lemma.

Lemma L.3. If $\varphi \in V'$ then $\operatorname{span}(\varphi) \subseteq (\operatorname{null} \varphi)^0$; if V is finite-dimensional then this containment is an equality.

Proof. For any $a \in \mathbf{F}$ and $v \in \operatorname{null} \varphi$ we have $a\varphi(v) = 0$. Thus $\operatorname{span}(\varphi) \subseteq (\operatorname{null} \varphi)^0$. Now observe that $\dim \operatorname{span}(\varphi) = \dim \operatorname{range} \varphi$, since

$$\varphi = 0 \quad \Rightarrow \quad \dim \operatorname{span}(\varphi) = 0 = \dim \operatorname{range} \varphi,$$

$$\varphi \neq 0 \quad \Rightarrow \quad \dim \operatorname{span}(\varphi) = 1 = \dim \operatorname{range} \varphi,$$

where we have used Exercise 3.F.1 for the second implication. Assuming that V is finite-dimensional, we can use (3.21) and (3.125) to obtain the equality

$$\dim \operatorname{span}(\varphi) = \dim \operatorname{range} \varphi = \dim (\operatorname{null} \varphi)^0.$$

2.39 allows us to conclude that $\operatorname{span}(\varphi) = (\operatorname{null} \varphi)^0$. □

Using Exercise 3.F.22 and Lemma L.3, observe that

$$\begin{aligned}((\operatorname{null} \varphi_1) \cap \cdots \cap (\operatorname{null} \varphi_m))^0 &= (\operatorname{null} \varphi_1)^0 + \cdots + (\operatorname{null} \varphi_m)^0 \\ &= \operatorname{span}(\varphi_1) + \cdots + \operatorname{span}(\varphi_m) = \operatorname{span}(\varphi_1, \ldots, \varphi_m).\end{aligned}$$

Thus (a) and (b) are equal.

Exercise 3.F.24. Suppose V is finite-dimensional and $v_1, \ldots, v_m \in V$. Define a linear map $\Gamma : V' \to \mathbf{F}^m$ by $\Gamma(\varphi) = (\varphi(v_1), \ldots, \varphi(v_m))$.

(a) Prove that $v_1, \ldots, v_m$ spans V if and only if Γ is injective.

(b) Prove that $v_1, \ldots, v_m$ is linearly independent if and only if Γ is surjective.

Solution. Let $e_1, \ldots, e_m$ be the standard basis of $\mathbf{F}^m$ and let $\psi_1, \ldots, \psi_m$ be the corresponding dual basis of $(\mathbf{F}^m)'$, so that the map $\Psi : \mathbf{F}^m \to (\mathbf{F}^m)'$ given by $\Psi(e_k) = \psi_k$ is an isomorphism. Define $T \in \mathcal{L}(\mathbf{F}^m, V)$ by $Te_k = v_k$, i.e.

$$T(x_1, \ldots, x_m) = x_1 v_1 + \cdots + x_m v_m.$$

For any $\varphi \in V'$ and $k \in \{1, \ldots, m\}$, observe that

$$\begin{aligned}[T'(\varphi)](e_k) = \varphi(Te_k) = \varphi(v_k) &= \sum_{j=1}^{m} \varphi(v_j)\psi_j(e_k) \\ &= [\Psi(\varphi(v_1), \ldots, \varphi(v_m))](e_k) = [\Psi(\Gamma(\varphi))](e_k).\end{aligned}$$

Thus $T' = \Psi \circ \Gamma$. Because Ψ is a bijection, it follows that the injectivity of Γ is equivalent to the injectivity of T' and the surjectivity of Γ is equivalent to the surjectivity of T'.

(a) Observe that

$$\operatorname{span}(v_1, \ldots, v_m) = V \quad \Leftrightarrow \quad T \text{ is surjective} \quad \Leftrightarrow \quad T' \text{ is injective} \quad \Leftrightarrow \quad \Gamma \text{ is injective},$$

where the first equivalence follows from Exercise 3.B.3, the second equivalence follows from 3.129, and the third equivalence follows from our previous discussion.

(b) Observe that

$$\begin{aligned} v_1, \ldots, v_m \text{ is linearly independent} \quad &\Leftrightarrow \quad T \text{ is injective} \\ &\Leftrightarrow \quad T' \text{ is surjective} \quad \Leftrightarrow \quad \Gamma \text{ is surjective}, \end{aligned}$$

where the first equivalence follows from Exercise 3.B.3, the second equivalence follows from 3.131, and the third equivalence follows from our previous discussion.

Exercise 3.F.25. Suppose V is finite-dimensional and $\varphi_1, \ldots, \varphi_m \in V'$. Define a linear map $\Gamma : V \to \mathbf{F}^m$ by $\Gamma(v) = (\varphi_1(v), \ldots, \varphi_m(v))$.

(a) Prove that $\varphi_1, \ldots, \varphi_m$ spans V' if and only if Γ is injective.

(b) Prove that $\varphi_1, \ldots, \varphi_m$ is linearly independent if and only if Γ is surjective.

Solution. Let $e_1, \ldots, e_m$ be the standard basis of $\mathbf{F}^m$ and let $\psi_1, \ldots, \psi_m$ be the corresponding dual basis of $(\mathbf{F}^m)'$, so that the map $\Psi : \mathbf{F}^m \to (\mathbf{F}^m)'$ given by $\Psi(e_k) = \psi_k$ is an isomorphism. For any $(x_1, \ldots, x_m) \in \mathbf{F}^m$ and $v \in V$, observe that

$$\begin{aligned} [\Gamma'(\Psi(x_1, \ldots, x_m))](v) &= [\Gamma'(x_1\psi_1 + \cdots + x_m\psi_m)](v) \\ &= [x_1\psi_1 + \cdots + x_m\psi_m](\Gamma(v)) \\ &= [x_1\psi_1 + \cdots + x_m\psi_m](\varphi_1(v), \ldots, \varphi_m(v)) \\ &= x_1\varphi_1(v) + \cdots + x_m\varphi_m(v) \\ &= [x_1\varphi_1 + \cdots + x_m\varphi_m](v). \end{aligned}$$

It follows that $\Gamma' \circ \Psi : \mathbf{F}^m \to V'$ is given by

$$\Gamma'(\Psi(x_1, \ldots, x_m)) = x_1\varphi_1 + \cdots + x_m\varphi_m.$$

Because Ψ is a bijection, the injectivity of Γ' is equivalent to the injectivity of $\Gamma' \circ \Psi$ and the surjectivity of Γ' is equivalent to the surjectivity of $\Gamma' \circ \Psi$.

(a) Observe that

$$\begin{aligned} \operatorname{span}(\varphi_1, \ldots, \varphi_m) = V' \quad &\Leftrightarrow \quad \Gamma' \circ \Psi \text{ is surjective} \\ &\Leftrightarrow \quad \Gamma' \text{ is surjective} \quad \Leftrightarrow \quad \Gamma \text{ is injective}, \end{aligned}$$

where the first equivalence follows from Exercise 3.B.3, the second equivalence follows from our previous discussion, and the third equivalence follows from 3.131.

(b) Observe that

$$\varphi_1, \dots, \varphi_m \text{ is linearly independent} \quad \Leftrightarrow \quad \Gamma' \circ \Psi \text{ is injective}$$
$$\Leftrightarrow \quad \Gamma' \text{ is injective} \quad \Leftrightarrow \quad \Gamma \text{ is surjective},$$

where the first equivalence follows from Exercise 3.B.3, the second equivalence follows from our previous discussion, and third equivalence follows from 3.129.

Exercise 3.F.26. Suppose V is finite-dimensional and Ω is a subspace of V'. Prove that

$$\Omega = \{v \in V : \varphi(v) = 0 \text{ for every } \varphi \in \Omega\}^0.$$

Solution. Let $U = \{v \in V : \varphi(v) = 0 \text{ for every } \varphi \in \Omega\}$ and let $\varphi_1, \dots, \varphi_m$ be a basis of Ω. Certainly $U \subseteq (\operatorname{null} \varphi_1) \cap \cdots \cap (\operatorname{null} \varphi_m)$. Suppose $v \in (\operatorname{null} \varphi_1) \cap \cdots \cap (\operatorname{null} \varphi_m)$ and let $\varphi \in \Omega$ be given. There are scalars $a_1, \dots, a_m$ such that $\varphi = a_1\varphi_1 + \cdots + a_m\varphi_m$, which gives us

$$\varphi(v) = (a_1\varphi_1 + \cdots + a_m\varphi_m)(v) = a_1\varphi_1(v) + \cdots + a_m\varphi_m(v) = 0.$$

Thus $v \in U$ and it follows that $U = (\operatorname{null} \varphi_1) \cap \cdots \cap (\operatorname{null} \varphi_m)$. We may now apply Exercise 3.F.23 to see that

$$U^0 = \left((\operatorname{null} \varphi_1) \cap \cdots \cap (\operatorname{null} \varphi_m)\right)^0 = \operatorname{span}(\varphi_1, \dots, \varphi_m) = \Omega.$$

Exercise 3.F.27. Suppose $T \in \mathcal{L}(\mathcal{P}_5(\mathbf{R}))$ and $\operatorname{null} T' = \operatorname{span}(\varphi)$, where φ is the linear functional on $\mathcal{P}_5(\mathbf{R})$ defined by $\varphi(p) = p(8)$. Prove that

$$\operatorname{range} T = \{p \in \mathcal{P}_5(\mathbf{R}) : p(8) = 0\}.$$

Solution. Observe that

$$\begin{aligned}
\operatorname{range} T &= \left\{p \in \mathcal{P}_5(\mathbf{R}) : \psi(p) = 0 \text{ for every } \psi \in (\operatorname{range} T)^0\right\} && \text{(Exercise 3.F.20)} \\
&= \{p \in \mathcal{P}_5(\mathbf{R}) : \psi(p) = 0 \text{ for every } \psi \in \operatorname{null} T'\} && (3.128) \\
&= \{p \in \mathcal{P}_5(\mathbf{R}) : \psi(p) = 0 \text{ for every } \psi \in \operatorname{span}(\varphi)\} \\
&= \{p \in \mathcal{P}_5(\mathbf{R}) : \lambda\varphi(p) = 0 \text{ for every } \lambda \in \mathbf{R}\} \\
&= \{p \in \mathcal{P}_5(\mathbf{R}) : \lambda p(8) = 0 \text{ for every } \lambda \in \mathbf{R}\} \\
&= \{p \in \mathcal{P}_5(\mathbf{R}) : p(8) = 0\}.
\end{aligned}$$

Exercise 3.F.28. Suppose V is finite-dimensional and $\varphi_1, \dots, \varphi_m$ is a linearly independent list in V'. Prove that

$$\dim((\operatorname{null} \varphi_1) \cap \cdots \cap (\operatorname{null} \varphi_m)) = (\dim V) - m.$$

Solution. By Exercise 3.F.23 and the linear independence of the list $\varphi_1, \dots, \varphi_m$ we have

$$\dim\left((\operatorname{null}\varphi_1)\cap\cdots\cap(\operatorname{null}\varphi_m)\right)^0=\dim\operatorname{span}(\varphi_1,\ldots,\varphi_m)=m.$$

It then follows from 3.125 that

$$\dim((\operatorname{null}\varphi_1)\cap\cdots\cap(\operatorname{null}\varphi_m))=(\dim V)-m.$$

Exercise 3.F.29. Suppose V and W are finite-dimensional and $T\in\mathcal{L}(V,W)$.

(a) Prove that if $\varphi\in W'$ and $\operatorname{null}T'=\operatorname{span}(\varphi)$, then $\operatorname{range}T=\operatorname{null}\varphi$.

(b) Prove that if $\psi\in V'$ and $\operatorname{range}T'=\operatorname{span}(\psi)$, then $\operatorname{null}T=\operatorname{null}\psi$.

Solution.

(a) Observe that

$$\begin{aligned}\operatorname{range}T&=\left\{w\in W:\psi(w)=0\text{ for every }\psi\in(\operatorname{range}T)^0\right\} &&(\text{Exercise 3.F.20})\\&=\{w\in W:\psi(w)=0\text{ for every }\psi\in\operatorname{null}T'\} &&(3.128)\\&=\{w\in W:\psi(w)=0\text{ for every }\psi\in\operatorname{span}(\varphi)\}\\&=\{w\in W:\lambda\varphi(w)=0\text{ for every }\lambda\in\mathbf{F}\}\\&=\{w\in W:\varphi(w)=0\}\\&=\operatorname{null}\varphi.\end{aligned}$$

(b) Observe that

$$\begin{aligned}\operatorname{null}T&=\left\{v\in V:\varphi(v)=0\text{ for every }\varphi\in(\operatorname{null}T)^0\right\} &&(\text{Exercise 3.F.20})\\&=\{v\in V:\varphi(v)=0\text{ for every }\varphi\in\operatorname{range}T'\} &&(3.130)\\&=\{v\in V:\varphi(v)=0\text{ for every }\varphi\in\operatorname{span}(\psi)\}\\&=\{v\in V:\lambda\psi(v)=0\text{ for every }\lambda\in\mathbf{F}\}\\&=\{v\in V:\psi(v)=0\}\\&=\operatorname{null}\psi.\end{aligned}$$

Exercise 3.F.30. Suppose V is finite-dimensional and $\varphi_1,\ldots,\varphi_n$ is a basis of V'. Show that there exists a basis of V whose dual basis is $\varphi_1,\ldots,\varphi_n$.

Solution. Let $k\in\{1,\ldots,n\}$ be given and define

$$U_k=\bigcap\left\{\operatorname{null}\varphi_j:j\in\{1,\ldots,n\}\setminus\{k\}\right\}.$$

By 3.111 and Exercise 3.F.28 we have $\dim U_k=1$ and thus there is some non-zero $u_k\in V$ such that $U_k=\operatorname{span}(u_k)$. Note that Exercise 3.F.28 also implies that

$$(\operatorname{null}\varphi_1) \cap \cdots \cap (\operatorname{null}\varphi_n) = \{0\}.$$

Since $u_k \neq 0$ it must then be the case that $u_k \notin \operatorname{null}\varphi_k$. Thus we can define $v_k = (\varphi_k(u_k))^{-1}u_k$. Notice that $\varphi_k(v_k) = 1$ and, for $j \neq k$,

$$u_k \in \operatorname{null}\varphi_j \quad \Rightarrow \quad \varphi_j(v_k) = 0.$$

If $a_1, \ldots, a_n$ are scalars such that $a_1v_1 + \cdots + a_nv_n = 0$ then, for each $j \in \{1, \ldots, n\}$, applying φ_j to both sides of this equation shows that $a_j = 0$. Thus $v_1, \ldots, v_n$ is a linearly independent list of length $n = \dim V$; it follows that $v_1, \ldots, v_n$ is a basis of V. Because

$$\varphi_j(v_k) = \begin{cases} 1 & \text{if } j = k, \\ 0 & \text{if } j \neq k, \end{cases}$$

the uniqueness part of the linear map lemma (3.4) shows that $\varphi_1, \ldots, \varphi_n$ is the dual basis of $v_1, \ldots, v_n$.

Exercise 3.F.31. Suppose U is a subspace of V. Let $i : U \to V$ be the inclusion map defined by $i(u) = u$. Thus $i' \in \mathcal{L}(V', U')$.

(a) Show that $\operatorname{null} i' = U^0$.

(b) Prove that if V is finite-dimensional, then $\operatorname{range} i' = U'$.

(c) Prove that if V is finite-dimensional, then $\tilde{i'}$ is an isomorphism from V'/U^0 onto U'.

The isomorphism in (c) is natural in that it does not depend on a choice of basis in either vector space.

Solution.

(a) For $\varphi \in V'$ the map $i'(\varphi) = \varphi \circ i$ is simply the restriction of φ to U. Thus

$$i'(\varphi) = 0 \quad \Leftrightarrow \quad \varphi(u) = 0 \text{ for all } u \in U.$$

It follows that $\operatorname{null} i' = U^0$.

(b) Let $\psi \in U'$ be given. By Exercise 3.A.13 we can extend ψ to a linear functional $\varphi \in V'$ such that the restriction of φ to U is equal to ψ. Thus $i'(\varphi) = \psi$ and it follows that i' is surjective.

(c) By 3.107 and parts (a) and (b), $\tilde{i'}$ is an isomorphism from $V'/\operatorname{null} i' = V'/U^0$ onto $\operatorname{range} i' = U'$.

Exercise 3.F.32. The *double dual space* of V, denoted by V'', is defined to be the dual space of V'. In other words, $V'' = (V')'$. Define $\Lambda : V \to V''$ by

$$(\Lambda v)(\varphi) = \varphi(v)$$

for each $v \in V$ and each $\varphi \in V'$.

(a) Show that Λ is a linear map from V to V''.

(b) Show that if $T \in \mathcal{L}(V)$, then $T'' \circ \Lambda = \Lambda \circ T$, where $T'' = (T')'$.

(c) Show that if V is finite-dimensional, then Λ is an isomorphism from V onto V''.

Suppose V is finite-dimensional. Then V and V' are isomorphic, but finding an isomorphism from V onto V' generally requires choosing a basis of V. In contrast, the isomorphism Λ from V onto V'' does not require a choice of basis and thus is considered more natural.

Solution.

(a) Suppose $u, v \in V$ and $\mu \in \mathbf{F}$. For any $\varphi \in V'$ we have

$$(\Lambda(u+v))(\varphi) = \varphi(u+v) = \varphi(u) + \varphi(v) = (\Lambda u)(\varphi) + (\Lambda v)(\varphi) = (\Lambda u + \Lambda v)(\varphi).$$

Thus $\Lambda(u+v) = \Lambda u + \Lambda v$. Similarly, for any $\varphi \in V'$,

$$(\Lambda(\mu v))(\varphi) = \varphi(\mu v) = \mu\varphi(v) = \mu(\Lambda v)(\varphi) = (\mu\Lambda v)(\varphi).$$

Thus $\Lambda(\mu v) = \mu\Lambda v$. It follows that Λ is linear.

(b) Let $v \in V$ be given and observe that

$$(T''(\Lambda v))(\varphi) = (\Lambda v)(T'(\varphi)) = (\Lambda v)(\varphi \circ T) = \varphi(Tv) \quad \text{and} \quad (\Lambda(Tv))(\varphi) = \varphi(Tv).$$

Thus $T'' \circ \Lambda = \Lambda \circ T$.

(c) Let $v_1, ..., v_n$ be a basis of V and let $\varphi_1, ..., \varphi_n$ be the corresponding dual basis of V'. Suppose $v = a_1 v_1 + \cdots + a_n v_n$ is such that $\Lambda v = 0$, i.e. $\varphi(v) = 0$ for every $\varphi \in V'$. For each $k \in \{1, ..., n\}$ it follows that

$$0 = \varphi_k(v) = \varphi_k(a_1 v_1 + \cdots + a_n v_n) = a_k.$$

Thus $v = 0$, so that $\operatorname{null}\Lambda = \{0\}$, i.e. Λ is injective. By 3.111 we have

$$\dim V'' = \dim V' = \dim V$$

and so, by 3.65, we may conclude that Λ is an isomorphism.

Exercise 3.F.33. Suppose U is a subspace of V. Let $\pi : V \to V/U$ be the usual quotient map. Thus $\pi' \in \mathcal{L}\big((V/U)', V'\big)$.

(a) Show that π' is injective.

(b) Show that $\operatorname{range} \pi' = U^0$.

(c) Conclude that π' is an isomorphism from $(V/U)'$ onto U^0.

The isomorphism in (c) is natural in that it does not depend on a choice of basis in either vector space. In fact, there is no assumption here that any of these vector spaces are finite-dimensional.

Solution.

(a) Suppose $\varphi \in (V/U)'$ is such that $\pi'(\varphi) = 0$, i.e. $\varphi(v + U) = 0$ for all $v + U \in V/U$. Thus $\varphi = 0$, so that $\operatorname{null} \pi' = \{0\}$. It follows that π' is injective.

(b) Note that $U^0 = \{\varphi \in V' : U \subseteq \operatorname{null} \varphi\}$ by Exercise 3.F.19. Taking $W = \mathbf{F}$ in Exercise 3.E.19 then shows that $\operatorname{range} \pi' = U^0$.

(c) This is immediate from parts (a) and (b).

Chapter 4. Polynomials

Exercise 4.1. Suppose $w, z \in \mathbf{C}$. Verify the following equalities and inequalities.

(a) $z + \overline{z} = 2 \operatorname{Re} z$

(b) $z - \overline{z} = 2(\operatorname{Im} z)i$

(c) $z\overline{z} = |z|^2$

(d) $\overline{w + z} = \overline{w} + \overline{z}$ and $\overline{wz} = \overline{w}\,\overline{z}$

(e) $\overline{\overline{z}} = z$

(f) $|\operatorname{Re} z| \le |z|$ and $|\operatorname{Im} z| \le |z|$

(g) $|\overline{z}| = |z|$

(h) $|wz| = |w||z|$

The results above are the parts of 4.4 that were left to the reader.

Solution. Suppose $w = a + bi$ and $z = x + yi$.

(a) Observe that

$$z + \overline{z} = (x + yi) + (x - yi) = 2x = 2 \operatorname{Re} z.$$

(b) Observe that

$$z - \overline{z} = (x + yi) - (x - yi) = 2yi = 2(\operatorname{Im} z)i.$$

(c) Observe that

$$z\overline{z} = (x + yi)(x - yi) = x^2 + y^2 = (\operatorname{Re} z)^2 + (\operatorname{Im} z)^2 = |z|^2.$$

(d) Observe that

$$\overline{w + z} = (a + x) - (b + y)i = (a - bi) + (x - yi) = \overline{w} + \overline{z},$$

$$\overline{wz} = (ax - by) - (ay + bx)i = (a - bi)(x - yi) = \overline{w}\,\overline{z}.$$

(e) Observe that

$$\overline{\overline{z}} = \overline{x - yi} = x + yi = z.$$

(f) Since each quantity involved is positive, it will suffice to show that $|\operatorname{Re} z|^2 \le |z|^2$ and $|\operatorname{Im}(z)|^2 \le |z|^2$; these inequalities are immediate from the equation $|z|^2 = |\operatorname{Re} z|^2 + |\operatorname{Im} z|^2$.

(g) Observe that

$$|\overline{z}| = (\operatorname{Re} \overline{z})^2 + (\operatorname{Im} \overline{z})^2 = (\operatorname{Re} z)^2 + (-\operatorname{Im} z)^2 = (\operatorname{Re} z)^2 + (\operatorname{Im} z)^2 = |z|.$$

(h) Since both sides are positive, it will suffice to show that $|wz|^2 = |w|^2|z|^2$. Indeed, using parts (c) and (d),

$$|wz|^2 = wz\overline{wz} = wz\overline{w}\,\overline{z} = w\overline{w}z\overline{z} = |w|^2|z|^2.$$

Exercise 4.2. Prove that if $w, z \in \mathbf{C}$, then $||w| - |z|| \leq |w - z|$.

The inequality above is called the ***reverse triangle inequality****.*

Solution. Notice that

$$|w| = |w - z + z| \leq |w - z| + |z| \quad \Rightarrow \quad |w| - |z| \leq |w - z|,$$
$$|z| = |z - w + w| \leq |z - w| + |w| \quad \Rightarrow \quad |z| - |w| \leq |w - z|.$$

Thus $||w| - |z|| \leq |w - z|$.

Exercise 4.3. Suppose V is a complex vector space and $\varphi \in V'$. Define $\sigma : V \to \mathbf{R}$ by $\sigma(v) = \operatorname{Re}\varphi(v)$ for each $v \in V$. Show that

$$\varphi(v) = \sigma(v) - i\sigma(iv)$$

for all $v \in V$.

Solution. For any $z = x + iy \in \mathbf{C}$, note that

$$\operatorname{Re}(iz) = \operatorname{Re}(-y + ix) = -y = -\operatorname{Im} z.$$

For any $v \in V$ it follows that

$$\sigma(v) - i\sigma(iv) = \operatorname{Re}\varphi(v) - i\operatorname{Re}(\varphi(iv)) = \operatorname{Re}\varphi(v) - i\operatorname{Re}(i\varphi(v)) = \operatorname{Re}\varphi(v) + \operatorname{Im}\varphi(v) = \varphi(v).$$

Exercise 4.4. Suppose m is a positive integer. Is the set

$$\{0\} \cup \{p \in \mathcal{P}(\mathbf{F}) : \deg p = m\}$$

a subspace of $\mathcal{P}(\mathbf{F})$?

Solution. Let U be the set in question and observe that $x^m, 1 - x^m \in U$ but

$$x^m + 1 - x^m = 1 \notin U.$$

Thus U is not a subspace of $\mathcal{P}(\mathbf{F})$.

Exercise 4.5. Is the set

$$\{0\} \cup \{p \in \mathcal{P}(\mathbf{F}) : \deg p \text{ is even}\}$$

a subspace of $\mathcal{P}(\mathbf{F})$?

Solution. Let U be the set in question and observe that $x^2, x - x^2 \in U$ but

$$x^2 + x - x^2 = x \notin U.$$

Thus U is not a subspace of $\mathcal{P}(\mathbf{F})$.

Exercise 4.6. Suppose that m and n are positive integers with $m \leq n$, and suppose $\lambda_1, \ldots, \lambda_m \in \mathbf{F}$. Prove that there exists a polynomial $p \in \mathcal{P}(\mathbf{F})$ with $\deg p = n$ such that $0 = p(\lambda_1) = \cdots = p(\lambda_m)$ and such that p has no other zeros.

Solution. Let $p(z) = (z - \lambda_1)(z - \lambda_2) \cdots (z - \lambda_m)^{n-m+1}$ and note that $\deg p = n$ and that each λ_k is a zero of p. The uniqueness parts of 4.13 and 4.16, together with 4.6, shows that p can have no other zeros.

Exercise 4.7. Suppose that m is a nonnegative integer, $z_1, \ldots, z_{m+1}$ are distinct elements of $\mathbf{F}$, and $w_1, \ldots, w_{m+1} \in \mathbf{F}$. Prove that there exists a unique polynomial $p \in \mathcal{P}_m(\mathbf{F})$ such that

$$p(z_k) = w_k$$

for each $k = 1, \ldots, m + 1$.

This result can be proved without using linear algebra. However, try to find the clearer, shorter proof that uses some linear algebra.

Solution. Note that the list $p_0, \ldots, p_{m+1} \in \mathcal{P}_m(\mathbf{F})$ given by

$$\begin{aligned} p_0 &= 1, \\ p_1 &= z - z_1, \\ p_2 &= (z - z_1)(z - z_2), \\ &\vdots \\ p_m &= (z - z_1)(z - z_2) \cdots (z - z_m) \end{aligned}$$

is a basis of $\mathcal{P}_m(\mathbf{F})$ by Exercise 2.C.9. Define a map $T \in \mathcal{L}(\mathcal{P}_m(\mathbf{F}), \mathbf{F}^{m+1})$ by

$$Tp = (p(z_1), \ldots, p(z_{m+1})).$$

Notice that the matrix of T with respect to the basis $p_0, \ldots, p_m$ of $\mathcal{P}_m(\mathbf{F})$ and the standard basis of $\mathbf{F}^{m+1}$ is

$$\begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & A_{1,1} & 0 & \cdots & 0 \\ 1 & A_{2,1} & A_{2,2} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & A_{m,1} & A_{m,2} & \cdots & A_{m,m} \end{pmatrix}, \quad \text{where } A_{j,k} = \prod_{i=1}^{k} (z_{j+1} - z_i) \text{ for } j \geq k.$$

Notice further that each $A_{j,k}$ is non-zero because the elements $z_1, \ldots, z_{m+1}$ are distinct. A straightforward calculation then shows that the rows of this matrix are linearly independent; it follows from Exercise 3.C.17 that T is injective and hence, because $\dim \mathcal{P}_m(\mathbf{F}) = \dim \mathbf{F}^{m+1}$, invertible. Thus there exists a unique $p \in \mathcal{P}_m(\mathbf{F})$ such that

$$Tp = (p(z_1), ..., p(z_{m+1})) = (w_1, w_2, ..., w_{m+1}).$$

Exercise 4.8. Suppose $p \in \mathcal{P}(\mathbf{C})$ has degree m. Prove that p has m distinct zeros if and only if p and its derivative p' have no zeros in common.

Solution. The cases $m = 0$ and $m = 1$ are straightforward to check. For $m \geq 2$, we will prove the equivalent statement

$$p \text{ has strictly less than } m \text{ distinct zeros} \quad \Leftrightarrow \quad p \text{ and } p' \text{ have a zero in common.}$$

If p has strictly less than m distinct zeros then it must be the case that p has a zero $\lambda \in \mathbf{C}$ such that $p(z) = (z - \lambda)^k q(z)$ for some positive integer $k \geq 2$ and some $q \in \mathcal{P}(\mathbf{C})$. It follows that

$$p'(z) = k(z - \lambda)^{k-1} q(z) + (z - \lambda)^k q'(z)$$

and hence that $p'(\lambda) = 0$, since $k \geq 2$. Thus p and p' have the zero λ in common.

Now suppose that p and p' have a zero in common, say $\lambda \in \mathbf{F}$, so that

$$p(z) = (z - \lambda)q(z) \quad \text{and} \quad p'(z) = (z - \lambda)r(z)$$

for some $q, r \in \mathcal{P}(\mathbf{C})$. The product rule gives us

$$q(z) + (z - \lambda)q'(z) = p'(z) = (z - \lambda)r(z).$$

Evaluating this expression at $z = \lambda$ shows that $q(\lambda) = 0$, so that $z - \lambda$ is a factor of q. It follows that p is of the form $p(z) = (z - \lambda)^2 t(z)$ for some $t \in \mathcal{P}(\mathbf{C})$ satisfying $\deg t = m - 2$. Thus p has strictly less than m zeros.

Exercise 4.9. Prove that every polynomial of odd degree with real coefficients has a real zero.

Solution. Let $p \in \mathcal{P}(\mathbf{R})$ be a polynomial of odd degree. By 4.16, p is of the form

$$p(x) = c(x - \lambda_1) \cdots (x - \lambda_m)(x^2 + b_1 x + c_1) \cdots (x^2 + b_M x + c_M),$$

where $c, \lambda_1, ..., \lambda_m, b_1, ..., b_M, c_1, ..., c_M \in \mathbf{R}$, $c > 0$, and $b_k^2 < 4c_k$ for each k. This implies that $\deg p = m + 2M$; since $\deg p$ is odd, it must be the case that $m > 0$. Thus p has at least one real zero.

Exercise 4.10. For $p \in \mathcal{P}(\mathbf{R})$, define $Tp : \mathbf{R} \to \mathbf{R}$ by

$$(Tp)(x) = \begin{cases} \dfrac{p(x) - p(3)}{x - 3} & \text{if } x \neq 3, \\ p'(3) & \text{if } x = 3 \end{cases}$$

for each $x \in \mathbf{R}$. Show that $Tp \in \mathcal{P}(\mathbf{R})$ for every polynomial $p \in \mathcal{P}(\mathbf{R})$ and also show that $T : \mathcal{P}(\mathbf{R}) \to \mathcal{P}(\mathbf{R})$ is a linear map.

Solution. Let $p \in \mathcal{P}(\mathbf{R})$ be given and notice that $p(x) - p(3)$ has a zero at $x = 3$, so that

$$p(x) - p(3) = (x - 3)q(x)$$

for some unique $q \in \mathcal{P}(\mathbf{R})$. It follows that for any $x \neq 3$ we have

$$q(x) = \frac{p(x) - p(3)}{x - 3}.$$

Differentiating the equality $p(x) - p(3) = (x - 3)q(x)$ gives us $p'(x) = q(x) + (x - 3)q'(x)$, whence $p'(3) = q(3)$. Thus $Tp = q \in \mathcal{P}(\mathbf{R})$.

Let $p_1, p_2, \in \mathcal{P}(\mathbf{R})$ and $\lambda \in \mathbf{R}$ be given. There are unique polynomials $q_1, q_2 \in \mathcal{P}(\mathbf{R})$ such that

$$p_1(x) - p_1(3) = (x - 3)q_1(x) \quad \text{and} \quad p_2(x) - p_2(3) = (x - 3)q_2(x).$$

As we showed above, it follows that $Tp_1 = q_1$ and $Tp_2 = q_2$. Notice that

$$(p_1 + p_2)(x) - (p_1 + p_2)(3) = (x - 3)(q_1 + q_2)(x)$$
$$\text{and} \quad (\lambda p_1)(x) - (\lambda p_1)(3) = (x - 3)(\lambda q_1)(x).$$

By uniqueness we must have $T(p_1 + p_2) = q_1 + q_2 = Tp_1 + Tp_2$ and $T(\lambda p_1) = \lambda q_1 = \lambda T p_1$. Thus T is linear.

Exercise 4.11. Suppose $p \in \mathcal{P}(\mathbf{C})$. Define $q : \mathbf{C} \to \mathbf{C}$ by

$$q(z) = p(z)\overline{p(\overline{z})}.$$

Prove that q is a polynomial with real coefficients.

Solution. Suppose $p(z) = a_0 + a_1 z + \dots + a_m z^m$ for some non-negative integer m. Observe that $\overline{p(\overline{z})} = \overline{a_0} + \overline{a_1} z + \dots + \overline{a_m} z^m$. It follows that

$$q(z) = p(z)\overline{p(\overline{z})} = \sum_{k=0}^{2m} \left(\sum_{i+j=k} a_i \overline{a_j} \right) z^k.$$

For any $k \in \{0, \ldots, 2m\}$ note that

$$\overline{\sum_{i+j=k} a_i \overline{a_j}} = \sum_{i+j=k} \overline{a_i \overline{a_j}} = \sum_{i+j=k} a_j \overline{a_i} = \sum_{i+j=k} a_i \overline{a_j},$$

where the last equality follows by reindexing. Thus $\sum_{i+j=k} a_i \overline{a_j} \in \mathbf{R}$ for each $k \in \{0, ..., 2m\}$, i.e. q has real coefficients.

Exercise 4.12. Suppose m is a nonnegative integer and $p \in \mathcal{P}_m(\mathbf{C})$ is such that there are distinct real numbers $x_0, x_1, ..., x_m$ with $p(x_k) \in \mathbf{R}$ for each $k = 0, 1, ..., m$. Prove that all coefficients of p are real.

Solution. By Exercise 4.7 there is a unique polynomial $q \in \mathcal{P}_m(\mathbf{R})$ such that $q(x_k) = p(x_k)$ for each $k \in \{0, ..., m\}$. It follows that the polynomial $p - q \in \mathcal{P}_m(\mathbf{C})$ has $m + 1$ distinct zeros and thus, by (4.8), $p = q \in \mathcal{P}_m(\mathbf{R})$.

Exercise 4.13. Suppose $p \in \mathcal{P}(\mathbf{F})$ with $p \neq 0$. Let $U = \{pq : q \in \mathcal{P}(\mathbf{F})\}$.

(a) Show that $\dim \mathcal{P}(\mathbf{F})/U = \deg p$.

(b) Find a basis of $\mathcal{P}(\mathbf{F})/U$.

Solution.

(a) Let $m = \deg p$. If $m = 0$, i.e. p is a non-zero constant polynomial, then $U = \mathcal{P}(\mathbf{F})$ and thus $\mathcal{P}(\mathbf{F})/U = \{0\}$, so that $\dim \mathcal{P}(\mathbf{F})/U = 0 = \deg p$. Suppose that $m \geq 1$. For any $s \in \mathcal{P}(\mathbf{F})$, the division algorithm for polynomials (4.9) implies that there are unique polynomials $q, r \in \mathcal{P}(\mathbf{F})$ such that $s = pq + r$ and $\deg r < \deg p$. Thus the map $T : \mathcal{P}(\mathbf{F}) \to \mathcal{P}_{m-1}(\mathbf{F})$ given by $Ts = r$ is well-defined. Let $s_1, s_2 \in \mathcal{P}(\mathbf{F})$ and $\lambda \in \mathbf{F}$ be given. There are unique polynomials $q_1, q_2, r_1, r_2 \in \mathcal{P}(\mathbf{F})$ such that

$$s_1 = pq_1 + r_1, \quad s_2 = pq_2 + r_2, \quad \deg r_1 < \deg p, \quad \text{and} \quad \deg r_2 < \deg p.$$

Thus $Ts_1 = r_1$ and $Ts_2 = r_2$. Observe that

$$s_1 + s_2 = p(q_1 + q_2) + (r_1 + r_2), \quad \deg(r_1 + r_2) \leq \max\{\deg r_1, \deg r_2\} < \deg p,$$
$$\lambda s_1 = p(\lambda q_1) + (\lambda r_1), \quad \deg(\lambda r_1) \leq \deg r_1 < \deg p.$$

It follows from the uniqueness part of the division algorithm that

$$T(s_1 + s_2) = r_1 + r_2 = Ts_1 + Ts_2 \quad \text{and} \quad T(\lambda s_1) = \lambda r_1 = \lambda T s_1.$$

Thus T is linear.

For any $r \in \mathcal{P}_{m-1}(\mathbf{F})$ we have $\deg r < \deg p$ and thus $Tr = r$. Hence T is surjective. Notice that $pq \in U$ has remainder zero upon division by p; it follows that $T(pq) = 0$. Conversely, if $s = pq + r$ is such that $Ts = r = 0$ then $s = pq \in U$. Thus $\operatorname{null} T = U$. It now follows from 3.107 that $\mathcal{P}(\mathbf{F})/U$ is isomorphic to $\mathcal{P}_{m-1}(\mathbf{F})$ via the map $\tilde{T} : \mathcal{P}(\mathbf{F})/U \to \mathcal{P}_{m-1}(\mathbf{F})$ given by $\tilde{T}(s + U) = Ts$. Thus

$$\dim \mathcal{P}(\mathbf{F})/U = \dim \mathcal{P}_{m-1}(\mathbf{F}) = m = \deg p.$$

(b) Notice that $\tilde{T}(1+U), \tilde{T}(z+U), ..., \tilde{T}(z^{m-1}+U)$ is the list $1, z, ..., z^{m-1}$, i.e. the standard basis of $\mathcal{P}_{m-1}(\mathbf{F})$. Because $\tilde{T}^{-1}$ is an isomorphism, it follows that

$$\tilde{T}^{-1}(1), \tilde{T}^{-1}(z), ..., \tilde{T}^{-1}(z^{m-1}) = 1+U, z+U, ..., z^{m-1}+U$$

is a basis of $\mathcal{P}(\mathbf{F})/U$.

Exercise 4.14. Suppose $p, q \in \mathcal{P}(\mathbf{C})$ are nonconstant polynomials with no zeros in common. Let $m = \deg p$ and $n = \deg q$. Use linear algebra as outlined below in (a)-(c) to prove that there exist $r \in \mathcal{P}_{n-1}(\mathbf{C})$ and $s \in \mathcal{P}_{m-1}(\mathbf{C})$ such that

$$rp + sq = 1.$$

(a) Define $T : \mathcal{P}_{n-1}(\mathbf{C}) \times \mathcal{P}_{m-1}(\mathbf{C}) \to \mathcal{P}_{m+n-1}(\mathbf{C})$ by

$$T(r, s) = rp + sq.$$

Show that the linear map T is injective.

(b) Show that the linear map T in (a) is surjective.

(c) Use (b) to conclude that there exist $r \in \mathcal{P}_{n-1}(\mathbf{C})$ and $s \in \mathcal{P}_{m-1}(\mathbf{C})$ such that $rp + sq = 1$.

Solution.

(a) Note that $m, n \geq 1$ since p, q are non-constant. Let $\lambda_1, ..., \lambda_m$ be the zeros of p and let $\mu_1, ..., \mu_n$ be the zeros of q; by assumption we have $p(\mu_k) \neq 0$ and $q(\lambda_k) \neq 0$ for all k. Suppose that $r \in \mathcal{P}_{n-1}(\mathbf{C})$ and $s \in \mathcal{P}_{m-1}(\mathbf{C})$ are such that $rp + sq = 0$. In particular, for each $k \in \{1, ..., n\}$,

$$r(\mu_k)p(\mu_k) + s(\mu_k)q(\mu_k) = r(\mu_k)p(\mu_k) = 0 \quad \Rightarrow \quad r(\mu_k) = 0,$$

where we have used that $q(\mu_k) = 0$ and $p(\mu_k) \neq 0$. Thus r is a polynomial of degree at most $n-1$ with n zeros; it follows from 4.8 that $r = 0$. A similar argument with the λ_k's shows that $s = 0$. Thus $\operatorname{null} T = \{0\}$, i.e. T is injective.

(b) Notice that

$$\begin{aligned} \dim(\mathcal{P}_{n-1}(\mathbf{C}) \times \mathcal{P}_{m-1}(\mathbf{C})) &= \dim \mathcal{P}_{n-1}(\mathbf{C}) + \dim \mathcal{P}_{m-1}(\mathbf{C}) \\ &= n + m = \dim \mathcal{P}_{m+n-1}(\mathbf{C}). \end{aligned}$$

Since T is injective, it follows from 3.65 that T is also surjective.

(c) This is immediate from the surjectivity of T.

Chapter 5. Eigenvalues and Eigenvectors

5.A. Invariant Subspaces

Exercise 5.A.1. Suppose $T \in \mathcal{L}(V)$ and U is a subspace of V.

(a) Prove that if $U \subseteq \operatorname{null} T$, then U is invariant under T.

(b) Prove that if $\operatorname{range} T \subseteq U$, then U is invariant under T.

Solution.

(a) Let $u \in U \subseteq \operatorname{null} T$ be given and observe that $Tu = 0 \in U$. Thus U is invariant under T.

(b) Let $u \in U$ be given and observe that $Tu \in \operatorname{range} T \subseteq U$. Thus U is invariant under T.

Exercise 5.A.2. Suppose that $T \in \mathcal{L}(V)$ and $V_1, ..., V_m$ are subspaces of V invariant under T. Prove that $V_1 + \cdots + V_m$ is invariant under T.

Solution. Let $v = a_1 v_1 + \cdots + a_m v_m \in V_1 + \cdots + V_m$ be given. By assumption each $Tv_k \in V_k$ and thus

$$Tv = a_1 T v_1 + \cdots + a_m T v_k \in V_1 + \cdots + V_m.$$

Hence $V_1 + \cdots + V_m$ is invariant under T.

Exercise 5.A.3. Suppose $T \in \mathcal{L}(V)$. Prove that the intersection of every collection of subspaces of V invariant under T is invariant under T.

Solution. Let $\mathcal{U}$ be a collection of subspaces of V invariant under T. For any $u \in \bigcap \mathcal{U}$ and any $U \in \mathcal{U}$, we have $u \in U$ and thus $Tu \in U$. It follows that $Tu \in \bigcap \mathcal{U}$ and hence that $\bigcap \mathcal{U}$ is invariant under T.

Exercise 5.A.4. Prove or give a counterexample: If V is finite-dimensional and U is a subspace of V that is invariant under every operator on V, then $U = \{0\}$ or $U = V$.

Solution. This is true. It will suffice to show that if $U \neq \{0\}$ then $U = V$. Suppose therefore that there exists some $v_1 \in U$ with $v_1 \neq 0$ and extend this to a basis $v_1, ..., v_m$ of V. For each $k \in \{1, ..., m\}$ define an operator $T_k \in \mathcal{L}(V)$ by $T_k v_1 = v_k$ and $T_k v_j = 0$ for $j \neq 1$. By assumption U is invariant under T_k and thus $T_k v_1 = v_k \in U$. It follows that U contains the basis $v_1, ..., v_m$ of V and hence that $U = V$.

Exercise 5.A.5. Suppose $T \in \mathcal{L}(\mathbf{R}^2)$ is defined by $T(x, y) = (-3y, x)$. Find the eigenvalues of T.

Solution. Geometrically, we can observe that T is a counterclockwise rotation by 90° about the origin followed by a dilation of the x-axis by a factor of 3. We can now argue as in 5.9(a) to see that T has no eigenvalues.

Algebraically, for $\lambda \in \mathbf{R}$ we can try to solve the equation $T(x, y) = (-3y, x) = (\lambda x, \lambda y)$. Substituting $x = \lambda y$ into $-3y = \lambda x$ gives us $-3y = \lambda^2 y$. Because $y = 0$ implies $x = 0$, and eigenvectors are non-zero, we may assume that $y \neq 0$ and thus obtain the equation $\lambda^2 + 3 = 0$. Since this equation has no real solutions, we see that T has no eigenvalues.

Exercise 5.A.6. Define $T \in \mathcal{L}(\mathbf{F}^2)$ by $T(w, z) = (z, w)$. Find all eigenvalues and eigenvectors of T.

Solution. T is a reflection in the line $w = z$. An appeal to our geometric intuition suggests that 1 is an eigenvalue with corresponding eigenvector $(1, 1)$ and that -1 is an eigenvalue with corresponding eigenvector $(-1, 1)$. To see this algebraically, suppose $\lambda \in \mathbf{F}$ and $(w, z) \neq (0, 0)$ are such that $T(w, z) = (z, w) = (\lambda w, \lambda z)$. Substituting $z = \lambda w$ into $w = \lambda z$ gives us $w = \lambda^2 w$. Since $w = 0$ implies $z = 0$, and eigenvectors are non-zero, we may assume that $w \neq 0$ and thus obtain the equation $\lambda^2 - 1 = 0$, which has solutions $\lambda = \pm 1$. These are both indeed eigenvalues, since

$$T(1, 1) = (1, 1) \quad \text{and} \quad T(-1, 1) = (1, -1) = -(-1, 1).$$

Since $\dim \mathbf{F}^2 = 2$, it follows from 5.11 and 5.12 that there are no other eigenvalues of T and no other eigenvectors of T linearly independent from the two given above. We may conclude that the eigenvalues and eigenvectors of T are precisely:

eigenvalue	corresponding eigenvectors
1	(w, w) for $w \in \mathbf{F} \setminus \{0\}$
-1	$(-w, w)$ for $w \in \mathbf{F} \setminus \{0\}$

Exercise 5.A.7. Define $T \in \mathcal{L}(\mathbf{F}^3)$ by $T(z_1, z_2, z_3) = (2z_2, 0, 5z_3)$. Find all eigenvalues and eigenvectors of T.

Solution. T can be thought of as the composition of the following transformations:

- a projection onto the $z_2 z_3$-plane;
- a clockwise rotation of 90° around the z_3-axis; after the projection onto the $z_2 z_3$-plane, this is equivalent to a reflection in the plane $z_1 = z_2$;
- a dilation of the z_1-axis by a factor of 2;

- a dilation of the z_3-axis by a factor of 5.

In other words, T maps $(z_1, z_2, z_3) \in \mathbf{F}^3$ like so:

$$(z_1, z_2, z_3) \mapsto (0, z_2, z_3) \mapsto (z_2, 0, z_3) \mapsto (2z_2, 0, z_3) \mapsto (2z_2, 0, 5z_3).$$

An appeal to our geometric intuition suggests that 5 is an eigenvalue with corresponding eigenvector $(0, 0, 1)$ and that 0 is an eigenvector with corresponding eigenvector $(1, 0, 0)$. To prove this algebraically, suppose that $\lambda \in \mathbf{F}$ and $(z_1, z_2, z_3) \neq (0, 0, 0)$ are such that

$$T(z_1, z_2, z_3) = (2z_2, 0, 5z_3) = (\lambda z_1, \lambda z_2, \lambda z_3).$$

If $\lambda \neq 0$ then the equation $\lambda z_2 = 0$ implies that $z_2 = 0$ and thus the equation $2z_2 = \lambda z_1$ gives us $z_1 = 0$. Since eigenvectors are non-zero, it must be the case that $z_3 \neq 0$ and so the equation $5z_3 = \lambda z_3$ implies that $\lambda = 5$. So the only possible eigenvalues are 0 and 5, which are indeed eigenvalues since

$$T(0, 0, 1) = (0, 0, 5) = 5(0, 0, 1) \quad \text{and} \quad T(1, 0, 0) = (0, 0, 0) = 0(1, 0, 0).$$

We claim that there are no other eigenvectors of T linearly independent from these two. As we just showed, any eigenvector of T corresponding to the eigenvalue 5 must satisfy $z_1 = z_2 = 0$ and thus each such eigenvector is a scalar multiple of $(0, 0, 1)$. If (z_1, z_2, z_3) is an eigenvector corresponding to the eigenvalue 0, i.e. $(z_1, z_2, z_3) \in \operatorname{null} T$, then

$$T(z_1, z_2, z_3) = (2z_2, 0, 5z_3) = (0, 0, 0) \quad \Rightarrow \quad z_2 = z_3 = 0 \quad \Rightarrow \quad (z_1, z_2, z_3) = z_1(1, 0, 0).$$

We may conclude that the eigenvalues and eigenvectors of T are precisely:

eigenvalue	corresponding eigenvectors
5	$(0, 0, w)$ for $w \in \mathbf{F} \setminus \{0\}$
0	$(w, 0, 0)$ for $w \in \mathbf{F} \setminus \{0\}$

Exercise 5.A.8. Suppose $P \in \mathcal{L}(V)$ is such that $P^2 = P$. Prove that if λ is an eigenvalue of P, then $\lambda = 0$ or $\lambda = 1$.

Solution. Suppose that λ is an eigenvalue of P, i.e. there is some $v \neq 0$ such that $Pv = \lambda v$. Notice that

$$\lambda v = Pv = P(Pv) = P(\lambda v) = \lambda Pv = \lambda^2 v.$$

Because $v \neq 0$, this implies that $\lambda = \lambda^2$. Thus $\lambda = 0$ or $\lambda = 1$.

Exercise 5.A.9. Define $T : \mathcal{P}(\mathbf{R}) \to \mathcal{P}(\mathbf{R})$ by $Tp = p'$. Find all eigenvalues and eigenvectors of T.

Solution. Note that $T(1) = 0 = 0(1)$, so that 0 is an eigenvalue of T. Note further that the only polynomials whose derivative is zero are the constant polynomials. Thus the eigenvectors corresponding to the eigenvalue 0 of T are precisely the non-zero constant polynomials.

Suppose $p \in \mathcal{P}(\mathbf{R})$ satisfies $\deg p \geq 1$. If $\lambda \neq 0$ then $\deg(\lambda p) = \deg p$, whereas $\deg p' = \deg p - 1$. Thus it cannot be the case that $Tp = p' = \lambda p$ and we may conclude that the eigenvalues and eigenvectors of T are precisely:

eigenvalue	corresponding eigenvectors
0	non-zero constant polynomials in $\mathcal{P}(\mathbf{R})$

Exercise 5.A.10. Define $T \in \mathcal{L}(\mathcal{P}_4(\mathbf{R}))$ by $(Tp)(x) = xp'(x)$ for all $x \in \mathbf{R}$. Find all eigenvalues and eigenvectors of T.

Solution. Letting $p_k \in \mathcal{P}_4(\mathbf{R})$ be given by $p_k(x) = x^k$ for $k \in \{0, ..., 4\}$, notice that

$$(Tp_k)(x) = kx^k = kp_k(x).$$

By 5.11 and 5.12 we may conclude that the eigenvalues and eigenvectors of T are precisely:

eigenvalue	corresponding eigenvectors
$k \in \{0, ..., 4\}$	αp_k for $\alpha \in \mathbf{R} \setminus \{0\}$

Exercise 5.A.11. Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and $\alpha \in \mathbf{F}$. Prove that there exists $\delta > 0$ such that $T - \lambda I$ is invertible for all $\lambda \in \mathbf{F}$ such that $0 < |\alpha - \lambda| < \delta$.

Solution. If T has no eigenvalues then 5.7 shows that $T - \lambda I$ is invertible for all $\lambda \in \mathbf{F}$, so that any $\delta > 0$ will suffice, say $\delta = 1$. Suppose therefore that T has at least one eigenvalue and let $\lambda_1, ..., \lambda_n$ be the distinct eigenvalues of T; there are only finitely many eigenvalues of T by 5.12. Let

$$\delta = \min\{|\alpha - \lambda_k| : k \in \{1, ..., n\} \text{ and } \lambda_k \neq \alpha\}.$$

It follows from this definition that δ is positive and furthermore that $\lambda \neq \lambda_k$ for any $k \in \{1, ..., n\}$ and any $\lambda \in \mathbf{F}$ such that $0 < |\alpha - \lambda| < \delta$. That is, λ is not an eigenvalue of T for any $\lambda \in \mathbf{F}$ such that $0 < |\alpha - \lambda| < \delta$. By 5.7 this is equivalent to saying that $T - \lambda I$ is invertible for any $\lambda \in \mathbf{F}$ such that $0 < |\alpha - \lambda| < \delta$.

Exercise 5.A.12. Suppose $V = U \oplus W$, where U and W are nonzero subspaces of V. Define $P \in \mathcal{L}(V)$ by $P(u + w) = u$ for each $u \in U$ and each $w \in W$. Find all eigenvalues and eigenvectors of P.

Solution. Notice that $P^2 = P$. It follows from Exercise 5.A.8 that the only possible eigenvalues of P are 1 and 0. Because U and W are non-zero, there is some non-zero $u \in U$ and some non-zero $w \in W$. Observe that

$$Pu = 1u \quad \text{and} \quad Pw = 0w.$$

Thus 1 and 0 are indeed eigenvalues of P. The above equations show that any non-zero elements of U are eigenvectors of T corresponding to the eigenvalue 1 and any non-zero elements of W are eigenvectors of T corresponding to the eigenvalue 0. If $v = u + w$ is an eigenvector of T corresponding to the eigenvalue 1, then observe that

$$u + w = P(u + w) = u \quad \Rightarrow \quad w = 0 \quad \Rightarrow \quad v \in U \setminus \{0\}.$$

Similarly, if $v = u + w$ is an eigenvector of T corresponding to the eigenvalue 0, then observe that

$$0 = P(u + w) = u \quad \Rightarrow \quad v \in W \setminus \{0\}.$$

We may conclude that the eigenvalues and eigenvectors of P are precisely:

eigenvalue	corresponding eigenvectors
1	$u \in U \setminus \{0\}$
0	$w \in W \setminus \{0\}$

Exercise 5.A.13. Suppose $T \in \mathcal{L}(V)$. Suppose $S \in \mathcal{L}(V)$ is invertible.

(a) Prove that T and $S^{-1}TS$ have the same eigenvalues.

(b) What is the relationship between the eigenvectors of T and the eigenvectors of $S^{-1}TS$?

Solution.

(a) Suppose $\lambda \in \mathbf{F}$ is an eigenvalue of T a corresponding eigenvector $v \in V$. Because S is surjective there is some $u \in V$ such that $v = Su$; notice that $u \neq 0$ since $v \neq 0$. It follows that

$$Tv = \lambda v \quad \Leftrightarrow \quad (TS)(u) = \lambda Su \quad \Leftrightarrow \quad (S^{-1}TS)(u) = \lambda u.$$

Thus λ is an eigenvalue of $S^{-1}TS$ with a corresponding eigenvector u.

Similarly, suppose $\lambda \in \mathbf{F}$ is an eigenvalue of $S^{-1}TS$ with a corresponding eigenvector $u \in V$. Because S^{-1} is surjective there is some $v \in V$ such that $u = S^{-1}v$; notice that $v \neq 0$ since $u \neq 0$. It follows that

$$(S^{-1}TS)(u) = \lambda u \quad \Leftrightarrow \quad (S^{-1}T)(v) = \lambda S^{-1}v \quad \Leftrightarrow \quad Tv = \lambda v.$$

Thus λ is an eigenvalue of T with a corresponding eigenvector v.

(b) Let $\lambda \in \mathbf{F}$ be an eigenvalue of T. As we showed in part (a), this is the case if and only if λ is an eigenvalue of $S^{-1}TS$. Define

$$E(\lambda, T) = \{v \in V : v \neq 0 \text{ and } Tv = \lambda v\}$$

$$\text{and} \quad E(\lambda, S^{-1}TS) = \{u \in V : u \neq 0 \text{ and } (S^{-1}TS)(u) = \lambda u\}.$$

That is, $E(\lambda, T)$ is the collection of eigenvectors of T corresponding to the eigenvalue λ and $E(\lambda, S^{-1}TS)$ is the collection of eigenvectors of $S^{-1}TS$ corresponding to the eigenvalue λ. Our calculations in part (a) show that

$$E(\lambda, T) = \{Su : u \in E(\lambda, S^{-1}TS)\} \quad \text{and} \quad E(\lambda, S^{-1}TS) = \{S^{-1}v : v \in E(\lambda, T)\}.$$

Exercise 5.A.14. Give an example of an operator on $\mathbf{R}^4$ that has no (real) eigenvalues.

Solution. Define $T \in \mathcal{L}(\mathbf{R}^4)$ by $T(x, y, z, t) = (y, z, t, -x)$ and suppose $\lambda \in \mathbf{R}$ is such that

$$T(x, y, z, t) = (y, z, t, -x) = \lambda(x, y, z, t).$$

We then have $-x = \lambda t = \lambda^2 z = \lambda^3 y = \lambda^4 x$. Notice that $x = 0$ implies $y = z = t = 0$. Since we are looking for eigenvectors, we can assume that $x \neq 0$ and thus arrive at the equation $\lambda^4 + 1 = 0$, which has no real solutions. It follows that T has no real eigenvalues.

Exercise 5.A.15. Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and $\lambda \in \mathbf{F}$. Show that λ is an eigenvalue of T if and only if λ is an eigenvalue of the dual operator $T' \in \mathcal{L}(V')$.

Solution. Observe that

$$\begin{aligned}
\lambda \text{ is an eigenvalue of } T \quad &\Leftrightarrow \quad T - \lambda I \text{ is not injective} && (5.7)\\
&\Leftrightarrow \quad (T - \lambda I)' \text{ is not surjective} && (3.131)\\
&\Leftrightarrow \quad T' - \lambda I' \text{ is not surjective} && (\text{Exercise 3.F.12})\\
&\Leftrightarrow \quad T' - \lambda I \text{ is not surjective} && (\text{Exercise 3.F.13})\\
&\Leftrightarrow \quad \lambda \text{ is an eigenvalue of } T'. && (5.7)
\end{aligned}$$

Exercise 5.A.16. Suppose $v_1, \ldots, v_n$ is a basis of V and $T \in \mathcal{L}(V)$. Prove that if λ is an eigenvalue of T, then

$$|\lambda| \leq n \max\Big\{\big|\mathcal{M}(T)_{j,k}\big| : 1 \leq j, k \leq n\Big\},$$

where $\mathcal{M}(T)_{j,k}$ denotes the entry in row j, column k of the matrix of T with respect to the basis $v_1, \ldots, v_n$.

See Exercise 19 in Section 6A for a different bound on $|\lambda|$.

Solution. Let $A_{j,k} = \mathcal{M}(T)_{j,k}$ and suppose that $v = b_1 v_1 + \cdots + b_n v_n$ is an eigenvector of T corresponding to the eigenvalue λ. Notice that

$$\begin{aligned}\lambda b_1 v_1 + \cdots + \lambda b_n v_n &= b_1 T v_1 + \cdots + b_n T v_n \\ &= b_1 \left(\sum_{j=1}^{n} A_{j,1} v_j \right) + \cdots + b_n \left(\sum_{j=1}^{n} A_{j,n} v_j \right) = \sum_{j=1}^{n} \left(\sum_{k=1}^{n} A_{j,k} b_k \right) v_j.\end{aligned}$$

It follows from unique representation that

$$\lambda b_j = \sum_{k=1}^{n} A_{j,k} b_k \quad \Rightarrow \quad |\lambda||b_j| \leq \sum_{k=1}^{n} |A_{j,k}||b_k|$$

for each $j \in \{1, \ldots, n\}$. Let $|b_i|$ be the largest amongst the values $|b_1|, \ldots, |b_n|$ and notice that $|b_i| > 0$, since $|b_i| = 0$ implies $b_1 = \cdots = b_n = 0$, so that v is zero—but v is an eigenvector and thus non-zero. Let $M = \max\left\{ \left|\mathcal{M}(T)_{j,k}\right| : 1 \leq j, k \leq n \right\}$ and observe that

$$|\lambda||b_i| \leq \sum_{k=1}^{n} |A_{i,k}||b_k| \quad \Rightarrow \quad |\lambda| \leq \sum_{k=1}^{n} |A_{i,k}| \frac{|b_k|}{|b_i|} \leq \sum_{k=1}^{n} |A_{i,k}| \leq \sum_{k=1}^{n} M = nM.$$

Exercise 5.A.17. Suppose $\mathbf{F} = \mathbf{R}$, $T \in \mathcal{L}(V)$, and $\lambda \in \mathbf{R}$. Prove that λ is an eigenvalue of T if and only if λ is an eigenvalue of the complexification $T_{\mathbf{C}}$.

See Exercise 33 in Section 3B for the definition of $T_{\mathbf{C}}$.

Solution. Suppose that λ is an eigenvalue of T with a corresponding eigenvector $v \in V$. Notice that $v + i0$ is non-zero since v is non-zero, and notice further that

$$T_{\mathbf{C}}(v + 0i) = Tv + iT(0) = \lambda v + 0i = \lambda(v + 0i).$$

Thus λ is an eigenvalue of $T_{\mathbf{C}}$ with a corresponding eigenvector $v + 0i$.

Now suppose that λ is an eigenvalue of $T_{\mathbf{C}}$ with a corresponding eigenvector $u + iv$. It follows that

$$T_{\mathbf{C}}(u + iv) = Tu + iTv = \lambda(u + iv) = \lambda u + i(\lambda v) \quad \Rightarrow \quad Tu = \lambda u \text{ and } Tv = \lambda v.$$

Since $u + iv \neq 0$, at least one of u, v is non-zero. Thus λ is an eigenvalue of T with u or v (or both) as a corresponding eigenvector.

Exercise 5.A.18. Suppose $\mathbf{F} = \mathbf{R}$, $T \in \mathcal{L}(V)$, and $\lambda \in \mathbf{C}$. Prove that λ is an eigenvalue of the complexification $T_{\mathbf{C}}$ if and only if $\overline{\lambda}$ is an eigenvalue of $T_{\mathbf{C}}$.

Solution. Suppose that $\lambda = a + bi$ and $u + iv$ is an eigenvector of $T_{\mathbf{C}}$ corresponding to λ. Observe that

$$\lambda(u + iv) = T_{\mathbf{C}}(u + iv) \quad \Leftrightarrow \quad (a + bi)(u + iv) = (au - bv) + (av + bu)i = Tu + iTv.$$

Thus $Tu = au - bv$ and $Tv = av + bu$. It follows that

$$T_{\mathbf{C}}(u-iv) = Tu - iTv = (au - bv) - (av + bu) = (a+bi)(u-iv) = \overline{\lambda}(u-iv).$$

Furthermore, $u - iv$ is non-zero since $u + iv$ is non-zero. Thus $\overline{\lambda}$ is an eigenvalue of $T_{\mathbf{C}}$ with a corresponding eigenvector $u - iv$.

We have now shown that

$$\lambda \in \mathbf{C} \text{ is an eigenvalue of } T_{\mathbf{C}} \quad \Rightarrow \quad \overline{\lambda} \text{ is an eigenvalue of } T_{\mathbf{C}}.$$

The converse can be obtained by replacing λ with $\overline{\lambda}$ in the implication above and using that $\overline{\overline{\lambda}} = \lambda$.

Exercise 5.A.19. Show that the forward shift operator $T \in \mathcal{L}(\mathbf{F}^\infty)$ defined by

$$T(z_1, z_2, ...) = (0, z_1, z_2, ...)$$

has no eigenvalues.

Solution. We are looking for solutions to the equation

$$(0, z_1, z_2, ...) = (\lambda z_1, \lambda z_2, \lambda z_3, ...),$$

where $(z_1, z_2, ...) \neq 0$ and $\lambda \in \mathbf{F}$. Notice that $\lambda = 0$ implies that each $z_k = 0$. If $\lambda \neq 0$ then the equation $0 = \lambda z_1$ implies that $z_1 = 0$, which gives us the equation $0 = \lambda z_2$, which implies that $z_2 = 0$, and so on. Thus both assumptions $\lambda = 0$ and $\lambda \neq 0$ imply that $(z_1, z_2, ...) = 0$. We may conclude that T has no eigenvalues.

Exercise 5.A.20. Define the backward shift operator $S \in \mathcal{L}(\mathbf{F}^\infty)$ by

$$S(z_1, z_2, z_3, ...) = (z_2, z_3, ...).$$

(a) Show that every element of $\mathbf{F}$ is an eigenvalue of S.

(b) Find all eigenvectors of S.

Solution.

(a) Observe that for any $\lambda \in \mathbf{F}$ and any $\alpha \in \mathbf{F} \setminus \{0\}$ we have $\alpha(1, \lambda, \lambda^2, ...) \neq 0$ and

$$S(\alpha(1, \lambda, \lambda^2, ...)) = \alpha(\lambda, \lambda^2, \lambda^3, ...) = \lambda\alpha(1, \lambda, \lambda^2, ...).$$

Thus each $\lambda \in \mathbf{F}$ is an eigenvalue of S.

(b) Let $\lambda \in \mathbf{F}$ be given and suppose that $(z_1, z_2, z_3, ...)$ is an eigenvector of S corresponding to λ, i.e.

$$(z_2, z_3, z_4, ...) = (\lambda z_1, \lambda z_2, \lambda z_3, ...).$$

This equation implies that $z_2 = \lambda z_1$, which gives us $z_3 = \lambda z_2 = \lambda^2 z_1$, and so on. Thus

$$(z_1, z_2, z_3, ...) = (z_1, \lambda z_1, \lambda^2 z_1, ...) = z_1(1, \lambda, \lambda^2, ...);$$

this implies $z_1 \neq 0$ since eigenvectors are non-zero. Conversely, any vector of the form $\alpha(1, \lambda, \lambda^2, \ldots)$ for $\alpha \in \mathbf{F} \setminus \{0\}$ is an eigenvector of S corresponding to λ, as we showed in part (a). We may conclude that the eigenvalues and eigenvectors of S are precisely:

eigenvalue	corresponding eigenvectors
$\lambda \in \mathbf{F}$	$\alpha(1, \lambda, \lambda^2, \ldots)$ for $\alpha \in \mathbf{F} \setminus \{0\}$

Exercise 5.A.21. Suppose $T \in \mathcal{L}(V)$ is invertible.

(a) Suppose $\lambda \in \mathbf{F}$ with $\lambda \neq 0$. Prove that λ is an eigenvalue of T if and only if $\frac{1}{\lambda}$ is an eigenvalue of T^{-1}.

(b) Prove that T and T^{-1} have the same eigenvectors.

Solution.

(a) For $\lambda \neq 0$ and $v \neq 0$, observe that

$$Tv = \lambda v \quad \Leftrightarrow \quad v = \lambda T^{-1} v \quad \Leftrightarrow \quad \lambda^{-1} v = T^{-1} v.$$

(b) Notice that T and T^{-1} are both injective and so neither has 0 as an eigenvalue. Thus any eigenvector of T or T^{-1} must correspond to a non-zero eigenvalue. It follows from part (a) that

$$v \text{ is an eigenvector of } T \quad \Leftrightarrow \quad v \text{ is an eigenvector of } T^{-1}.$$

Exercise 5.A.22. Suppose $T \in \mathcal{L}(V)$ and there exist nonzero vectors u and w in V such that

$$Tu = 3w \quad \text{and} \quad Tw = 3u.$$

Prove that 3 or -3 is an eigenvalue of T.

Solution. Applying T to both sides of the equation $Tu = 3w$ shows that $T^2 u = 9u$ or equivalently $(T^2 - 9I)(u) = 0$. Because u is non-zero, this demonstrates that the operator $T^2 - 9I = (T - 3I)(T + 3I)$ is not injective. It must then be the case that at least one of the operators $T - 3I, T + 3I$ is not injective and thus 3 or -3 is an eigenvalue of T.

Exercise 5.A.23. Suppose V is finite-dimensional and $S, T \in \mathcal{L}(V)$. Prove that ST and TS have the same eigenvalues.

Solution. Exercise 3.D.11 shows that ST is invertible if and only if TS is invertible; by 5.7, this is equivalent to ST having 0 as an eigenvalue if and only if TS has 0 as an eigenvalue.

Suppose that $\lambda \neq 0$ is an eigenvalue of ST with corresponding eigenvector $v \in V$, i.e. $S(Tv) = \lambda v$, and note that $Tv \neq 0$ since $\lambda \neq 0$ and $v \neq 0$. Note further that

$$(TS)(Tv) = T(S(Tv)) = T(\lambda v) = \lambda Tv.$$

Thus λ is an eigenvalue of TS with a corresponding eigenvector Tv. Swapping the roles of S and T in this argument shows that any non-zero eigenvalue of TS must also be an eigenvalue of ST.

Exercise 5.A.24. Suppose A is an n-by-n matrix with entries in $\mathbf{F}$. Define $T \in \mathcal{L}(\mathbf{F}^n)$ by $Tx = Ax$, where elements of $\mathbf{F}^n$ are thought of as n-by-1 column vectors.

(a) Suppose the sum of the entries in each row of A equals 1. Prove that 1 is an eigenvalue of T.

(b) Suppose the sum of the entries in each column of A equals 1. Prove that 1 is an eigenvalue of T.

Solution.

(a) Let $A_{j,k}$ be the entries of A; our assumption is that $\sum_{k=1}^{n} A_{j,k} = 1$ for each $j \in \{1, ..., n\}$. Observe that

$$T(1, ..., 1) = \begin{pmatrix} A_{1,1} & \cdots & A_{1,n} \\ \vdots & \ddots & \vdots \\ A_{n,1} & \cdots & A_{n,n} \end{pmatrix} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^{n} A_{1,k} \\ \vdots \\ \sum_{k=1}^{n} A_{n,k} \end{pmatrix} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}.$$

Thus 1 is an eigenvalue of T.

(b) Here are two arguments.

(1) Let $e_1, ..., e_n$ be the standard basis of $\mathbf{F}^n$ and let $\psi \in (\mathbf{F}^n)'$ be the linear functional given by $\psi(e_k) = 1$. Certainly the matrix of T with respect to $e_1, ..., e_n$ is A. It follows that

$$(\psi \circ (T - I))(e_k) = \psi\left(\left(\sum_{j=1}^{n} A_{j,k} e_j\right) - e_k\right) = \left(\sum_{j=1}^{n} A_{j,k}\right) - 1 = 0,$$

where we have used our assumption that the sum of the entries in each column of A equals 1. Thus $\psi \circ (T - I) \in (\mathbf{F}^n)'$ is the zero map. If the operator $T - I$ were invertible then it would have to be the case that $\psi = 0$; because ψ is non-zero, we see that $T - I$ is not invertible. That is, 1 is an eigenvalue of T.

(2) Define $S \in \mathcal{L}(\mathbf{F}^n)$ by $Sx = A^{\mathrm{t}}x$. Because the rows of A^{t} are the columns of A, part (a) shows that 1 is an eigenvalue of S. Let $e_1, ..., e_n$ be the standard basis of $\mathbf{F}^n$, let $\varphi_1, ..., \varphi_n$ be the corresponding dual basis, and define an isomorphism $\Phi \in \mathcal{L}(\mathbf{F}^n, (\mathbf{F}^n)')$ by $\Phi e_k = \varphi_k$. Certainly the matrix of T with respect to $e_1, ..., e_n$ is A and thus, by 3.132, the matrix of T' with respect to $\varphi_1, ..., \varphi_n$ is A^{t}. It follows that

$$\begin{aligned}(\Phi^{-1} T' \Phi)(e_k) = (\Phi^{-1} T')(\varphi_k) = \Phi^{-1}(A_{k,1}\psi_1 + \cdots + A_{k,n}\psi_n) \\ = A_{k,1}e_1 + \cdots + A_{k,n}e_n = A^{\mathrm{t}}e_k = Se_k.\end{aligned}$$

Thus $\Phi^{-1}T'\Phi = S$, so that $\Phi^{-1}T'\Phi$ has 1 as an eigenvalue. A small modification of the argument given in Exercise 5.A.13 (a) shows that $\Phi^{-1}T'\Phi$ and T' have the same eigenvalues. It follows that 1 is an eigenvalue of T' and hence, by Exercise 5.A.15, 1 is an eigenvalue of T.

Exercise 5.A.25. Suppose $T \in \mathcal{L}(V)$ and u, w are eigenvectors of T such that $u + w$ is also an eigenvector of T. Prove that u and w are eigenvectors of T corresponding to the same eigenvalue.

Solution. Suppose that the eigenvectors $u, w, u + w$ correspond to the eigenvalues α, β, λ, respectively. Observe that

$$\alpha u + \beta w = Tu + Tw = T(u + w) = \lambda(u + w) = \lambda u + \lambda w.$$

It follows that $(\alpha - \lambda)u + (\beta - \lambda)w = 0$. If we suppose that $\alpha \neq \beta$ then 5.11 shows that u and w are linearly independent and the equation $(\alpha - \lambda)u + (\beta - \lambda)w = 0$ then implies that $\alpha = \beta = \lambda$, contradicting our assumption. Thus $\alpha = \beta$.

Exercise 5.A.26. Suppose $T \in \mathcal{L}(V)$ is such that every nonzero vector in V is an eigenvector of T. Prove that T is a scalar multiple of the identity operator.

Solution. The case where $V = \{0\}$ is easily handled, so assume that $V \neq \{0\}$ and fix some non-zero $u \in V$; by assumption we have $Tu = \lambda u$ for some $\lambda \in \mathbf{F}$. Suppose that $v \in V$ is non-zero. If $u + v \neq 0$ then by assumption v and $u + v$ are both eigenvectors of T. It then follows from Exercise 5.A.25 that u and v correspond to the same eigenvalue λ, so that $Tv = \lambda v$. If $u + v = 0$ then $Tv = -Tu = -\lambda u = \lambda v$. Thus we have $Tv = \lambda v$ for all $v \in V$, i.e. $T = \lambda I$.

Exercise 5.A.27. Suppose that V is finite-dimensional and $k \in \{1, ..., \dim V - 1\}$. Suppose $T \in \mathcal{L}(V)$ is such that every subspace of V of dimension k is invariant under T. Prove that T is a scalar multiple of the identity operator.

Solution. If $\dim V = 0$ then $T = 0I$ and if $\dim V = 1$ then Exercise 3.A.7 shows that $T = \lambda I$ for some $\lambda \in \mathbf{F}$. Suppose that $\dim V \geq 2$. For $k \in \{1, ..., \dim V - 1\}$, let $P(k)$ be the statement that if every subspace of V of dimension k is invariant under T, then T is a scalar multiple of the identity operator. We will prove that $P(1)$ holds and that $P(k)$ implies $P(k + 1)$, provided $k + 1 \in \{1, ..., \dim V - 1\}$.

Suppose that every one-dimensional subspace of V is invariant under T and let $v \in V$ be non-zero. By assumption the subspace $\operatorname{span}(v)$ is invariant under T and thus $Tv = \lambda v$ for some $\lambda \in \mathbf{F}$, i.e. v is an eigenvector of T. It follows from Exercise 5.A.26 that T is a scalar multiple of the identity operator. Thus $P(1)$ holds.

Now suppose that $P(k)$ holds for some k such that $k + 1 \in \{1, ..., \dim V - 1\}$ and suppose that every subspace of dimension $k + 1$ is invariant under T. Let U be a subspace of dimension k. Because $k \leq \dim V - 2$, we can find linearly independent vectors $v, w \in V$ such that $v, w \notin U$. Let $u \in U$ be given. By assumption the subspaces $U \oplus \operatorname{span}(v)$ and $U \oplus \operatorname{span}(w)$, which have dimension $k+1$, are invariant under T. Since u belongs to each of these subspaces, it follows that

$$Tu = a_1 u_1 + bv \quad \text{and} \quad Tu = a_2 u_2 + cw$$

for some $u_1, u_2 \in U$ and some $a_1, a_2, b, c \in \mathbf{F}$. This implies that $a_1 u_1 - a_2 u_2 = cw - bv$, so that $cw - bv \in U \cap \operatorname{span}(v, w) = \{0\}$. Thus $cw - bv = 0$ and the linear independence of v, w then gives us $b = c = 0$, so that $Tu \in U$. That is, U is invariant under T. Because we assumed that $P(k)$ holds, it now follows that T is a scalar multiple of the identity, i.e. $P(k+1)$ holds.

Thus, by induction, $P(k)$ holds for each $k \in \{1, ..., \dim V - 1\}$.

Exercise 5.A.28. Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Prove that T has at most $1 + \dim \operatorname{range} T$ distinct eigenvalues.

Solution. Suppose $\lambda_1, ..., \lambda_n$ are n distinct eigenvalues of T with corresponding eigenvectors $v_1, ..., v_n$ and note that each $Tv_k = \lambda_k v_k \in \operatorname{range} T$. The list $v_1, ..., v_n$ is linearly independent by 5.11 and thus the list $\lambda_1 v_1, ..., \lambda_n v_n$ is also linearly independent, provided each λ_k is non-zero; if some $\lambda_k = 0$ (since the eigenvalues are distinct there can be at most one such λ_k), we can discard $\lambda_k v_k$ from the list and be left with a linearly independent list of $n-1$ vectors. In either case, there are at least $n-1$ linearly independent vectors in $\operatorname{range} T$ and thus $n \leq 1 + \dim \operatorname{range} T$.

Exercise 5.A.29. Suppose $T \in \mathcal{L}(\mathbf{R}^3)$ and $-4, 5$, and $\sqrt{7}$ are eigenvalues of T. Prove that there exists $x \in \mathbf{R}^3$ such that $Tx - 9x = \left(-4, 5, \sqrt{7}\right)$.

Solution. Because T has $3 = \dim \mathbf{R}^3$ distinct eigenvalues, 5.12 shows that 9 cannot be an eigenvalue of T. By 5.7 this is equivalent to the operator $T - 9I$ being invertible. Thus the desired $x \in \mathbf{R}^3$ is $(T - 9I)^{-1}\left(-4, 5, \sqrt{7}\right)$.

Exercise 5.A.30. Suppose $T \in \mathcal{L}(V)$ and $(T - 2I)(T - 3I)(T - 4I) = 0$. Suppose λ is an eigenvalue of T. Prove that $\lambda = 2$ or $\lambda = 3$ or $\lambda = 4$.

Solution. We have $Tv = \lambda v$ for some $v \neq 0$. Observe that

$$0 = (T - 2I)(T - 3I)(T - 4I)v = (\lambda - 2)(\lambda - 3)(\lambda - 4)v.$$

Since v is non-zero, this equation implies that $\lambda \in \{2, 3, 4\}$.

Exercise 5.A.31. Give an example of $T \in \mathcal{L}(\mathbf{R}^2)$ such that $T^4 = -I$.

Solution. Define $T \in \mathcal{L}(\mathbf{R}^2)$ by

$$T(1,0) = \left(\cos\frac{\pi}{4}, \sin\frac{\pi}{4}\right) = \frac{1}{\sqrt{2}}(1,1) \quad \text{and} \quad T(0,1) = \left(\cos\frac{3\pi}{4}, \sin\frac{3\pi}{4}\right) = \frac{1}{\sqrt{2}}(-1,1).$$

Note that T is a counterclockwise rotation about the origin by 45°. It follows that T^4 is a counterclockwise rotation about the origin by 180°, i.e.

$$T^4(1,0) = (\cos\pi, \sin\pi) = (-1,0) \quad \text{and} \quad T(0,1) = \left(\cos\frac{3\pi}{2}, \sin\frac{3\pi}{2}\right) = (0,-1).$$

Thus $T^4 = -I$.

Exercise 5.A.32. Suppose $T \in \mathcal{L}(V)$ has no eigenvalues and $T^4 = I$. Prove that $T^2 = -I$.

Solution. Observe that $T^4 = I$ is equivalent to $(T^2 - I)(T^2 + I) = 0$. Because T has no eigenvalues, the operators $T - I$ and $T + I$ must be injective and thus their product $(T - I)(T + I) = T^2 - I$ is also injective. It follows that for any $v \neq 0$,

$$0 = (T^2 - I)(T^2 + I)v \quad \Leftrightarrow \quad 0 = (T^2 + I)v \quad \Leftrightarrow \quad T^2 v = -v.$$

Thus $T^2 = -I$.

Exercise 5.A.33. Suppose $T \in \mathcal{L}(V)$ and m is a positive integer.

(a) Prove that T is injective if and only if T^m is injective.

(b) Prove that T is surjective if and only if T^m is surjective.

Solution. Certainly these results are true if $m = 1$, so suppose that $m \geq 2$.

(a) The composition of injective functions is again injective, so T^m is injective if T is injective. Suppose that T^m is injective and let $v \in \operatorname{null} T$ be given. It follows that

$$Tv = 0 \quad \Rightarrow \quad T^m v = T^{m-1}(0) = 0 \quad \Rightarrow \quad v \in \operatorname{null} T^m = \{0\} \quad \Rightarrow \quad v = 0.$$

Thus $\operatorname{null} T = \{0\}$, i.e. T is injective.

(b) Suppose that T^m is surjective and let $w \in V$ be given. There exists some $v \in V$ such that $T^m v = w$, which is equivalent to $T(T^{m-1}v) = w$. Thus T is surjective.

Suppose that T is surjective and let $w \in V$ be given. There exist vectors $v_1, ..., v_m$ such that:

$$Tv_1 = w \quad \text{and} \quad Tv_k = v_{k-1} \text{ for } k \geq 2.$$

It follows that $T^m v_m = w$ and hence that T^m is surjective.

Exercise 5.A.34. Suppose V is finite-dimensional and $v_1, ..., v_m \in V$. Prove that the list $v_1, ..., v_m$ is linearly independent if and only if there exists $T \in \mathcal{L}(V)$ such that $v_1, ..., v_m$ are eigenvectors of T corresponding to distinct eigenvalues.

Solution. Suppose $v_1, \dots, v_m$ is linearly independent and extend this to a basis $v_1, \dots, v_m, w_1, \dots, w_n$ of V. Define $T \in \mathcal{L}(V)$ by

$$Tv_k = kv_k \text{ for } k \in \{1, \dots, m\} \quad \text{and} \quad Tw_k = 0 \text{ for } k \in \{1, \dots, n\}.$$

It follows that $v_1, \dots, v_m$ are eigenvectors of T corresponding to distinct eigenvalues.

The converse implication is the content of 5.11.

Exercise 5.A.35. Suppose that $\lambda_1, \dots, \lambda_n$ is a list of distinct real numbers. Prove that the list $e^{\lambda_1 x}, \dots, e^{\lambda_n x}$ is linearly independent in the vector space of real-valued functions on $\mathbf{R}$.

> *Hint: Let $V = \operatorname{span}(e^{\lambda_1 x}, \dots, e^{\lambda_n x})$, and define an operator $D \in \mathcal{L}(V)$ by $Df = f'$. Find eigenvalues and eigenvectors of D.*

Solution. Let $V = \operatorname{span}(e^{\lambda_1 x}, \dots, e^{\lambda_n x})$, and define an operator $D \in \mathcal{L}(V)$ by $Df = f'$. For each $k \in \{1, \dots, n\}$,

$$D(e^{\lambda_k x}) = (e^{\lambda_k x})' = \lambda_k e^{\lambda_k x}.$$

This demonstrates that D indeed maps V into V and also that each λ_k is an eigenvalue of D with a corresponding eigenvector $e^{\lambda_k x}$. Because the eigenvalues $\lambda_1, \dots, \lambda_n$ are distinct, 5.11 shows that the corresponding eigenvectors $e^{\lambda_1 x}, \dots, e^{\lambda_n x}$ are linearly independent.

Exercise 5.A.36. Suppose that $\lambda_1, \dots, \lambda_n$ is a list of distinct positive numbers. Prove that the list $\cos(\lambda_1 x), \dots, \cos(\lambda_n x)$ is linearly independent in the vector space of real-valued functions on $\mathbf{R}$.

Solution. Let $V = \operatorname{span}(\cos(\lambda_1 x), \dots, \cos(\lambda_n x))$ and define $D \in \mathcal{L}(V)$ by $Df = \frac{\mathrm{d}^4}{\mathrm{d}x^4} f$. For each $k \in \{1, \dots, n\}$,

$$D(\cos(\lambda_k x)) = \frac{\mathrm{d}^4}{\mathrm{d}x^4} \cos(\lambda_k x) = \lambda_k^4 \cos(\lambda_k x).$$

This demonstrates that D indeed maps V into V and also that each λ_k^4 is an eigenvalue of D with a corresponding eigenvector $\cos(\lambda_k x)$. Because $\lambda_1, \dots, \lambda_n$ are distinct and positive, the real numbers $\lambda_1^4, \dots, \lambda_n^4$ are also distinct. It then follows from 5.11 that the list $\cos(\lambda_1 x), \dots, \cos(\lambda_n x)$ is linearly independent.

Exercise 5.A.37. Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Define $\mathcal{A} \in \mathcal{L}(\mathcal{L}(V))$ by

$$\mathcal{A}(S) = TS$$

for each $S \in \mathcal{L}(V)$. Prove that the set of eigenvalues of T equals the set of eigenvalues of $\mathcal{A}$.

Solution. Suppose that $\lambda \in \mathbf{F}$ is an eigenvalue of T, so that there is some non-zero $v_1 \in V$ such that $Tv_1 = \lambda v_1$. Extend v_1 to a basis $v_1, ..., v_m$ of V and define $S \in \mathcal{L}(V)$ by

$$Sv_1 = v_1 \quad \text{and} \quad Sv_k = 0 \text{ for } k \geq 2.$$

Notice that $S \neq 0$ and that

$$(T - \lambda I)Sv_1 = (T - \lambda I)v_1 = 0 \quad \text{and} \quad (T - \lambda I)Sv_k = 0 \text{ for } k \geq 2.$$

Thus $(T - \lambda I)S = 0$, i.e. $TS = \lambda S$. It follows that λ is an eigenvalue of $\mathcal{A}$ with a corresponding eigenvector S.

Now suppose that $\lambda \in \mathbf{F}$ is an eigenvalue of $\mathcal{A}$, i.e. there is some non-zero $S \in \mathcal{L}(V)$ such that $TS = \lambda S$, or equivalently $(T - \lambda I)S = 0$. If $T - \lambda I$ were injective then the equation $(T - \lambda I)S = 0$ would imply that $S = 0$. Given that S is non-zero, it must be the case that $T - \lambda I$ is not injective and thus λ is an eigenvalue of T.

Exercise 5.A.38. Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and U is a subspace of V invariant under T. The *quotient operator* $T/U \in \mathcal{L}(V/U)$ is defined by

$$(T/U)(v + U) = Tv + U$$

for each $v \in V$.

(a) Show that the definition of T/U makes sense (which requires using the condition that U is invariant under T) and show that T/U is an operator on V/U.

(b) Show that each eigenvalue of T/U is an eigenvalue of T.

Solution.

(a) Suppose that $v, w \in V$ are such that $v + U = w + U$, which by 3.101 is equivalent to $v - w \in U$. Because U is invariant under T it follows that $Tv - Tw \in U$ and another application of 3.101 gives us $Tv + U = Tw + U$. Thus the definition of T/U makes sense.

Certainly T/U maps V/U into V/U. Let $v + U, w + U \in V/U$ and $\lambda \in \mathbf{F}$ be given. Observe that

$$\begin{aligned}(T/U)((v + U) + (w + U)) &= (T/U)((v + w) + U) = T(v + w) + U \\ &= (Tv + Tw) + U = (Tv + U) + (Tw + U) = (T/U)(v + U) + (T/U)(w + U),\end{aligned}$$

$$\begin{aligned}(T/U)(\lambda(v + U)) &= (T/U)(\lambda v + U) = T(\lambda v) + U \\ &= \lambda Tv + U = \lambda(Tv + U) = \lambda(T/U)(v + U).\end{aligned}$$

Thus T/U is a linear operator on V/U.

(b) Suppose $\lambda \in \mathbf{F}$ is an eigenvalue of T/U, i.e. there exists a non-zero $v + U \in V/U$ such that

$$(T/U)(v + U) = Tv + U = \lambda(v + U) = \lambda v + U.$$

Note that U is invariant under $T - \lambda I$: for any $u \in U$ we have $Tu - \lambda u \in U$ since U is invariant under T. Here are two ways to proceed; notice that in each argument we only require that U is finite-dimensional.

(1) Consider the restriction operator $R := (T - \lambda I)|_U \in \mathcal{L}(U)$. There are two cases.

Suppose that R is not surjective; by 3.65 this is equivalent to $\operatorname{null} R \neq \{0\}$, so that there exists some non-zero $u \in U$ such that $Ru = 0$, i.e. $Tu = \lambda u$. Thus λ is an eigenvalue of T.

Now suppose that R is surjective. Because $Tv + U = \lambda v + U$, we have $Tv = \lambda v + w$ for some $w \in U$. The surjectivity of R implies that there exists some $u \in U$ such that $-w = Ru = Tu - \lambda u$. Observe that

$$T(v + u) = Tv + Tu = \lambda v + w + Tu = \lambda v + \lambda u = \lambda(v + u).$$

Note that $v + u$ must be non-zero, otherwise v would belong to U, contradicting that $v + U$ is non-zero. Thus λ is an eigenvalue of T.

(2) Let $u_1, \ldots, u_n$ be a basis of U. Because $Tv + U = \lambda v + U$ and U is invariant under $T - \lambda I$, the list

$$(T - \lambda I)v, (T - \lambda I)u_1, \ldots, (T - \lambda I)u_n$$

is contained in U. This is a list of $n + 1$ vectors in an n-dimensional space and hence must be linearly dependent, i.e. there are scalars $a_0, a_1, \ldots, a_n$, not all zero, such that $w := a_0 v + a_1 u_1 + \cdots + a_n u_n$ satisfies $(T - \lambda I)w = 0$. Note that w must be non-zero: if $w = 0$ and $a_0 \neq 0$ then $v \in U$, contradicting $v + U \neq 0$, and if $w = 0$ and $a_0 = 0$ then $a_1 u_1 + \cdots + a_n u_n$ is a non-trivial linear combination, contradicting the linear independence of $u_1, \ldots, u_n$. Thus λ is an eigenvalue of T.

Exercise 5.A.39. Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Prove that T has an eigenvalue if and only if there exists a subspace of V of dimension $\dim V - 1$ that is invariant under T.

Solution. Suppose that T has an eigenvalue $\lambda \in \mathbf{F}$ with a corresponding eigenvector $v \in V$ it follows that $\dim \operatorname{null}(T - \lambda I) \geq 1$ and hence that $\dim \operatorname{range}(T - \lambda I) \leq \dim V - 1$. By taking a basis of $\operatorname{range}(T - \lambda I)$ and, if necessary, extending it to a linearly independent list of length $\dim V - 1$, we can obtain a subspace U of V satisfying $\operatorname{range}(T - \lambda I) \subseteq U$ and $\dim U = \dim V - 1$. Exercise 5.A.39 shows that U is invariant under $T - \lambda I$, which implies that U is invariant under T:

$$u \in U \quad \Rightarrow \quad Tu - \lambda u \in U \quad \Rightarrow \quad Tu \in U.$$

Now suppose that there exists a subspace U of V such that U is invariant under T and $\dim U = \dim V - 1$, and consider the quotient operator $T/U \in \mathcal{L}(V/U)$. Since $\dim V/U = 1$, Exercise 3.A.7 shows that $T/U = \lambda I$ for some $\lambda \in \mathbf{F}$; it follows that λ is an eigenvalue of T/U and thus, by Exercise 5.A.38, λ is an eigenvalue of T.

Exercise 5.A.40. Suppose $S, T \in \mathcal{L}(V)$ and S is invertible. Suppose $p \in \mathcal{P}(\mathbf{F})$ is a polynomial. Prove that

$$p(STS^{-1}) = Sp(T)S^{-1}.$$

Solution. Notice that $(STS^{-1})^0 = I = SS^{-1} = SIS^{-1} = ST^0S^{-1}$. For a non-negative integer k, suppose that $(STS^{-1})^k = ST^kS^{-1}$ and observe that

$$(STS^{-1})^{k+1} = (STS^{-1})^k STS^{-1} = ST^kS^{-1}STS^{-1} = ST^kITS^{-1} = ST^{k+1}S^{-1}.$$

Thus $(STS^{-1})^k = ST^kS^{-1}$ for all non-negative integers k. Suppose $p = \sum_{k=0}^n a_k z^k$ and observe that

$$p(STS^{-1}) = \sum_{k=0}^n a_k (STS^{-1})^k = \sum_{k=0}^n a_k ST^kS^{-1} = S\left(\sum_{k=0}^n a_k T^k\right)S^{-1} = Sp(T)S^{-1}.$$

Exercise 5.A.41. Suppose $T \in \mathcal{L}(V)$ and U is a subspace of V invariant under T. Prove that U is invariant under $p(T)$ for every polynomial $p \in \mathcal{P}(\mathbf{F})$.

Solution. Certainly U is invariant under T^k for any non-negative integer k. Suppose $p = \sum_{k=0}^n a_k z^k$ and observe that

$$p(T)u = \left(\sum_{k=0}^n a_k T^k\right)u = \sum_{k=0}^n a_k T^k u;$$

this belongs to U since each $T^k u \in U$ and U is closed under vector addition and scalar multiplication.

Exercise 5.A.42. Define $T \in \mathcal{L}(\mathbf{F}^n)$ by $T(x_1, x_2, x_3, ..., x_n) = (x_1, 2x_2, 3x_3, ..., nx_n)$.

(a) Find all eigenvalues and eigenvectors of T.

(b) Find all subspaces of $\mathbf{F}^n$ that are invariant under T.

Solution.

(a) Let $e_1, ..., e_n$ be the standard basis of $\mathbf{F}^n$ and notice that $Te_k = ke_k$ for each $k \in \{1, ..., n\}$. Thus k is an eigenvalue of T with a corresponding eigenvector e_k. By 5.11 and 5.12 we can conclude that the eigenvalues and eigenvectors of T are precisely:

eigenvalue	corresponding eigenvectors
$k \in \{1, ..., n\}$	$\operatorname{span}(e_k) \setminus \{0\}$

(b) First, let us prove some useful results.

Lemma L.4. Suppose $T \in \mathcal{L}(V)$ and U is a subspace of V invariant under T. If $\lambda_1, \ldots, \lambda_m$ are distinct eigenvalues of T with corresponding eigenvectors $v_1, \ldots, v_m$, then

$$v_1 + \cdots + v_m \in U \quad \Leftrightarrow \quad v_k \in U \text{ for each } k \in \{1, \ldots, m\}.$$

Proof. If $v_k \in U$ for each $k \in \{1, \ldots, m\}$ then $v_1 + \cdots + v_m \in U$ since U is a subspace of V.

For the converse, we will proceed by induction on m. The base case $m = 1$ is clear, so suppose that the result holds for some positive integer m and let $\lambda_1, \ldots, \lambda_{m+1}$ be distinct eigenvalues of T with corresponding eigenvectors $v_1, \ldots, v_{m+1}$. Suppose that $v = v_1 + \cdots + v_{m+1} \in U$. Because U is invariant under T we have

$$\begin{aligned} &Tv = \lambda_1 v_1 + \cdots + \lambda_{m+1} v_{m+1} \in U \\ &\qquad \Rightarrow \quad Tv - \lambda_{m+1} v = (\lambda_1 - \lambda_{m+1}) v_1 + \cdots + (\lambda_m - \lambda_{m+1}) v_m \in U \end{aligned}$$

Let $k \in \{1, \ldots, m\}$ be given. By assumption the eigenvalues $\lambda_1, \ldots, \lambda_{m+1}$ are distinct and thus $\lambda_k - \lambda_{m+1} \neq 0$. It follows that $(\lambda_k - \lambda_{m+1}) v_k$ is an eigenvector of T corresponding to the eigenvalue λ_k. Our induction hypothesis then guarantees that $(\lambda_k - \lambda_{m+1}) v_k$ belongs to U. Thus v_k belongs to U for each $k \in \{1, \ldots, m\}$, which gives us

$$v_{m+1} = v - v_1 - \cdots - v_m \in U.$$

This completes the induction step and the proof. □

Lemma L.5. Suppose $T \in \mathcal{L}(V), \dim V = n$, and $\lambda_1, \ldots, \lambda_n$ are distinct eigenvalues of T with corresponding eigenvectors $v_1, \ldots, v_n$. If U is a subspace of V invariant under T, then

$$U = (U \cap E_1) \oplus \cdots \oplus (U \cap E_n) \quad \text{where} \quad E_k = \operatorname{span}(v_k).$$

Proof. By 5.11 the eigenvectors $v_1, \ldots, v_n$ are linearly independent and hence form a basis of V, so that $V = E_1 \oplus \cdots \oplus E_n$. For any $u \in U$ we have $u = e_1 + \cdots + e_n$, where each $e_k \in E_k$. If any $e_k = 0$ then certainly $e_k \in U$; otherwise, e_k is an eigenvector of T corresponding to the eigenvalue λ_k and thus, by Lemma L.4, the non-zero e_k's belong to U also. It follows that $u \in (U \cap E_1) + \cdots + (U \cap E_n)$ and hence that

$$U = (U \cap E_1) + \cdots + (U \cap E_n).$$

The directness of this sum follows immediately from the directness of the sum $V = E_1 \oplus \cdots \oplus E_n$. □

Lemma L.6. If $T \in \mathcal{L}(V), \dim V = n \geq 1$, and $\lambda_1, \ldots, \lambda_n$ are distinct eigenvalues of T with corresponding eigenvectors $v_1, \ldots, v_n$, then the non-zero subspaces of V which are invariant under T are precisely those of the form

$$\operatorname{span}(v_{k_1}, \ldots, v_{k_m})$$

for some choice of integers $1 \leq k_1 < \cdots < k_m \leq n$ with $1 \leq m \leq n$.

Proof. It is straightforward to verify that each $\operatorname{span}(v_{k_1}, \ldots, v_{k_m})$ is indeed a subspace of V invariant under T. For $k \in \{1, \ldots, n\}$ let $E_k = \operatorname{span}(v_k)$ and suppose U is a non-zero subspace of V invariant under T. By Lemma L.5 we have

$$U = (U \cap E_1) \oplus \cdots \oplus (U \cap E_n).$$

For each k, since $\dim E_k = 1$, we can either have $U \cap E_k = \{0\}$ or $U \cap E_k = E_k$. Because U is non-zero, there must be at least one k such that $U \cap E_k = E_k$; let $1 \leq k_1 < \cdots < k_m \leq n$ be those indices for which $U \cap E_k = E_k$. It follows that

$$U = (U \cap E_1) \oplus \cdots \oplus (U \cap E_n) = E_{k_1} \oplus \cdots \oplus E_{k_m} = \operatorname{span}(v_{k_1}, \ldots, v_{k_m}). \qquad \square$$

Now let us return to the exercise. As we showed in part (a), the eigenvalues of T are $1, \ldots, n$ with corresponding eigenvectors $e_1, \ldots, e_n$. It then follows from Lemma L.6 that the non-zero subspaces of $\mathbf{F}^n$ which are invariant under T are precisely those of the form

$$\operatorname{span}(e_{k_1}, \ldots, e_{k_m})$$

for some choice of integers $1 \leq k_1 < \cdots < k_m \leq n$ with $1 \leq m \leq n$.

Exercise 5.A.43. Suppose that V is finite-dimensional, $\dim V > 1$, and $T \in \mathcal{L}(V)$. Prove that $\{p(T) : p \in \mathcal{P}(\mathbf{F})\} \neq \mathcal{L}(V)$.

Solution. If every operator in $\mathcal{L}(V)$ could be realized as $p(T)$ for some $p \in \mathcal{P}(\mathbf{F})$, then each pair of operators in $\mathcal{L}(V)$ would commute with each other by 5.17. However, because $\dim V > 1$, Exercise 3.A.16 shows that there exist two operators in $\mathcal{L}(V)$ which do not commute with each other.

5.B. The Minimal Polynomial

Exercise 5.B.1. Suppose $T \in \mathcal{L}(V)$. Prove that 9 is an eigenvalue of T^2 if and only if 3 or -3 is an eigenvalue of T.

Solution. We showed in the solution to Exercise 5.A.22 that if 9 is an eigenvalue of T^2 then 3 or -3 is an eigenvalue of T. Conversely, suppose there is some non-zero $v \in V$ such that $Tv = \pm 3v$. It follows that $T^2 v = (\pm 3)^2 v = 9v$ and thus 9 is an eigenvalue of T^2.

Exercise 5.B.2. Suppose V is a complex vector space and $T \in \mathcal{L}(V)$ has no eigenvalues. Prove that every subspace of V invariant under T is either $\{0\}$ or infinite-dimensional.

Solution. Suppose U is a non-zero subspace of V invariant under T and consider the restriction operator $T|_U$. If U were finite-dimensional then 5.19 would imply that $T|_U$ has an eigenvalue, which would also be an eigenvalue of T. Since T has no eigenvalues it must be the case that U is infinite-dimensional.

Exercise 5.B.3. Suppose n is a positive integer and $T \in \mathcal{L}(\mathbf{F}^n)$ is defined by

$$T(x_1, \dots, x_n) = (x_1 + \dots + x_n, \dots, x_1 + \dots + x_n).$$

(a) Find all eigenvalues and eigenvectors of T.

(b) Find the minimal polynomial of T.

The matrix of T with respect to the standard basis of $\mathbf{F}^n$ consists of all 1's.

Solution.

(a) If $n = 1$ then T is the identity operator on $\mathbf{F}$, whose only eigenvalue is 1 with corresponding eigenvectors $x \in \mathbf{F} \setminus \{0\}$.

Suppose that $n \geq 2$. Some straightforward calculations reveal that

$$\operatorname{null} T = \{(-x_2 - \dots - x_n, x_2, \dots, x_n) \in \mathbf{F}^n : x_2, \dots, x_n \in \mathbf{F}\}$$
$$\text{and} \quad \operatorname{range} T = \operatorname{span}((1, \dots, 1)).$$

Note that $\dim \operatorname{null} T = n - 1$ and $\dim \operatorname{range} T = 1$. Since $\dim \operatorname{null} T \geq 1$, it follows that 0 is an eigenvalue of T. Notice that n is also an eigenvalue of T, since

$$T(1, \dots, 1) = (n, \dots, n) = n(1, \dots, 1).$$

We claim that these are the only eigenvalues of T. Indeed, if $x \neq 0$ and $\lambda \neq 0$ are such that $Tx = \lambda x$, then since $\operatorname{range} T = \operatorname{span}((1, \dots, 1))$ there must exist some $\alpha \in \mathbf{F}$ such that

$$Tx = \lambda x = \alpha(1, \dots, 1) \quad \Rightarrow \quad x = \lambda^{-1}\alpha(1, \dots, 1).$$

Thus the eigenvector x, which corresponds to the eigenvalue λ, and the eigenvector $(1, \dots, 1)$, which corresponds to the eigenvalue n, are linearly dependent. It follows from the contrapositive of 5.11 that $\lambda = n$.

Certainly $\operatorname{null} T \setminus \{0\}$ is the collection of eigenvectors of T corresponding to the eigenvalue 0. Because $\operatorname{range} T = \operatorname{span}((1, \dots, 1))$, the collection of eigenvectors of T corresponding to the eigenvalue n must be $\operatorname{span}((1, \dots, 1)) \setminus \{0\}$.

(b) If $n = 1$ then T is the identity operator on $\mathbf{F}$ and it is then clear that the minimal polynomial of T is $p(z) = z - 1$.

Suppose that $n \geq 2$ and let $p \in \mathcal{P}(\mathbf{F})$ be the minimal polynomial of T. As we showed in part (a) the eigenvalues of T are 0 and n, which must be the zeros of p by 5.27(a), i.e. p is divisible by z and $z - n$. This implies that $\deg p \geq 2$. A straightforward calculation shows that $T(T - nI) = 0$. The uniqueness of p and the minimality of its degree allow us to conclude that $p(z) = z(z - n)$.

Exercise 5.B.4. Suppose $\mathbf{F} = \mathbf{C}, T \in \mathcal{L}(V), p \in \mathcal{P}(\mathbf{C})$, and $\alpha \in \mathbf{C}$. Prove that α is an eigenvalue of $p(T)$ if and only if $\alpha = p(\lambda)$ for some eigenvalue λ of T.

Solution. Suppose that $\alpha = p(\lambda)$ for some eigenvalue λ of T. There is some non-zero $v \in V$ such that $Tv = \lambda v$; as shown in the proof of 5.27 we then have $p(T)v = p(\lambda)v = \alpha v$ and thus α is an eigenvalue of $p(T)$.

For the converse, we must assume that p is non-constant (see the errata for the third edition of LADR), or that V is finite-dimensional. To demonstrate this, let $V = \mathbf{C}^\infty, p(z) = \alpha \in \mathbf{C}$, and let $T \in \mathcal{L}(\mathbf{C}^\infty)$ be the forward shift operator. Certainly α is an eigenvalue of $p(T) = \alpha I$, but we may not express α as $p(\lambda)$ for some eigenvalue λ of T because T has no eigenvalues, as shown in Exercise 5.A.19. Since any operator on a non-zero complex vector space has an eigenvalue (5.19), we will not encounter this issue if we assume that V is finite-dimensional. The result as stated is true if we assume that p is non-constant, as we now show.

Suppose that $\deg p \geq 1$ and that α is an eigenvalue of $p(T)$, i.e. there is some non-zero $v \in V$ such that $p(T)v = \alpha v$. Let $q \in \mathcal{P}(\mathbf{C})$ be given by $q(z) = p(z) - \alpha$. Because q is a polynomial over $\mathbf{C}$, 4.13 shows that there is a factorization

$$q(z) = c(z - \lambda_1) \cdots (z - \lambda_m)$$

for some $c, \lambda_1, \dots, \lambda_m \in \mathbf{C}$. Since $\deg q = \deg p \geq 1$, it must be the case that $c \neq 0$ and $m \geq 1$. Note that $q(T)v = 0$ since $p(T)v = \alpha v$ and thus

$$0 = q(T)v = c(T - \lambda_1 I) \cdots (T - \lambda_m I)v.$$

Because $c \neq 0$ and $v \neq 0$, the equation above implies that there is some $k \in \{1, \dots, m\}$ such that $T - \lambda_k I$ is not injective, i.e. λ_k is an eigenvalue of T. Furthermore, $p(\lambda_k) = q(\lambda_k) + \alpha = \alpha$ since λ_k is a zero of q.

Exercise 5.B.5. Give an example of an operator on $\mathbf{R}^2$ that shows the result in Exercise 4 does not hold if $\mathbf{C}$ is replaced with $\mathbf{R}$.

Solution. Let $T \in \mathcal{L}(\mathbf{R}^2)$ be given by $T(x, y) = (-y, x)$, i.e. a counterclockwise rotation about the origin by 90°. Let $p(t) = t^2$ and notice that $p(T) = T^2 = -I$, since T^2 is a counterclockwise rotation about the origin by 180°. Thus $p(T)$ has the eigenvalue -1. However, we cannot possibly express -1 as $p(\lambda)$ for some eigenvalue λ of T because T has no eigenvalues, as shown in 5.9(a).

Exercise 5.B.6. Suppose $T \in \mathcal{L}(\mathbf{F}^2)$ is defined by $T(w, z) = (-z, w)$. Find the minimal polynomial of T.

Solution. Let e_1, e_2 be the standard basis of $\mathbf{F}^2$ and note that $Te_1 = e_2$ and $T^2e_1 = -e_1$. Note further that the system

$$c_0e_1 + c_1Te_1 = -T^2e_1, \quad \text{i.e.} \quad c_0e_1 + c_1e_2 = e_1,$$

has the unique solution $c_0 = 1$ and $c_1 = 0$. As shown in the textbook (see the discussion after 5.24), this implies that the minimal polynomial of T is $p(t) = 1 + t^2$.

Exercise 5.B.7.

(a) Give an example of $S, T \in \mathcal{L}(\mathbf{F}^2)$ such that the minimal polynomial of ST does not equal the minimal polynomial of TS.

(b) Suppose V is finite-dimensional and $S, T \in \mathcal{L}(V)$. Prove that if at least one of S, T is invertible, then the minimal polynomial of ST equals the minimal polynomial of TS.

Hint: Show that if S is invertible and $p \in \mathcal{P}(\mathbf{F})$, then $p(TS) = S^{-1}p(ST)S$.

Solution.

(a) Let $S, T \in \mathcal{L}(\mathbf{F}^2)$ be the operators whose matrices with respect to the standard basis of $\mathbf{F}^2$ are

$$\mathcal{M}(S) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad \mathcal{M}(T) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

A simple computation shows that

$$\mathcal{M}(ST) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad \mathcal{M}(TS) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

It follows that the minimal polynomial of TS is $p(z) = z$. Since $p(ST) = ST \neq 0$, it cannot be the case that p is the minimal polynomial of ST.

(b) Suppose at least one of S, T is invertible, say S (the case where T is invertible is handled similarly). Notice that $(TS)^0 = I = S^{-1}IS = S^{-1}(ST)^0 S$. Suppose that $(TS)^k = S^{-1}(ST)^k S$ for some non-negative integer k and observe that

$$(TS)^{k+1} = (TS)^k TS = S^{-1}(ST)^k STS = S^{-1}(ST)^{k+1} S.$$

It follows by induction that $(TS)^k = S^{-1}(ST)^k S$ for all non-negative integers k. Now suppose $p(z) = \sum_{k=0}^{n} a_k z^k$ is some polynomial in $\mathcal{P}(\mathbf{F})$ and observe that

$$p(TS) = \sum_{k=0}^{n} a_k (TS)^k = \sum_{k=0}^{n} a_k S^{-1}(ST)^k S = S^{-1}\left(\sum_{k=0}^{n} a_k (ST)^k\right) S = S^{-1}p(ST)S.$$

We can now show that the minimal polynomial of ST equals the minimal polynomial of TS. Let p be the minimal polynomial of ST and let q be the minimal polynomial of TS. Notice that

$$p(TS) = S^{-1}p(ST)S = 0 \quad \text{and} \quad 0 = q(TS) = S^{-1}q(ST)S \quad \Rightarrow \quad q(ST) = 0.$$

It follows from 5.29 that p is a polynomial multiple of q and q is a polynomial multiple of p. This implies $p = cq$ for some $c \in \mathbf{F}$ (as the next lemma shows); because p and q are monic we must have $c = 1$ and thus $p = q$.

Lemma L.7. Suppose $p, q \in \mathcal{P}(\mathbf{F})$ are non-zero. If p is a polynomial multiple of q and q is a polynomial multiple of p then $p = cq$ for some $c \in \mathbf{F} \setminus \{0\}$.

Proof. There are polynomials $r, s \in \mathcal{P}(\mathbf{F})$ such that $p = rq$ and $q = sp$, which gives us $p = rsp$. This implies that $r(x)s(x) = 1$ for all $x \in \mathbf{F}$ such that $p(x) \neq 0$; there are finitely many such x because p is non-zero. Thus $r(x)s(x) = 1$ holds for infinitely many x and so must hold for all $x \in \mathbf{F}$ (otherwise $rs - 1$ is a non-zero polynomial with infinitely many roots). This equation forces $\deg r = 1$, so that $r = c$ for some $c \in \mathbf{F} \setminus \{0\}$. Thus $p = cq$. $\square$

Exercise 5.B.8. Suppose $T \in \mathcal{L}(\mathbf{R}^2)$ is the operator of counterclockwise rotation by $1°$. Find the minimal polynomial of T.

Because $\dim \mathbf{R}^2 = 2$, *the degree of the minimal polynomial of* T *is at most* 2. *Thus the minimal polynomial of* T *is not the tempting polynomial* $x^{180} + 1$, *even though* $T^{180} = -I$.

Solution. Let e_1, e_2 be the standard basis of $\mathbf{R}^2$ and observe that

$$Te_1 = \cos\left(\tfrac{\pi}{180}\right)e_1 + \sin\left(\tfrac{\pi}{180}\right)e_2 \quad \text{and} \quad T^2 e_1 = \cos\left(\tfrac{\pi}{90}\right)e_1 + \sin\left(\tfrac{\pi}{90}\right)e_2.$$

Thus, solving the system of equations $c_0 e_1 + c_1 Te_1 = -Te_2$ amounts to solving the system

$$\begin{pmatrix} 1 & \cos\left(\frac{\pi}{180}\right) \\ 0 & \sin\left(\frac{\pi}{180}\right) \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \end{pmatrix} = \begin{pmatrix} -\cos\left(\frac{\pi}{90}\right) \\ -\sin\left(\frac{\pi}{90}\right) \end{pmatrix}.$$

This system of equations has the unique solution $c_0 = 1$ and $c_1 = -2\cos\left(\frac{\pi}{180}\right)$. As shown in the textbook (see the discussion after 5.24), this implies that the minimal polynomial of T is $p(x) = 1 - 2\cos\left(\frac{\pi}{180}\right)x + x^2$.

Exercise 5.B.9. Suppose $T \in \mathcal{L}(V)$ is such that with respect to some basis of V, all entries of the matrix of T are rational numbers. Explain why all coefficients of the minimal polynomial of T are rational numbers.

Solution. By 5.22, there exists a minimal positive integer $m \leq \dim V$ such that the equation

$$c_0 I + c_1 T + \cdots + c_{m-1} T^{m-1} = -T^m \tag{1}$$

has a unique solution $c_0, c_1, ..., c_{m-1} \in \mathbf{F}$. Moreover, the numbers $c_0, c_1, ..., c_{m-1}, 1$ are the coefficients of the minimal polynomial of T. Thus it will suffice to show that each c_i is a rational number.

By assumption there is a basis $v_1, ..., v_n$ of V such that the entries of the matrix of T with respect to this basis are rational numbers. Let A denote this matrix and consider the matrix equation

$$x_0 I + x_1 A + \cdots + x_{m-1} A^{m-1} = -A^m.$$

As noted in the textbook (see the discussion after 5.24), this equation can be thought of as a system of n^2 equations in the m unknowns $x_0, x_1, ..., x_{m-1}$. That is, letting $(A^i)_{j,k}$ be the entry in the j^{th} row and k^{th} column of A^i, for each $j, k \in \{1, ..., n\}$ we have a linear equation

$$\sum_{i=0}^{m-1} (A^i)_{j,k}\, x_i = (-A^m)_{j,k}.$$

Because the entries of A are rational, it follows from the definition of matrix multiplication that each $(A^i)_{j,k}$ is also rational. Thus each coefficient in this system of equations is a rational number. If this system has a solution $(x_0, x_1, ..., x_{m-1})$, then Gaussian elimination (or some other method) shows that each x_i is a rational function of the coefficients of the system; it follows that each x_i is rational. Now observe that equation (1) implies that $(c_0, c_1, ..., c_{m-1})$ is a solution of this system of equations. Thus each c_i is rational.

Exercise 5.B.10. Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and $v \in V$. Prove that

$$\operatorname{span}(v, Tv, ..., T^m v) = \operatorname{span}\left(v, Tv, ..., T^{\dim V - 1} v\right)$$

for all integers $m \geq \dim V - 1$.

Solution. Let $n = \dim V$ and let $U_m = \operatorname{span}(v, Tv, \ldots, T^m v)$ for a non-negative integer m; define $U_{-1} = \{0\}$. Our goal is to show that $U_m = U_{n-1}$ for all $m \geq n-1$.

First we will use induction to show that if $T^{k+1}v \in U_k$ for some $k \geq -1$ then $T^m v \in U_k$ for all $m \geq k+1$. The base case $m = k+1$ is clear, so suppose that $T^m v \in U_k$ for some $m \geq k+1$, i.e.

$$T^m v = a_0 v + a_1 Tv + \cdots + a_{k-1}T^{k-1}v + a_k T^k v.$$

Now observe that

$$T^{m+1}v = a_0 Tv + a_1 T^2 v + \cdots + a_{k-1}T^k v + a_k T^{k+1}v.$$

Certainly $a_0 Tv + a_1 T^2 v + \cdots + a_{k-1}T^k v \in U_k$, and since $T^{k+1}v$ belongs to U_k by assumption, it follows that $T^{m+1}v \in U_k$. This completes the induction step.

The previous result implies that if $T^{k+1}v \in U_k$ for some k then $U_m = U_k$ for all $m \geq k$. For a non-negative integer m, note that if $T^{k+1}v \notin U_k$ for all $k \in \{-1, \ldots, m-1\}$ then the linear dependence lemma (2.19) shows that the list $v, Tv, \ldots, T^m v$ is linearly independent. Because the list $v, Tv, \ldots, T^n v$ is linearly dependent (it has length $n+1$ and $\dim V = n$), it follows that there exists some $k \in \{-1, \ldots, n-1\}$ such that $T^{k+1}v \in U_k$, which implies that $U_m = U_k$ for all $m \geq k$. In particular, $U_m = U_{n-1}$ for all $m \geq n-1$.

Exercise 5.B.11. Suppose V is a two-dimensional vector space, $T \in \mathcal{L}(V)$, and the matrix of T with respect to some basis of V is $\begin{pmatrix} a & c \\ b & d \end{pmatrix}$.

(a) Show that $T^2 - (a+d)T + (ad-bc)I = 0$.

(b) Show that the minimal polynomial of T equals

$$\begin{cases} z - a & \text{if } b = c = 0 \text{ and } a = d, \\ z^2 - (a+d)z + (ad-bc) & \text{otherwise.} \end{cases}$$

Solution.

(a) Letting $A = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$, a straightforward calculation shows that

$$A^2 - (a+d)A + (ad-bc)I = 0.$$

It follows that $T^2 - (a+d)T + (ad-bc)I = 0$.

(b) If $b = c = 0$ and $a = d$ then $T = aI$ and it is then clear that the minimal polynomial of T is $z - a$. If $b \neq 0$, or $c \neq 0$, or $a \neq d$, then T is not a scalar multiple of the identity. It follows that the equation $xI = -T$ has no solution for $x \in \mathbf{F}$ and thus the degree of the minimal polynomial of T must be at least 2. Since the degree of the minimal polynomial of T can be at most $\dim V = 2$, we see that the degree of the minimal polynomial of T must equal 2. Thus, because $p(z) = z^2 - (a+d)z + (ad-bc)$ is monic, satisfies $p(T) = 0$, and has degree equal to the degree of the minimal polynomial of T, it follows from the uniqueness of the minimal polynomial of an operator that p is the minimal polynomial of T.

Exercise 5.B.12. Define $T \in \mathcal{L}(\mathbf{F}^n)$ by $T(x_1, x_2, x_3, ..., x_n) = (x_1, 2x_2, 3x_3, ..., nx_n)$. Find the minimal polynomial of T.

Solution. Let $p \in \mathcal{P}(\mathbf{F})$ be the minimal polynomial of T. As we showed in Exercise 5.A.42 (a), each $k \in \{1, ..., n\}$ is an eigenvalue of T. It follows from 5.27 that each $k \in \{1, ..., n\}$ is a zero of p, which implies $\deg p \geq n$. Since $\deg p \leq n$ by 5.22, it must be the case that

$$p(z) = (z-1)(z-2)(z-3)\cdots(z-n).$$

Exercise 5.B.13. Suppose $T \in \mathcal{L}(V)$ and $p \in \mathcal{P}(\mathbf{F})$. Prove that there exists a unique $r \in \mathcal{P}(\mathbf{F})$ such that $p(T) = r(T)$ and $\deg r$ is less than the degree of the minimal polynomial of T.

Solution. Let $q \in \mathcal{P}(\mathbf{F})$ be the minimal polynomial of T, which must be non-zero. The division algorithm for polynomials (4.9) shows that there exist unique polynomials $s, r \in \mathcal{P}(\mathbf{F})$ such that $p = sq + r$ and $\deg r < \deg q$. Because $q(T) = 0$, it follows that $p(T) = r(T)$.

Exercise 5.B.14. Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$ has minimal polynomial $4 + 5z - 6z^2 - 7z^3 + 2z^4 + z^5$. Find the minimal polynomial of T^{-1}.

Solution. We will use the following useful lemma.

Lemma L.8. Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$ is invertible. If the minimal polynomial of T is $p \in \mathcal{P}(\mathbf{F})$ with $p = \sum_{k=0}^{m} c_k z^k$, then the minimal polynomial of T^{-1} is

$$\frac{1}{c_0} \sum_{k=0}^{m} c_{m-k} z^k.$$

(Note that $c_0 \neq 0$ since T is invertible.)

Proof. Let $q = \frac{1}{c_0} \sum_{k=0}^{m} c_{m-k} z^k$ and note that q is monic and $\deg q = m$. Note further that

$$\begin{aligned} 0 = T^{-m} p(T) = T^{-m} \sum_{k=0}^{m} c_k T^k &= \sum_{k=0}^{m} c_k T^{k-m} \\ &= \sum_{k=0}^{m} c_k \left(T^{-1}\right)^{m-k} = \sum_{k=0}^{m} c_{m-k} \left(T^{-1}\right)^k = q\left(T^{-1}\right). \end{aligned}$$

It follows that the degree of the minimal polynomial of T^{-1} is at most $\deg q = m$. By replacing T with T^{-1} in the previous argument and using that $\left(T^{-1}\right)^{-1} = T$, we see that the minimal polynomials of T and T^{-1} must have the same degree. Because q is monic, satisfies $q\left(T^{-1}\right) = 0$, and has the same degree as the minimal polynomial of T^{-1}, it follows from the uniqueness of the minimal polynomial of an operator that q is the minimal polynomial of T^{-1}. $\square$

It is now immediate from Lemma L.8 that the minimal polynomial of T^{-1} is

$$\frac{1}{4} + \frac{1}{2} z - \frac{7}{4} z^2 - \frac{3}{2} z^3 + \frac{5}{4} z^4 + z^5.$$

Exercise 5.B.15. Suppose V is a finite-dimensional complex vector space with $\dim V > 0$ and $T \in \mathcal{L}(V)$. Define $f : \mathbf{C} \to \mathbf{R}$ by

$$f(\lambda) = \dim \operatorname{range}(T - \lambda I).$$

Prove that f is not a continuous function.

Solution. Let $m = \dim V$ and note that, by 5.19, there exists an eigenvalue $\lambda \in \mathbf{C}$ of T. It follows from 5.7 that $T - \lambda I$ is not surjective, so that $f(\lambda) < m$. Consider the sequence $(\lambda_n)_{n=1}^{\infty}$ of distinct complex numbers given by $\lambda_n = \lambda + \frac{1}{n}$, which satisfies $\lim_{n \to \infty} \lambda_n = \lambda$. Because T can have at most m distinct eigenvalues 5.13, we may choose a subsequence $(\lambda_{n_k})_{k=1}^{\infty}$ such that each λ_{n_k} is not an eigenvalue of T. By 5.7 each operator $T - \lambda_{n_k} I$ must be surjective. It follows that $f(\lambda_{n_k}) = m$ for each positive integer k, which implies

$$\lim_{k \to \infty} f(\lambda_{n_k}) = m > f(\lambda) = f\left(\lim_{k \to \infty} \lambda_{n_k}\right).$$

Thus f is not continuous at λ.

Exercise 5.B.16. Suppose $a_0, \ldots, a_{n-1} \in \mathbf{F}$. Let T be the operator on $\mathbf{F}^n$ whose matrix (with respect to the standard basis) is

$$\begin{pmatrix} 0 & & & & -a_0 \\ 1 & 0 & & & -a_1 \\ & 1 & \ddots & & -a_2 \\ & & \ddots & & \vdots \\ & & & 0 & -a_{n-2} \\ & & & 1 & -a_{n-1} \end{pmatrix},$$

Here all entries of the matrix are 0 except for all 1's on the line under the diagonal and the entries in the last column (some of which might also be 0). Show that the minimal polynomial of T is the polynomial

$$a_0 + a_1 z + \cdots + a_{n-1} z^{n-1} + z^n.$$

The matrix above is called the ***companion matrix*** *of the polynomial above. This exercise shows that every monic polynomial is the minimal polynomial of some operator. Hence a formula or an algorithm that could produce exact eigenvalues for each operator on each* $\mathbf{F}^n$ *could then produce exact zeros for each polynomial [by 5.27(a)]. Thus there is no such formula or algorithm. However, efficient numeric methods exist for obtaining very good approximations for the eigenvalues of an operator.*

Solution. Let $e_0, \ldots, e_{n-1}$ be the standard basis of $\mathbf{F}^n$ and observe that

$$Te_0 = e_1, \quad T^2 e_0 = e_2, \quad \ldots, \quad T^{n-1} e_0 = e_{n-1}, \quad T^n e_0 = -(a_0 e_0 + \cdots + a_{n-1} e_{n-1}).$$

It follows that the equation $c_0 e_0 + c_1 T e_0 + c_2 T^2 e_0 + \cdots + c_{n-1} T^{n-1} e_0 = -T^n e_0$ is equivalent to

$$c_0 e_0 + c_1 e_1 + c_2 e_2 + \cdots + c_{n-1} e_{n-1} = a_0 e_0 + a_1 e_1 + a_2 e_2 + \cdots + a_{n-1} e_{n-1}.$$

By unique representation, this equation has the unique solution $(c_0, \ldots, c_{n-1}) = (a_0, \ldots, a_{n-1})$. It follows (see the discussion in the textbook after 5.24) that the minimal polynomial of T is $a_0 + a_1 z + \cdots + a_{n-1} z^{n-1} + z^n$.

Exercise 5.B.17. Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and p is the minimal polynomial of T. Suppose $\lambda \in \mathbf{F}$. Show that the minimal polynomial of $T - \lambda I$ is the polynomial q defined by $q(z) = p(z + \lambda)$.

Solution. Let s be the minimal polynomial of $T - \lambda I$. Notice that q is monic and satisfies $\deg q = \deg p$. Notice further that

$$q(T - \lambda I) = p(T - \lambda I + \lambda I) = p(T) = 0.$$

It follows that $\deg s \le \deg q = \deg p$. Now define a polynomial r by $r(z) = s(z - \lambda)$ and observe that $\deg r = \deg s$ and that

$$r(T) = s(T - \lambda I) = 0.$$

It follows that $\deg p \le \deg r = \deg s$. Thus $\deg s = \deg p$. Because q is a monic polynomial satisfying $q(T - \lambda I) = 0$, and the degree of q equals the degree of the minimal polynomial of $T - \lambda I$, it follows from the uniqueness of the minimal polynomial of an operator that q is the minimal polynomial of $T - \lambda I$.

Exercise 5.B.18. Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and p is the minimal polynomial of T. Suppose $\lambda \in \mathbf{F} \setminus \{0\}$. Show that the minimal polynomial of λT is the polynomial q defined by $q(z) = \lambda^{\deg p} p\big(\frac{z}{\lambda}\big)$.

Solution. Let s be the minimal polynomial of λT. Observe that $\deg q = \deg p$ and that

$$q(\lambda T) = \lambda^{\deg p} p(T) = 0.$$

Thus $\deg s \le \deg q$. Let r be given by $r(z) = s(\lambda z)$ and notice that $\deg r = \deg s$ since $\lambda \neq 0$. Furthermore, $r(T) = s(\lambda T) = 0$, which implies $\deg p \le \deg r$, i.e. $\deg q \le \deg s$. Thus $\deg s = \deg q$. Because p is monic, the highest degree term of $p\big(\frac{z}{\lambda}\big)$ is $\lambda^{-\deg p} z^{\deg p}$ and it follows that q is monic. Thus q is a monic polynomial satisfying $q(\lambda T) = 0$ whose degree equals the degree of the minimal polynomial of λT. The uniqueness of the minimal polynomial of an operator then implies that q is the minimal polynomial of λT.

Exercise 5.B.19. Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Let $\mathcal{E}$ be the subspace of $\mathcal{L}(V)$ defined by

$$\mathcal{E} = \{q(T) : q \in \mathcal{P}(\mathbf{F})\}.$$

Prove that $\dim \mathcal{E}$ equals the degree of the minimal polynomial of T.

Solution. Let p be the minimal polynomial of T. Define $\Phi \in \mathcal{L}(\mathcal{P}(\mathbf{F}), \mathcal{L}(V))$ by $\Phi q = q(T)$ and notice that $\operatorname{range} \Phi = \mathcal{E}$. Notice further that, by 5.29, $\operatorname{null} \Phi = \{pq : q \in \mathcal{P}(\mathbf{F})\}$. Exercise 4.13 shows that $\dim \mathcal{P}(\mathbf{F})/(\operatorname{null} \Phi) = \deg p$ and 3.107 shows that Φ is an isomorphism from $\mathcal{P}(\mathbf{F})/(\operatorname{null} \Phi)$ onto $\mathcal{E}$. Thus $\dim \mathcal{E} = \deg p$.

Exercise 5.B.20. Suppose $T \in \mathcal{L}(\mathbf{F}^4)$ is such that the eigenvalues of T are $3, 5, 8$. Prove that $(T - 3I)^2 (T - 5I)^2 (T - 8I)^2 = 0$.

Solution. Let p be the minimal polynomial of T and let $q(z) = (z - 3)^2 (z - 5)^2 (z - 8)^2$. 5.27 shows that $3, 5, 8$ are zeros of p, so that p is of the form

$$p(z) = s(z)(z - 3)(z - 5)(z - 8)$$

for some polynomial s. Because $\deg p \leq \dim \mathbf{F}^4 = 4$ and $3, 5, 8$ are the only zeros of p, we must have $s \in \{1, z-3, z-5, z-8\}$. It follows that q is a polynomial multiple of p and thus, by 5.29, $q(T) = 0$.

Exercise 5.B.21. Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Prove that the minimal polynomial of T has degree at most $1 + \dim \operatorname{range} T$.

If $\dim \operatorname{range} T < \dim V - 1$, *then this exercise gives a better upper bound than 5.22 for the degree of the minimal polynomial of* T.

Solution. Let $p \in \mathcal{P}(\mathbf{F})$ be the minimal polynomial of T and let $q \in \mathcal{P}(\mathbf{F})$ be the minimal polynomial of $T|_{\operatorname{range} T}$. For any $v \in V$ observe that

$$q(T)Tv = q\big(T|_{\operatorname{range} T}\big)Tv = 0.$$

Thus $q(T)T = 0$. It follows from the minimality of $\deg p$ that

$$\deg p \leq \deg(xq(x)) = 1 + \deg q \leq 1 + \dim \operatorname{range} T.$$

Exercise 5.B.22. Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Prove that T is invertible if and only if $I \in \operatorname{span}\big(T, T^2, \ldots, T^{\dim V}\big)$.

Solution. Let $p(z) = c_0 + c_1 z + \cdots + c_{m-1} z^{m-1} + z^m$ be the minimal polynomial of T. If T is invertible then 5.32 shows that $c_0 \neq 0$ and thus

$$\begin{aligned} &c_0 I + c_1 T + \cdots + c_{m-1} T^{m-1} + T^m = 0 \\ &\qquad\qquad \Rightarrow \quad I = -c_0^{-1}\big(c_1 T + \cdots + c_{m-1} T^{m-1} + T^m\big) \in \operatorname{span}\big(T, \ldots, T^{\dim V}\big). \end{aligned}$$

Now suppose that $I \in \operatorname{span}(T, T^2, \ldots, T^n)$, where $n = \dim V$, so that

$$I = a_1 T + a_2 T^2 + \cdots + a_n T^n$$

for some $a_1, a_2, \ldots, a_n \in \mathbf{F}$. Let $q \in \mathcal{P}(\mathbf{F})$ be given by $q(z) = -1 + a_1 z + a_2 z^2 + \cdots + a_n z^n$ and note that $q(T) = 0$. It follows from 5.29 that q is a polynomial multiple of p. Because 0 is not a root of q, it must be that 0 is not a root of p either, i.e. the constant term of p is not zero. 5.32 allows us to conclude that T is invertible.

Exercise 5.B.23. Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Let $n = \dim V$. Prove that if $v \in V$, then $\operatorname{span}(v, Tv, \ldots, T^{n-1}v)$ is invariant under T.

Solution. It will suffice to show that $T^n v \in \operatorname{span}(v, Tv, \ldots, T^{n-1}v)$. This is immediate from Exercise 5.B.10.

Exercise 5.B.24. Suppose V is a finite-dimensional complex vector space. Suppose $T \in \mathcal{L}(V)$ is such that 5 and 6 are eigenvalues of T and that T has no other eigenvalues. Prove that $(T - 5I)^{\dim V - 1}(T - 6I)^{\dim V - 1} = 0$.

Solution. Let $n = \dim V$ and note that because T has 2 distinct eigenvalues, 5.11 implies $n \geq 2$. Let $q(z) = (z-5)^{n-1}(z-6)^{n-1}$, and let p be the minimal polynomial of T. Since 5 and 6 are the only eigenvalues of T, 5.27(b) shows that p is of the form $p(z) = (z-5)^k(z-6)^\ell$ for some positive integers k, ℓ. Thus q is a polynomial multiple of p and it then follows from 5.29 that $q(T) = 0$.

Exercise 5.B.25. Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and U is a subspace of V that is invariant under T.

(a) Prove that the minimal polynomial of T is a polynomial multiple of the minimal polynomial of the quotient operator T/U.

(b) Prove that

$$(\text{minimal polynomial of } T|_U) \times (\text{minimal polynomial of } T/U)$$

is a polynomial multiple of the minimal polynomial of T.

The quotient operator T/U was defined in Exercise 38 in Section 5A.

Solution.

(a) We will use the following lemma.

Lemma L.9. Suppose $T \in \mathcal{L}(V), U$ is a subspace of V invariant under T, and $\pi : V \to V/U$ is the quotient map. If $p \in \mathcal{P}(\mathbf{F})$ then $\pi p(T) = p(T/U)\pi$.

Proof. Suppose $p = \sum_{k=0}^m c_k z^k$ and let $\pi : V \to V/U$ be the quotient map. For any non-negative integer k and any $v \in V$, the definition of the quotient operator implies that $(T/U)^k(v + U) = T^k v + U$. Thus, for any $v \in V$,

$$\pi p(T)v = \sum_{k=0}^m c_k \pi\left(T^k v\right) = \sum_{k=0}^m c_k\left(T^k v + U\right)$$
$$= \sum_{k=0}^m c_k (T/U)^k (v + U) = p(T/U)(\pi(v)).$$

It follows that $\pi p(T) = p(T/U)\pi$. □

Let $p = \sum_{k=0}^m c_k z^k$ be the minimal polynomial of T and let $\pi : V \to V/U$ be the quotient map. For any $v + U \in V/U$, it follows from Lemma L.9 that

$$p(T/U)(v + U) = p(T/U)(\pi(v)) = \pi p(T)v = \pi(0) = 0.$$

Thus $p(T/U) = 0$. It then follows from 5.29 that p is a multiple of the minimal polynomial of T/U.

(b) Let r and s be the minimal polynomials of $T|_U$ and T/U, and let $\pi : V \to V/U$ be the quotient map. By 2.33 there is a subspace W of V such that $V = U \oplus W$. For any $u \in U$, notice that $r(T)u = r(T|_U)u = 0$. It then follows from Lemma L.9 that, for any $w \in W$,

$$\pi(s(T)w) = s(T/U)(w + U) = 0 \quad \Rightarrow \quad s(T)w \in U \quad \Rightarrow \quad r(T)s(T)w = 0.$$

Let $v = u + w \in V$ be given and observe that

$$r(T)s(T)v = r(T)s(T)u + r(T)s(T)w = s(T)r(T)u = 0.$$

Thus $r(T)s(T) = 0$ and it then follows from 5.29 that rs is a polynomial multiple of the minimal polynomial of T.

Exercise 5.B.26. Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and U is a subspace of V that is invariant under T. Prove that the set of eigenvalues of T equals the union of the set of eigenvalues of $T|_U$ and the set of eigenvalues of T/U.

Solution. Let p, r, and s be the minimal polynomials of $T, T|_U$, and T/U, and let Z_p, Z_r, and Z_s be the collection of zeros of p, r, and s. By 5.27(a), it will suffice to show that $Z_p = Z_r \cup Z_s$. It follows from Exercise 5.B.25 and 5.31 that there exist polynomials q, a, b such that:

(1) $r(x)s(x) = p(x)q(x)$;

(2) $p(x) = a(x)r(x)$;

(3) $p(x) = b(x)s(x)$.

Equation (1) shows that if $\lambda \in \mathbf{F}$ is such that $p(\lambda) = 0$ then $r(\lambda) = 0$ or $s(\lambda) = 0$. That is, $Z_p \subseteq Z_r \cup Z_s$. Equations (2) and (3) show that if $\lambda \in \mathbf{F}$ is such that $r(\lambda) = 0$ or $s(\lambda) = 0$ then $p(\lambda) = 0$. That is, $Z_r \cup Z_s \subseteq Z_p$. Thus $Z_p = Z_r \cup Z_s$.

Exercise 5.B.27. Suppose $\mathbf{F} = \mathbf{R}$, V is finite-dimensional, and $T \in \mathcal{L}(V)$. Prove that the minimal polynomial of $T_{\mathbf{C}}$ equals the minimal polynomial of T.

The complexification $T_{\mathbf{C}}$ was defined in Exercise 33 of Section 3B.

Solution. Let $p \in \mathcal{P}(\mathbf{R})$ be the minimal polynomial of T and let $q \in \mathcal{P}(\mathbf{C})$ be the minimal polynomial of $T_{\mathbf{C}}$. Because we can identify $a \in \mathbf{R}$ with $a + 0i \in \mathbf{C}$, we can think of p as a polynomial with complex coefficients. Thus it makes sense to consider $p(T_{\mathbf{C}})$. From the definitions of $T_{\mathbf{C}}$ and of scalar multiplication in $V_{\mathbf{C}}$, note that, for any non-negative integer k, any $a \in \mathbf{R}$, and any $u + iv \in V_{\mathbf{C}}$,

$$T_{\mathbf{C}}^k(u + iv) = \left(T^k u\right) + i\left(T^k v\right) \quad \text{and} \quad aT_{\mathbf{C}}(u + iv) = (aTu) + i(aTv).$$

Combining this with the definition of vector addition in $V_{\mathbf{C}}$, for all $u + iv \in V_{\mathbf{C}}$ it follows that

$$p(T_{\mathbf{C}})(u + iv) = (p(T)u) + i(p(T)v) = 0 + 0i.$$

Thus $p(T_{\mathbf{C}}) = 0$ and hence, by 5.29, p is a polynomial multiple of q, where we think of p as an element of $\mathcal{P}(\mathbf{C})$.

Let $\mathcal{B} := v_1, \ldots, v_n$ be a basis of V and let $\mathcal{B}_{\mathbf{C}} := v_1 + 0i, \ldots, v_n + 0i$; it follows from Exercise 2.B.11 that $\mathcal{B}_{\mathbf{C}}$ is a basis of $V_{\mathbf{C}}$. Because $T_{\mathbf{C}}(v_k + 0i) = Tv_k + 0i$, the matrix of $T_{\mathbf{C}}$ with respect to $\mathcal{B}_{\mathbf{C}}$ must be equal to the matrix of T with respect to $\mathcal{B}$, where we think of $\mathcal{M}(T, \mathcal{B})$ as a matrix with complex entries. Letting A denote this matrix, it follows that each entry of A is a real number and hence that $\overline{A} = A$, where $\overline{A}$ is the matrix obtained by taking the complex conjugate of each entry of A. Note that, for any non-negative integer k, each entry of A^k must also be a real number and thus $\overline{A^k} = A^k$. Suppose that $q = \sum_{k=0}^{m} a_k z^k$ and observe that

$$0 = q(A) \quad \Rightarrow \quad 0 = \overline{q(A)} = \overline{\sum_{k=0}^{m} a_k A^k} = \sum_{k=0}^{m} \overline{a_k A^k} = \sum_{k=0}^{m} \overline{a_k}\,\overline{A^k} = \sum_{k=0}^{m} \overline{a_k} A^k;$$

the algebraic properties of "matrix complex conjugation" used here follow quickly from 4.4. Thus, letting $\overline{q} = \sum_{k=0}^{m} \overline{a_k} z^k$, we have $\overline{q}(A) = 0$, which implies that $(q - \overline{q})(A) = 0$ and hence that $(q - \overline{q})(T_{\mathbf{C}}) = 0$. Note that, because q and $\overline{q}$ are monic, we have $\deg(q - \overline{q}) < \deg q$. It must then be the case that $q - \overline{q}$ is the zero polynomial, since q is the minimal polynomial of $T_{\mathbf{C}}$. This gives us $a_k = \overline{a_k}$, i.e. $a_k \in \mathbf{R}$, for each k and thus q can be thought of as a polynomial with real coefficients, so that $q(T)$ makes sense. Because A is also the matrix of T with respect to $\mathcal{B}$, the equation $0 = q(A)$ shows that $0 = q(T)$ and thus, by 5.29, q must be a polynomial multiple of p. We may now appeal to Lemma L.7 and the fact that p and q are both monic to conclude that $p = q$.

Exercise 5.B.28. Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Prove that the minimal polynomial of $T' \in \mathcal{L}(V')$ equals the minimal polynomial of T.

The dual map V' was defined in Section 3F.

Solution. Let p be the minimal polynomial of T and let q be the minimal polynomial of T'. For any $s \in \mathcal{P}(\mathbf{F})$, 3.120 shows that $s(T)' = s(T')$, and Exercise 3.F.16 shows that $0' = 0$. Thus

$$p(T) = 0 \quad \Rightarrow \quad p(T)' = 0' \quad \Leftrightarrow \quad p(T') = 0.$$

It follows from 5.29 that p is a polynomial multiple of q. Let $\Lambda : V \to V''$ be the isomorphism defined in Exercise 3.F.32 and observe that

$$0 = q(T') = q(T)' \quad \Rightarrow \quad q(T)'' = 0 \quad \Rightarrow \quad q(T)'' \circ \Lambda = 0 \quad \Rightarrow \quad \Lambda \circ q(T) = 0 \quad \Rightarrow \quad q(T) = 0,$$

where we have used Exercise 3.F.32 (b) for the third implication and the injectivity of Λ for the last implication. Thus, by 5.29, q must be a polynomial multiple of p. We may now appeal to Lemma L.7 and the fact that p and q are both monic to conclude that $p = q$.

Exercise 5.B.29. Show that every operator on a finite-dimensional vector space of dimension at least two has an invariant subspace of dimension two.

Exercise 6 in Section 5C will give an improvement of this result when $\mathbf{F} = \mathbf{C}$.

Solution. For an integer $k \geq 2$, let $P(k)$ be the statement that any operator on a vector space of dimension k has an invariant subspace of dimension two. We will proceed by induction on k. For the base case $P(2)$, we can take the invariant subspace to be the vector space itself.

Now suppose that $P(k)$ holds for some $k \geq 2$, let T be an operator on some vector space V satisfying $\dim V = k + 1$, and let $p \in \mathcal{P}(\mathbf{F})$ be the minimal polynomial of T; note that $\deg p \geq 1$ since $V \neq \{0\}$.

If p has a linear factor, i.e. if p has a zero, then 5.27 shows that T has an eigenvalue. It then follows from Exercise 5.A.39 that there exists a subspace U of V which is invariant under T and satisfies $\dim U = k$. Our induction hypothesis guarantees that there is a subspace W of U which is invariant under $T|_U$ and such that $\dim W = 2$. It follows that W is a two-dimensional subspace of V which is invariant under T.

If p has no linear factor then note that T has no eigenvalues by 5.27(a). Note further that, by the fundamental theorem of algebra (4.12/4.13), we must have $\mathbf{F} = \mathbf{R}$. It then follows from 4.16 that p has a factorization $p = f_1 \cdots f_m$ where each $f_j \in \mathcal{P}(\mathbf{R})$ is quadratic. Because $0 = p(T) = f_1(T) \cdots f_m(T)$, there must exist some $j \in \{1, \ldots, m\}$ such that $f_j(T)$ is not injective, i.e. there exists some non-zero $v \in V$ such that $f_j(T)v = 0$. Since f_j is quadratic we have $f_j(x) = ax^2 + bx + c$ for some $a, b, c \in \mathbf{R}$ such that $a \neq 0$. Thus

$$0 = f_j(T)v = aT^2v + bTv + cv \quad \Rightarrow \quad T^2v \in \operatorname{span}(v, Tv).$$

It follows that $\operatorname{span}(v, Tv)$ is invariant under T. Furthermore, because $v \neq 0$ and T has no eigenvalues, we must have $\dim \operatorname{span}(v, Tv) = 2$. Thus V has a two-dimensional subspace which is invariant under T.

In either case, $P(k + 1)$ holds. This completes the induction step and the proof.

5.C. Upper-Triangular Matrices

Exercise 5.C.1. Prove or give a counterexample: If $T \in \mathcal{L}(V)$ and T^2 has an upper-triangular matrix with respect to some basis of V, then T has an upper triangular matrix with respect to some basis of V.

Solution. This is false. For a counterexample, take $T \in \mathcal{L}(\mathbf{R}^2)$ to be a counterclockwise rotation about the origin by 90°, i.e. $T(x, y) = (-y, x)$. Then $T^2 = -I$, so that the matrix of T^2 with respect to the standard basis of $\mathbf{R}^2$ is the upper-triangular matrix

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

However, note that if T had an upper-triangular matrix with respect to some basis of $\mathbf{R}^2$ then T would have an eigenvalue—but T has no eigenvalues, as shown in 5.9(a).

Exercise 5.C.2. Suppose A and B are upper-triangular matrices of the same size, with $\alpha_1, ..., \alpha_n$ on the diagonal of A and $\beta_1, ..., \beta_n$ on the diagonal of B.

(a) Show that $A + B$ is an upper-triangular matrix with $\alpha_1 + \beta_1, ..., \alpha_n + \beta_n$ on the diagonal.

(b) Show that AB is an upper-triangular matrix with $\alpha_1\beta_1, ..., \alpha_n\beta_n$ on the diagonal.

The results in this exercise are used in the proof of 5.81.

Solution. Note that A and B satisfy

$$A_{j,k} = \begin{cases} \alpha_j & \text{if } j = k, \\ 0 & \text{if } j > k, \end{cases} \qquad B_{j,k} = \begin{cases} \beta_j & \text{if } j = k, \\ 0 & \text{if } j > k. \end{cases}$$

(a) From the definition of matrix addition, we have

$$(A + B)_{j,k} = A_{j,k} + B_{j,k} = \begin{cases} \alpha_j + \beta_j & \text{if } j = k, \\ 0 & \text{if } j > k. \end{cases}$$

Thus $A + B$ is upper-triangular with $\alpha_1 + \beta_1, ..., \alpha_n + \beta_n$ on the diagonal.

(b) By the definition of matrix multiplication, we have

$$(AB)_{j,j} = \sum_{r=1}^{n} A_{j,r}B_{r,j} = A_{j,j}B_{j,j} + \sum_{r \neq j} A_{j,r}B_{r,j} = \alpha_j\beta_j + \sum_{r \neq j} A_{j,r}B_{r,j}.$$

For $r \neq j$ we either have $j > r$, in which case $A_{j,r} = 0$, or we have $r > j$, in which case $B_{r,j} = 0$. Thus $\sum_{r \neq j} A_{j,r}B_{r,j} = 0$ and it follows that $(AB)_{j,j} = \alpha_j\beta_j$.

For $j > k$, observe that

$$(AB)_{j,k} = \sum_{r=1}^{n} A_{j,r}B_{r,k} = \sum_{r=1}^{j-1} A_{j,r}B_{r,k} + \sum_{r=j}^{n} A_{j,r}B_{r,k}.$$

For $1 \le r \le j-1$ we have $j > r$ and thus $A_{j,r} = 0$; it follows that $\sum_{r=1}^{j-1} A_{j,r}B_{r,k} = 0$. For $j \le r \le n$ we have $r \ge j > k$ and thus $B_{r,k} = 0$; it follows that $\sum_{r=j}^{n} A_{j,r}B_{r,k} = 0$. Hence $(AB)_{j,k} = 0$ and we may conclude that AB is upper-triangular with $\alpha_1\beta_1, ..., \alpha_n\beta_n$ on the diagonal.

Exercise 5.C.3. Suppose $T \in \mathcal{L}(V)$ is invertible and $v_1, ..., v_n$ is a basis of V with respect to which the matrix of T is upper triangular, with $\lambda_1, ..., \lambda_n$ on the diagonal. Show that the matrix of T^{-1} is also upper triangular with respect to the basis $v_1, ..., v_n$, with

$$\frac{1}{\lambda_1}, ..., \frac{1}{\lambda_n}$$

on the diagonal.

Solution. Because the matrix of T with respect to $v_1, ..., v_n$ is upper-triangular with $\lambda_1, ..., \lambda_n$ on the diagonal, we have $Tv_1 = \lambda_1 v_1$ and, for each $k \ge 2, Tv_k = u_k + \lambda_k v_k$ for some $u_k \in \operatorname{span}(v_1, ..., v_{k-1})$. Observe that

$$Tv_1 = \lambda_1 v_1 \quad \Rightarrow \quad T^{-1}v_1 = \lambda_1^{-1} v_1.$$

This shows that $\operatorname{span}(v_1)$ is invariant under T^{-1} and that the first diagonal entry of $\mathcal{M}(T^{-1})$ is λ_1^{-1}. Now observe that

$$Tv_2 = u_2 + \lambda_2 v_2 \quad \Rightarrow \quad T^{-1}v_2 = \lambda_2^{-1}T^{-1}u_2 + \lambda_2^{-1}v_2;$$

note that $\lambda_2^{-1}T^{-1}u_2 \in \operatorname{span}(v_1)$ because $u_2 \in \operatorname{span}(v_1)$ and $\operatorname{span}(v_1)$ is invariant under T^{-1}. Thus $T^{-1}v_2 \in \operatorname{span}(v_1, v_2)$, so that $\operatorname{span}(v_1, v_2)$ is invariant under T^{-1}, and the second diagonal entry of $\mathcal{M}(T^{-1})$ is λ_2^{-1}. Continuing in this manner, we see that each $\operatorname{span}(v_1, ..., v_k)$ is invariant under T^{-1}, whence $\mathcal{M}(T^{-1})$ is upper-triangular, and that the diagonal entries of $\mathcal{M}(T^{-1})$ are $\lambda_1^{-1}, ..., \lambda_n^{-1}$.

Exercise 5.C.4. Give an example of an operator whose matrix with respect to some basis contains only 0's on the diagonal, but the operator is invertible.

This exercise and the exercise below show that 5.41 fails without the hypothesis that an upper-triangular matrix is under consideration.

Solution. Let $T \in \mathcal{L}(\mathbf{R}^2)$ be given by $T(x, y) = (y, x)$. The matrix of T with respect to the standard basis of $\mathbf{R}^2$ is

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Furthermore, T is invertible since $T^2 = I$.

Exercise 5.C.5. Give an example of an operator whose matrix with respect to some basis contains only nonzero numbers on the diagonal, but the operator is not invertible.

Solution. Let $T \in \mathcal{L}(\mathbf{R}^2)$ be given by $T(x, y) = (x + y, x + y)$. The matrix of T with respect to the standard basis of $\mathbf{R}^2$ is

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

Furthermore, T is not invertible because T is not injective: $T(1, -1) = (0, 0)$.

Exercise 5.C.6. Suppose $\mathbf{F} = \mathbf{C}, V$ is finite-dimensional, and $T \in \mathcal{L}(V)$. Prove that if $k \in \{1, ..., \dim V\}$, then V has a k-dimensional subspace invariant under T.

Solution. By 5.47 there is a basis $v_1, ..., v_n$ of V such that the matrix of T with respect to $v_1, ..., v_n$ is upper-triangular. It follows from 5.39 that for each $k \in \{1, ..., n\}$, the subspace $\operatorname{span}(v_1, ..., v_k)$ is k-dimensional and invariant under T.

Exercise 5.C.7. Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and $v \in V$.

(a) Prove that there exists a unique monic polynomial p_v of smallest degree such that $p_v(T)v = 0$.

(b) Prove that the minimal polynomial of T is a polynomial multiple of p_v.

Solution.

(a) If $v = 0$ then we can take $p_v = 1$, so suppose that $v \neq 0$. The list $v, Tv, ..., T^{\dim V} v$ has length $1 + \dim V$ and hence must be linearly dependent. By the linear dependence lemma (2.19), there exists a least integer $k \in \{1, ..., \dim V\}$ such that $T^k v \in U \coloneqq \operatorname{span}\big(v, Tv, ..., T^{k-1}v\big)$ and such that $v, Tv, ..., T^{k-1}v$ is linearly independent, so that $\dim U = k$. Note that U is invariant under T because $T^k v \in U$. Let p_v be the minimal polynomial of $T|_U$ and observe that, since $v \in U$,

$$p_v(T)v = p_v(T|_U)v = 0.$$

If q is a polynomial of degree $\ell < \deg p_v \leq \dim U = k$ such that $q(T)v = 0$ then $T^\ell v \in \operatorname{span}\big(v, Tv, ..., T^{\ell-1}v\big)$, contradicting the minimality of k. Thus the degree of p_v is minimal. If r is a monic polynomial of degree $\deg p_v$ satisfying $r(T)v = 0$ then $p_v - r$ satisfies $(p_v - r)(T)v = 0$ and $\deg(p_v - r) < \deg p_v$. If $p_v - r$ were not zero then we could divide by the leading coefficient to obtain a monic polynomial s satisfying $s(T)v = 0$ and $\deg s < \deg p_v$, which contradicts the minimality of $\deg p_v$. Thus p_v is unique.

(b) This is immediate from part (a) and 5.31.

Exercise 5.C.8. Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and there exists a nonzero vector $v \in V$ such that $T^2v + 2Tv = -2v$.

(a) Prove that if $\mathbf{F} = \mathbf{R}$, then there does not exist a basis of V with respect to which T has an upper triangular matrix.

(b) Prove that if $\mathbf{F} = \mathbf{C}$ and A is an upper-triangular matrix that equals the matrix of T with respect to some basis of V, then $-1+i$ or $-1-i$ appears on the diagonal of A.

Solution.

(a) Let $q(x) = x^2 + 2x + 2$, so that $q(T)v = 0$, and let p_v be defined as in Exercise 5.C.7. Note that $\deg p_v \geq 1$ since $v \neq 0$. A straightforward modification of 5.29 shows that q must be a polynomial multiple of p_v. It follows that $\deg p_v \geq 2$ since q has no real roots. The minimality of $\deg p_v$ then implies that $p_v = q$ and thus, by Exercise 5.C.7 (b), the minimal polynomial of T must be a polynomial multiple of q. Thus the minimal polynomial of T does not split into linear factors and it then follows from 5.44 that there is no basis of V with respect to which T has an upper-triangular matrix.

(b) Let $q(z) = z^2 + 2z + 2 = (z+1-i)(z+1+i)$, so that $q(T)v = 0$, and let p_v be defined as in Exercise 5.C.7. Note that $\deg p_v \geq 1$ since $v \neq 0$. A straightforward modification of 5.29 shows that q must be a polynomial multiple of p_v. There are then three possibilities:

$$p_v = z + 1 - i, \quad p_v = z + 1 + i, \quad \text{or} \quad p_v = q.$$

In any case, at least one of $-1+i, -1-i$ is a root of p_v. It follows from Exercise 5.C.7 (b) that at least one of $-1+i, -1-i$ is a root of the minimal polynomial of T and hence, by 5.27, at least one of $-1+i, -1-i$ is an eigenvalue of T. 5.41 allows us to conclude that at least one of $-1+i, -1-i$ appears on the diagonal of A.

Exercise 5.C.9. Suppose B is a square matrix with complex entries. Prove that there exists an invertible square matrix A with complex entries such that $A^{-1}BA$ is an upper-triangular matrix.

Solution. Suppose B is an n-by-n matrix and let $T \in \mathcal{L}(\mathbf{C}^n)$ be given by $Tx = Bx$, where we think of elements of $\mathbf{C}^n$ as column vectors. Evidently, the matrix of T with respect to the standard basis $e_1, \ldots, e_n$ of $\mathbf{C}^n$ is B. By 5.47 there is a basis $v_1, \ldots, v_n$ of $\mathbf{C}^n$ such that $\mathcal{M}(T, (v_1, \ldots, v_n))$ is upper-triangular. Let $A = \mathcal{M}(I, (v_1, \ldots, v_n), (e_1, \ldots, e_n))$; as 3.84 shows, it follows that $A^{-1}BA$ equals the upper-triangular matrix $\mathcal{M}(T, (v_1, \ldots, v_n))$.

Exercise 5.C.10. Suppose $T \in \mathcal{L}(V)$ and $v_1, \dots, v_n$ is a basis of B. Show that the following are equivalent.

(a) The matrix of T with respect to $v_1, \dots, v_n$ is lower triangular.

(b) $\operatorname{span}(v_k, \dots, v_n)$ is invariant under T for each $k = 1, \dots, n$.

(c) $Tv_k \in \operatorname{span}(v_k, \dots, v_n)$ for each $k = 1, \dots, n$.

A square matrix is called ***lower triangular*** *if all entries above the diagonal are* 0.

Solution. Suppose that (a) holds and let $k \in \{1, \dots, n\}$ be given. For any $j \in \{k, \dots, n\}$ we have $Tv_j \in \operatorname{span}(v_j, \dots, v_n)$ since the matrix of T with respect to $v_1, \dots, v_n$ is lower-triangular. Because $\operatorname{span}(v_j, \dots, v_n) \subseteq \operatorname{span}(v_k, \dots, v_n)$ for $j \geq k$, it follows that $Tv_j \in \operatorname{span}(v_k, \dots, v_n)$ for each $j \in \{k, \dots n\}$. Thus $\operatorname{span}(v_k, \dots, v_n)$ is invariant under T, i.e. (b) holds.

Now suppose that (b) holds. For any $k \in \{1, \dots, n\}$ we have $v_k \in \operatorname{span}(v_k, \dots, v_n)$, which is invariant under T by assumption. Thus $Tv_k \in \operatorname{span}(v_k, \dots, v_n)$, i.e. (c) holds.

Suppose that (c) holds, so that each Tv_k can be written as a linear combination of the basis vectors $v_k, \dots, v_n$ only. It follows that each entry above the diagonal of $\mathcal{M}(T)$ is zero, i.e. $\mathcal{M}(T)$ is lower-triangular. Thus (a) holds.

Exercise 5.C.11. Suppose $\mathbf{F} = \mathbf{C}$ and V is finite-dimensional. Prove that if $T \in \mathcal{L}(V)$, then there exists a basis of V with respect to which T has a lower-triangular matrix.

Solution. By 5.47 there is a basis $v_1, \dots, v_n$ of V with respect to which the matrix of T is upper-triangular. For each $k \in \{1, \dots, n\}$ define $u_k = v_{n-k+1}$ and observe that, using 5.39,

$$Tu_k = Tv_{n-k+1} \in \operatorname{span}(v_1, v_2, \dots, v_{n-k}, v_{n-k+1}) = \operatorname{span}(u_n, u_{n-1}, \dots, u_{k+1}, u_k).$$

It follows from Exercise 5.C.10 that the matrix of T with respect to $u_1, \dots, u_n$ is lower-triangular.

Exercise 5.C.12. Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$ has an upper-triangular matrix with respect to some basis of V, and U is a subspace of V that is invariant under T.

(a) Prove that $T|_U$ has an upper-triangular matrix with respect to some basis of U.

(b) Prove that the quotient operator T/U has an upper-triangular matrix with respect to some basis of V/U.

The quotient operator T/U was defined in Exercise 38 in Section 5A.

Solution.

(a) Let p be the minimal polynomial of T and let q be the minimal polynomial of $T|_U$. By 5.31 and 5.44, p is a product of linear factors and also a polynomial multiple of

q. It follows that q is a product of linear factors and thus, by 5.44, $T|_U$ has an upper-triangular matrix with respect to some basis of U.

(b) Similarly to part (a), let p be the minimal polynomial of T and let q be the minimal polynomial of T/U. By Exercise 5.B.25 (a) and 5.44, p is a product of linear factors and also a polynomial multiple of q. It follows that q is a product of linear factors and thus, by 5.44, T/U has an upper-triangular matrix with respect to some basis of V/U.

Exercise 5.C.13. Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Suppose there exists a subspace U of V that is invariant under T such that $T|_U$ has an upper-triangular matrix with respect to some basis of U and also T/U has an upper-triangular matrix with respect to some basis of V/U. Prove that T has an upper-triangular matrix with respect to some basis of V.

Solution. Let p, q, r be the minimal polynomials of $T, T|_U$, and T/U. By Exercise 5.B.25 (b) and 5.44, q and r are products of linear factors and qr is a polynomial multiple of p. It follows that p is a product of linear factors and thus, by 5.44, T has an upper-triangular matrix with respect to some basis of V.

Exercise 5.C.14. Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Prove that T has an upper-triangular matrix with respect to some basis of V if and only if the dual operator T' has an upper-triangular matrix with respect to some basis of the dual space V'.

Solution. Exercise 5.B.28 shows that T and T' have the same minimal polynomial. The desired equivalence now follows from 5.44.

5.D. Diagonalizable Operators

Exercise 5.D.1. Suppose V is a finite-dimensional complex vector space and $T \in \mathcal{L}(V)$.

(a) Prove that if $T^4 = I$, then T is diagonalizable.

(b) Prove that if $T^4 = T$, then T is diagonalizable.

(c) Give an example of an operator $T \in \mathcal{L}(\mathbf{C}^2)$ such that $T^4 = T^2$ and T is not diagonalizable.

Solution.

(a) Let $q(z) = z^4 - 1 = (z-1)(z+1)(z-i)(z+i)$, let p be the minimal polynomial of T, and note that $q(T) = 0$. It follows from 5.29 that q is a polynomial multiple of p and so p must be a product of distinct linear factors. Thus, by 5.62, T is diagonalizable.

(b) Let $q(z) = z^4 - z = z(z-1)\left(z + \frac{1}{2} - \frac{\sqrt{3}}{2}i\right)\left(z + \frac{1}{2} + \frac{\sqrt{3}}{2}i\right)$, let p be the minimal polynomial of T, and note that $q(T) = 0$. It follows from 5.29 that q is a polynomial multiple of p and so p must be a product of distinct linear factors. Thus, by 5.62, T is diagonalizable.

(c) Let $T \in \mathcal{L}(\mathbf{C}^2)$ be given by $T(w,z) = (z,0)$ and notice that $T^2 = T^4 = 0$. The matrix of T with respect to the standard basis of $\mathbf{C}^2$ is

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Thus, by 5.41, the only eigenvalue of T is 0. Letting p be the minimal polynomial of T, 5.27 shows that p is of the form z^k for some $k \in \{1,2\}$. Since $T \neq 0$, it must be that $p(z) = z^2$. It follows from 5.62 that T is not diagonalizable.

Exercise 5.D.2. Suppose $T \in \mathcal{L}(V)$ has a diagonal matrix A with respect to some basis of V. Prove that if $\lambda \in \mathbf{F}$, then λ appears on the diagonal of A precisely $\dim E(\lambda, T)$ times.

Solution. Suppose $v_1, \ldots, v_n$ is a basis of V such that $A := \mathcal{M}(T, (v_1, \ldots, v_n))$ is diagonal and let $\lambda_1, \ldots, \lambda_n$ denote the diagonal entries of A. Note that the list of those $Tv_k = \lambda_k v_k$ such that $\lambda_k \neq 0$ is linearly independent since $v_1, \ldots, v_n$ is linearly independent. Thus, letting d be the number of indices $k \in \{1, \ldots, n\}$ such that $\lambda_k = 0$, we have $\dim \operatorname{range} T \geq n - d$. It follows that

$$\dim \operatorname{null} T \leq d. \tag{1}$$

For $\lambda \in \mathbf{F}$, let d_λ be the number of times λ appears on the diagonal of A. By replacing T with $T - \lambda I$ in (1), we find that $\dim E(\lambda, T) \leq d_\lambda$. Now observe that, by 5.55(d),

$$\sum_{\lambda \in \mathbf{F}} \dim E(\lambda, T) = n = \sum_{\lambda \in \mathbf{F}} d_\lambda. \tag{2}$$

(Both of these are finite sums since $\dim E(\lambda, T) = d_\lambda = 0$ for all but finitely many $\lambda \in \mathbf{F}$.) It follows that each inequality $\dim E(\lambda, T) \le d_\lambda$ must in fact be an equality, otherwise the left-hand side of (2) would be strictly less than the right-hand side.

Exercise 5.D.3. Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Prove that if the operator T is diagonalizable, then $V = \operatorname{null} T \oplus \operatorname{range} T$.

Solution. Let $\lambda_1, \ldots, \lambda_m$ be the distinct non-zero eigenvalues of T (this list may be empty). By 5.55, we have

$$V = E(0, T) \oplus E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T) = \operatorname{null} T \oplus W,$$

where $W = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)$; if the list $\lambda_1, \ldots, \lambda_m$ is empty then take $W = \{0\}$. Let $Tv \in \operatorname{range} T$ be given. The direct sum expression above shows that v is of the form

$$v = u + w_1 + \cdots + w_m \in \operatorname{null} T \oplus W \quad \Rightarrow \quad Tv = \lambda_1 w_1 + \cdots + \lambda_m w_m \in W.$$

Thus $\operatorname{range} T \subseteq W$. Now let $w_1 + \cdots + w_m \in W$ be given. Because each $\lambda_k \ne 0$, it follows that

$$w_1 + \cdots + w_m = T(\lambda_1^{-1} w_1 + \cdots + \lambda_m^{-1} w_m) \in \operatorname{range} T.$$

Thus $W = \operatorname{range} T$ and we may conclude that $V = \operatorname{null} T \oplus \operatorname{range} T$.

Exercise 5.D.4. Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Prove that the following are equivalent.

(a) $V = \operatorname{null} T \oplus \operatorname{range} T$.

(b) $V = \operatorname{null} T + \operatorname{range} T$.

(c) $\operatorname{null} T \cap \operatorname{range} T = \{0\}$.

Solution. Certainly (a) implies (b). Suppose that (b) holds. By 2.43, we have

$$\dim(\operatorname{null} T \cap \operatorname{range} T) = \dim \operatorname{null} T + \dim \operatorname{range} T - \dim(\operatorname{null} T + \operatorname{range} T).$$

By assumption $\dim(\operatorname{null} T + \operatorname{range} T) = \dim V$ and the fundamental theorem of linear maps (3.21) shows that $\dim \operatorname{null} T + \dim \operatorname{range} T = \dim V$ also. Thus $\dim(\operatorname{null} T \cap \operatorname{range} T) = 0$, so that $\operatorname{null} T \cap \operatorname{range} T = \{0\}$, i.e. (c) holds.

Suppose that (c) holds. It follows from 1.46 that the sum $\operatorname{null} T \oplus \operatorname{range} T$ is direct. Furthermore,

$$\dim(\operatorname{null} T \oplus \operatorname{range} T) = \dim \operatorname{null} T + \dim \operatorname{range} T = \dim V$$

by 1.46 and 2.43. Thus, by 2.39, $V = \operatorname{null} T \oplus \operatorname{range} T$, i.e. (a) holds.

Exercise 5.D.5. Suppose V is a finite-dimensional complex vector space and $T \in \mathcal{L}(V)$. Prove that T is diagonalizable if and only if

$$V = \operatorname{null}(T - \lambda I) \oplus \operatorname{range}(T - \lambda I)$$

for every $\lambda \in \mathbf{C}$.

Solution. Suppose that T is diagonalizable, so that there is some basis of V with respect to which the matrix of T is diagonal:

$$\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

For any $\lambda \in \mathbf{C}$, the matrix of the operator $T - \lambda I$ with respect to this same basis is also diagonal:

$$\begin{pmatrix} \lambda_1 - \lambda & & 0 \\ & \ddots & \\ 0 & & \lambda_n - \lambda \end{pmatrix}.$$

So $T - \lambda I$ is also diagonalizable and thus by Exercise 5.D.3 we have

$$V = \operatorname{null}(T - \lambda I) \oplus \operatorname{range}(T - \lambda I).$$

Here are two proofs of the converse.

(1) By contrapositive: suppose that T is not diagonalizable. Let $\lambda_1, \ldots, \lambda_m$ be the distinct eigenvalues of T and let p be the minimal polynomial of T; by 5.27(b) we have

$$p(z) = (z - \lambda_1)^{n_1} \cdots (z - \lambda_m)^{n_m}$$

for some positive integers $n_1, \ldots, n_m$. Because T is not diagonalizable, 5.62 shows that there must be some $k \in \{1, \ldots, m\}$ such that $n_k \geq 2$. Let q be the polynomial given by

$$q(z) = (z - \lambda_k)^{n_k - 1} \prod_{j \neq k} (z - \lambda_j)^{n_j}.$$

Notice that $p(z) = (z - \lambda_k) q(z)$, so that $\deg q < \deg p$, and, for any $v \in V$,

$$0 = p(T)v = (T - \lambda_k I) q(T) v \quad \Rightarrow \quad q(T)v \in \operatorname{null}(T - \lambda_k I),$$

$$n_k - 1 \geq 1 \quad \Rightarrow \quad q(T)v \in \operatorname{range}(T - \lambda_k I).$$

Thus $q(T)v \in \operatorname{null}(T - \lambda_k I) \cap \operatorname{range}(T - \lambda_k I)$ for all $v \in V$. Since $\deg q < \deg p$ it must be the case that $q(T) \neq 0$, i.e. there exists some $v \in V$ such that $q(T)v \neq 0$. It follows that

$$\operatorname{null}(T - \lambda_k I) \cap \operatorname{range}(T - \lambda_k I) \neq \{0\}$$

and hence, by Exercise 5.D.4, $V \neq \operatorname{null}(T - \lambda_k I) \oplus \operatorname{range}(T - \lambda_k I)$.

(2) By strong induction on $\dim V$. Let $P(n)$ be the following statement: if V is an n-dimensional complex vector space, $T \in \mathcal{L}(V)$, and

$$V = \operatorname{null}(T - \lambda I) \oplus \operatorname{range}(T - \lambda I)$$

for all $\lambda \in \mathbf{C}$, then T is diagonalizable. The truth of $P(0)$ is clear, so suppose that $P(0), \ldots, P(n)$ all hold for some $n \geq 1$. Let V be an $(n+1)$-dimensional vector space and suppose that $T \in \mathcal{L}(V)$ satisfies

$$V = \operatorname{null}(T - \lambda I) \oplus \operatorname{range}(T - \lambda I)$$

for all $\lambda \in \mathbf{C}$. By 5.19 there exists an eigenvalue $\lambda_0 \in \mathbf{C}$ of T; let $U = \operatorname{range}(T - \lambda_0 I)$ and note that U is invariant under T by (5.18). Note further that, for any $\lambda \in \mathbf{C}$,

$$\operatorname{null}(T|_U - \lambda I|_U) \cap \operatorname{range}(T|_U - \lambda I|_U) \subseteq \operatorname{null}(T - \lambda I) \cap \operatorname{range}(T - \lambda I) = \{0\}.$$

It follows that $U = \operatorname{null}(T|_U - \lambda I|_U) \oplus \operatorname{range}(T|_U - \lambda I|_U)$ for every $\lambda \in \mathbf{C}$, where we have used the equivalence of (a) and (c) in Exercise 5.D.4. By assumption $V = E(\lambda_0, T) \oplus U$; since λ_0 is an eigenvalue of T we have $\dim E(\lambda_0, T) \geq 1$ and thus $\dim U < \dim V$. Our induction hypothesis now guarantees that there is a basis of U consisting of eigenvectors of $T|_U$, which must also be eigenvectors of T. Combining this basis with a basis of $E(\lambda_0, T)$, we obtain a basis of V consisting of eigenvectors of T. It follows from 5.55 that T is diagonalizable. This completes the induction step and the proof.

Exercise 5.D.6. Suppose $T \in \mathcal{L}(\mathbf{F}^5)$ and $\dim E(8, T) = 4$. Prove that $T - 2I$ or $T - 6I$ is invertible.

Solution. We will prove the contrapositive statement. Suppose that neither $T - 2I$ nor $T - 6I$ is invertible, so that $\dim E(2, T) \geq 1$ and $\dim E(6, T) \geq 1$. If 8 is an eigenvalue of T then 5.54 shows that $\dim E(8, T) + \dim E(2, T) + \dim E(6, T) \leq \dim \mathbf{F}^5$, and if 8 is not an eigenvalue of T then 5.54 together with $\dim E(8, T) = 0$ gives us the same inequality. In either case,

$$\begin{aligned} &\dim E(8, T) + \dim E(2, T) + \dim E(6, T) \leq \dim \mathbf{F}^5 = 5 \\ &\qquad\Rightarrow \quad \dim E(8, T) \leq 5 - \dim E(2, T) - \dim E(6, T) \leq 3 < 4. \end{aligned}$$

Thus $\dim E(8, T) \neq 4$.

Exercise 5.D.7. Suppose $T \in \mathcal{L}(V)$ is invertible. Prove that

$$E(\lambda, T) = E\left(\frac{1}{\lambda}, T^{-1}\right)$$

for every $\lambda \in \mathbf{F}$ with $\lambda \neq 0$.

Solution. This follows from the equivalence, for $\lambda \neq 0$ and any $v \in V$,

$$Tv = \lambda v \quad \Leftrightarrow \quad T^{-1}v = \lambda^{-1}v.$$

Exercise 5.D.8. Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Let $\lambda_1, ..., \lambda_m$ denote the distinct nonzero eigenvalues of T. Prove that

$$\dim E(\lambda_1, T) + \cdots + \dim E(\lambda_m, T) \leq \dim \operatorname{range} T.$$

Solution. Note that

$$\dim E(0, T) + \dim E(\lambda_1, T) + \cdots + \dim E(\lambda_m, T) \leq \dim V \tag{1}$$

follows from 5.54: if 0 is an eigenvalue of T then the inequality is immediate from 5.54, and if 0 is not an eigenvalue of T then $\dim E(0, T) = 0$ and thus we can add $\dim E(0, T)$ to the left-hand side of the inequality in 5.54. The fundamental theorem of linear maps (3.21) shows that

$$\dim V = \dim \operatorname{null} T + \dim \operatorname{range} T = \dim E(0, T) + \dim \operatorname{range} T.$$

Combining this with inequality (1) gives us the desired inequality.

Exercise 5.D.9. Suppose $R, T \in \mathcal{L}(\mathbf{F}^3)$ each have $2, 6, 7$ as eigenvalues. Prove that there exists invertible operator $S \in \mathcal{L}(\mathbf{F}^3)$ such that $R = S^{-1}TS$.

Solution. Since R and T both have $3 = \dim \mathbf{F}^3$ distinct eigenvalues, 5.58 shows that they are both diagonalizable, i.e. there exists a basis u_1, u_2, u_3 and a basis v_1, v_2, v_3 of V such that

$$Ru_1 = 2u_1, \quad Ru_2 = 6u_2, \quad Ru_3 = 7u_3, \quad Tv_1 = 2v_1, \quad Tv_2 = 6v_2, \quad Tv_3 = 7v_3.$$

Define $S \in \mathcal{L}(\mathbf{F}^3)$ by $Su_k = v_k$ and note that S is invertible since it maps a basis to a basis. Furthermore,

$$S^{-1}TSu_1 = S^{-1}Tv_1 = 2S^{-1}v_1 = 2u_1 = Ru_1.$$

Similarly, $S^{-1}TSu_k = Ru_k$ for $k \in \{2, 3\}$. Thus $S^{-1}TS = R$.

Exercise 5.D.10. Find $R, T \in \mathcal{L}(\mathbf{F}^4)$ such that R and T each have $2, 6, 7$ as eigenvalues, R and T have no other eigenvalues, and there does not exist an invertible operator $S \in \mathcal{L}(\mathbf{F}^4)$ such that $R = S^{-1}TS$.

Solution. Let R and T be the operators which have the matrices

$$\mathcal{M}(R) = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 6 & 0 \\ 0 & 0 & 0 & 7 \end{pmatrix} \quad \text{and} \quad \mathcal{M}(T) = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 6 & 0 & 0 \\ 0 & 0 & 6 & 0 \\ 0 & 0 & 0 & 7 \end{pmatrix}$$

with respect to the standard basis e_1, e_2, e_3, e_4 of $\mathbf{F}^4$. Since these matrices are upper-triangular, 5.41 shows that the eigenvalues of R and T are precisely $2, 6, 7$. To disprove the existence of an invertible operator $S \in \mathcal{L}(\mathbf{F}^4)$ such that $R = S^{-1}TS$, let $S \in \mathcal{L}(\mathbf{F}^4)$ be any invertible operator. By Exercise 5.A.13, $S^{-1}TS$ also has 2 as an eigenvalue. Furthermore, the eigenspace $E(2,T)$ is the image under S of the eigenspace $E(2,S^{-1}TS)$. A restriction of S thus provides us with an isomorphism between $E(2,T)$ and $E(2,S^{-1}TS)$; in particular, these eigenspaces must have the same dimension. However, note that

$$\dim E(2,T) = \dim \operatorname{span}(e_1) = 1 \neq 2 = \dim \operatorname{span}(e_1, e_2) = \dim E(2,R).$$

Thus there cannot exist an invertible $S \in \mathcal{L}(\mathbf{F}^4)$ such that $R = S^{-1}TS$.

Exercise 5.D.11. Find $T \in \mathcal{L}(\mathbf{C}^3)$ such that 6 and 7 are eigenvalues of T and such that T does not have a diagonal matrix with respect to any basis of $\mathbf{C}^3$.

Solution. Let T be the operator which has the matrix

$$\mathcal{M}(T) = \begin{pmatrix} 6 & 1 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 7 \end{pmatrix}$$

with respect to the standard basis e_1, e_2, e_3 of $\mathbf{C}^3$. Since this matrix is upper-triangular, 5.41 shows that the eigenvalues of T are precisely 6 and 7. Some routine calculations reveal that $E(6,T) = \operatorname{span}(e_1)$ and $E(7,T) = \operatorname{span}(e_3)$. It follows that

$$\dim E(6,T) + \dim E(7,T) = 2 \neq 3 = \dim \mathbf{C}^3.$$

Thus, by 5.55, T is not diagonalizable.

Exercise 5.D.12. Suppose $T \in \mathcal{L}(\mathbf{C}^3)$ is such that 6 and 7 are eigenvalues of T. Furthermore, suppose T does not have a diagonal matrix with respect to any basis of $\mathbf{C}^3$. Prove that there exists $(z_1, z_2, z_3) \in \mathbf{C}^3$ such that

$$T(z_1, z_2, z_3) = (6 + 8z_1, 7 + 8z_2, 13 + 8z_3).$$

Solution. Since $\dim \mathbf{C}^3 = 3$, it must be the case that 6 and 7 are the only eigenvalues of T; if T had another distinct eigenvalue then, by 5.58, T would be diagonalizable. It follows from 5.7 that $T - 8I$ is surjective and thus there exists $(z_1, z_2, z_3) \in \mathbf{C}^3$ such that

$$(T - 8I)(z_1, z_2, z_3) = (6, 7, 13) \quad \Leftrightarrow \quad T(z_1, z_2, z_3) = (6 + 8z_1, 7 + 8z_2, 13 + 8z_3).$$

Exercise 5.D.13. Suppose A is a diagonal matrix with distinct entries on the diagonal and B is a matrix of the same size as A. Show that $AB = BA$ if and only if B is a diagonal matrix.

Solution. Suppose A has diagonal entries $a_1, \dots, a_n$. If B is also diagonal with diagonal entries $b_1, \dots, b_n$, then

$$AB = \begin{pmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & a_n \end{pmatrix} \begin{pmatrix} b_1 & & 0 \\ & \ddots & \\ 0 & & b_n \end{pmatrix} = \begin{pmatrix} a_1 b_1 & & 0 \\ & \ddots & \\ 0 & & a_n b_n \end{pmatrix} = \begin{pmatrix} b_1 & & 0 \\ & \ddots & \\ 0 & & b_n \end{pmatrix} \begin{pmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & a_n \end{pmatrix} = BA.$$

Now suppose that B is not diagonal, i.e. there exist $j, k \in \{1, ..., n\}$ such that $j \neq k$ and $B_{j,k} \neq 0$. Observe that

$$(AB)_{j,k} = \sum_{r=1}^{n} A_{j,r} B_{r,k} = a_j B_{j,k} \neq B_{j,k} a_k = \sum_{r=1}^{n} B_{j,r} A_{r,k} = (BA)_{j,k},$$

where $a_j B_{j,k} \neq B_{j,k} a_k$ follows since $B_{j,k} \neq 0$ and $a_j \neq a_k$. Thus $AB \neq BA$.

Exercise 5.D.14.

(a) Give an example of a finite-dimensional complex vector space and an operator T on that vector space such that T^2 is diagonalizable but T is not diagonalizable.

(b) Suppose $\mathbf{F} = \mathbf{C}$, k is a positive integer, and $T \in \mathcal{L}(V)$ is invertible. Prove that T is diagonalizable if and only if T^k is diagonalizable.

Solution.

(a) Define T as in Exercise 5.D.1 (c). As we showed there, T is not diagonalizable. However, $T^2 = 0$ is certainly diagonalizable.

(b) Suppose that T is diagonalizable, so that there exists a basis $v_1, ..., v_n$ of V with respect to which the matrix of T is diagonal, say

$$\mathcal{M}(T) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

for some $\lambda_1, ..., \lambda_n \in \mathbf{C}$. It follows from 3.43 and direct calculation that

$$\mathcal{M}(T^k) = [\mathcal{M}(T)]^k = \begin{pmatrix} \lambda_1^k & & 0 \\ & \ddots & \\ 0 & & \lambda_n^k \end{pmatrix}.$$

Thus T^k is diagonalizable.

Now suppose that T^k is diagonalizable and let p be the minimal polynomial of T^k. By 5.27 and 5.62 we have $p(z) = (z - \lambda_1) \cdots (z - \lambda_m)$, where $\lambda_1, ..., \lambda_m$ are the distinct eigenvalues of T^k. Note that T^k is invertible since T is invertible; it follows that each λ_j is non-zero. It can be shown that any non-zero complex number has exactly k distinct k^{th} roots. For each $j \in \{1, ..., m\}$, let $\mu_{1,j}, ..., \mu_{k,j}$ be the k distinct solutions to $z^k = \lambda_j$. Observe that, for any $a, i \in \{1, ..., k\}$,

$$b \neq j \quad \Rightarrow \quad \lambda_b \neq \lambda_j \quad \Leftrightarrow \quad \mu_{a,b}^k \neq \mu_{i,j}^k \quad \Rightarrow \quad \mu_{a,b} \neq \mu_{i,j}.$$

Thus, if we let q be the polynomial given by

$$q(z) = \left(z^k - \lambda_1\right) \cdots \left(z^k - \lambda_m\right) = \prod_{j=1}^{m} \prod_{i=1}^{k} \left(z - \mu_{i,j}\right),$$

then q is a product of distinct linear factors. Notice that $q(T) = p\left(T^k\right) = 0$; it follows from 5.29 that q is a polynomial multiple of the minimal polynomial of T. Thus the minimal polynomial of T is a product of distinct linear factors. 5.62 allows us to conclude that T is diagonalizable.

Exercise 5.D.15. Suppose V is a finite-dimensional complex vector space, $T \in \mathcal{L}(V)$, and p is the minimal polynomial of T. Prove that the following are equivalent.

(a) T is diagonalizable.

(b) There does not exist $\lambda \in \mathbf{C}$ such that p is a polynomial multiple of $(z - \lambda)^2$.

(c) p and its derivative p' have no zeros in common.

(d) The greatest common divisor of p and p' is the constant polynomial 1.

The ***greatest common divisor*** *of p and p′ is the monic polynomial q of largest degree such that p and p′ are both polynomial multiples of q. The Euclidean algorithm for polynomials (look it up) can quickly determine the greatest common divisor of two polynomials, without requiring any information about the zeros of the polynomials. Thus the equivalence of (a) and (d) above shows that we can determine whether T is diagonalizable without knowing anything about the zeros of p.*

Solution. By 5.27(b), p is of the form $p(z) = (z - \lambda_1)^{n_1} \cdots (z - \lambda_m)^{n_m}$, where $\lambda_1, \ldots, \lambda_m$ is a list of the distinct eigenvalues of T and each n_k is a positive integer.

Note that (b) is equivalent to $n_1 = \cdots = n_m = 1$, which by 5.62 is equivalent to (a). Thus (a) and (b) are equivalent.

Certainly p has m distinct zeros, and observe that $n_1 = \cdots = n_m = 1$ if and only if $\deg p = m$. Exercise 4.8 now shows that (b) and (c) are equivalent.

Suppose the negation of (b) holds, so that $p(z) = (z - \lambda)^2 s(z)$ for some polynomial s; the product rule gives us $p'(z) = (z - \lambda)(2s(z) + (z - \lambda)s'(z))$. Thus the greatest common divisor of p and p' must have degree at least 1, so that (d) does not hold. The contrapositive of this and the equivalence of (b) and (c) shows that (d) implies (c).

Let q be the greatest common divisor of p and p' and suppose the negation of (d) holds, so that $\deg q \geq 1$. The fundamental theorem of algebra shows that q has some zero, which must also be a zero of p and p', i.e. the negation of (c) holds. Thus (c) and (d) are equivalent.

Exercise 5.D.16. Suppose that $T \in \mathcal{L}(V)$ is diagonalizable. Let $\lambda_1, \ldots, \lambda_m$ denote the distinct eigenvalues of T. Prove that a subspace U of V is invariant under T if and only if there exist subspaces $U_1, \ldots, U_m$ of V such that $U_k \subseteq E(\lambda_k, T)$ for each k and $U = U_1 \oplus \cdots \oplus U_m$.

Solution. Suppose there exist such subspaces and let $u = u_1 + \cdots + u_m \in U$ be given. Observe that

$$Tu = Tu_1 + \cdots + Tu_m = \lambda_1 u_1 + \cdots + \lambda_m u_m \in U_1 \oplus \cdots \oplus U_m = U.$$

Thus U is invariant under T.

Now suppose that U is invariant under T. Because T is diagonalizable, (5.55) shows that

$$V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T).$$

For each $k \in \{1, \ldots, m\}$, let $U_k = U \cap E(\lambda_k, T) \subseteq E(\lambda_k, T)$. The directness of the sum $U_1 \oplus \cdots \oplus U_m$ is immediate from the directness of the sum $E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)$, and since each $U_k \subseteq U$ we have $U_1 \oplus \cdots \oplus U_m \subseteq U$. For any $u \in U$ we have $u = v_1 + \cdots + v_m$, where each $v_k \in E(\lambda_k, T)$. Lemma L.4 shows that each $v_k \in U$ and thus u belongs to $U_1 \oplus \cdots \oplus U_m$, so that $U \subseteq U_1 \oplus \cdots \oplus U_m$. We may conclude that $U = U_1 \oplus \cdots \oplus U_m$.

Exercise 5.D.17. Suppose V is finite-dimensional. Prove that $\mathcal{L}(V)$ has a basis consisting of diagonalizable operators.

Solution. Let $v_1, \ldots, v_n$ be a basis of V; in what follows, all matrices of operators are with respect to this basis. For $i, j \in \{1, \ldots, n\}$ such that $i \neq j$, define $T_{i,j} \in \mathcal{L}(V)$ by

$$T_{i,j} v_k = \begin{cases} k v_k & \text{if } k \neq j, \\ k v_k + v_i & \text{if } k = j. \end{cases}$$

Thus the matrix of $T_{i,j}$ has diagonal entries $1, \ldots, n$, a 1 in the i^{th} row and j^{th} column, and 0's elsewhere. Notice that this matrix is either upper- or lower-triangular. It follows that the eigenvalues of $T_{i,j}$ are precisely $1, \ldots, n$ and hence $T_{i,j}$ is diagonalizable by 5.58. For each $j \in \{1, \ldots, n\}$, define $S_j \in \mathcal{L}(V)$ by

$$S_j v_k = \begin{cases} v_k & \text{if } k = j, \\ 0 & \text{if } k \neq j. \end{cases}$$

Notice that the matrix of S_j has a 1 in the j^{th} row and j^{th} column, and 0's elsewhere. Thus each S_j is diagonalizable.

Suppose we have a linear combination of the list

$$\mathcal{B} := T_{1,2}, \ldots, T_{1,n}, \ldots, T_{n,1}, \ldots, T_{n,n-1}, S_1, \ldots, S_n$$

which equals zero. For $i \neq j$, only the operator $T_{i,j}$ has a non-zero entry in the i^{th} row and j^{th} column of its matrix. It follows that the coefficient of $T_{i,j}$ in the linear combination must be zero and we are left with a linear combination $a_1 S_1 + \cdots + a_n S_n = 0$, which clearly implies $a_1 = \cdots = a_n = 0$. Thus $\mathcal{B}$ is linearly independent. A straightforward counting argument shows that $\mathcal{B}$ has length n^2 and so we may conclude that $\mathcal{B}$ is a basis of $\mathcal{L}(V)$ consisting of diagonalizable operators.

Exercise 5.D.18. Suppose that $T \in \mathcal{L}(V)$ is diagonalizable and U is a subspace of V that is invariant under T. Prove that the quotient operator T/U is a diagonalizable operator on V/U.

The quotient operator T/U was defined in Exercise 38 in Section 5A.

Solution. By 5.62 and Exercise 5.B.25 (a), the minimal polynomial of T is a product of distinct linear factors and also a polynomial multiple of the minimal polynomial of T/U; it follows that the minimal polynomial of T/U is a product of distinct linear factors and 5.62 allows us to conclude that T/U is diagonalizable.

Exercise 5.D.19. Prove or give a counterexample: If $T \in \mathcal{L}(V)$ and there exists a subspace U of V that is invariant under T such that $T|_U$ and T/U are both diagonalizable, then T is diagonalizable.

See Exercise 13 in Section 5C for an analogous statement about upper-triangular matrices.

Solution. This is false. For a counterexample, consider the operator $T \in \mathcal{L}(\mathbf{F}^2)$ given by $Te_1 = 0$ and $Te_2 = e_1$, where e_1, e_2 is the standard basis of $\mathbf{F}^2$. It is straightforward to verify that the only eigenvalue of T is 0 and that $E(0, T) = \operatorname{span}(e_1)$. It then follows from 5.55 that T is not diagonalizable. However, if we let U be the T-invariant subspace $E(0, T)$, then $T|_U$ and T/U are both operators on 1-dimensional vector spaces and hence are diagonalizable.

Exercise 5.D.20. Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Prove that T is diagonalizable if and only if the dual operator T' is diagonalizable.

Solution. By Exercise 5.B.28, T and T' have the same minimal polynomial. The desired equivalence is then immediate from 5.62.

Exercise 5.D.21. The *Fibonacci sequence* $F_0, F_1, F_2, \ldots$ is defined by

$$F_0 = 0,\ F_1 = 1, \text{ and } F_n = F_{n-2} + F_{n-1} \text{ for } n \geq 2.$$

Define $T \in \mathcal{L}(\mathbf{R}^2)$ by $T(x, y) = (y, x + y)$.

(a) Show that $T^n(0, 1) = (F_n, F_{n+1})$ for each nonnegative integer n.

(b) Find the eigenvalues of T.

(c) Find a basis of $\mathbf{R}^2$ consisting of eigenvectors of T.

(d) Use the solution to (c) to compute $T^n(0, 1)$. Conclude that

$$F_n = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right]$$

for each nonnegative integer n.

(e) Use (d) to conclude that if n is a nonnegative integer, then the Fibonacci number F_n is the integer that is closest to

$$\frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n.$$

Each F_n is a nonnegative integer, even though the right side of the formula in (d) does not look like an integer. The number

$$\frac{1+\sqrt{5}}{2}$$

is called the ***golden ratio****.*

Solution.

(a) We will proceed by induction. The base case $n = 0$ is clear, so suppose that $T^n(0, 1) = (F_n, F_{n+1})$ holds for some non-negative integer n and observe that

$$T^{n+1}(0, 1) = T(T^n(0, 1)) = T(F_n, F_{n+1}) = (F_{n+1}, F_n + F_{n+1}) = (F_{n+1}, F_{n+2}).$$

This completes the induction step and the proof.

(b) We are looking for solutions $(x, y) \neq (0, 0)$ and $\lambda \in \mathbf{R}$ of the equation

$$T(x, y) = (y, x + y) = (\lambda x, \lambda y).$$

From the equation $y = \lambda x$ we see that $x = 0$ implies $y = 0$, so we may assume that x is non-zero. Substituting $y = \lambda x$ into the equation $x + y = \lambda y$ and cancelling x gives us the equation $\lambda^2 - \lambda - 1 = 0$, which has two distinct real solutions:

$$\lambda_1 = \frac{1+\sqrt{5}}{2} \quad \text{and} \quad \lambda_2 = \frac{1-\sqrt{5}}{2}.$$

These are indeed eigenvalues, since

$$T(1, \lambda_1) = (\lambda_1, \lambda_1 + 1) = (\lambda_1, \lambda_1^2) = \lambda_1(1, \lambda_1),$$

where we have used that λ_1 satisfies the equation $\lambda_1^2 - \lambda_1 - 1 = 0$ for the second equality. Similarly, $T(1, \lambda_2) = \lambda_2(1, \lambda_2)$. Since $\dim \mathbf{R}^2 = 2$, 5.12 allows us to conclude that the eigenvalues of T are precisely λ_1 and λ_2.

(c) Since $\lambda_1 \neq \lambda_2$, the eigenvectors $v_1 = (1, \lambda_1)$ and $v_2 = (1, \lambda_2)$ found in part (b) are linearly independent by 5.11 and thus form a basis of $\mathbf{R}^2$.

(d) Observe that

$$v_1 - v_2 = (0, \lambda_1 - \lambda_2) = \big(0, \sqrt{5}\big).$$

Thus $(0, 1) = \frac{1}{\sqrt{5}}(v_1 - v_2)$. For any positive integer n, it follows that

$$T^n(0, 1) = \frac{1}{\sqrt{5}}(T^n v_1 - T^n v_2) = \frac{1}{\sqrt{5}}(\lambda_1^n v_1 - \lambda_2^n v_2) = \frac{1}{\sqrt{5}}\big(\lambda_1^n - \lambda_2^n, \lambda_1^{n+1} - \lambda_2^{n+1}\big).$$

Given the result of part (a), we may conclude that

$$F_n = \frac{1}{\sqrt{5}}(\lambda_1^n - \lambda_2^n) = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right].$$

(e) Certainly $F_0 = 0$ is the closest integer to $\frac{1}{\sqrt{5}}$. For any positive integer n, observe that

$$\begin{aligned} 2 < \sqrt{5} < 3 \quad &\Rightarrow \quad -1 < \frac{1-\sqrt{5}}{2} < -\frac{1}{2} \\ &\Rightarrow \quad -1 < \left(\frac{1-\sqrt{5}}{2}\right)^n < 1 \\ &\Rightarrow \quad -\frac{1}{\sqrt{5}} < -\frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^n < \frac{1}{\sqrt{5}} \\ &\Rightarrow \quad -\frac{1}{2} < -\frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^n < \frac{1}{2}. \end{aligned}$$

It then follows from part (d) that

$$\frac{1}{\sqrt{5}}\lambda_1^n - \frac{1}{2} < F_n < \frac{1}{\sqrt{5}}\lambda_1^n + \frac{1}{2},$$

i.e. F_n is an integer belonging to the open interval $\left(\frac{1}{\sqrt{5}}\lambda_1^n - \frac{1}{2}, \frac{1}{\sqrt{5}}\lambda_1^n + \frac{1}{2}\right)$, which has length 1. We may conclude that F_n is the integer closest to $\frac{1}{\sqrt{5}}\lambda_1^n$.

Exercise 5.D.22. Suppose $T \in \mathcal{L}(V)$ and A is an n-by-n matrix that is the matrix of T with respect to some basis of V. Prove that if

$$|A_{j,j}| > \sum_{\substack{k=1 \\ k \neq j}}^{n} |A_{j,k}|$$

for each $j \in \{1, \ldots, n\}$, then T is invertible.

This exercise states that if the diagonal entries of the matrix of T are large compared to the nondiagonal entries, then T is invertible.

Solution. If T has no eigenvalues then certainly 0 is not an eigenvalue of T and thus T is invertible. Otherwise, let $\lambda \in \mathbf{F}$ be an eigenvalue of T. 5.67 shows that there exists a $j \in \{1, \ldots, n\}$ such that

$$|\lambda - A_{j,j}| \leq \sum_{\substack{k=1 \\ k \neq j}}^{n} |A_{j,k}| < |A_{j,j}| \quad \Rightarrow \quad |\lambda| > 0 \quad \Rightarrow \quad \lambda \neq 0.$$

Thus 0 is not an eigenvalue of T and it follows that T is invertible.

Exercise 5.D.23. Suppose the definition of the Gershgorin disks is changed so that the radius of the k^{th} disk is the sum of the absolute values of the entries in column (instead of row) k of A, excluding the diagonal entry. Show that the Gershgorin disk theorem (5.67) still holds with this changed definition.

Solution. Suppose $T \in \mathcal{L}(V)$ and $v_1, \ldots, v_n$ is a basis of V. Let A be the matrix of T with respect to $v_1, \ldots, v_n$; it follows from 3.132 that A^{t} is the matrix of T' with respect to the dual basis $\varphi_1, \ldots, \varphi_n$ of V'. Let $\lambda \in \mathbf{F}$ be an eigenvalue of T. Exercise 5.A.15 shows that λ is also an eigenvalue of T' and thus, by 5.67, there exists a $k \in \{1, \ldots, n\}$ such that

$$|\lambda - A_{k,k}| \leq \sum_{\substack{j=1 \\ j \neq k}}^{n} |A^{\mathrm{t}}_{k,j}| = \sum_{\substack{j=1 \\ j \neq k}}^{n} |A_{j,k}|.$$

Thus λ is contained in the k^{th} Gershgorin "column-disk" of T with respect to $v_1, \ldots, v_n$.

5.E. Commuting Operators

Exercise 5.E.1. Give an example of two commuting operators S, T on $\mathbf{F}^4$ such that there is a subspace of $\mathbf{F}^4$ that is invariant under S but not under T and there is a subspace of $\mathbf{F}^4$ that is invariant under T but not under S.

Solution. Let $S, T \in \mathcal{L}(\mathbf{F}^4)$ be given by

$$S(x_1, x_2, x_3, x_4) = (x_2, x_1, 0, 0) \quad \text{and} \quad T(x_1, x_2, x_3, x_4) = (0, 0, x_4, x_3).$$

Notice that S and T commute, since for any $x \in \mathbf{F}^4$ we have $STx = TSx = 0$. Notice further that, for any $\lambda \in \mathbf{F}$,

$$S(\lambda, 0, 0, 0) = (0, \lambda, 0, 0), \quad T(\lambda, 0, 0, 0) = 0, \quad S(0, 0, \lambda, 0) = 0, \quad T(0, 0, \lambda, 0) = (0, 0, 0, \lambda).$$

It follows that $\operatorname{span}((1, 0, 0, 0))$ is invariant under T but not under S, and $\operatorname{span}((0, 0, 1, 0))$ is invariant under S but not under T.

Exercise 5.E.2. Suppose $\mathcal{E}$ is a subset of $\mathcal{L}(V)$ and every element of $\mathcal{E}$ is diagonalizable. Prove that there exists a basis of V with respect to which every element of $\mathcal{E}$ has a diagonal matrix if and only if every pair of elements of $\mathcal{E}$ commutes.

This exercise extends 5.76, which considers the case in which $\mathcal{E}$ contains only two elements. For this exercise, $\mathcal{E}$ may contain any number of elements, and $\mathcal{E}$ may even be an infinite set.

Solution. Suppose there exists such a basis $v_1, ..., v_n$ and let $S, T \in \mathcal{E}$ be given. The matrices of S and T with respect to $v_1, ..., v_n$ are diagonal and hence commute. It follows from 5.74 that S and T commute.

Suppose that every pair of elements of $\mathcal{E}$ commutes and suppose that $\dim V = n$. Because $\dim \mathcal{L}(V) = n^2$, there must exist a subset $\mathcal{F} \subseteq \mathcal{E}$ of cardinality at most n^2 such that every operator in $\mathcal{E}$ is a linear combination of operators in $\mathcal{F}$. Suppose that $\mathcal{F} = \{T_1, ..., T_m\}$ for some $m \leq n^2$. Since T_1 is diagonalizable, 5.55 shows that

$$V = \bigoplus_{\lambda_1 \in \mathbf{F}} E(\lambda_1, T_1);$$

note that, since T_1 has at most n distinct eigenvalues, all but finitely many of the summands $E(\lambda_1, T_1)$ are equal to $\{0\}$, so that this direct sum is finite. Let $\lambda_1 \in \mathbf{F}$ be given. Because T_1 and T_2 commute, 5.75 shows that $E(\lambda_1, T_1)$ is invariant under T_2. It then follows from 5.65 that $T_2|_{E(\lambda_1, T_1)}$ is diagonalizable and thus, by 5.55,

$$E(\lambda_1, T_1) = \bigoplus_{\lambda_2 \in \mathbf{F}} E\Big(\lambda_2, T_2|_{E(\lambda_1, T_1)}\Big);$$

again, this direct sum is finite since T_2 has at most n distinct eigenvalues. Notice that

$$E\big(\lambda_2, T_2|_{E(\lambda_1,T_1)}\big) = \{v \in E(\lambda_1, T_1) : T_2 v = \lambda_2 v\}$$
$$= \{v \in V : T_1 v = \lambda_1 v \text{ and } T_2 v = \lambda_2 v\} = E(\lambda_1, T_1) \cap E(\lambda_2, T_2).$$

Combining this with $V = \bigoplus_{\lambda_1 \in \mathbf{F}} E(\lambda_1, T_1)$ and $E(\lambda_1, T_1) = \bigoplus_{\lambda_2 \in \mathbf{F}} E\big(\lambda_2, T_2|_{E(\lambda_1,T_1)}\big)$, we see that

$$V = \bigoplus_{\lambda_1, \lambda_2 \in \mathbf{F}} (E(\lambda_1, T_1) \cap E(\lambda_2, T_2)).$$

If we continue this process, we find that

$$V = \bigoplus_{\lambda_1, \dots, \lambda_m \in \mathbf{F}} (E(\lambda_1, T_1) \cap \dots \cap E(\lambda_m, T_m)),$$

where this direct sum is finite because each T_k has at most n distinct eigenvalues. If we take a basis for each non-zero summand $E(\lambda_1, T_1) \cap \dots \cap E(\lambda_m, T_m)$ and combine these bases, we obtain a basis $v_1, \dots, v_n$ of V such that each basis vector is an eigenvector of each T_k. Thus the matrix of each T_k with respect to $v_1, \dots, v_n$ is diagonal. Because each $T \in \mathcal{E}$ is a linear combination of $T_1, \dots, T_m$, and a linear combination of diagonal matrices is a diagonal matrix, we see that the matrix of each $T \in \mathcal{E}$ is diagonal with respect to $v_1, .., v_n$.

Exercise 5.E.3. Suppose $S, T \in \mathcal{L}(V)$ are such that $ST = TS$. Suppose $p \in \mathcal{P}(\mathbf{F})$.

(a) Prove that $\operatorname{null} p(S)$ is invariant under T.

(b) Prove that $\operatorname{range} p(S)$ is invariant under T.

See 5.18 for the special case $S = T$.

Solution. Suppose $p = \sum_{k=0}^{m} c_k z^k$.

(a) Let $v \in \operatorname{null} p(S)$ be given and observe that

$$p(S)(Tv) = \sum_{k=0}^{m} c_k S^k (Tv) = T\left(\sum_{k=0}^{m} c_k S^k v\right) = T(p(S)v) = T(0) = 0,$$

where we have used that T is linear and that T commutes with S for the second equality. Thus $\operatorname{null} p(S)$ is invariant under T.

(b) Let $v \in \operatorname{range} p(S)$ be given, so that $v = p(S)w$ for some $w \in V$, and observe that

$$Tv = T\left(\sum_{k=0}^{m} c_k S^k w\right) = \sum_{k=0}^{m} c_k S^k (Tw) = p(S)(Tw) \in \operatorname{range} p(S),$$

where we have used that T is linear and that T commutes with S for the second equality. Thus $\operatorname{range} p(S)$ is invariant under T.

Exercise 5.E.4. Prove or give a counterexample: If A is a diagonal matrix and B is an upper-triangular matrix of the same size as A, then A and B commute.

Solution. This is false:

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Exercise 5.E.5. Prove that a pair of operators on a finite-dimensional vector space commute if and only if their dual operators commute.

See 3.118 for the definition of the dual of an operator.

Solution. Suppose V is a finite-dimensional vector space and $S, T \in \mathcal{L}(V)$. Let $v_1, \ldots, v_n$ be a basis of V, let $\varphi_1, \ldots, \varphi_n$ be the corresponding dual basis of V', and let

$$A = \mathcal{M}(S, (v_1, \ldots, v_n)) \quad \text{and} \quad B = \mathcal{M}(T, (v_1, \ldots, v_n)).$$

It follows from 3.132 that the matrices of S' and T' with respect to $\varphi_1, \ldots, \varphi_n$ are A^{t} and B^{t}. To show that S and T commute if and only if S' and T' commute, it will suffice, by 5.74, to show that A and B commute if and only if A^{t} and B^{t} commute. Indeed, using Exercise 3.C.15,

$$AB = BA \quad \Leftrightarrow \quad (AB)^{\mathrm{t}} = (BA)^{\mathrm{t}} \quad \Leftrightarrow \quad B^{\mathrm{t}}A^{\mathrm{t}} = A^{\mathrm{t}}B^{\mathrm{t}}.$$

Exercise 5.E.6. Suppose V is a finite-dimensional complex vector space and $S, T \in \mathcal{L}(V)$ commute. Prove that there exist $\alpha, \lambda \in \mathbf{C}$ such that

$$\operatorname{range}(S - \alpha I) + \operatorname{range}(T - \lambda I) \neq V.$$

Solution. By 5.80 there is a basis $v_1, \ldots, v_n$ such that the matrices $\mathcal{M}(S)$ and $\mathcal{M}(T)$ are both upper-triangular, say

$$\mathcal{M}(S) = \begin{pmatrix} * & \cdots & * \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \alpha \end{pmatrix} \quad \text{and} \quad \mathcal{M}(T) = \begin{pmatrix} * & \cdots & * \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda \end{pmatrix}$$

for some $\alpha, \lambda \in \mathbf{C}$. It follows that

$$\mathcal{M}(S - \alpha I) = \begin{pmatrix} * & \cdots & * \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix} \quad \text{and} \quad \mathcal{M}(T - \lambda I) = \begin{pmatrix} * & \cdots & * \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix}$$

and hence that $\operatorname{range}(S - \alpha I)$ and $\operatorname{range}(T - \lambda I)$ are both contained in $\operatorname{span}(v_1, \ldots, v_{n-1})$. Thus $\operatorname{range}(S - \alpha I) + \operatorname{range}(T - \lambda I) \subseteq \operatorname{span}(v_1, \ldots, v_{n-1}) \neq V$.

Exercise 5.E.7. Suppose V is a complex vector space, $S \in \mathcal{L}(V)$ is diagonalizable, and $T \in \mathcal{L}(V)$ commutes with S. Prove that there is a basis of V such that S has a diagonal matrix with respect to this basis and T has an upper-triangular matrix with respect to this basis.

Solution. We will proceed by induction on $\dim V$. Certainly the result is true for $\dim V = 1$, since all 1-by-1 matrices are diagonal. Let V be a complex vector space of dimension $n > 1$, suppose the result holds for all complex vector spaces of smaller dimension, let $S \in \mathcal{L}(V)$ be diagonalizable, and suppose $T \in \mathcal{L}(V)$ commutes with S. By 5.19 there exists an eigenvalue $\lambda \in \mathbf{C}$ of S and it then follows from Exercise 5.D.5 that $V = U \oplus W$, where

$$U = \operatorname{null}(S - \lambda I) \quad \text{and} \quad W = \operatorname{range}(S - \lambda I).$$

If $S = \lambda I$ then the matrix of S with respect to any basis of V is diagonal and thus the desired basis of V is given by 5.47. If $S \neq \lambda I$ then $1 \leq \dim U < n$ and $1 \leq \dim W < n$. Furthermore, Exercise 5.E.3 shows that U and W are invariant under both S and T. Because S and T commute, their restrictions to any subspace of V will also commute and thus we can apply our induction hypothesis to both U and W to obtain a basis $v_1, ..., v_m$ of U and a basis $v_{m+1}, ..., v_n$ of W such that each v_k is an eigenvector of S and such that

$$Tv_k \in \operatorname{span}(v_1, ..., v_k) \text{ for } k \in \{1, ..., m\},$$

$$\text{and} \quad Tv_k \in \operatorname{span}(v_{m+1}, ..., v_k) \subseteq \operatorname{span}(v_1, ..., v_k) \text{ for } k \in \{m+1, ..., n\}.$$

Thus $v_1, ..., v_n$ is a basis of V such that $\mathcal{M}(S, (v_1, ..., v_n))$ is diagonal and such that $\mathcal{M}(T, (v_1, ..., v_n))$ is upper-triangular. This completes the induction step and the proof.

Exercise 5.E.8. Suppose $m = 3$ in Example 5.72 and D_x, D_y are the commuting partial differentiation operators on $\mathcal{P}_3(\mathbf{R}^2)$ from that example. Find a basis of $\mathcal{P}_3(\mathbf{R}^2)$ with respect to which D_x and D_y each have an upper-triangular matrix.

Solution. Consider the list

$$B := 1, x, y, x^2, xy, y^2, x^3, x^2y, xy^2, y^3.$$

From the definition of $\mathcal{P}_3(\mathbf{R}^2)$, it is clear that B spans $\mathcal{P}_3(\mathbf{R}^2)$. Suppose we have a linear combination

$$a_{0,0} + a_{1,0}x + a_{0,1}y + a_{2,0}x^2 + a_{1,1}xy + a_{0,2}y^2 + a_{3,0}x^3 + a_{2,1}x^2y + a_{1,2}xy^2 + a_{0,3}y^3 = 0.$$

Taking $y = 0$ shows that $a_{0,0} = a_{1,0} = a_{2,0} = a_{3,0} = 0$, given the linear independence of $1, x, x^2, x^3$ in $\mathcal{P}(\mathbf{R})$, and similarly taking $x = 0$ gives us $a_{0,1} = a_{0,2} = a_{0,3} = 0$. Thus we are left with the linear combination

$$a_{1,1}xy + a_{2,1}x^2y + a_{1,2}xy^2 = 0.$$

Taking $(x, y) \in \{(1,1), (2,1), (1,2)\}$ gives us the system of linear equations

$$\begin{pmatrix} 1 & 1 & 1 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{pmatrix} \begin{pmatrix} a_{1,1} \\ a_{2,1} \\ a_{1,2} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

which has the unique solution $a_{1,1} = a_{2,1} = a_{1,2} = 0$. Thus B is linearly independent and hence forms a basis of $\mathcal{P}_3(\mathbf{R}^2)$. Now observe that applying D_x to each vector in B gives us the list

$$0, 1, 0, 2x, y, 0, 3x^2, 2xy, y^2, 0.$$

It follows that the matrix of D_x with respect to B is upper-triangular:

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Similarly, we find that the matrix of D_y with respect to B is upper-triangular.

Exercise 5.E.9. Suppose V is a finite-dimensional nonzero complex vector space. Suppose that $\mathcal{E} \subseteq \mathcal{L}(V)$ is such that S and T commute for all $S, T \in \mathcal{E}$.

(a) Prove that there is a vector in V that is an eigenvector for every element of $\mathcal{E}$.

(b) Prove that there is a basis of V with respect to which every element of $\mathcal{E}$ has an upper-triangular matrix.

This exercise extends 5.78 and 5.80, which consider the case in which $\mathcal{E}$ contains only two elements. For this exercise, $\mathcal{E}$ may contain any number of elements, and $\mathcal{E}$ may even be an infinite set.

Solution.

(a) Suppose that $\dim V = n$. Because $\dim \mathcal{L}(V) = n^2$, there must exist a subset $\mathcal{F} \subseteq \mathcal{E}$ of cardinality at most n^2 such that every operator in $\mathcal{E}$ is a linear combination of operators in $\mathcal{F}$. Suppose that $\mathcal{F} = \{T_1, ..., T_m\}$ for some $m \leq n^2$. By 5.19 T_1 has an eigenvalue λ_1, and because T_1 and T_2 commute, 5.75 shows that $E(\lambda_1, T_1)$ is invariant under T_2. Another application of 5.19 shows that $T_2|_{E(\lambda_1, T_1)}$ has an eigenvector, which must also be an eigenvector of T_1. Thus $E(\lambda_1, T_1) \cap E(\lambda_2, T_2) \neq \{0\}$. Since T_3 commutes with both T_1 and T_2, 5.75 shows that $E(\lambda_1, T_1)$ and $E(\lambda_2, T_2)$ are both invariant under T_3 and thus, by Exercise 5.A.3, the intersection $E(\lambda_1, T_1) \cap E(\lambda_2, T_2)$ is also invariant under T_3. It then follows from 5.19 that T_3 restricted to $E(\lambda_1, T_1) \cap E(\lambda_2, T_2)$ has an eigenvector, which must also be an eigenvector of T_1 and T_2. By continuing in this manner, we obtain a $v \in V$ that is an eigenvector of each T_k, say $T_k v = \lambda_k v$ for some $\lambda_k \in \mathbf{C}$.

Let $T \in \mathcal{E}$ be given. As noted above, T must be a linear combination of operators in $\mathcal{F}$, say $T = \sum_{k=1}^{m} c_k T_k$. It follows that

$$Tv = \left(\sum_{k=1}^{m} c_k T_k\right) v = \sum_{k=1}^{m} c_k \lambda_k v = \left(\sum_{k=1}^{m} c_k \lambda_k\right) v.$$

Thus v is an eigenvector of T.

(b) Let us first consider the special case where $\mathcal{E}$ is finite. Our proof here is a generalization of the proof of 5.80. For a positive integer n, let $P(n)$ be the statement that if V is an n-dimensional complex vector space and $\{T_1, ..., T_m\}$ is a collection of pairwise commuting operators on V for some $m \geq 2$, then there is a basis of V with respect to which each T_k has an upper-triangular matrix.

The truth of $P(1)$ is clear. For some $n > 1$, suppose that $P(n-1)$ holds, let V be an n-dimensional complex vector space, and let $\{T_1, ..., T_m\}$ be a collection of pairwise commuting operators on V for some $m \geq 2$. By part (a) there exists a $v_1 \in V$ which is an eigenvector of each T_k, so that $T_k v_1 \in \operatorname{span}(v_1)$. Using 2.33, let W be such that $V = \operatorname{span}(v_1) \oplus W$ and define $P \in \mathcal{L}(V, W)$ by $P(a v_1 + w) = w$. For each $k \in \{1, ..., m\}$, define $\hat{T}_k \in \mathcal{L}(W)$ by $\hat{T}_k w = P(T_k w)$. Because each pair of operators in $\{T_1, ..., T_m\}$ commutes, the proof of 5.80 shows that each pair of operators in $\left\{\hat{T}_1, ..., \hat{T}_m\right\}$ also commutes. We can now apply our induction hypothesis to obtain a basis $v_2, ..., v_n$ of W with respect to which the matrix of each $\hat{T}_k$ is upper-triangular. The list $v_1, ..., v_n$ is a basis of V. For each $j \in \{2, ..., n\}$ and each $k \in \{1, ..., m\}$, there exists $a_{j,k} \in \mathbf{C}$ such that

$$T_k v_j = a_{j,k} v_1 + \hat{T}_k v_j.$$

Because $\hat{T}_k v_j \in \operatorname{span}(v_2, ..., v_j)$, this equation implies that $T_k v_j \in \operatorname{span}(v_1, ..., v_j)$. Thus the matrix of each T_k with respect to the basis $v_1, ..., v_n$ is upper triangular. This completes the induction step and the proof.

Now let us consider the general case where $\mathcal{E}$ may be infinite. Because $\dim \mathcal{L}(V) = n^2$, there must exist a subset $\mathcal{F} \subseteq \mathcal{E}$ of cardinality at most n^2 such that every operator in $\mathcal{E}$ is a linear combination of operators in $\mathcal{F}$. Suppose that $\mathcal{F} = \{T_1, ..., T_m\}$ for some $m \leq n^2$. The special case we just proved implies that there is a basis $v_1, ..., v_n$ of V with respect to which each T_k has an upper-triangular matrix. Because a linear combination of upper-triangular matrices is again an upper-triangular matrix and each $T \in \mathcal{E}$ is a linear combination of the operators $\{T_1, ..., T_m\}$, we see that the matrix of each $T \in \mathcal{E}$ with respect to $v_1, ..., v_n$ is upper-triangular.

Exercise 5.E.10. Give an example of two commuting operators S, T on a finite-dimensional real vector space such that $S + T$ has an eigenvalue that does not equal an eigenvalue of S plus an eigenvalue of T and ST has an eigenvalue that does not equal an eigenvalue of S times an eigenvalue of T.

This exercise shows that 5.81 does not hold on real vector spaces.

Solution. Let $S, T \in \mathcal{L}(\mathbf{R}^2)$ be given by $S(x, y) = (-y, x)$ and $T = -S$, i.e. S is a counter-clockwise rotation about the origin by 90° and T is a clockwise rotation about the origin by 90°. It follows that $S + T = 0$ and $ST = I$, so that 0 is an eigenvalue of $S + T$ and 1 is an eigenvalue of ST. However, we may not express either of these eigenvalues as a sum or product of eigenvalues of S and T, because S and T do not have eigenvalues (see 5.9(a)).

Chapter 6. Inner Product Spaces

6.A. Inner Products and Norms

Exercise 6.A.1. Prove or give a counterexample: If $v_1, \dots, v_m \in V$, then

$$\sum_{j=1}^{m} \sum_{k=1}^{m} \langle v_j, v_k \rangle \geq 0.$$

Solution. Suppose we have a sequence of vectors $v_1, v_2, v_3, \dots$ in V. We will use induction on m to prove that

$$\|v_1 + \cdots + v_m\|^2 = \sum_{j=1}^{m} \sum_{k=1}^{m} \langle v_j, v_k \rangle$$

for each positive integer m. The base case is clear, so suppose that the result holds for some positive integer m and observe that

$$\begin{aligned}
\left\|v_1 + \cdots + v_m + v_{m+1}\right\|^2 &= \langle v_1 + \cdots + v_m + v_{m+1}, v_1 + \cdots + v_m + v_{m+1} \rangle \\
&= \langle v_1 + \cdots + v_m, v_1 + \cdots + v_m \rangle + \langle v_1 + \cdots + v_m, v_{m+1} \rangle \\
&\quad + \langle v_{m+1}, v_1 + \cdots + v_m \rangle + \langle v_{m+1}, v_{m+1} \rangle \\
&= \sum_{j=1}^{m} \sum_{k=1}^{m} \langle v_j, v_k \rangle + \sum_{j=1}^{m} \langle v_j, v_{m+1} \rangle + \sum_{k=1}^{m} \langle v_{m+1}, v_k \rangle + \langle v_{m+1}, v_{m+1} \rangle \\
&= \sum_{j=1}^{m+1} \sum_{k=1}^{m+1} \langle v_j, v_k \rangle.
\end{aligned}$$

This completes the induction step. The desired inequality is now immediate:

$$\sum_{j=1}^{m} \sum_{k=1}^{m} \langle v_j, v_k \rangle = \|v_1 + \cdots + v_m\|^2 \geq 0.$$

Exercise 6.A.2. Suppose $S \in \mathcal{L}(V)$. Define $\langle \cdot, \cdot \rangle_1$ by

$$\langle u, v \rangle_1 = \langle Su, Sv \rangle$$

for all $u, v \in V$. Show that $\langle \cdot, \cdot \rangle_1$ is an inner product on V if and only if S is injective.

Solution. If S is not injective then there exists some non-zero $v \in V$ such that $Sv = 0$. It follows that

$$\langle v, v\rangle_1 = \langle Sv, Sv\rangle = \langle 0, 0\rangle = 0.$$

Thus $\langle\cdot,\cdot\rangle_1$ fails to have the definiteness property required by 6.2 and hence is not an inner product on V.

Now suppose that S is injective. We verify each property required by 6.2.

Positivity. We have $\langle v, v\rangle_1 = \langle Sv, Sv\rangle \geq 0$ for all $v \in V$ by the positivity of $\langle\cdot,\cdot\rangle$.

Definiteness. We have $\langle v, v\rangle_1 = \langle Sv, Sv\rangle = 0$ if and only if $Sv = 0$ by the definiteness of $\langle\cdot,\cdot\rangle$, and $Sv = 0$ if and only if $v = 0$ by the injectivity of S. Thus $\langle v, v\rangle_1 = 0$ if and only if $v = 0$.

Additivity in the first slot. Let $u, v, w \in V$ be given and observe that

$$\langle u + v, w\rangle_1 = \langle S(u + v), Sw\rangle = \langle Su + Sv, Sw\rangle = \langle Su, Sw\rangle + \langle Sv, Sw\rangle = \langle u, w\rangle_1 + \langle v, w\rangle_1,$$

where we have used the linearity of S and the additivity in the first slot of $\langle\cdot,\cdot\rangle$.

Homogeneity in the first slot. Let $\lambda \in \mathbf{F}$ and $u, v \in V$ be given and observe that

$$\langle \lambda u, v\rangle_1 = \langle S(\lambda u), Sv\rangle = \langle \lambda Su, Sv\rangle = \lambda\langle Su, Sv\rangle = \lambda\langle u, v\rangle_1,$$

where we have used the linearity of S and the homogeneity in the first slot of $\langle\cdot,\cdot\rangle$.

Conjugate symmetry. Let $u, v \in V$ be given and observe that

$$\overline{\langle v, u\rangle_1} = \overline{\langle Sv, Su\rangle} = \langle Su, Sv\rangle = \langle u, v\rangle_1,$$

where we have used the conjugate symmetry of $\langle\cdot,\cdot\rangle$.

Exercise 6.A.3.

(a) Show that the function taking an ordered pair $((x_1, x_2), (y_1, y_2))$ of elements of $\mathbf{R}^2$ to $|x_1y_1| + |x_2y_2|$ is not an inner product on $\mathbf{R}^2$.

(b) Show that the function taking an ordered pair $((x_1, x_2, x_3), (y_1, y_2, y_3))$ of elements of $\mathbf{R}^3$ to $x_1y_1 + x_3y_3$ is not an inner product on $\mathbf{R}^3$.

Solution.

(a) Let f be the function in question, i.e. $f : \mathbf{R}^2 \times \mathbf{R}^2 \to \mathbf{R}$ is given by

$$f((x_1, x_2), (y_1, y_2)) = |x_1y_1| + |x_2y_2|,$$

and notice that

$$f((-1, 0), (1, 0)) = 1 \neq -1 = -f((1, 0), (1, 0)).$$

Thus f is not homogeneous in the first slot and hence is not an inner product on $\mathbf{R}^2$.

(b) Let f be the function in question, i.e. $f : \mathbf{R}^3 \times \mathbf{R}^3 \to \mathbf{R}$ is given by

$$f((x_1, x_2, x_3), (y_1, y_2, y_3)) = x_1y_1 + x_3y_3,$$

and notice that

$$f((0,1,0),(0,1,0)) = 0.$$

Thus f fails to have the definiteness property required by 6.2 and hence is not an inner product on $\mathbf{R}^3$.

Exercise 6.A.4. Suppose $T \in \mathcal{L}(V)$ is such that $\|Tv\| \leq \|v\|$ for every $v \in V$. Prove that $T - \sqrt{2}I$ is injective.

Solution. We will prove the contrapositive statement. If $T - \sqrt{2}I$ is not injective then there is some non-zero $v \in V$ such that $Tv = \sqrt{2}v$. Because $v \neq 0$ we have $\|v\| \neq 0$ and thus

$$\|Tv\| = \left\|\sqrt{2}v\right\| = \sqrt{2}\|v\| > \|v\|.$$

Exercise 6.A.5. Suppose V is a real inner product space.

(a) Show that $\langle u+v, u-v\rangle = \|u\|^2 - \|v\|^2$ for every $u, v \in V$.

(b) Show that if $u, v \in V$ have the same norm, then $u+v$ is orthogonal to $u-v$.

(c) Use (b) to show that the diagonals of a rhombus are perpendicular to each other.

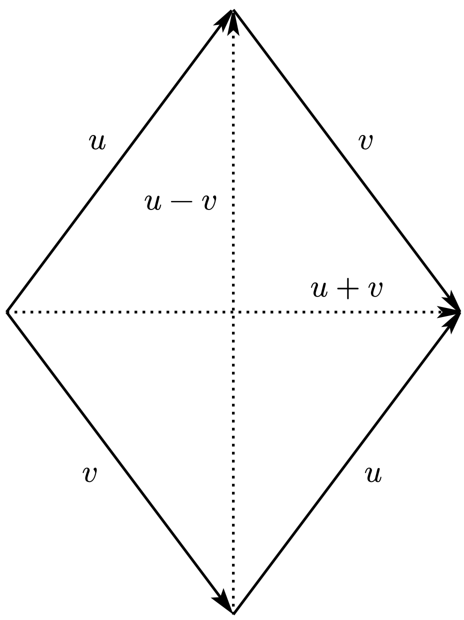
Solution.

(a) For any $u, v \in V$ we have

$$\langle u+v, u-v\rangle = \langle u, u\rangle - \langle u, v\rangle + \langle v, u\rangle - \langle v, v\rangle = \|u\|^2 - \|v\|^2.$$

(b) This is immediate from part (a).

(c) In plane geometry, a rhombus is a quadrilateral whose four sides have the same length. Letting u and v denote the two non-parallel sides, the diagonals are given by $u+v$ and $u-v$. Since $\|u\| = \|v\|$, part (b) shows that $u+v$ and $u-v$ are perpendicular to each other.



Exercise 6.A.6. Suppose $u, v \in V$. Prove that $\langle u, v\rangle = 0 \Leftrightarrow \|u\| \leq \|u + av\|$ for all $a \in \mathbf{F}$.

Solution. Suppose that $\langle u, v\rangle = 0$ and let $a \in \mathbf{F}$ be given. Observe that

$$\langle u, av\rangle = \overline{a}\langle u, v\rangle = 0.$$

Thus u and av are orthogonal. It follows from the Pythagorean Theorem (6.12) that

$$\|u + av\|^2 = \|u\|^2 + \|av\|^2 \geq \|u\|^2.$$

Taking square roots gives the desired inequality.

Now suppose that $\langle u, v\rangle \neq 0$. By 6.11 it must be the case that $v \neq 0$. Thus we can define c and w as in 6.13, so that $\langle w, v\rangle = 0$ and $u = cv + w$. Since w and v are orthogonal, the Pythagorean Theorem (6.12) shows that

$$\|u\|^2 = \|cv + w\|^2 = |c|^2\|v\|^2 + \|w\|^2 > \|w\|^2 = \|u - cv\|^2;$$

the inequality is strict here because $c \neq 0$ and $v \neq 0$. Taking square roots gives us $\|u\| > \|u - cv\|$ and thus a choice of $a = -c$ gives us the desired result.

Exercise 6.A.7. Suppose $u, v \in V$. Prove that $\|au + bv\| = \|bu + av\|$ for all $a, b \in \mathbf{R}$ if and only if $\|u\| = \|v\|$.

Solution. For any $a, b \in \mathbf{R}$, note that

$$\|au + bv\| = \|bu + av\| \quad \Leftrightarrow \quad \|au + bv\|^2 = \|bu + av\|^2.$$

Note further that

$$\|au + bv\|^2 = a^2\|u\|^2 + 2ab\langle u, v\rangle + b^2\|v\|^2$$
$$\text{and} \quad \|bu + av\|^2 = b^2\|u\|^2 + 2ab\langle u, v\rangle + a^2\|v\|^2.$$

Thus $\|au + bv\|^2 = \|bu + av\|^2$ holds if and only if

$$(a^2 - b^2)\left(\|u\|^2 - \|v\|^2\right) = 0.$$

Given this, it will suffice to show that $(a^2 - b^2)\left(\|u\|^2 - \|v\|^2\right) = 0$ for all $a, b \in \mathbf{R}$ if and only if $\|u\| = \|v\|$. The reverse implication is clear; for the forward implication, simply take $a = 1$ and $b = 0$.

Exercise 6.A.8. Suppose $a, b, c, x, y \in \mathbf{R}$ and $a^2 + b^2 + c^2 + x^2 + y^2 \leq 1$. Prove that $a + b + c + 4x + 9y \leq 10$.

Solution. Let $u, v \in \mathbf{R}^5$ be given by $u = (a, b, c, x, y)$ and $v = (1, 1, 1, 4, 9)$. The Cauchy-Schwarz inequality (6.14) shows that

$$
\begin{aligned}
a+b+c+4x+9y &\le |a+b+c+4x+9y| = |\langle u,v\rangle| \\
&\le \|u\|\|v\| = \sqrt{a^2+b^2+c^2+x^2+y^2}\sqrt{1+1+1+16+81} \le 10.
\end{aligned}
$$

Exercise 6.A.9. Suppose $u, v \in V$ and $\|u\| = \|v\| = 1$ and $\langle u, v\rangle = 1$. Prove that $u = v$.

Solution. Observe that

$$\langle u-v, u-v\rangle = \langle u,u\rangle - \langle u,v\rangle - \langle v,u\rangle + \langle v,v\rangle = \|u\|^2 - 2\operatorname{Re}\langle u,v\rangle + \|v\|^2 = 0.$$

It follows from definiteness that $u - v = 0$.

Exercise 6.A.10. Suppose $u, v \in V$ and $\|u\| \le 1$ and $\|v\| \le 1$. Prove that

$$\sqrt{1-\|u\|^2}\sqrt{1-\|v\|^2} \le 1 - |\langle u,v\rangle|.$$

Solution. Observe that

$$
\begin{aligned}
0 \le (\|u\| - \|v\|)^2 \quad &\Leftrightarrow \quad 0 \le \|u\|^2 - 2\|u\|\|v\| + \|v\|^2 \\
&\Leftrightarrow \quad -\|u\|^2 - \|v\|^2 \le -2\|u\|\|v\| \\
&\Leftrightarrow \quad 1 - \|u\|^2 - \|v\|^2 + \|u\|^2\|v\|^2 \le 1 - 2\|u\|\|v\| + \|u\|^2\|v\|^2 \\
&\Leftrightarrow \quad \left(1-\|u\|^2\right)\left(1-\|v\|^2\right) \le (1 - \|u\|\|v\|)^2.
\end{aligned}
$$

Since $\|u\| \le 1$ and $\|v\| \le 1$, the quantities $1-\|u\|^2, 1-\|v\|^2$, and $1 - \|u\|\|v\|$ are non-negative. Thus we may take square roots to obtain the inequality

$$\sqrt{1-\|u\|^2}\sqrt{1-\|v\|^2} \le 1 - \|u\|\|v\|.$$

The Cauchy-Schwarz inequality (6.14) shows that $1 - \|u\|\|v\| \le 1 - |\langle u,v\rangle|$ and thus

$$\sqrt{1-\|u\|^2}\sqrt{1-\|v\|^2} \le 1 - |\langle u,v\rangle|.$$

Exercise 6.A.11. Find vectors $u, v \in \mathbf{R}^2$ such that u is a scalar multiple of $(1,3)$, v is orthogonal to $(1,3)$, and $(1,2) = u + v$.

Solution. Let $x = (1,2), y = (1,3)$, and let

$$c = \frac{\langle x,y\rangle}{\|y\|^2} = \frac{7}{10}, \quad u = cy, \quad \text{and} \quad v = x - cy.$$

Then u is a scalar multiple of y and, as 6.13 shows, $\langle v, y\rangle = 0$ and $x = u + v$.

Exercise 6.A.12. Suppose a, b, c, d are positive numbers.

(a) Prove that $(a+b+c+d)\left(\frac{1}{a}+\frac{1}{b}+\frac{1}{c}+\frac{1}{d}\right) \geq 16$.

(b) For which positive numbers a, b, c, d is the inequality above an equality?

Solution.

(a) If we let

$$u = \left(\sqrt{a}, \sqrt{b}, \sqrt{c}, \sqrt{d}\right) \quad \text{and} \quad v = \left(\frac{1}{\sqrt{a}}, \frac{1}{\sqrt{b}}, \frac{1}{\sqrt{c}}, \frac{1}{\sqrt{d}}\right),$$

then

$$\langle u, v\rangle = 4, \quad \|u\| = \sqrt{a+b+c+d}, \quad \text{and} \quad \|v\| = \sqrt{\frac{1}{a}+\frac{1}{b}+\frac{1}{c}+\frac{1}{d}}.$$

Squaring both sides of the Cauchy-Schwarz inequality (6.14) gives the desired inequality.

(b) For positive numbers a, b, c, d, we claim that

$$(a+b+c+d)\left(\frac{1}{a}+\frac{1}{b}+\frac{1}{c}+\frac{1}{d}\right) = 16 \quad \Leftrightarrow \quad a = b = c = d.$$

The reverse implication is straightforward to check. For the forward implication, define u and v as in part (a) and note that the Cauchy-Schwarz inequality is an equality if and only if one of u, v is a scalar multiple of the other. If $u = \lambda v$ for some $\lambda \in \mathbf{R}$, then necessarily $\lambda > 0$ since $a > 0$ and

$$\sqrt{a} = \frac{\lambda}{\sqrt{a}} \quad \Rightarrow \quad a = \lambda.$$

Similarly we find that $b = c = d = \lambda$. If $v = \lambda u$ for some $\lambda \in \mathbf{R}$ then again λ must be positive and

$$\frac{1}{\sqrt{a}} = \lambda\sqrt{a} \quad \Rightarrow \quad a = \frac{1}{\lambda}.$$

Similarly we find that $b = c = d = \frac{1}{\lambda}$. In either case we have $a = b = c = d$.

Exercise 6.A.13. Show that the square of an average is less than or equal to the average of the squares. More precisely, show that if $a_1, \ldots, a_n \in \mathbf{R}$, then the square of the average of $a_1, \ldots, a_n$ is less than or equal to the average of $a_1^2, \ldots, a_n^2$.

Solution. For a positive integer n and real numbers $a_1, \ldots, a_n$, let

$$u = (a_1, \ldots, a_n) \in \mathbf{R}^n \quad \text{and} \quad v = \left(\frac{1}{n}, \ldots, \frac{1}{n}\right) \in \mathbf{R}^n.$$

Observe that

$$\langle u, v\rangle^2 = \left(\frac{a_1 + \cdots + a_n}{n}\right)^2, \quad \|u\|^2 = a_1^2 + \cdots + a_n^2, \quad \text{and} \quad \|v\|^2 = \frac{1}{n}.$$

Squaring both sides of the Cauchy-Schwarz inequality (6.14) gives us the desired inequality.

Exercise 6.A.14. Suppose $v \in V$ and $v \neq 0$. Prove that $v/\|v\|$ is the unique closest element on the unit sphere of V to v. More precisely, prove that if $u \in V$ and $\|u\| = 1$, then

$$\left\| v - \frac{v}{\|v\|} \right\| \leq \|v - u\|,$$

with equality only if $u = v/\|v\|$.

Solution. Some routine calculations show that

$$\left\| v - \frac{v}{\|v\|} \right\|^2 = \|v\|^2 + 1 - 2\|v\| \quad \text{and} \quad \|v - u\|^2 = \|v\|^2 + 1 - 2\operatorname{Re}\langle v, u\rangle.$$

Thus

$$\left\| v - \frac{v}{\|v\|} \right\| \leq \|v - u\| \quad \Leftrightarrow \quad \left\| v - \frac{v}{\|v\|} \right\|^2 \leq \|v - u\|^2 \quad \Leftrightarrow \quad \operatorname{Re}\langle v, u\rangle \leq \|v\|.$$

Indeed, using the Cauchy-Schwarz inequality,

$$\operatorname{Re}\langle v, u\rangle \leq |\langle v, u\rangle| \leq \|v\|\|u\| = \|v\|.$$

As the proof of 6.17 shows, we have equality here if and only if one of u, v is a non-negative real multiple of the other. If $u = \lambda v$ for some $\lambda \geq 0$, then

$$1 = \|u\| = |\lambda|\|v\| = \lambda\|v\| \quad \Rightarrow \quad \lambda = \frac{1}{\|v\|} \quad \Rightarrow \quad u = \frac{v}{\|v\|},$$

and if $v = \lambda u$ for some $\lambda \geq 0$ then

$$\|v\| = |\lambda|\|u\| = \lambda \quad \Rightarrow \quad u = \frac{v}{\|v\|}.$$

Exercise 6.A.15. Suppose u, v are nonzero vectors in $\mathbf{R}^2$. Prove that

$$\langle u, v\rangle = \|u\|\|v\| \cos\theta,$$

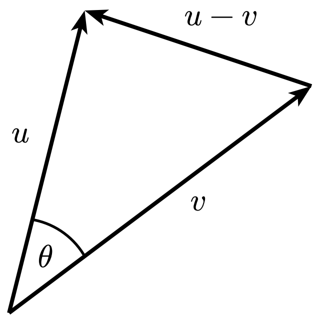
where θ is the angle between u and v (thinking of u and v as arrows with initial point at the origin).

Hint: Use the law of cosines on the triangle formed by u, v, and $u - v$.

Solution. The law of cosines applied to the triangle formed by u, v, and $u - v$ states that

$$\|u - v\|^2 = \|u\|^2 + \|v\|^2 - 2\|u\|\|v\| \cos\theta.$$

This expression together with the identity $\|u-v\|^2 = \|u\|^2 + \|v\|^2 - 2\langle u, v\rangle$ gives us the desired equality.



Exercise 6.A.16. The angle between two vectors (thought of as arrows with initial point at the origin) in $\mathbf{R}^2$ or $\mathbf{R}^3$ can be defined geometrically. However, geometry is not as clear in $\mathbf{R}^n$ for $n > 3$. Thus the angle between two nonzero vectors $x, y \in \mathbf{R}^n$ is defined to be

$$\arccos \frac{\langle x, y\rangle}{\|x\|\|y\|},$$

where the motivation for this definition comes from Exercise 15. Explain why the Cauchy-Schwarz inequality is needed to show that this definition makes sense.

Solution. The arccos function is only defined on the interval $[-1, 1]$; for the definition in question to make sense, we must have

$$\frac{\langle x, y\rangle}{\|x\|\|y\|} \in [-1, 1] \quad \Leftrightarrow \quad \frac{|\langle x, y\rangle|}{\|x\|\|y\|} \in [0, 1]$$

for any non-zero $x, y \in \mathbf{R}^n$. The Cauchy-Schwarz inequality ensures this.

Exercise 6.A.17. Prove that

$$\left(\sum_{k=1}^{n} a_k b_k\right)^2 \leq \left(\sum_{k=1}^{n} k a_k^2\right)\left(\sum_{k=1}^{n} \frac{b_k^2}{k}\right)$$

for all real numbers $a_1, \ldots, a_n$ and $b_1, \ldots, b_n$.

Solution. Let n be a positive integer and let $a_1, \ldots, a_n, b_1, \ldots, b_n$ be real numbers. Define

$$u = \left(a_1, \sqrt{2}a_2, \sqrt{3}a_3, \ldots, \sqrt{n}a_n\right) \quad \text{and} \quad v = \left(b_1, \frac{b_2}{\sqrt{2}}, \frac{b_3}{\sqrt{3}}, \ldots, \frac{b_n}{\sqrt{n}}\right).$$

Then

$$\langle u, v\rangle = \sum_{k=1}^{n} a_k b_k, \quad \|u\| = \left(\sum_{k=1}^{n} k a_k^2\right)^{1/2}, \quad \text{and} \quad \|v\| = \left(\sum_{k=1}^{n} \frac{b_k^2}{k}\right)^{1/2}.$$

Squaring both sides of the Cauchy-Schwarz inequality gives us the desired inequality.

Exercise 6.A.18.

(a) Suppose $f : [1, \infty) \to [0, \infty)$ is continuous. Show that

$$\left(\int_1^\infty f\right)^2 \leq \int_1^\infty x^2 (f(x))^2 \, dx.$$

(b) For which continuous functions $f : [1, \infty) \to [0, \infty)$ is the inequality in (a) an equality with both sides finite?

Solution.

(a) For $t \geq 1$ consider the vector space of continuous real-valued functions on the interval $[1, t]$ equipped with the inner product

$$\langle g, h \rangle = \int_1^t g(x)h(x) \, \mathrm{d}x.$$

The Cauchy-Schwarz inequality shows that

$$\begin{aligned}\left(\int_1^t f(x) \, \mathrm{d}x\right)^2 = \left(\int_1^t \frac{x}{x} f(x) \, \mathrm{d}x\right)^2 &\leq \left(\int_1^t x^2 (f(x))^2 \, \mathrm{d}x\right)\left(\int_1^t \frac{1}{x^2} \, \mathrm{d}x\right) \\ &= \left(\int_1^t x^2 (f(x))^2 \, \mathrm{d}x\right)\left(1 - \frac{1}{t}\right).\end{aligned}$$

Because f is non-negative, both integrals $\int_1^\infty f(x) \, \mathrm{d}x$ and $\int_1^\infty x^2 (f(x))^2 \, \mathrm{d}x$ either converge or diverge to infinity. If $\int_1^\infty x^2 (f(x))^2 \, \mathrm{d}x = \infty$ then the desired inequality certainly holds, and if $\int_1^\infty x^2 (f(x))^2 \, \mathrm{d}x$ converges then the inequality

$$\left(\int_1^t f(x) \, \mathrm{d}x\right)^2 \leq \left(\int_1^t x^2 (f(x))^2 \, \mathrm{d}x\right)\left(1 - \frac{1}{t}\right)$$

shows that $\int_1^\infty f(x) \, \mathrm{d}x$ also converges and furthermore that

$$\left(\int_1^\infty f(x) \, \mathrm{d}x\right)^2 \leq \int_1^\infty x^2 (f(x))^2 \, \mathrm{d}x.$$

(b) The Cauchy-Schwarz inequality used in part (a) is an equality if and only if $xf(x)$ and x^{-1} are linearly dependent as functions on $[1, \infty)$, i.e. if and only if $f(x) = \lambda x^{-2}$ for all $x \geq 1$ and some $\lambda \geq 0$. In this case we obtain

$$\left(\int_1^\infty f(x) \, \mathrm{d}x\right)^2 = \int_1^\infty x^2 (f(x))^2 \, \mathrm{d}x = \lambda^2.$$

Exercise 6.A.19. Suppose $v_1, \ldots, v_n$ is a basis of V and $T \in \mathcal{L}(V)$. Prove that if λ is an eigenvalue of T, then

$$|\lambda|^2 \leq \sum_{j=1}^{n} \sum_{k=1}^{n} \left|\mathcal{M}(T)_{j,k}\right|^2,$$

where $\mathcal{M}(T)_{j,k}$ denotes the entry in row j, column k of the matrix of T with respect to the basis $v_1, \ldots, v_n$.

Solution. It is straightforward to verify that

$$\langle a_1 v_1 + \cdots + a_n v_n, b_1 v_1 + \cdots + b_n v_n \rangle = a_1 \overline{b_1} + \cdots + a_n \overline{b_n}$$

is an inner product on V (this is essentially the Euclidean inner product after identifying V with $\mathbf{F}^n$). Because λ is an eigenvalue of T, there is a non-zero $v \in V$ such that $Tv = \lambda v$; by replacing v with $v/\|v\|$ if necessary, we may assume that $\|v\| = 1$. Suppose that $v = a_1 v_1 + \cdots + a_n v_n$ and observe that

$$Tv = \sum_{k=1}^{n} a_k T v_k = \sum_{k=1}^{n} a_k \sum_{j=1}^{n} \mathcal{M}(T)_{j,k} v_j = \sum_{j=1}^{n} \left(\sum_{k=1}^{n} a_k \mathcal{M}(T)_{j,k} \right) v_j$$

$$\Rightarrow \quad \|Tv\|^2 = \sum_{j=1}^{n} \left| \sum_{k=1}^{n} a_k \mathcal{M}(T)_{j,k} \right|^2.$$

For each $j \in \{1, \ldots, n\}$, applying the Cauchy-Schwarz inequality to the vectors

$$(a_1, \ldots, a_n) \quad \text{and} \quad \left(\overline{\mathcal{M}(T)_{j,1}}, \ldots, \overline{\mathcal{M}(T)_{j,n}} \right)$$

in $\mathbf{F}^n$ with the Euclidean inner product shows that

$$\left| \sum_{k=1}^{n} a_k \mathcal{M}(T)_{j,k} \right|^2 \leq \left(\sum_{k=1}^{n} |a_k|^2 \right) \left(\sum_{k=1}^{n} \left|\mathcal{M}(T)_{j,k}\right|^2 \right)$$

$$= \|v\|^2 \left(\sum_{k=1}^{n} \left|\mathcal{M}(T)_{j,k}\right|^2 \right) = \sum_{k=1}^{n} \left|\mathcal{M}(T)_{j,k}\right|^2.$$

Thus

$$|\lambda|^2 = \|\lambda v\|^2 = \|Tv\|^2 = \sum_{j=1}^{n} \left| \sum_{k=1}^{n} a_k \mathcal{M}(T)_{j,k} \right|^2 \leq \sum_{j=1}^{n} \sum_{k=1}^{n} \left|\mathcal{M}(T)_{j,k}\right|^2.$$

Exercise 6.A.20. Prove that if $u, v \in V$ then $\big|\|u\| - \|v\|\big| \leq \|u - v\|$.

The inequality above is called the ***reverse triangle inequality****. For the reverse triangle inequality when $V = \mathbf{C}$, see* *Exercise 2 in Chapter 4*.

Solution. Notice that

$$\|u\| = \|u - v + v\| \leq \|u - v\| + \|v\| \quad \Rightarrow \quad \|u\| - \|v\| \leq \|u - v\|,$$
$$\|v\| = \|v - u + u\| \leq \|u - v\| + \|u\| \quad \Rightarrow \quad \|v\| - \|u\| \leq \|u - v\|.$$

Thus $|\|u\| - \|v\|| \leq \|u - v\|$.

Exercise 6.A.21. Suppose $u, v \in V$ are such that

$$\|u\| = 3, \quad \|u + v\| = 4, \quad \|u - v\| = 6.$$

What number does $\|v\|$ equal?

Solution. Rearranging the parallelogram equality (6.21) for $\|v\|$ gives

$$\|v\| = \left(\frac{\|u + v\|^2 + \|u - v\|^2}{2} - \|u\|^2 \right)^{1/2}.$$

Substituting the given values, we find $\|v\| = \sqrt{17}$.

Exercise 6.A.22. Show that if $u, v \in V$, then

$$\|u + v\|\|u - v\| \leq \|u\|^2 + \|v\|^2.$$

Solution. Notice that

$$0 \leq (\|u + v\| - \|u - v\|)^2 \quad \Leftrightarrow \quad 4\|u + v\|\|u - v\| \leq (\|u + v\| + \|u - v\|)^2$$
$$\Leftrightarrow \quad \|u + v\|\|u - v\| \leq \tfrac{1}{2}(\|u + v\| + \|u - v\|)^2 - \|u + v\|\|u - v\| = \|u\|^2 + \|v\|^2,$$

where the last equality is the parallelogram equality (6.21).

Exercise 6.A.23. Suppose $v_1, ..., v_m \in V$ are such that $\|v_k\| \leq 1$ for each $k = 1, ..., m$. Show that there exist $a_1, ..., a_m \in \{1, -1\}$ such that

$$\|a_1 v_1 + \cdots + a_m v_m\| \leq \sqrt{m}.$$

Solution. We will inductively define the integers $a_1, ..., a_m$. To begin, simply take $a_1 = 1$. For $k \in \{1, ..., m - 1\}$, suppose we have chosen $a_1, ..., a_k \in \{1, -1\}$ such that

$$\|u\| \leq \sqrt{k}, \quad \text{where} \quad u = a_1 v_1 + \cdots + a_k v_k.$$

It follows from Exercise 6.A.22 that

$$\left\|u + v_{k+1}\right\|\left\|u - v_{k+1}\right\| \leq \|u\|^2 + \left\|v_{k+1}\right\|^2 \leq k + 1.$$

Thus at least one of $\left\|u + v_{k+1}\right\|, \left\|u - v_{k+1}\right\|$ is less than or equal to $\sqrt{k+1}$. Let $a_{k+1} = 1$ if $\left\|u + v_{k+1}\right\| \leq \sqrt{k+1}$ and let $a_{k+1} = -1$ otherwise. By repeating this process until $k = m - 1$, we obtain the desired integers $a_1, ..., a_m$.

Exercise 6.A.24. Prove or give a counterexample: If $\|\cdot\|$ is the norm associated with an inner product on $\mathbf{R}^2$, then there exists $(x,y) \in \mathbf{R}^2$ such that $\|(x,y)\| \neq \max\{|x|,|y|\}$.

Solution. Let $f : \mathbf{R}^2 \to \mathbf{R}$ be given by $f(x,y) = \max\{|x|,|y|\}$, let $u = (1,0)$ and $v = (1,1)$, and observe that

$$[f(u+v)]^2 + [f(u-v)]^2 = 5 \neq 4 = 2\Big([f(u)]^2 + [f(v)]^2\Big).$$

Since a norm associated with an inner product must satisfy the parallelogram equality (6.21), it follows that any norm $\|\cdot\|$ associated with an inner product on $\mathbf{R}^2$ cannot be given by f. That is, there must exist some $(x,y) \in \mathbf{R}^2$ such that $\|(x,y)\| \neq f(x,y)$.

Exercise 6.A.25. Suppose $p > 0$. Prove that there is an inner product on $\mathbf{R}^2$ such that the associated norm is given by

$$\|(x,y)\| = \big(|x|^p + |y|^p\big)^{1/p}$$

for all $(x,y) \in \mathbf{R}^2$ if and only if $p = 2$.

Solution. Let $f : \mathbf{R}^2 \to \mathbf{R}$ be given by $f(x,y) = \big(|x|^p + |y|^p\big)^{1/p}$, let $u = (1,0)$ and $v = (0,1)$, and observe that

$$[f(u+v)]^2 + [f(u-v)]^2 = 2^{1+2/p} \quad \text{and} \quad 2\Big([f(u)]^2 + [f(v)]^2\Big) = 4.$$

If f was indeed a norm arising from an inner product then f would satisfy the parallelogram equality (6.21). Since the quantities above are equal if and only if $2^{1+2/p} = 4$, i.e. if and only if $p = 2$, the only possible value for p is 2, which indeed gives the norm associated with the Euclidean inner product on $\mathbf{R}^2$, as 6.8(a) shows.

Exercise 6.A.26. Suppose V is a real inner product space. Prove that

$$\langle u, v\rangle = \frac{\|u+v\|^2 - \|u-v\|^2}{4}$$

for all $u, v \in V$.

Solution. For any $u, v \in V$, observe that

$$\begin{aligned}
\|u+v\|^2 - \|u-v\|^2 &= \langle u+v, u+v\rangle - \langle u-v, u-v\rangle \\
&= \langle u,u\rangle + 2\langle u,v\rangle + \langle v,v\rangle - \langle u,u\rangle + 2\langle u,v\rangle - \langle v,v\rangle \\
&= 4\langle u,v\rangle.
\end{aligned}$$

Exercise 6.A.27. Suppose V is a complex inner product space. Prove that

$$\langle u,v\rangle = \frac{\|u+v\|^2 - \|u-v\|^2 + \|u+iv\|^2 i - \|u-iv\|^2 i}{4}$$

for all $u, v \in V$.

Solution. For any $u, v \in V$, observe that

$$\begin{aligned}
\|u+v\|^2 - \|u-v\|^2 &= \langle u+v, u+v\rangle - \langle u-v, u-v\rangle \\
&= \langle u,u\rangle + \langle u,v\rangle + \overline{\langle u,v\rangle} + \langle v,v\rangle \\
&\qquad - \langle u,u\rangle + \langle u,v\rangle + \overline{\langle u,v\rangle} - \langle v,v\rangle \\
&= 2\Big(\langle u,v\rangle + \overline{\langle u,v\rangle}\Big) \\
&= 4\operatorname{Re}\langle u,v\rangle.
\end{aligned}$$

Furthermore,

$$\begin{aligned}
i\|u+iv\|^2 - i\|u-iv\|^2 &= i\langle u+iv, u+iv\rangle - i\langle u-iv, u-iv\rangle \\
&= i\langle u,u\rangle - i^2\langle u,v\rangle + i^2\overline{\langle u,v\rangle} + i\langle v,v\rangle \\
&\qquad - i\langle u,u\rangle - i^2\langle u,v\rangle + i^2\overline{\langle u,v\rangle} - i\langle v,v\rangle \\
&= 2\Big(\langle u,v\rangle - \overline{\langle u,v\rangle}\Big) \\
&= 4\operatorname{Im}\langle u,v\rangle.
\end{aligned}$$

It follows that

$$\|u+v\|^2 - \|u-v\|^2 + \|u+iv\|^2 i - \|u-iv\|^2 i = 4\langle u,v\rangle.$$

Exercise 6.A.28. A norm on a vector space U is a function

$$\|\cdot\| : U \to [0,\infty)$$

such that $\|u\| = 0$ if and only if $u = 0$, $\|\alpha u\| = |\alpha|\|u\|$ for all $\alpha \in \mathbf{F}$ and all $u \in U$, and $\|u+v\| \le \|u\| + \|v\|$ for all $u, v \in U$. Prove that a norm satisfying the parallelogram equality comes from an inner product (in other words, show that if $\|\cdot\|$ is a norm on U satisfying the parallelogram equality, then there is an inner product $\langle\cdot,\cdot\rangle$ on U such that $\|u\| = \langle u,u\rangle^{1/2}$ for all $u \in U$).

Solution. Let us first consider the case where U is a real vector space. Define $\langle\cdot,\cdot\rangle : U \times U \to \mathbf{R}$ by

$$\langle u,v\rangle = \frac{\|u+v\|^2 - \|u-v\|^2}{4}.$$

For any $u \in U$ we have

$$\langle u, u \rangle = \frac{\|2u\|^2}{4} = \|u\|^2.$$

Thus the norm is given by $\|u\| = \langle u, u \rangle^{1/2}$. We now show that $\langle \cdot, \cdot \rangle$ is an inner product on U.

Positive-definiteness. Combining the identity $\langle u, u \rangle = \|u\|^2$ with the properties of the norm $\| \cdot \|$ shows that $\langle \cdot, \cdot \rangle$ is positive-definite.

Symmetry. For any $u, v \in V$, observe that $\|v - u\| = |-1|\|u - v\| = \|u - v\|$; it follows that

$$\langle u, v \rangle = \frac{\|u + v\|^2 - \|u - v\|^2}{4} = \frac{\|v + u\|^2 - \|v - u\|^2}{4} = \langle v, u \rangle.$$

Additivity in the first slot. Let $u, v, w \in U$ be given. Since $\| \cdot \|$ satisfies the parallelogram equality, we have

$$\|v + 2w\|^2 + \|v\|^2 = 2\|v + w\|^2 + 2\|w\|^2,$$
$$\|v - 2w\|^2 + \|v\|^2 = 2\|v - w\|^2 + 2\|w\|^2.$$

Subtracting the latter of these equations from the former gives us

$$\|v + 2w\|^2 - \|v - 2w\|^2 = 2\|v + w\|^2 - 2\|v - w\|^2. \tag{1}$$

Now we use the parallelogram equality two more times:

$$2\|u + v + w\|^2 + 2\|u - w\|^2 = \|v + 2u\|^2 + \|v + 2w\|^2,$$
$$2\|u + v - w\|^2 + 2\|u + w\|^2 = \|v + 2u\|^2 + \|v - 2w\|^2.$$

Subtracting the latter of these equations from the former gives us

$$2\Big(\|u + v + w\|^2 + \|u - w\|^2\Big) - 2\Big(\|u + v - w\|^2 + \|u + w\|^2\Big) = \|v + 2w\|^2 - \|v - 2w\|^2.$$

Combining this with equation (1), we see that

$$2\Big(\|u + v + w\|^2 + \|u - w\|^2\Big) - 2\Big(\|u + v - w\|^2 + \|u + w\|^2\Big) = 2\|v + w\|^2 - 2\|v - w\|^2.$$

Equivalently,

$$\frac{\|u + v + w\|^2 - \|u + v - w\|^2}{4} = \frac{\|u + w\|^2 - \|u - w\|^2 + \|v + w\|^2 - \|v - w\|^2}{4},$$

which is exactly the statement $\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$.

Homogeneity in the first slot. Suppose $u, v \in U$. First, we will use induction to show that $\langle nu, v \rangle = n\langle u, v \rangle$ for all positive integers n. The base case $n = 1$ is clear, so suppose that the result holds for some positive integer n and observe that

$$\langle (n + 1)u, v \rangle = \langle nu + u, v \rangle = \langle nu, v \rangle + \langle u, v \rangle = n\langle u, v \rangle + \langle u, v \rangle = (n + 1)\langle u, v \rangle,$$

where we have used additivity in the first slot and the induction hypothesis. This completes the induction step and thus $\langle nu, v \rangle = n\langle u, v \rangle$ for all positive integers n. Certainly $\langle 0, v \rangle = 0$

and so we may extend this result to all non-negative integers. If n is a positive integer then observe that

$$\langle -nu, v\rangle + n\langle u, v\rangle = \langle -nu, v\rangle + \langle nu, v\rangle = \langle 0, v\rangle = 0,$$

where we have used additivity in the first slot and homogeneity in the first slot for positive integers. It follows that $\langle -nu, v\rangle = -n\langle u, v\rangle$ and thus we have homogeneity in the first slot for all integers.

To extend homogeneity in the first slot to rational numbers, let n be a positive integer. By additivity in the first slot, we have

$$n\langle n^{-1}u, v\rangle = \textstyle\sum_{j=1}^{n}\langle n^{-1}u, v\rangle = \left\langle \sum_{j=1}^{n} n^{-1}u, v\right\rangle = \langle u, v\rangle,$$

which implies that $\langle n^{-1}u, v\rangle = n^{-1}\langle u, v\rangle$. Combining this with homogeneity in the first slot for integers allows us to extend homogeneity in the first slot to rational numbers.

Finally, to obtain homogeneity in the first slot for all real numbers, let $\lambda \in \mathbf{R}$ be given. There exists a sequence (r_n) of rational numbers satisfying $\lim_{n\to\infty} r_n = \lambda$. The reverse triangle inequality (Exercise 6.A.20) shows that the function $U \to \mathbf{R}$ given by $u \mapsto \|u\|$ is continuous. Combining this with standard results on compositions and linear combinations of continuous functions, we have

$$\begin{aligned}
\lambda\langle u, v\rangle &= \lim_{n\to\infty} r_n\langle u, v\rangle \\
&= \lim_{n\to\infty} \langle r_n u, v\rangle \\
&= \lim_{n\to\infty} \frac{\|r_n u + v\|^2 - \|r_n u - v\|^2}{4} \\
&= \frac{\|\lambda u + v\|^2 - \|\lambda u - v\|^2}{4} \\
&= \langle \lambda u, v\rangle.
\end{aligned}$$

Now let us consider the case where U is a complex vector space. Define $B : U \times U \to \mathbf{R}$ by

$$B(u, v) = \frac{\|u + v\|^2 - \|u - v\|^2}{4}$$

and define $\langle \cdot, \cdot\rangle : U \times U \to \mathbf{C}$ by

$$\langle u, v\rangle = B(u, v) + iB(u, iv) = \frac{\|u + v\|^2 - \|u - v\|^2}{4} + i\frac{\|u + iv\|^2 - \|u - iv\|^2}{4}.$$

Observe that for any $u \in U$ we have $B(u, u) = \|u\|^2$ and

$$B(u, iu) = \frac{\|u + iu\|^2 - \|u - iu\|^2}{4} = \frac{|1 + i|^2\|u\|^2 - |1 - i|^2\|u\|^2}{4} = 0.$$

Thus $\|u\| = \langle u, u\rangle^{1/2}$. We now show that $\langle \cdot, \cdot\rangle$ is an inner product on U.

Positive-definiteness. Combining the identity $\langle u, u\rangle = \|u\|^2$ with the properties of the norm $\|\cdot\|$ shows that $\langle\cdot,\cdot\rangle$ is positive-definite.

Conjugate symmetry. For any $u, v \in U$, note that

$$\overline{\langle v, u\rangle} = \overline{B(v,u) + iB(v,iu)} = B(u,v) - iB(v,iu),$$

where we have used that B is real-valued and symmetric (we showed this in the case where U is a real vector space.) Given the expression above, to verify conjugate symmetry of $\langle\cdot,\cdot\rangle$ it will suffice to show that $B(u,iv) = -B(v,iu)$. Indeed,

$$\begin{aligned}-4B(v,iu) &= \|v-iu\|^2 - \|v+iu\|^2 = \|-i(u+iv)\|^2 - \|i(u-iv)\|^2\\ &= |-i|^2\|u+iv\|^2 - \|i\|^2\|u-iv\|^2 = \|u+iv\|^2 - \|u-iv\|^2 = 4B(u,iv).\end{aligned}$$

Additivity in the first slot. Let $u, v, w \in U$ be given. The proof of additivity in the first slot we gave for the case where U is a real vector space equally shows that B is additive in the first slot. It follows that

$$\begin{aligned}\langle u+v, w\rangle &= B(u+v,w) + iB(u+v,iw)\\ &= B(u,w) + iB(u,iw) + B(v,w) + iB(v,iw) = \langle u,w\rangle + \langle v,w\rangle.\end{aligned}$$

Homogeneity in the first slot. Let $u, v \in U$ be given. The proof of homogeneity in the first slot we gave for the case where U is a real vector space equally shows that B is homogeneous in the first slot with respect to real numbers. It follows that for any $\lambda \in \mathbf{R}$ we have $\langle \lambda u, v\rangle = \lambda\langle u, v\rangle$. Now observe that

$$\begin{aligned}4\langle iu, v\rangle &= \|iu+v\|^2 - \|iu-v\|^2 + i\left(\|iu+iv\|^2 - \|iu-iv\|^2\right)\\ &= \|i(u-iv)\|^2 - \|i(u+iv)\|^2 + i\left(\|i(u+v)\|^2 - \|i(u-v)\|^2\right)\\ &= \|u-iv\|^2 - \|u+iv\|^2 + i\left(\|u+v\|^2 - \|u-v\|^2\right)\\ &= i\left(\|u+v\|^2 - \|u-v\|^2 + i\left(\|u+iv\|^2 - \|u-iv\|^2\right)\right)\\ &= 4i\langle u, v\rangle.\end{aligned}$$

Thus $\langle iu, v\rangle = i\langle u, v\rangle$. It follows that, for any $x + iy \in \mathbf{C}$,

$$\langle (x+iy)u, v\rangle = \langle xu + iyu, v\rangle = \langle xu, v\rangle + \langle iyu, v\rangle = x\langle u,v\rangle + iy\langle u,v\rangle = (x+iy)\langle u, v\rangle.$$

Exercise 6.A.29. Suppose $V_1, \ldots, V_m$ are inner product spaces. Show that the equation

$$\langle (u_1, \ldots, u_m), (v_1, \ldots, v_m)\rangle = \langle u_1, v_1\rangle + \cdots + \langle u_m, v_m\rangle$$

defines an inner product on $V_1 \times \cdots \times V_m$.

In the expression above on the right, for each $k = 1, \ldots, m$, the inner product $\langle u_k, v_k\rangle$ denotes the inner product on V_k. Each of the spaces $V_1, \ldots, V_m$ may have a different inner product, even though the same notation is used here.

Solution. For convenience, let $\mathbf{V} = V_1 \times \cdots \times V_m$. We verify each property in definition 6.2.

Positivity. Let $(v_1, ..., v_m) \in \mathbf{V}$ be given and observe that

$$\langle (v_1, ..., v_m), (v_1, ..., v_m) \rangle = \langle v_1, v_1 \rangle + \cdots + \langle v_m, v_m \rangle.$$

Since each $\langle v_k, v_k \rangle$ is non-negative, it follows that $\langle (v_1, ..., v_m), (v_1, ..., v_m) \rangle$ is non-negative.

Definiteness. For $(v_1, ..., v_m) \in \mathbf{V}$, note that $(v_1, ..., v_m) = 0$ if and only if each $v_k = 0$. By the definiteness of the inner product on each V_k, this is the case if and only if each $\langle v_k, v_k \rangle = 0$. From the non-negativity of the expression

$$\langle (v_1, ..., v_m), (v_1, ..., v_m) \rangle = \langle v_1, v_1 \rangle + \cdots + \langle v_m, v_m \rangle,$$

we see that each $\langle v_k, v_k \rangle = 0$ if and only if $\langle (v_1, ..., v_m), (v_1, ..., v_m) \rangle = 0$.

Additivity in the first slot. Let $(u_1, ..., u_m), (v_1, ..., v_m), (w_1, ..., w_m) \in \mathbf{V}$ be given and observe that

$$\begin{aligned}
&\langle (u_1, ..., u_m) + (v_1, ..., v_m), (w_1, ..., w_m) \rangle \\
&= \langle (u_1 + v_1, ..., u_m + v_m), (w_1, ..., w_m) \rangle \\
&= \langle u_1 + v_1, w_1 \rangle + \cdots + \langle u_m + v_m, w_m \rangle \\
&= \langle u_1, w_1 \rangle + \langle v_1, w_1 \rangle + \cdots + \langle u_m, w_m \rangle + \langle v_m, w_m \rangle \\
&= \langle u_1, w_1 \rangle + \cdots + \langle u_m, w_m \rangle + \langle v_1, w_1 \rangle + \cdots + \langle v_m, w_m \rangle \\
&= \langle (u_1, ..., u_m), (w_1, ..., w_m) \rangle + \langle (v_1, ..., v_m), (w_1, ..., w_m) \rangle,
\end{aligned}$$

where we have used the additivity in the first slot of the inner product on each V_k.

Homogeneity in the first slot. Let $\lambda \in \mathbf{F}$ and $(u_1, ..., u_m), (v_1, ..., v_m) \in \mathbf{V}$ be given, and observe that

$$\begin{aligned}
\langle \lambda(u_1, ..., u_m), (v_1, ..., v_m) \rangle &= \langle (\lambda u_1, ..., \lambda u_m), (v_1, ..., v_m) \rangle \\
&= \langle \lambda u_1, v_1 \rangle + \cdots + \langle \lambda u_m, v_m \rangle \\
&= \lambda \langle u_1, v_1 \rangle + \cdots + \lambda \langle u_m, v_m \rangle \\
&= \lambda (\langle u_1, v_1 \rangle + \cdots + \langle u_m, v_m \rangle) \\
&= \lambda \langle (u_1, ..., u_m), (v_1, ..., v_m) \rangle,
\end{aligned}$$

where we have used homogeneity in the first slot of the inner product on each V_k.

Conjugate symmetry. Let $(u_1, ..., u_m), (v_1, ..., v_m) \in \mathbf{V}$ be given and observe that

$$\begin{aligned}
&\overline{\langle (v_1, ..., v_m), (u_1, ..., u_m) \rangle} = \overline{\langle v_1, u_1 \rangle + \cdots + \langle v_m, u_m \rangle} \\
&\quad = \overline{\langle v_1, u_1 \rangle} + \cdots + \overline{\langle v_m, u_m \rangle} = \langle u_1, v_1 \rangle + \cdots + \langle u_m, v_m \rangle = \langle (u_1, ..., u_m), (v_1, ..., v_m) \rangle,
\end{aligned}$$

where we have used the conjugate symmetry of the inner product on each V_k.

Exercise 6.A.30. Suppose V is a real inner product space. For $u, v, w, x \in V$, define

$$\langle u + iv, w + ix \rangle_{\mathbf{C}} = \langle u, w \rangle + \langle v, x \rangle + (\langle v, w \rangle - \langle u, x \rangle)i.$$

(a) Show that $\langle \cdot, \cdot \rangle_{\mathbf{C}}$ makes $V_{\mathbf{C}}$ into a complex inner product space.

(b) Show that if $u, v \in V$, then

$$\langle u, v \rangle_{\mathbf{C}} = \langle u, v \rangle \quad \text{and} \quad \|u + iv\|_{\mathbf{C}}^2 = \|u\|^2 + \|v\|^2.$$

See Exercise 8 in Section 1B for the definition of the complexification $V_{\mathbf{C}}$.

Solution.

(a) We verify each property in definition 6.2.

Positive-definiteness. For any $u + iv \in V_{\mathbf{C}}$,

$$\langle u + iv, u + iv \rangle_{\mathbf{C}} = \langle u, u \rangle + \langle v, v \rangle + (\langle v, u \rangle - \langle u, v \rangle)i = \langle u, u \rangle + \langle v, v \rangle,$$

where we have used the symmetry of $\langle \cdot, \cdot \rangle$. The positivity of $\langle \cdot, \cdot \rangle$ and the expression above gives us the positivity of $\langle \cdot, \cdot \rangle_{\mathbf{C}}$. Moreover,

$$\langle u + iv, u + iv \rangle_{\mathbf{C}} = 0 \quad \Leftrightarrow \quad \langle u, u \rangle = 0 \text{ and } \langle v, v \rangle = 0 \quad \Leftrightarrow \quad u = v = 0 \quad \Leftrightarrow \quad u + iv = 0.$$

Conjugate symmetry. For any $u + iv, w + ix \in V_{\mathbf{C}}$, observe that

$$\begin{aligned}\overline{\langle w + ix, u + iv \rangle_{\mathbf{C}}} &= \overline{\langle w, u \rangle + \langle x, v \rangle + (\langle x, u \rangle - \langle w, v \rangle)i} \\ &= \langle u, w \rangle + \langle v, x \rangle + (\langle v, w \rangle - \langle u, x \rangle)i = \langle u + iv, w + ix \rangle_{\mathbf{C}},\end{aligned}$$

where we have used the symmetry of $\langle \cdot, \cdot \rangle$.

Additivity in the first slot. Let $u + iv, w + ix, y + iz \in V_{\mathbf{C}}$ be given and observe that

$$\begin{aligned}\langle (u + iv) + (w + ix), y + iz \rangle_{\mathbf{C}} &= \langle (u + w) + i(v + x), y + iz \rangle_{\mathbf{C}} \\ &= \langle u + w, y \rangle + \langle v + x, z \rangle + (\langle v + x, y \rangle - \langle u + w, z \rangle)i \\ &= \langle u, y \rangle + \langle w, y \rangle + \langle v, z \rangle + \langle x, z \rangle \\ &\qquad + (\langle v, y \rangle + \langle x, y \rangle - \langle u, z \rangle - \langle w, z \rangle)i \\ &= \langle u, y \rangle + \langle v, z \rangle + (\langle v, y \rangle - \langle u, z \rangle)i \\ &\qquad + \langle w, y \rangle + \langle x, z \rangle + (\langle x, y \rangle - \langle w, z \rangle)i \\ &= \langle u + iv, y + iz \rangle_{\mathbf{C}} + \langle w + ix, y + iz \rangle_{\mathbf{C}},\end{aligned}$$

where we have used the additivity in the first slot of $\langle \cdot, \cdot \rangle$.

Homogeneity in the first slot. Let $u + iv, w + ix \in V_{\mathbf{C}}$ and $a + bi \in \mathbf{C}$ be given and observe that

$$
\begin{aligned}
\langle (a+bi)(u+iv), w+ix\rangle_{\mathbf{C}} &= \langle (au-bv)+i(av+bu), w+ix\rangle_{\mathbf{C}} \\
&= \langle au-bv, w\rangle + \langle av+bu, x\rangle \\
&\quad + (\langle av+bu, w\rangle - \langle au-bv, x\rangle)i \\
&= a\langle u,w\rangle - b\langle v,w\rangle + a\langle v,x\rangle + b\langle u,x\rangle \\
&\quad + (a\langle v,w\rangle + b\langle u,w\rangle - a\langle u,x\rangle + b\langle v,x\rangle)i \\
&= [a(\langle u,w\rangle + \langle v,x\rangle) - b(\langle v,w\rangle - \langle u,x\rangle)] \\
&\quad + [a(\langle v,w\rangle - \langle u,x\rangle) + b(\langle u,w\rangle + \langle v,x\rangle)]i \\
&= (a+bi)[\langle u,w\rangle + \langle v,x\rangle + (\langle v,w\rangle - \langle u,x\rangle)i] \\
&= (a+bi)\langle u+iv, w+ix\rangle_{\mathbf{C}},
\end{aligned}
$$

where we have used the homogeneity in the first slot of $\langle\cdot,\cdot\rangle$.

(b) For $u, v \in V$ we have

$$\langle u,v\rangle_{\mathbf{C}} = \langle u,v\rangle + \langle 0,0\rangle + (\langle 0,v\rangle - \langle u,0\rangle)i = \langle u,v\rangle.$$

Furthermore, as we showed in part (a) when we verified the positive-definiteness of $\langle\cdot,\cdot\rangle_{\mathbf{C}}$,

$$\|u+iv\|_{\mathbf{C}}^2 = \langle u+iv, u+iv\rangle_{\mathbf{C}} = \langle u,u\rangle + \langle v,v\rangle = \|u\|^2 + \|v\|^2.$$

Exercise 6.A.31. Suppose $u, v, w \in V$. Prove that

$$\left\|w - \tfrac{1}{2}(u+v)\right\|^2 = \frac{\|w-u\|^2 + \|w-v\|^2}{2} - \frac{\|u-v\|^2}{4}.$$

Solution. It will suffice to prove that

$$4\left\|w - \tfrac{1}{2}(u+v)\right\|^2 = 2\Big(\|w-u\|^2 + \|w-v\|^2\Big) - \|u-v\|^2,$$

which is equivalent to

$$\|2w-u-v\|^2 + \|u-v\|^2 = 2\Big(\|w-u\|^2 + \|w-v\|^2\Big),$$

which follows immediately from the parallelogram equality (6.21).

Exercise 6.A.32. Suppose that E is a subset of V with the property that $u, v \in E$ implies $\frac{1}{2}(u+v) \in E$. Let $w \in V$. Show that there is at most one point in E that is closest to w. In other words, show that there is at most one $u \in E$ such that

$$\|w-u\| \le \|w-x\|$$

for all $x \in E$.

Solution. Suppose $u, v \in E$ are both closest to w, so that $\|w-u\| = \|w-v\|$. It follows from Exercise 6.A.31 that

$$\frac{\|u-v\|^2}{4} = \|w-u\|^2 - \left\|w - \tfrac{1}{2}(u+v)\right\|^2.$$

Since u and v belong to E we must have $\frac{1}{2}(u+v) \in E$, and then since u is a point in E closest to w we must have $\|w-u\| \le \left\|w - \frac{1}{2}(u+v)\right\|$. It follows that $\|u-v\|^2 \le 0$, which is the case if and only if $u = v$.

Exercise 6.A.33. Suppose f, g are differentiable functions from $\mathbf{R}$ to $\mathbf{R}^n$.

(a) Show that

$$\langle f(t), g(t)\rangle' = \langle f'(t), g(t)\rangle + \langle f(t), g'(t)\rangle.$$

(b) Suppose c is a positive number and $\|f(t)\| = c$ for every $t \in \mathbf{R}$. Show that $\langle f'(t), f(t)\rangle = 0$ for every $t \in \mathbf{R}$.

(c) Interpret the result in (b) geometrically in terms of the tangent vector to a curve lying on a sphere in $\mathbf{R}^n$ centered at the origin.

A function $f : \mathbf{R} \to \mathbf{R}^n$ is called differentiable if there exist differentiable functions $f_1, \dots, f_n$ from $\mathbf{R}$ to $\mathbf{R}$ such that $f(t) = (f_1(t), \dots, f_n(t))$ for each $t \in \mathbf{R}$. Furthermore, for each $t \in \mathbf{R}$, the derivative $f'(t) \in \mathbf{R}^n$ is defined by $f'(t) = (f_1'(t), \dots, f_n'(t))$.

Solution.

(a) Suppose that $f(t) = (f_1(t), \dots, f_n(t))$ and $g(t) = (g_1(t), \dots, g_n(t))$ for some differentiable functions $f_1, \dots, f_n, g_1, \dots, g_n : \mathbf{R} \to \mathbf{R}$. By the usual rules of differentiation we have

$$\begin{aligned}
\langle f(t), g(t)\rangle' &= (f_1(t)g_1(t) + \dots + f_n(t)g_n(t))' \\
&= f_1'(t)g_1(t) + f_1(t)g_1'(t) + \dots + f_n'(t)g_n(t) + f_n(t)g_n'(t) \\
&= f_1'(t)g_1(t) + \dots + f_n'(t)g_n(t) + f_1(t)g_1'(t) + \dots + f_n(t)g_n'(t) \\
&= \langle f'(t), g(t)\rangle + \langle f(t), g'(t)\rangle.
\end{aligned}$$

(b) By part (a) we have

$$0 = (c^2)' = \left(\|f(t)\|^2\right)' = \langle f(t), f(t)\rangle' = 2\langle f'(t), f(t)\rangle.$$

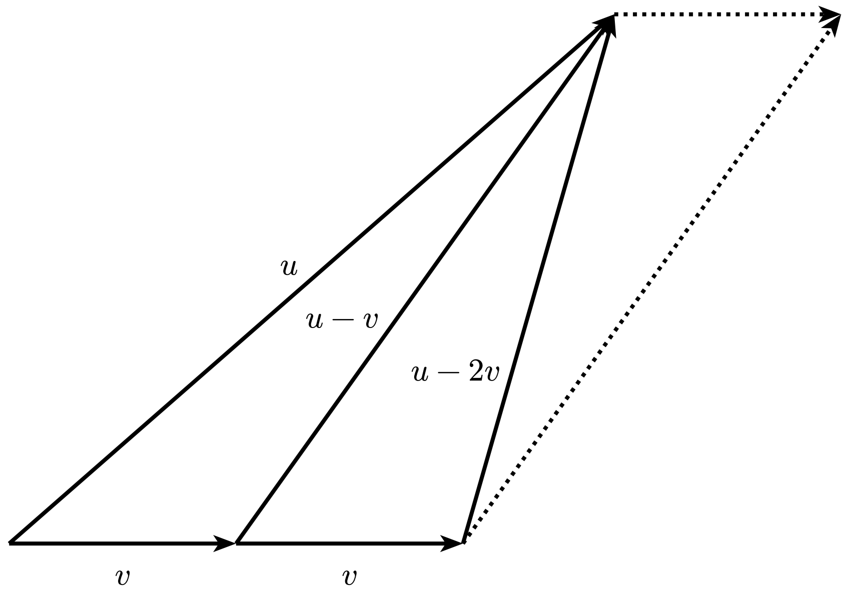
Thus $\langle f'(t), f(t)\rangle = 0$.

(c) Suppose $c > 0$. A differentiable function $f : \mathbf{R} \to \mathbf{R}^n$ satisfying $\|f(t)\| = c$ for every $t \in \mathbf{R}$ traces out a curve which lies on an $(n-1)$-sphere of radius c centered at the origin in $\mathbf{R}^n$. The tangent vector to this curve is given by f'; the result of part (b) states that this tangent vector is always orthogonal to the curve f.

Exercise 6.A.34. Use inner products to prove Apollonius's identity: In a triangle with sides of length a, b, and c, let d be the length of the line segment from the midpoint of the side of length c to the opposite vertex. Then

$$a^2 + b^2 = \tfrac{1}{2}c^2 + 2d^2.$$

Solution. Set up the triangle as follows.



Thus

$$\|u\| = a, \quad \|v\| = \tfrac{1}{2}c, \quad \|u - v\| = d, \quad \|u - 2v\| = b.$$

Consider the parallelogram formed by the vectors $u - v$ and v. The parallelogram equality states that

$$\|u\|^2 = \|u - 2v\|^2 = 2\|v\|^2 + 2\|u - v\|^2.$$

Substituting the given side lengths, we obtain

$$a^2 + b^2 = \tfrac{1}{2}c^2 + 2d^2.$$

Exercise 6.A.35. Fix a positive integer n. The *Laplacian* Δp of a twice differentiable real-valued function p on $\mathbf{R}^n$ is the function on $\mathbf{R}^n$ defined by

$$\Delta p = \frac{\partial^2 p}{\partial x_1^2} + \cdots + \frac{\partial^2 p}{\partial x_n^2}.$$

The function p is called *harmonic* if $\Delta p = 0$.

A *polynomial* on $\mathbf{R}^n$ is a linear combination (with coefficients in $\mathbf{R}$) of functions of the form $x_1^{m_1} \cdots x_n^{m_n}$, where $m_1, ..., m_n$ are nonnegative integers.

Suppose q is a polynomial on $\mathbf{R}^n$. Prove that there exists a harmonic polynomial p on $\mathbf{R}^n$ such that $p(x) = q(x)$ for every $x \in \mathbf{R}^n$ with $\|x\| = 1$.

> *The only fact about harmonic functions that you need for this exercise is that if p is a harmonic function on $\mathbf{R}^n$ and $p(x) = 0$ for all $x \in \mathbf{R}^n$ with $\|x\| = 1$, then $p = 0$.*
>
> *Hint: A reasonable guess is that the desired harmonic polynomial p is of the form $q + \left(1 - \|x\|^2\right)r$ for some polynomial r. Prove that there is a polynomial r on $\mathbf{R}^n$ such that $q + \left(1 - \|x\|^2\right)r$ is harmonic by defining an operator T on a suitable vector space by*
>
> $$Tr = \Delta\Big(\left(1 - \|x\|^2\right)r\Big)$$
>
> *and then showing that T is injective and hence surjective.*

Solution. Let us first provide a few definitions. A monomial on $\mathbf{R}^n$ is a polynomial on $\mathbf{R}^n$ of the form $x_1^{m_1} \cdots x_n^{m_n}$, where $m_1, ..., m_n$ are non-negative integers. The degree of such a monomial is the sum $m_1 + \cdots + m_n$. The degree of a non-zero polynomial p on $\mathbf{R}^n$ is the greatest degree amongst its monomial terms $x_1^{m_1} \cdots x_n^{m_n}$ and the degree of the zero polynomial is defined to be $-\infty$.

Define $\mathcal{P}_m^n(\mathbf{R})$ to be the collection of all polynomials on $\mathbf{R}^n$ of degree at most m and note that $\mathcal{P}_m^n(\mathbf{R})$ is a subset of the vector space $\mathbf{R}^{\mathbf{R}^n}$; in fact, size the zero polynomial is simply the zero function, and addition and scalar multiplication of polynomials of degree at most m will not result in a polynomial of degree greater than $m, \mathcal{P}_m^n(\mathbf{R})$ is a vector subspace of $\mathbf{R}^{\mathbf{R}^n}$.

It is straightforward to verify that the collection of all monomials of degree at most m forms a basis of $\mathcal{P}_m^n(\mathbf{R})$. This collection is finite and thus $\mathcal{P}_m^n(\mathbf{R})$ is a finite-dimensional vector space.

Clearly, the Laplacian Δp of a polynomial p on $\mathbf{R}^n$ is itself a polynomial and furthermore satisfies either $\deg \Delta p = -\infty$ or $\deg \Delta p = \deg p - 2$. Thus the Laplacian defines an operator $\Delta \in \mathcal{L}(\mathcal{P}_m^n(\mathbf{R}))$; the linearity of Δ follows from the linearity of partial differentiation.

The function

$$x = (x_1, ..., x_n) \in \mathbf{R}^n \mapsto \|x\|^2 = x_1^2 + \cdots + x_n^2$$

is a polynomial on $\mathbf{R}^n$ of degree 2. Given this, if r is a non-zero polynomial on $\mathbf{R}^n$, then $\left(1 - \|x\|^2\right)r$ is also a polynomial on $\mathbf{R}^n$ of degree $r + 2$; if r is the zero polynomial then so is $\left(1 - \|x\|^2\right)r$. It follows that $\Delta\left(1 - \|x\|^2\right)r$ is a polynomial on $\mathbf{R}^n$ of degree at most $\deg r$.

Let q be a polynomial on $\mathbf{R}^n$ and let $m = \deg q$. By our previous discussion, the operator $T \in \mathcal{L}(\mathcal{P}_m^n(\mathbf{R}))$ given by

$$T(r) = \Delta\left(\left(1 - \|x\|^2\right)r\right)$$

is well-defined; the linearity of T follows from the linearity of Δ and distributivity on $\mathbf{R}$.

We claim that T is injective. If $T(r) = 0$ for some $r \in \mathcal{P}_m^n(\mathbf{R})$ then $\left(1 - \|x\|^2\right)r$ is a harmonic polynomial on $\mathbf{R}^n$ which satisfies $\left(1 - \|x\|^2\right)r = 0$ for all $x \in \mathbf{R}^n$ such that $\|x\| = 1$; the fact about harmonic functions given in the exercise then implies that $\left(1 - \|x\|^2\right)r = 0$. It follows that r is identically zero on the open set $\{x \in \mathbf{R}^n : \|x\| = 1\}^{\mathrm{c}}$ and hence that r is the zero polynomial. Thus $\operatorname{null} T = \{0\}$, i.e. T is injective.

By 3.65 T must be surjective. Hence there exists some $r \in \mathcal{P}_m^n(\mathbf{R})$ such that $T(r) = \Delta(-q)$, which by linearity is equivalent to

$$\Delta\left(q + \left(1 - \|x\|^2\right)r\right) = 0.$$

Thus $p = q + \left(1 - \|x\|^2\right)r$ is a harmonic polynomial on $\mathbf{R}^n$ which satisfies $p(x) = q(x)$ for all $x \in \mathbf{R}^n$ with $\|x\| = 1$.

6.B. Orthonormal Bases

Exercise 6.B.1. Suppose $e_1, \ldots, e_m$ is a list of vectors in V such that

$$\|a_1 e_1 + \cdots + a_m e_m\|^2 = |a_1|^2 + \cdots + |a_m|^2$$

for all $a_1, \ldots, a_m \in \mathbf{F}$. Show that $e_1, \ldots, e_m$ is an orthonormal list.

This exercise provides a converse to 6.24.

Solution. For each $k \in \{1, \ldots, m\}$, taking each of $a_1, \ldots, a_m$ to be 0 except $a_k = 1$ shows that $\|e_k\| = 1$. Suppose $j, k \in \{1, \ldots, m\}$ are such that $j \neq k$ and let $a \in \mathbf{F}$ be given. Observe that

$$\|e_j\|^2 = 1 \leq 1 + |a|^2 = \|e_j + a e_k\|^2 \quad \Rightarrow \quad \|e_j\| \leq \|e_j + a e_k\|.$$

It follows from Exercise 6.A.6 that $\langle e_j, e_k \rangle = 0$. Thus $e_1, \ldots, e_m$ is an orthonormal list.

Exercise 6.B.2.

(a) Suppose $\theta \in \mathbf{R}$. Show that both

$$(\cos\theta, \sin\theta), (-\sin\theta, \cos\theta) \quad \text{and} \quad (\cos\theta, \sin\theta), (\sin\theta, -\cos\theta)$$

are orthonormal bases of $\mathbf{R}^2$.

(b) Show that each orthonormal basis of $\mathbf{R}^2$ is of the form given by one of the two possibilities in (a).

Solution.

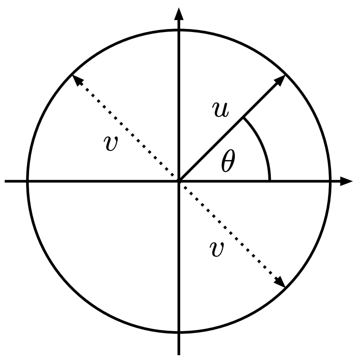
(a) Observe that

$$\langle(\cos\theta, \sin\theta), (\cos\theta, \sin\theta)\rangle = \langle(-\sin\theta, \cos\theta), (-\sin\theta, \cos\theta)\rangle = \cos^2\theta + \sin^2\theta = 1,$$

$$\langle(\cos\theta, \sin\theta), (-\sin\theta, \cos\theta)\rangle = \cos\theta\sin\theta - \cos\theta\sin\theta = 0.$$

Thus $(\cos\theta, \sin\theta), (-\sin\theta, \cos\theta)$ is an orthonormal basis of $\mathbf{R}^2$. A similar calculation shows that $(\cos\theta, \sin\theta), (\sin\theta, -\cos\theta)$ is also an orthonormal basis of $\mathbf{R}^2$.

(b) Suppose u, v is an orthonormal basis of $\mathbf{R}^2$. Let θ be the angle that u makes with the positive x-axis, as shown below.



Note that $\|u\| = 1$, so that u lies on the circle of radius 1 centered at the origin in $\mathbf{R}^2$. It follows that $u = (\cos\theta, \sin\theta)$.

Since u and v are orthogonal, plane geometry tells us that v either makes an angle of $\theta + \frac{\pi}{2}$ or $\theta - \frac{\pi}{2}$ with the positive x-axis, and since $\|v\| = 1$ we know that v also lies on the circle of radius 1 centered at the origin. It follows that

$$v = \big(\cos(\theta + \tfrac{\pi}{2}), \sin(\theta + \tfrac{\pi}{2})\big) = (-\sin\theta, \cos\theta)$$
$$\text{or} \quad v = \big(\cos(\theta - \tfrac{\pi}{2}), \sin(\theta - \tfrac{\pi}{2})\big) = (\sin\theta, -\cos\theta).$$

Exercise 6.B.3. Suppose $e_1, ..., e_m$ is an orthonormal list in V and $v \in V$. Prove that

$$\|v\|^2 = |\langle v, e_1\rangle|^2 + \cdots + |\langle v, e_m\rangle|^2 \quad \Leftrightarrow \quad v \in \operatorname{span}(e_1, ..., e_m).$$

Solution. Suppose $v \in \operatorname{span}(e_1, ..., e_m)$. Note that $e_1, ..., e_m$ is linearly independent by 6.25 and hence is a basis of $\operatorname{span}(e_1, ..., e_m)$. It follows from 6.30 that

$$\|v\|^2 = |\langle v, e_1\rangle|^2 + \cdots + |\langle v, e_m\rangle|^2.$$

Now suppose that $\|v\|^2 = |\langle v, e_1\rangle|^2 + \cdots + |\langle v, e_m\rangle|^2$, let $u = \langle v, e_1\rangle e_1 + \cdots + \langle v, e_m\rangle e_m$, and note that $\|u\|^2 = \|v\|^2$ by 6.24. Observe that

$$\langle u, v\rangle = \Big\langle \textstyle\sum_{k=1}^m \langle v, e_k\rangle e_k, v\Big\rangle = \displaystyle\sum_{k=1}^m \langle \langle v, e_k\rangle e_k, v\rangle = \sum_{k=1}^m |\langle v, e_k\rangle|^2 = \|v\|^2.$$

Thus

$$\|u - v\|^2 = \|u\|^2 + \|v\|^2 - 2\operatorname{Re}\langle u, v\rangle = 2\|v\|^2 - 2\|v\|^2 = 0,$$

from which it follows that $v = u \in \operatorname{span}(e_1, ..., e_m)$.

Exercise 6.B.4. Suppose n is a positive integer. Prove that

$$\frac{1}{\sqrt{2\pi}}, \frac{\cos x}{\sqrt{\pi}}, \frac{\cos 2x}{\sqrt{\pi}}, \dots, \frac{\cos nx}{\sqrt{\pi}}, \frac{\sin x}{\sqrt{\pi}}, \frac{\sin 2x}{\sqrt{\pi}}, \dots, \frac{\sin nx}{\sqrt{\pi}}$$

is an orthonormal list of vectors in $C[-\pi, \pi]$, the vector space of continuous real-valued functions on $[-\pi, \pi]$ with inner product

$$\langle f, g \rangle = \int_{-\pi}^{\pi} fg.$$

Hint: The following formulas should help.

$$(\sin x)(\cos y) = \frac{\sin(x-y) + \sin(x+y)}{2}$$

$$(\sin x)(\sin y) = \frac{\cos(x-y) - \cos(x+y)}{2}$$

$$(\cos x)(\cos y) = \frac{\cos(x-y) + \cos(x+y)}{2}$$

Solution. We calculate, for $k \in \{1, \dots, n\}$,

$$\left\langle \frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{2\pi}} \right\rangle = \int_{-\pi}^{\pi} \left(\frac{1}{\sqrt{2\pi}} \right)^2 \mathrm{d}x = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathrm{d}x = 1;$$

$$\left\langle \frac{\cos kx}{\sqrt{\pi}}, \frac{\cos kx}{\sqrt{\pi}} \right\rangle = \int_{-\pi}^{\pi} \left(\frac{\cos kx}{\sqrt{\pi}} \right)^2 \mathrm{d}x = \frac{2}{k\pi} \int_{0}^{k\pi} \cos^2 y \, \mathrm{d}y = \frac{1}{k\pi} [y + \sin y \cos y]_{y=0}^{y=k\pi} = 1;$$

$$\left\langle \frac{\sin kx}{\sqrt{\pi}}, \frac{\sin kx}{\sqrt{\pi}} \right\rangle = \int_{-\pi}^{\pi} \left(\frac{\sin kx}{\sqrt{\pi}} \right)^2 \mathrm{d}x = \frac{2}{k\pi} \int_{0}^{k\pi} \sin^2 y \, \mathrm{d}y = \frac{1}{k\pi} [y - \sin y \cos y]_{y=0}^{y=k\pi} = 1;$$

$$\left\langle \frac{1}{\sqrt{2\pi}}, \frac{\cos kx}{\sqrt{\pi}} \right\rangle = \int_{-\pi}^{\pi} \frac{1}{\sqrt{2\pi}} \frac{\cos kx}{\sqrt{\pi}} \, \mathrm{d}x = \frac{\sqrt{2}}{k\pi} \int_{0}^{k\pi} \cos y \, \mathrm{d}y = \frac{\sqrt{2}}{k\pi} [\sin y]_{y=0}^{y=k\pi} = 0;$$

$$\left\langle \frac{1}{\sqrt{2\pi}}, \frac{\sin kx}{\sqrt{\pi}} \right\rangle = \int_{-\pi}^{\pi} \frac{1}{\sqrt{2\pi}} \frac{\sin kx}{\sqrt{\pi}} \, \mathrm{d}x = 0,$$

where we have used that $\sin kx$ is an odd function for the last equality. For $j, k \in \{1, \dots, n\}$ such that $j \neq k$, we have

$$\left\langle \frac{\cos jx}{\sqrt{\pi}}, \frac{\cos kx}{\sqrt{\pi}} \right\rangle = \int_{-\pi}^{\pi} \frac{\cos jx}{\sqrt{\pi}} \frac{\cos kx}{\sqrt{\pi}} \, \mathrm{d}x = \frac{1}{\pi} \int_{0}^{\pi} \cos((j-k)x) + \cos((j+k)x) \, \mathrm{d}x$$

$$= \frac{1}{\pi} \left[\frac{\sin((j-k)x)}{j-k} \right]_{x=0}^{x=\pi} + \frac{1}{\pi} \left[\frac{\sin((j+k)x)}{j+k} \right]_{x=0}^{x=\pi} = 0;$$

$$\left\langle \frac{\sin jx}{\sqrt{\pi}}, \frac{\sin kx}{\sqrt{\pi}} \right\rangle = \int_{-\pi}^{\pi} \frac{\sin jx}{\sqrt{\pi}} \frac{\sin kx}{\sqrt{\pi}} \, \mathrm{d}x = \frac{1}{\pi} \int_{0}^{\pi} \cos((j-k)x) - \cos((j+k)x) \, \mathrm{d}x$$

$$= \frac{1}{\pi} \left[\frac{\sin((j-k)x)}{j-k} \right]_{x=0}^{x=\pi} - \frac{1}{\pi} \left[\frac{\sin((j+k)x)}{j+k} \right]_{x=0}^{x=\pi} = 0.$$

Finally, for any $j, k \in \{1, ..., n\}$ we have

$$\left\langle \frac{\cos jx}{\sqrt{\pi}}, \frac{\sin kx}{\sqrt{\pi}} \right\rangle = \int_{-\pi}^{\pi} \frac{\cos jx}{\sqrt{\pi}} \frac{\sin kx}{\sqrt{\pi}} \, \mathrm{d}x = 0,$$

where we have used that $\cos jx \sin kx$ is an odd function for the last equality.

Exercise 6.B.5. Suppose $f : [-\pi, \pi] \to \mathbf{R}$ is continuous. For each nonnegative integer k, define

$$a_k = \frac{1}{\sqrt{\pi}} \int_{-\pi}^{\pi} f(x) \cos(kx) \, dx \quad \text{and} \quad b_k = \frac{1}{\sqrt{\pi}} \int_{-\pi}^{\pi} f(x) \sin(kx) \, dx.$$

Prove that

$$\frac{a_0^2}{2} + \sum_{k=1}^{\infty} (a_k^2 + b_k^2) \leq \int_{-\pi}^{\pi} f^2.$$

The inequality above is actually an equality for all continuous functions $f : [-\pi, \pi] \to \mathbf{R}$. However, proving that this inequality is an equality involves Fourier series techniques beyond the scope of this book.

Solution. Consider $C[-\pi, \pi]$, the vector space of continuous real-valued functions on $[-\pi, \pi]$, with inner product

$$\langle f, g \rangle = \int_{-\pi}^{\pi} fg.$$

As we showed in Exercise 6.B.5,

$$\frac{1}{\sqrt{2\pi}}, \frac{\cos x}{\sqrt{\pi}}, \frac{\cos 2x}{\sqrt{\pi}}, \ldots, \frac{\cos nx}{\sqrt{\pi}}, \frac{\sin x}{\sqrt{\pi}}, \frac{\sin 2x}{\sqrt{\pi}}, \ldots, \frac{\sin nx}{\sqrt{\pi}}$$

is an orthonormal list of vectors in $C[-\pi, \pi]$ for any $n \geq 1$. Observe that, for $k \geq 1$,

$$a_0 = \frac{1}{\sqrt{\pi}} \int_{-\pi}^{\pi} f(x) \, \mathrm{d}x = \sqrt{2} \left\langle f, \frac{1}{\sqrt{2\pi}} \right\rangle;$$

$$a_k = \frac{1}{\sqrt{\pi}} \int_{-\pi}^{\pi} f(x) \cos(kx) \, \mathrm{d}x = \left\langle f, \frac{\cos kx}{\sqrt{\pi}} \right\rangle;$$

$$b_k = \frac{1}{\sqrt{\pi}} \int_{-\pi}^{\pi} f(x) \sin(kx) \, \mathrm{d}x = \left\langle f, \frac{\sin kx}{\sqrt{\pi}} \right\rangle.$$

Thus, by Bessel's inequality (6.26),

$$\frac{a_0^2}{2} + \sum_{k=1}^{n}(a_k^2 + b_k^2) \leq \|f\|^2 = \int_{-\pi}^{\pi} f^2.$$

for all $n \geq 1$. It then follows from the monotone convergence theorem that $\sum_{k=1}^{\infty}(a_k^2 + b_k^2)$ is a convergent series. Furthermore,

$$\frac{a_0^2}{2} + \sum_{k=1}^{\infty}(a_k^2 + b_k^2) \leq \int_{-\pi}^{\pi} f^2.$$

Exercise 6.B.6. Suppose $e_1, \ldots, e_n$ is an orthonormal basis of V.

(a) Prove that if $v_1, \ldots, v_n$ are vectors in V such that

$$\|e_k - v_k\| < \frac{1}{\sqrt{n}}$$

for each k, then $v_1, \ldots, v_n$ is a basis of V.

(b) Show that there exist $v_1, \ldots, v_n \in V$ such that

$$\|e_k - v_k\| \leq \frac{1}{\sqrt{n}}$$

for each k, but $v_1, \ldots, v_n$ is not linearly independent.

This exercise states in (a) that an appropriately small perturbation of an orthonormal basis is a basis. Then (b) shows that the number $1/\sqrt{n}$ on the right side of the inequality in (a) cannot be improved upon.

Solution.

(a) It will suffice to show that $v_1, \ldots, v_n$ is linearly independent, so suppose that $\sum_{k=1}^{n} a_k v_k = 0$ and observe that

$$\begin{aligned}
\textstyle\sum_{k=1}^{n} |a_k|^2 &= \left\|\textstyle\sum_{k=1}^{n} a_k e_k\right\|^2 && (6.24)\\
&= \left\|\textstyle\sum_{k=1}^{n} a_k (e_k - v_k)\right\|^2 && \left(\textstyle\sum_{k=1}^{n} a_k v_k = 0\right)\\
&\leq \left(\textstyle\sum_{k=1}^{n} |a_k| \|e_k - v_k\|\right)^2 && \text{(triangle inequality)}\\
&\leq \left(\textstyle\sum_{k=1}^{n} |a_k|^2\right)\left(\textstyle\sum_{k=1}^{n} \|e_k - v_k\|^2\right). && \text{(Cauchy-Schwarz inequality)}
\end{aligned}$$

By assumption we have $\sum_{k=1}^{n} \|e_k - v_k\|^2 < 1$. It follows that $\sum_{k=1}^{n} |a_k|^2 = 0$, which is the case if and only if each $a_k = 0$. Thus $v_1, \ldots, v_n$ is linearly independent.

(b) For each $k \in \{1, \ldots, n\}$ let

$$v_k = e_k - \frac{e_1 + \cdots + e_n}{n} = \left(-\frac{1}{n}\right)e_1 + \cdots + \left(1 - \frac{1}{n}\right)e_k + \cdots + \left(-\frac{1}{n}\right)e_n.$$

Notice that $v_1 + \cdots + v_n = 0$, so that $v_1, ..., v_n$ is linearly dependent. Furthermore, for each $k \in \{1, ..., n\}$, 6.24 shows that

$$\|e_k - v_k\| = \left\| \frac{1}{n} e_1 + \cdots + \frac{1}{n} e_k + \cdots + \frac{1}{n} e_n \right\| = \frac{1}{n} \|e_1 + \cdots + e_n\| = \frac{1}{n} \cdot \sqrt{n} = \frac{1}{\sqrt{n}}.$$

Exercise 6.B.7. Suppose $T \in \mathcal{L}(\mathbf{R}^3)$ has an upper-triangular matrix with respect to the basis $(1,0,0), (1,1,1), (1,1,2)$. Find an orthonormal basis of $\mathbf{R}^3$ with respect to which T has an upper-triangular matrix.

Solution. Performing the Gram-Schmidt procedure on the basis $(1,0,0), (1,1,1), (1,1,2)$ yields the orthonormal basis

$$(1,0,0), \quad \left(0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right), \quad \left(0, -\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right).$$

As the proof of 6.37 shows, the matrix of T with respect to this orthonormal basis must also be upper-triangular.

Exercise 6.B.8. Make $\mathcal{P}_2(\mathbf{R})$ into an inner product space by defining $\langle p, q \rangle = \int_0^1 pq$ for all $p, q \in \mathcal{P}_2(\mathbf{R})$.

(a) Apply the Gram-Schmidt procedure to the basis $1, x, x^2$ to produce an orthonormal basis of $\mathcal{P}_2(\mathbf{R})$.

(b) The differentiation operator (the operator that takes p to p') on $\mathcal{P}_2(\mathbf{R})$ has an upper-triangular matrix with respect to the basis $1, x, x^2$, which is not an orthonormal basis. Find the matrix of the differentiation operator on $\mathcal{P}_2(\mathbf{R})$ with respect to the orthonormal basis produced in (a) and verify that this matrix is upper triangular, as expected from the proof of 6.37.

Solution.

(a) By applying the Gram-Schmidt procedure we obtain the orthonormal basis

$$1, \quad 2\sqrt{3}\left(x - \tfrac{1}{2}\right), \quad 6\sqrt{5}\left(x^2 - x + \tfrac{1}{6}\right).$$

(b) Some routine calculations reveal that the matrix of T with respect to the orthonormal basis found in part (a) is

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 2\sqrt{3} & 0 \\ 0 & 0 & 6\sqrt{\frac{5}{3}} \end{pmatrix},$$

which is indeed upper-triangular.

Exercise 6.B.9. Suppose $e_1, ..., e_m$ is the result of applying the Gram-Schmidt procedure to a linearly independent list $v_1, ..., v_m \in V$. Prove that $\langle v_k, e_k \rangle > 0$ for each $k = 1, ..., m$.

Solution. Let $u_1, ..., u_m$ be a list in V and let $f_1, ..., f_m$ be an orthonormal list in V with the property that $\text{span}(f_1, ..., f_k) = \text{span}(u_1, ..., u_k)$ for each $k \in \{1, ..., m\}$. We will prove that

$$\langle u_k, f_k \rangle = 0 \text{ for some } k \in \{1, ..., m\} \quad \Rightarrow \quad u_1, ..., u_m \text{ is linearly dependent.}$$

The contrapositive of this implication will give us the desired result.

Suppose that there exists some $k \in \{1, ..., m\}$ such that $\langle u_k, f_k \rangle = 0$. Notice that $u_k \in \text{span}(u_1, ..., u_k) = \text{span}(f_1, ..., f_k)$; it follows from Exercise 6.B.3 that

$$\|u_k\|^2 = |\langle u_k, f_1 \rangle|^2 + \cdots + |\langle u_k, f_{k-1} \rangle|^2 + |\langle u_k, f_k \rangle|^2 = |\langle u_k, f_1 \rangle|^2 + \cdots + |\langle u_k, f_{k-1} \rangle|^2.$$

Another application of Exercise 6.B.3 shows that $u_k \in \text{span}(f_1, ..., f_{k-1}) = \text{span}(u_1, ..., u_{k-1})$. Thus $u_1, ..., u_m$ is linearly dependent.

Exercise 6.B.10. Suppose $v_1, ..., v_m$ is a linearly independent list in V. Explain why the orthonormal list produced by the formulas of the Gram-Schmidt procedure (6.32) is the only orthonormal list $e_1, ..., e_m$ in V such that $\langle v_k, e_k \rangle > 0$ and $\text{span}(v_1, ..., v_k) = \text{span}(e_1, ..., e_k)$ for each $k = 1, ..., m$.

The result in this exercise is used in the proof of 7.58.

Solution. Here is a useful lemma.

Lemma L.10. Suppose $v_1, \dots, v_m$ is a linearly independent list in V and let $e_1, \dots, e_m$ be the orthonormal list obtained by applying the Gram-Schmidt procedure to $v_1, \dots, v_m$. Let $S = \{\lambda \in \mathbf{F} : |\lambda| = 1\}$ (if $\mathbf{F} = \mathbf{R}$ then $S = \{-1, 1\}$ and if $\mathbf{F} = \mathbf{C}$ then S is the unit circle in the complex plane) and let S^m be the collection of functions $\{1, \dots, m\} \to S$. The orthonormal lists $u_1, \dots, u_m$ satisfying $\operatorname{span}(u_1, \dots, u_k) = \operatorname{span}(v_1, \dots, v_k)$ for each $k \in \{1, \dots, m\}$ are exactly those of the form $f(1)e_1, \dots, f(m)e_m$ for some $f \in S^m$.

Proof. Let $f \in S^m$ be given and suppose $j, k \in \{1, \dots, m\}$ are such that $j \neq k$. Observe that

$$\|f(k)e_k\| = |f(k)|\|e_k\| = 1 \quad \text{and} \quad \langle f(j)e_j, f(k)e_k \rangle = f(j)\overline{f(k)}\langle e_j, e_k \rangle = 0.$$

Furthermore, since $0 \notin S$ we have for each $k \in \{1, \dots, m\}$,

$$\operatorname{span}(f(1)e_1, \dots, f(k)e_k) = \operatorname{span}(e_1, \dots, e_k) = \operatorname{span}(v_1, \dots, v_k).$$

Thus $f(1)e_1, \dots, f(m)e_m$ is an orthonormal list satisfying

$$\operatorname{span}(f(1)e_1, \dots, f(k)e_k) = \operatorname{span}(v_1, \dots, v_k)$$

for each $k \in \{1, \dots, m\}$.

Now suppose that $u_1, \dots, u_m$ is an orthonormal list satisfying

$$\operatorname{span}(u_1, \dots, u_k) = \operatorname{span}(v_1, \dots, v_k) = \operatorname{span}(e_1, \dots, e_k)$$

for each $k \in \{1, \dots, m\}$. In particular $\operatorname{span}(e_1) = \operatorname{span}(u_1)$, from which it follows that $u_1 = \lambda_1 e_1$ for some $\lambda_1 \in \mathbf{F}$. Because $\|u_1\| = \|e_1\| = 1$ we see that $|\lambda_1| = 1$, so that $\lambda_1 \in S$; let $f(1) = \lambda_1$.

Given that $\operatorname{span}(e_1, e_2) = \operatorname{span}(u_1, u_2)$, notice that e_1, e_2 is an orthonormal basis of $\operatorname{span}(e_1, e_2)$ and that $u_2 \in \operatorname{span}(e_1, e_2)$. It follows from 6.30 that

$$u_2 = \langle u_2, e_1 \rangle e_1 + \langle u_2, e_2 \rangle e_2.$$

The orthonormality of the list u_1, u_2 shows that

$$0 = \langle u_1, u_2 \rangle = \langle f(1)e_1, u_2 \rangle = f(1)\langle e_1, u_2 \rangle,$$

which implies $\langle e_1, u_2 \rangle = 0$ since $f(1) \neq 0$. Thus $u_2 = \lambda_2 e_2$ for some $\lambda \in \mathbf{F}$. Because $\|u_2\| = \|e_2\| = 1$ we see that $|\lambda_2| = 1$, so that $\lambda_2 \in S$; let $f(2) = \lambda_2$.

By continuing in this manner we obtain an $f \in S^m$ such that $u_k = f(k)e_k$ for each $k \in \{1, \dots, m\}$. $\square$

Returning to the exercise, 6.32 and Exercise 6.B.9 show that the orthonormal list $e_1, \dots, e_m$ produced by the Gram-Schmidt procedure indeed satisfies $\langle v_k, e_k \rangle > 0$ and

$$\operatorname{span}(v_1, \dots, v_k) = \operatorname{span}(e_1, \dots, e_k)$$

for each $k \in \{1, \dots, m\}$.

Conversely, suppose that $u_1, \ldots, u_m$ is an orthonormal list in V such that $\langle v_k, u_k \rangle > 0$ and

$$\operatorname{span}(v_1, \ldots, v_k) = \operatorname{span}(u_1, \ldots, u_k)$$

for each $k \in \{1, \ldots, m\}$. Lemma L.10 shows that $u_1, \ldots, u_m$ is of the form $f(1)e_1, \ldots, f(m)e_m$ for some $f : \{1, \ldots, m\} \to \mathbf{F}$ satisfying $|f(k)| = 1$ for each $k \in \{1, \ldots, m\}$. For such a k, observe that

$$\langle v_k, u_k \rangle > 0 \quad \Leftrightarrow \quad f(k)\langle v_k, e_k \rangle > 0.$$

Since $\langle v_k, e_k \rangle > 0$ by Exercise 6.B.9, $f(k)$ must be a positive real number. Combining this with $|f(k)| = 1$, we see that $f(k) = 1$ for each $k \in \{1, \ldots, m\}$. Thus $u_1, \ldots, u_m$ is nothing but $e_1, \ldots, e_m$.

Exercise 6.B.11. Find a polynomial $q \in \mathcal{P}_2(\mathbf{R})$ such that $p\left(\frac{1}{2}\right) = \int_0^1 pq$ for every $p \in \mathcal{P}_2(\mathbf{R})$.

Solution. Equip $\mathcal{P}_2(\mathbf{R})$ with the inner product $\langle p, q \rangle = \int_0^1 pq$ and define $\varphi \in (\mathcal{P}_2(\mathbf{R}))'$ by $\varphi(p) = p\left(\frac{1}{2}\right)$. As the proof of 6.42 shows, if we take

$$q = \varphi(e_1)e_1 + \varphi(e_2)e_2 + \varphi(e_3)e_3 = -15x^2 + 15x - \tfrac{3}{2},$$

where e_1, e_2, e_3 is the orthonormal basis of $\mathcal{P}_2(\mathbf{R})$ found in Exercise 6.B.8 (a), then

$$\varphi(p) = p\left(\tfrac{1}{2}\right) = \langle p, q \rangle = \int_0^1 pq$$

for every $p \in \mathcal{P}_2(\mathbf{R})$.

Exercise 6.B.12. Find a polynomial $q \in \mathcal{P}_2(\mathbf{R})$ such that

$$\int_0^1 p(x)\cos(\pi x)\,dx = \int_0^1 pq$$

for every $p \in \mathcal{P}_2(\mathbf{R})$.

Solution. Equip $\mathcal{P}_2(\mathbf{R})$ with the inner product $\langle p, q \rangle = \int_0^1 pq$ and define $\varphi \in (\mathcal{P}_2(\mathbf{R}))'$ by $\varphi(p) = \int_0^1 p(x)\cos(\pi x)\,\mathrm{d}x$. As the proof of 6.42 shows, if we take

$$q = \varphi(e_1)e_1 + \varphi(e_2)e_2 + \varphi(e_3)e_3 = -\frac{24}{\pi^2}\left(x - \tfrac{1}{2}\right),$$

where e_1, e_2, e_3 is the orthonormal basis of $\mathcal{P}_2(\mathbf{R})$ found in Exercise 6.B.8 (a), then

$$\varphi(p) = \int_0^1 p(x)\cos(\pi x)\,\mathrm{d}x = \langle p, q \rangle = \int_0^1 pq$$

for every $p \in \mathcal{P}_2(\mathbf{R})$.

Exercise 6.B.13. Show that a list $v_1, \ldots, v_m$ of vectors in V is linearly dependent if and only if the Gram-Schmidt formula in 6.32 produces $f_k = 0$ for some $k \in \{1, \ldots, m\}$.

This exercise gives an alternative to Gaussian elimination techniques for determining whether a list of vectors in an inner product space is linearly dependent.

Solution. If $v_1, \ldots, v_m$ is linearly independent then each f_k must be non-zero, as the proof of 6.32 shows.

Suppose that $v_1, \ldots, v_m$ is linearly dependent. If $v_1 = 0$ then $f_1 = 0$; otherwise, let $k \in \{2, \ldots, m\}$ be the least integer such that $v_k \in \operatorname{span}(v_1, \ldots, v_{k-1})$ and note that $v_1, \ldots, v_{k-1}$ is linearly independent. This linear independence allows us to construct $f_1, \ldots, f_{k-1}$ as in 6.32 so that:

- each $f_i \neq 0$;
- $\operatorname{span}(v_1, \ldots, v_i) = \operatorname{span}(f_1, \ldots, f_i)$ for each $i \in \{1, \ldots, k-1\}$;
- $f_1, \ldots, f_{i-1}$ is pairwise orthogonal.

It follows that $v_k \in \operatorname{span}(f_1, \ldots, f_{k-1})$, say $v_k = a_1 f_1 + \cdots + a_{k-1} f_{k-1}$. Notice that, for each $i \in \{1, \ldots, k-1\}$,

$$\langle v_k, f_i \rangle = \langle a_i f_i, f_i \rangle = a_i \|f_i\|^2 \quad \Rightarrow \quad a_i = \frac{\langle v_k, f_i \rangle}{\|f_i\|^2}.$$

Thus, using the formula for f_k in 6.32,

$$f_k = v_k - \sum_{i=1}^{k-1} \frac{\langle v_k, f_i \rangle}{\|f_i\|^2} f_i = v_k - \sum_{i=1}^{k-1} a_i f_i = v_k - v_k = 0.$$

Exercise 6.B.14. Suppose V is a real inner product space and $v_1, \ldots, v_m$ is a linearly independent list of vectors in V. Prove that there exist exactly 2^m orthonormal lists $e_1, \ldots, e_m$ of vectors in V such that

$$\operatorname{span}(v_1, \ldots, v_k) = \operatorname{span}(e_1, \ldots, e_k)$$

for each $k \in \{1, \ldots, m\}$.

Solution. Let $e_1, \ldots, e_m$ be the orthonormal list obtained by applying the Gram-Schmidt procedure to $v_1, \ldots, v_m$. Lemma L.10 shows that the orthonormal lists $u_1, \ldots, u_m$ satisfying $\operatorname{span}(v_1, \ldots, v_k) = \operatorname{span}(u_1, \ldots, u_k)$ for each $k \in \{1, \ldots, m\}$ are precisely those of the form

$$f(1)e_1, \ldots, f(m)e_m$$

for some $f : \{1, \ldots, m\} \to \{-1, 1\}$. It is straightforward to verify that there are 2^m such functions, and that each $f : \{1, \ldots, m\} \to \{-1, 1\}$ gives a distinct orthonormal list $f(1)e_1, \ldots, f(m)e_m$. Thus there 2^m orthonormal lists $u_1, \ldots, u_m$ of vectors in V such that

$$\operatorname{span}(v_1, \ldots, v_k) = \operatorname{span}(u_1, \ldots, u_k)$$

for each $k \in \{1, ..., m\}$.

Exercise 6.B.15. Suppose $\langle \cdot, \cdot \rangle_1$ and $\langle \cdot, \cdot \rangle_2$ are inner products on V such that $\langle u, v \rangle_1 = 0$ if and only if $\langle u, v \rangle_2 = 0$. Prove that there is a positive number c such that $\langle u, v \rangle_1 = \langle u, v \rangle_2$ for every $u, v \in V$.

This exercise shows that if two inner products have the same pairs of orthogonal vectors, then each of the inner products is a scalar multiple of the other inner product.

Solution. If $V = \{0\}$ then we may take any $c > 0$ we like, since the only inner product on V is the map $(0, 0) \mapsto 0$. Suppose therefore that $V \neq \{0\}$ and for each non-zero $v \in V$ define

$$c_v = \frac{\langle v, v \rangle_1}{\langle v, v \rangle_2};$$

notice that c_v is positive. Suppose $u, v \in V$ are non-zero. Using orthogonal decomposition 6.13, we have

$$\left\langle u - \frac{\langle u, v \rangle_2}{\langle v, v \rangle_2} v, v \right\rangle_2 = 0.$$

Our assumption is that orthogonality with respect to $\langle \cdot, \cdot \rangle_2$ is equivalent to orthogonality with respect to $\langle \cdot, \cdot \rangle_1$ and thus

$$\left\langle u - \frac{\langle u, v \rangle_2}{\langle v, v \rangle_2} v, v \right\rangle_1 = 0 \quad \Leftrightarrow \quad \langle u, v \rangle_1 - \frac{\langle u, v \rangle_2}{\langle v, v \rangle_2} \langle v, v \rangle_1 = 0 \quad \Leftrightarrow \quad \langle u, v \rangle_1 = c_v \langle u, v \rangle_2.$$

Reversing the roles of u and v shows that $\langle v, u \rangle_1 = c_u \langle v, u \rangle_2$ and combining this with conjugate symmetry gives us

$$c_v \langle u, v \rangle_2 = \langle u, v \rangle_1 = \overline{\langle v, u \rangle_1} = \overline{c_u \langle v, u \rangle_2} = c_u \langle u, v \rangle_2.$$

Thus, for all non-zero $u, v \in V$, we have

(1) $\langle u, v \rangle_1 = c_v \langle u, v \rangle_2$;

(2) $c_v \langle u, v \rangle_2 = c_u \langle u, v \rangle_2$.

Given non-zero $u, v \in V$, there exists a non-zero $w \in V$ such that $\langle w, u \rangle_2 \neq 0$ and $\langle v, w \rangle_2 \neq 0$: if $\langle u, v \rangle_2 \neq 0$ then take $w = u$ and if $\langle u, v \rangle_2 = 0$ then take $w = u + v$. Using (2), it follows that

$$c_u \langle w, u \rangle_2 = c_w \langle w, u \rangle_2 \quad \text{and} \quad c_w \langle v, w \rangle_2 = c_v \langle v, w \rangle_2.$$

Since $\langle w, u \rangle_2 \neq 0$ and $\langle v, w \rangle_2 \neq 0$, these two equations imply that $c_u = c_w = c_v$. If we denote this common value by c (noting that $c > 0$), then we have shown that $c_v = c$ for all non-zero $v \in V$. It follows from (1) that $\langle u, v \rangle_1 = c \langle u, v \rangle_2$ for all non-zero $u, v \in V$. Certainly this equation also holds if $u = 0$ or $v = 0$ and thus $\langle u, v \rangle_1 = c \langle u, v \rangle_2$ for all $u, v \in V$.

Exercise 6.B.16. Suppose V is finite-dimensional. Suppose $\langle\cdot,\cdot\rangle_1, \langle\cdot,\cdot\rangle_2$ are inner products on V with corresponding norms $\|\cdot\|_1$ and $\|\cdot\|_2$. Prove that there exists a positive number c such that $\|v\|_1 \le c\|v\|_2$ for every $v \in V$.

Solution. By 6.35 there exists an orthonormal basis $e_1, \dots, e_n$ of V with respect to $\langle\cdot,\cdot\rangle_2$. Let $v \in V$ be given, so that $v = a_1 e_1 + \cdots + a_n e_n$ for some scalars $a_1, \dots, a_n$. Observe that

$$|a_1| + \cdots + |a_n| \le n \max\{|a_1|, \dots, |a_n|\} \le n\sqrt{|a_1|^2 + \cdots + |a_n|^2} = n\|v\|_2, \tag{1}$$

where the last equality follows from 6.24. If we let $M = \max\left\{\|e_1\|_1, \dots, \|e_n\|_1\right\}$, which is positive since each $e_k \ne 0$, then it follows from (1) and the triangle inequality that

$$\|v\|_1 \le |a_1|\|e_1\|_1 + \cdots + |a_n|\|e_n\|_1 \le M(|a_1| + \cdots + |a_n|) \le nM\|v\|_2.$$

Thus the desired positive constant is $c = nM$.

Exercise 6.B.17. Suppose $\mathbf{F} = \mathbf{C}$ and V is finite-dimensional. Prove that if T is an operator on V such that 1 is the only eigenvalue of T and $\|Tv\| \le \|v\|$ for all $v \in V$, then T is the identity operator.

Solution. By Schur's theorem (6.38), there is an orthonormal basis $e_1, \dots, e_n$ of V with respect to which the matrix of T is upper-triangular; because the only eigenvalue of T is 1, the diagonal entries of this matrix must equal 1. Thus, for each $k \in \{1, \dots, n\}$,

$$Te_k = A_{1,k}e_1 + \cdots + A_{1,k-1}e_{k-1} + e_k \quad \Rightarrow \quad \|Te_k\|^2 = \left|A_{1,k}\right|^2 + \cdots + \left|A_{1,k-1}\right|^2 + 1,$$

where we have used 6.24. By assumption $\|Te_k\|^2 \le \|e_k\|^2 = 1$ and thus

$$\left|A_{1,k}\right|^2 + \cdots + \left|A_{1,k-1}\right|^2 \le 0 \quad \Rightarrow \quad A_{1,k} = \cdots = A_{1,k-1} = 0.$$

It follows that the matrix of T with respect to $e_1, \dots, e_n$ is diagonal. Since each diagonal entry is equal to 1, we may conclude that T is the identity operator.

Exercise 6.B.18. Suppose $u_1, \dots, u_m$ is a linearly independent list in V. Show that there exists $v \in V$ such that $\langle u_k, v\rangle = 1$ for all $k \in \{1, \dots, m\}$.

Solution. Let $U = \operatorname{span}(u_1, \dots, u_m)$, so that $u_1, \dots, u_m$ is a basis of U, let $\varphi_1, \dots, \varphi_m$ be the dual basis of U', and let $\varphi = \varphi_1 + \cdots + \varphi_m$. The Riesz representation theorem (6.42) shows that there is a unique $v \in U$ such that $\varphi(u) = \langle u, v\rangle$ for each $u \in U$. In particular, $\langle u_k, v\rangle = \varphi(u_k) = 1$ for each $k \in \{1, \dots, m\}$.

Exercise 6.B.19. Suppose $v_1, \dots, v_n$ is a basis of V. Prove that there exists a basis $u_1, \dots, u_n$ of V such that

$$\langle v_j, u_k \rangle = \begin{cases} 0 & \text{if } j \neq k, \\ 1 & \text{if } j = k. \end{cases}$$

Solution. Let $\varphi_1, \dots, \varphi_n$ be the dual basis of $v_1, \dots, v_n$. For each $k \in \{1, \dots, n\}$, the Riesz representation theorem (6.42) shows that there is some $u_k \in V$ such that $\varphi_k(v) = \langle v, u_k \rangle$ for every $v \in V$. It follows that

$$\langle v_j, u_k \rangle = \varphi_k(v_j) = \begin{cases} 0 & \text{if } j \neq k, \\ 1 & \text{if } j = k. \end{cases}$$

Suppose $a_1, \dots, a_n$ are scalars such that $a_1 u_1 + \dots + a_n u_n = 0$. For each $k \in \{1, \dots, n\}$, observe that

$$0 = \langle a_1 u_1 + \dots + a_n u_n, v_k \rangle = a_1 \langle u_1, v_k \rangle + \dots + a_n \langle u_n, v_k \rangle = a_k.$$

It follows that $u_1, \dots, u_n$ is linearly independent and hence forms a basis of V.

Exercise 6.B.20. Suppose $\mathbf{F} = \mathbf{C}$, V is finite-dimensional, and $\mathcal{E} \subseteq \mathcal{L}(V)$ is such that

$$ST = TS$$

for all $S, T \in \mathcal{E}$. Prove that there is an orthonormal basis of V with respect to which every element of $\mathcal{E}$ has an upper-triangular matrix.

This exercise strengthens Exercise 9(b) in Section 5E (in the context of inner product spaces) by asserting that the basis in that exercise can be chosen to be orthonormal.

Solution. By Exercise 5.E.9 (b) there is a basis $v_1, \dots, v_n$ of V with respect to which every element of $\mathcal{E}$ has an upper-triangular matrix, i.e. such that

$$Te_k \in \operatorname{span}(v_1, \dots, v_k)$$

for every $k \in \{1, \dots, m\}$ and every $T \in \mathcal{E}$. Let $e_1, \dots, e_n$ be the orthonormal basis obtained by applying the Gram-Schmidt procedure (6.32) to $v_1, \dots, v_n$ and note that

$$Te_k \in \operatorname{span}(v_1, \dots, v_k) = \operatorname{span}(e_1, \dots, e_k)$$

for every $k \in \{1, \dots, m\}$ and every $T \in \mathcal{E}$. Thus $e_1, \dots, e_n$ is an orthonormal basis of V with respect to which every element of $\mathcal{E}$ has an upper-triangular matrix.

Exercise 6.B.21. Suppose $\mathbf{F} = \mathbf{C}$, V is finite-dimensional, $T \in \mathcal{L}(V)$, and all eigenvalues of T have absolute value less than 1. Let $\varepsilon > 0$. Prove that there exists a positive integer m such that $\|T^m v\| \leq \varepsilon \|v\|$ for every $v \in V$.

Solution. We will first translate the problem into a statement involving column vectors in $\mathbf{C}^{n,1}$ and matrices in $\mathbf{C}^{n,n}$, where $n = \dim V$. We shall use the following notation, which differs from the notation of the question.

- $\langle \cdot, \cdot \rangle_V : V \times V \to \mathbf{C}$. This is the given inner product on V.
- $\| \cdot \|_V : V \to [0, \infty)$. This is the norm on V arising from $\langle \cdot, \cdot \rangle_V$, i.e.
$$\|v\|_V = \sqrt{\langle v, v \rangle_V}.$$
- $\langle \cdot, \cdot \rangle : \mathbf{C}^{n,1} \times \mathbf{C}^{n,1} \to \mathbf{C}$. This is the Euclidean inner product on $\mathbf{C}^{n,1}$, i.e.
$$\langle x, y \rangle = \sum_{k=1}^{n} x_k \overline{y_k}, \quad \text{where } x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \text{ and } y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$
- $\| \cdot \| : \mathbf{C}^{n,1} \to [0, \infty)$. This is the Euclidean norm on $\mathbf{C}^{n,1}$ arising from $\langle \cdot, \cdot \rangle$, i.e.
$$\|x\| = \sqrt{\langle x, x \rangle} = \left(\sum_{k=1}^{n} |x_k|^2 \right)^{1/2}, \quad \text{where } x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Schur's theorem (6.38) implies that there is an orthonormal (with respect to $\langle \cdot, \cdot \rangle_V$) basis $e_1, ..., e_n$ of V such that the matrix $A \in \mathbf{C}^{n,n}$ of T with respect to $e_1, ..., e_n$ is upper-triangular. Given $v = x_1 e_1 + \cdots + x_n e_n \in V$, observe that
$$\|v\|_V = \left(\sum_{k=1}^{n} |x_k|^2 \right)^{1/2} = \|x\|, \quad \text{where } x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$
Thus it will suffice to show that there exists a positive integer m such that
$$\|A^m x\| \leq \varepsilon \|x\| \text{ for all } x \in \mathbf{C}^{n,1}.$$
In what follows, by a strictly upper-triangular matrix we mean a matrix that is upper-triangular and whose diagonal entries are zero. We shall use the following two easily verified facts about strictly upper-triangular matrices:

(i) if $D \in \mathbf{C}^{n,n}$ is diagonal and $N \in \mathbf{C}^{n,n}$ is strictly upper-triangular then DN and ND are both strictly upper-triangular;

(ii) if $N \in \mathbf{C}^{n,n}$ is strictly upper-triangular then $N^n = 0$.

Let D be the diagonal matrix whose diagonal entries are exactly those of U and let N be the strictly upper-triangular matrix whose entries above the diagonal are exactly those of U, so that $U = D + N$. Let
$$\rho = \max\{|D_{1,1}|, ..., |D_{n,n}|\} = \max\{|U_{1,1}|, ..., |U_{n,n}|\}$$
and note that $0 \leq \rho < 1$ since the diagonal elements of D and U are precisely the eigenvalues of T. Note further that
$$\|Dx\| = \left(|D_{1,1}|^2 |x_1|^2 + \cdots + |D_{n,n}|^2 |x_n|^2 \right)^{1/2} \leq \rho \|x\|$$

for any $x \in \mathbf{C}^{n,1}$. Let

$$C = \sum_{j=1}^{n} \sum_{k=1}^{n} |N_{j,k}|^2;$$

a calculation similar to the one given in Exercise 6.A.19 shows that $\|Nx\| \leq C\|x\|$ for all $x \in \mathbf{C}^{n,1}$. Putting everything together, for any integer $m \geq n$ we have the inequality

$$\|U^m x\| = \|(D+N)^m x\| \leq \sum_{k=0}^{n-1} \binom{m}{k} \rho^{m-k} C^k \|x\|.$$

This can be shown using induction, but is best illustrated by example. If $m = 3$ and $n = 2$ then (i) and (ii) show that $DNN = NDN = NND = NNN = 0$. Thus

$$\begin{aligned}
\|U^m x\| &= \|(D+N)^m x\| \\
&\leq \|(DDD + DDN + DND + NDD + DNN + NDN + NND + NNN)x\| \\
&= \|(DDD + DDN + DND + NDD)x\| \\
&\leq \|DDDx\| + \|DDNx\| + \|DNDx\| + \|NDDx\| \\
&\leq \rho(\rho(\rho\|x\|)) + \rho(\rho(C\|x\|)) + \rho(C(\rho\|x\|)) + C(\rho(\rho\|x\|)) \\
&= \sum_{k=0}^{1} \binom{3}{k} \rho^{3-k} C^k \|x\|.
\end{aligned}$$

For any $0 \leq k \leq n-1$ we have $\binom{m}{k} \leq m^{n-1}$ and $\rho^{m-k} \leq \rho^{m-n+1}$, since $0 \leq \rho < 1$. Thus, letting $\mu = \max\{1, C, ..., C^{n-1}\}$, we have the inequality

$$\|U^m x\| \leq \sum_{k=0}^{n-1} \binom{m}{k} \rho^{m-k} C^k \|x\| \leq \mu n m^{n-1} \rho^{m-n+1} \|x\|$$

for any $x \in \mathbf{C}^{n,1}$. Because $0 \leq \rho < 1$ we have $\lim_{m\to\infty} m^{n-1}\rho^{m-n+1} = 0$. Thus there exists a positive integer m such that $\|U^m x\| \leq \varepsilon\|x\|$ for every $x \in \mathbf{C}^{n,1}$.

Exercise 6.B.22. Suppose $C[-1,1]$ is the vector space of continuous real-valued functions on the interval $[-1,1]$ with inner product given by

$$\langle f, g\rangle = \int_{-1}^{1} fg$$

for all $f, g \in C[-1,1]$. Let φ be the linear functional on $C[-1,1]$ defined by $\varphi(f) = f(0)$. Show that there does not exist $g \in C[-1,1]$ such that

$$\varphi(f) = \langle f, g\rangle$$

for every $f \in C[-1,1]$.

This exercise shows that the Riesz representation theorem (6.42) does not hold on infinite-dimensional vector spaces without additional hypotheses on V and φ.

Solution. Suppose such a g exists and define $h \in C[-1,1]$ by $h(x) = x^2 g(x)$. Observe that

$$0 = h(0) = \varphi(h) = \langle h, g\rangle = \int_{-1}^{1} [xg(x)]^2 \, \mathrm{d}x.$$

Because the integrand is non-negative and continuous, we have $\int_{-1}^{1} [xg(x)]^2 \, \mathrm{d}x = 0$ if and only if $xg(x) = 0$ for all $x \in [-1,1]$. This implies that $g(x) = 0$ for all $x \in [-1,1] \setminus \{0\}$; the continuity of g then implies that $g(0) = 0$ also. Thus

$$f(0) = \int_{-1}^{1} f(x)g(x) \, \mathrm{d}x = \int_{-1}^{1} 0 \, \mathrm{d}x = 0$$

for every $f \in C[-1,1]$, which is certainly not true. We may conclude that no such g exists.

Exercise 6.B.23. For all $u, v \in V$, define $d(u,v) = \|u - v\|$.

(a) Show that d is a metric on V.

(b) Show that if V is finite-dimensional, then d is a complete metric on V (meaning that every Cauchy sequence converges).

(c) Show that every finite-dimensional subspace of V is a closed subset of V (with respect to the metric d).

This exercise requires familiarity with metric spaces.

Solution.

(a) Certainly d is non-negative. The equivalence of $d(u,v) = 0$ and $u = v$ follows from 6.9(a). The symmetry of d follows from 6.9(b), and the triangle inequality for d is immediate from 6.17.

(b) Let $v_1, \ldots, v_n$ be a basis of V and consider the 1-norm with respect to this basis:

$$\|v\|_1 = |a_1| + \cdots + |a_n|,$$

where $v = a_1 v_1 + \cdots + a_n v_n$. By Exercise 6.B.16, it will suffice to show that V is complete with respect to $\|\cdot\|_1$.

Let $(v_m)_{m=1}^{\infty}$ be a Cauchy sequence in $\left(V, \|\cdot\|_1\right)$, where $v_m = a_{m,1} v_1 + \cdots + a_{m,n} v_n$. For any $j \in \{1, \ldots, n\}$ and any positive integers k and m, we have the inequality

$$\left|a_{m,j} - a_{k,j}\right| \leq \|v_m - v_k\|_1.$$

It follows that $\left(a_{m,j}\right)_{m=1}^{\infty}$ is a Cauchy sequence in the complete metric space $\mathbf{F}$ and thus there exists some $a_j \in \mathbf{F}$ such that $\lim_{m \to \infty} a_{m,j} = a_j$. Define $v = a_1 v_1 + \cdots + a_n v_n$ and observe that

$$\|v_m - v\|_1 = \left|a_{m,1} - a_1\right| + \cdots + \left|a_{m,n} - a_n\right| \to 0 \text{ as } m \to \infty.$$

Thus $(v_m)_{m=1}^{\infty}$ is convergent. We may conclude that V is complete with respect to $\|\cdot\|_1$.

(c) Suppose U is a finite-dimensional subspace of V and $(u_m)_{m=1}^{\infty}$ is a sequence contained in U satisfying $\lim_{m \to \infty} \|u_m - v\|$ for some $v \in V$. We need to show that $v \in U$. The norm $\|\cdot\|$ restricts to a norm on U; by part (b), this normed space $(U, \|\cdot\|)$ must be complete since U is finite-dimensional. Because convergent sequences are necessarily Cauchy, completeness implies that the sequence $(u_m)_{m=1}^{\infty}$ converges to some $u \in U$. Since limits of sequences are unique it follows that $v = u \in U$, as desired.

6.C. Orthogonal Complements and Minimization Problems

Exercise 6.C.1. Suppose $v_1, ..., v_m \in V$. Prove that

$$\{v_1, ..., v_m\}^\perp = (\operatorname{span}(v_1, ..., v_m))^\perp.$$

Solution. Suppose that $v \in (\operatorname{span}(v_1, ..., v_m))^\perp$. In particular, for every $k \in \{1, ..., m\}$,

$$v \in \operatorname{span}(v_1, ..., v_m) \quad \Rightarrow \quad \langle v, v_k \rangle = 0.$$

It follows that $v \in \{v_1, ..., v_m\}^\perp$ and hence that $(\operatorname{span}(v_1, ..., v_m))^\perp \subseteq \{v_1, ..., v_m\}^\perp$.

Now suppose that $v \in \{v_1, ..., v_m\}^\perp$ and let $a_1 v_1 + \cdots + a_m v_m \in \operatorname{span}(v_1, ..., v_m)$ be given. Observe that

$$\langle v, a_1 v_1 + \cdots + a_m v_m \rangle = \overline{a_1} \langle v, v_1 \rangle + \cdots + \overline{a_m} \langle v, v_m \rangle = 0.$$

It follows that $v \in (\operatorname{span}(v_1, ..., v_m))^\perp$ and hence that $\{v_1, ..., v_m\}^\perp \subseteq (\operatorname{span}(v_1, ..., v_m))^\perp$. We may conclude that

$$\{v_1, ..., v_m\}^\perp = (\operatorname{span}(v_1, ..., v_m))^\perp.$$

Exercise 6.C.2. Suppose U is a subspace of V with basis $u_1, ..., u_m$ and

$$u_1, ..., u_m, v_1, ..., v_n$$

is a basis of V. Prove that if the Gram-Schmidt procedure is applied to the basis of V above, producing a list $e_1, ..., e_m, f_1, ..., f_n$, then $e_1, ..., e_m$ is an orthonormal basis of U and $f_1, ..., f_n$ is an orthonormal basis of $U^\perp$.

Solution. The Gram-Schmidt procedure guarantees that

$$\operatorname{span}(e_1, ..., e_m) = \operatorname{span}(u_1, ..., u_m) = U,$$

and 6.25 shows that $e_1, ..., e_m$ is linearly independent. Thus $e_1, ..., e_m$ is an orthonormal basis of U.

The Gram-Schmidt procedure also guarantees that for any $k \in \{1, ..., n\}$ the vector f_k is orthogonal to each vector in the list $e_1, ..., e_n$. By Exercise 6.C.1, this implies that

$$f_k \in (\operatorname{span}(e_1, ..., e_m))^\perp = U^\perp.$$

Note that $f_1, ..., f_n$ is linearly independent by 6.25 and $\dim U^\perp = \dim V - \dim U = n$ by 6.51. Thus, by 2.38, $f_1, ..., f_n$ is an orthonormal basis of $U^\perp$.

Exercise 6.C.3. Suppose U is the subspace of $\mathbf{R}^4$ defined by

$$U = \operatorname{span}((1,2,3,-4),(-5,4,3,2)).$$

Find an orthonormal basis of U and an orthonormal basis of $U^{\perp}$.

Solution. It is straightforward to verify that

$$u_1 = (1,2,3,-4), \quad u_2 = (-5,4,3,2), \quad v_1 = (1,0,0,0), \quad v_2 = (0,1,0,0)$$

is a basis of $\mathbf{R}^4$. Certainly u_1, u_2 is a basis of U. Performing the Gram-Schmidt procedure on this list yields the orthonormal list

$$e_1 = \frac{1}{\sqrt{30}}(1,2,3,-4), \quad e_2 = \frac{1}{\sqrt{12030}}(-77,56,39,38),$$
$$f_1 = \frac{1}{\sqrt{76190}}(190,117,60,151), \quad f_2 = \frac{1}{9\sqrt{190}}(0,81,-90,27).$$

As we showed in Exercise 6.C.3, e_1, e_2 must be an orthonormal basis of U and f_1, f_2 must be an orthonormal basis of $U^{\perp}$.

Exercise 6.C.4. Suppose $e_1, \dots, e_n$ is a list of vectors in V with $\|e_k\| = 1$ for each $k = 1, \dots, n$ and

$$\|v\|^2 = |\langle v, e_1\rangle|^2 + \cdots + |\langle v, e_n\rangle|^2$$

for all $v \in V$. Prove that $e_1, \dots, e_n$ is an orthonormal basis of V.

This exercise provides a converse to 6.30(b).

Solution. Let $k \in \{1, \dots, n\}$ be given and observe that

$$\|e_k\|^2 = |\langle e_k, e_k\rangle|^2 + \sum_{j=1,\, j\neq k}^{n} \left|\langle e_k, e_j\rangle\right|^2 \quad \Leftrightarrow \quad 1 = 1 + \sum_{j=1,\, j\neq k}^{n} \left|\langle e_k, e_j\rangle\right|^2$$
$$\Rightarrow \quad \sum_{j=1,\, j\neq k}^{n} \left|\langle e_k, e_j\rangle\right|^2 = 0.$$

Thus $\langle e_k, e_j\rangle = 0$ for each $j \neq k$. It follows that $e_1, \dots, e_n$ is an orthonormal list and hence, by 6.25, $e_1, \dots, e_n$ is linearly independent. Suppose that $v \in \{e_1, \dots, e_n\}^{\perp}$ and observe that

$$\|v\|^2 = |\langle v, e_1\rangle|^2 + \cdots + |\langle v, e_n\rangle|^2 = 0 \quad \Leftrightarrow \quad v = 0.$$

Thus $\{e_1, \dots, e_n\}^{\perp} = \{0\}$. It follows from Exercise 6.C.1 and 6.54 that $\operatorname{span}(e_1, \dots, e_n) = V$. We may conclude that $e_1, \dots, e_n$ is an orthonormal basis of V.

Exercise 6.C.5. Suppose that V is finite-dimensional and U is a subspace of V. Show that $P_{U^{\perp}} = I - P_U$, where I is the identity operator on V.

Solution. By 6.49 and 6.52, for any $v \in V$ we can write $v = u + w$ for unique vectors $u \in U = (U^\perp)^\perp$ and $w \in U^\perp$. It follows that

$$P_U(v) + P_{U^\perp}(v) = u + w = v.$$

Thus $P_U + P_{U^\perp} = I$.

Exercise 6.C.6. Suppose V is finite-dimensional and $T \in \mathcal{L}(V, W)$. Show that

$$T = TP_{(\operatorname{null} T)^\perp} = P_{\operatorname{range} T} T.$$

Solution. By 6.49 we have $V = \operatorname{null} T \oplus (\operatorname{null} T)^\perp$. Thus, for any $v \in V$, we can write $v = x + y$, where $x \in \operatorname{null} T$ and $y \in (\operatorname{null} T)^\perp$. It follows that

$$Tv = Ty = TP_{(\operatorname{null} T)^\perp}(v)$$

and hence that $T = TP_{(\operatorname{null} T)^\perp}$.

Because $\operatorname{range} T$ is finite-dimensional, 6.49 allows us to write $W = \operatorname{range} T \oplus (\operatorname{range} T)^\perp$. For any $v \in V$, observe that $Tv = Tv + 0 \in \operatorname{range} T \oplus (\operatorname{range} T)^\perp$. It follows that

$$Tv = P_{\operatorname{range} T} Tv$$

and hence that $T = P_{\operatorname{range} T} T$.

Exercise 6.C.7. Suppose that X and Y are finite-dimensional subspaces of V. Prove that $P_X P_Y = 0$ if and only if $\langle x, y \rangle = 0$ for all $x \in X$ and all $y \in Y$.

Solution. Suppose that $\langle x, y \rangle = 0$ for all $x \in X$ and $y \in Y$. For $v \in V$ write $v = y + z$, where $y \in Y$ and $z \in Y^\perp$, so that $P_Y v = y$. Our hypothesis ensures that $y \in X^\perp$ and thus

$$P_X P_Y v = P_X v = 0$$

by 6.57(c). Hence $P_X P_Y = 0$.

For the converse, suppose that $P_X P_Y = 0$ and let $x \in X$ and $y \in Y$ be given. Using 6.57(b) and 6.57(e), observe that

$$P_X y = P_X P_Y y = 0 \quad \Rightarrow \quad y \in \operatorname{null} P_X \quad \Rightarrow \quad y \in X^\perp \quad \Rightarrow \quad \langle x, y \rangle = 0.$$

Exercise 6.C.8. Suppose U is a finite-dimensional subspace of V and $v \in V$. Define a linear functional $\varphi : U \to \mathbf{F}$ by

$$\varphi(u) = \langle u, v \rangle$$

for all $u \in U$. By the Riesz representation theorem (6.42) as applied to the inner product space U, there exists a unique vector $w \in U$ such that

$$\varphi(u) = \langle u, w \rangle$$

for all $u \in U$. Show that $w = P_U v$.

Solution. We have $\langle u, v \rangle = \langle u, w \rangle$ for every $u \in U$. Equivalently, $\langle u, v - w \rangle = 0$ for every $u \in U$, so that $v - w \in U^\perp$. It follows that $v = w + v - w$, where $w \in U$ and $v - w \in U^\perp$. Thus $P_U v = w$.

Exercise 6.C.9. Suppose V is finite-dimensional. Suppose $P \in \mathcal{L}(V)$ is such that $P^2 = P$ and every vector in $\operatorname{null} P$ is orthogonal to every vector in $\operatorname{range} P$. Prove that there exists a subspace U of V such that $P = P_U$.

Solution. By Exercise 3.B.27 and 6.49 we have the decompositions

$$V = \operatorname{range} P \oplus \operatorname{null} P \quad \text{and} \quad V = \operatorname{range} P \oplus (\operatorname{range} P)^\perp,$$

which implies $\dim \operatorname{null} P = \dim (\operatorname{range} P)^\perp$. Combining this with the hypothesis $\operatorname{null} P \subseteq (\operatorname{range} P)^\perp$, we see that $\operatorname{null} P = (\operatorname{range} P)^\perp$. Let $U = \operatorname{range} P$; we claim that $P = P_U$. Let $v = Px + w \in V$ be given, where $Px \in \operatorname{range} P$ and $w \in (\operatorname{range} P)^\perp = \operatorname{null} P$. Observe that

$$P_U v = Px = P(Px + w) = Pv,$$

where we have used $P^2 = P$ and $w \in \operatorname{null} P$ for the second equality. Thus $P = P_U$.

Exercise 6.C.10. Suppose V is finite-dimensional and $P \in \mathcal{L}(V)$ is such that $P^2 = P$ and

$$\|Pv\| \leq \|v\|$$

for every $v \in V$. Prove that there exists a subspace U of V such that $P = P_U$.

Solution. Suppose $w \in \operatorname{null} P$ and $Px \in \operatorname{range} P$. Our hypothesis gives us the inequality

$$\|Px\| = \|P(Px + \lambda w)\| \leq \|Px + \lambda w\|$$

for any $\lambda \in \mathbf{F}$. It follows from Exercise 6.A.6 that $\langle w, Px \rangle = 0$ and hence that $\operatorname{null} P$ is contained in $(\operatorname{range} P)^\perp$. We can now let $U = \operatorname{range} P$ and proceed as in Exercise 6.C.9 to see that $P = P_U$.

Exercise 6.C.11. Suppose $T \in \mathcal{L}(V)$ and U is a finite-dimensional subspace of V. Prove that

$$U \text{ is invariant under } T \quad \Leftrightarrow \quad P_U T P_U = T P_U.$$

Solution. Suppose that U is invariant under T and let $v \in V$ be given. Observe that

$$P_U v \in U \quad \Rightarrow \quad T P_U v \in U \quad \Rightarrow \quad P_U T P_U v = T P_U v,$$

where the last implication follows from 6.57(b). Now suppose that U is not invariant under T, i.e. there is some $u \in U$ such that $Tu \notin U$. Note that

$$T P_U u = Tu \notin U \quad \text{and} \quad P_U T P_U u \in U,$$

where we have used 6.57(b) and 6.57(d). It follows that

$$P_U T P_U u \neq T P_U u \quad \Rightarrow \quad P_U T P_U \neq P_U T.$$

Exercise 6.C.12. Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and U is a subspace of V. Prove that

$$U \text{ and } U^{\perp} \text{ are both invariant under } T \quad \Leftrightarrow \quad P_U T = T P_U.$$

Solution. Suppose that U and $U^{\perp}$ are both invariant under T and let $v = u + w \in V$ be given, where $u \in U$ and $w \in U^{\perp}$. By assumption we have $Tu \in U$ and $Tw \in U^{\perp}$; using 6.57, it follows that

$$P_U T v = P_U(Tu + Tw) = Tu = T P_U u = T P_U(u + w) = T P_U v.$$

If U is not invariant under T then there exists some $u \in U$ such that $Tu \notin U$. It follows that

$$T P_U u = Tu \notin U \quad \text{and} \quad P_U T u \in U,$$

so that $T P_U \neq P_U T$. Similarly, if $U^{\perp}$ is not invariant under T then there exists some $w \in U^{\perp}$ such that $Tw \notin U^{\perp}$. It follows that

$$T P_U w = T(0) = 0 \quad \text{and} \quad P_U T w \neq 0,$$

so that $T P_U \neq P_U T$.

Exercise 6.C.13. Suppose $\mathbf{F} = \mathbf{R}$ and V is finite-dimensional. For each $v \in V$, let φ_v denote the linear functional on V defined by

$$\varphi_v(u) = \langle u, v \rangle$$

for all $u \in V$.

(a) Show that $v \mapsto \varphi_v$ is an injective linear map from V to V'.

(b) Use (a) and a dimension-counting argument to show that $v \mapsto \varphi_v$ is an isomorphism from V onto V'.

The purpose of this exercise is to give an alternative proof of the Riesz representation theorem (6.42 and 6.58) when $\mathbf{F} = \mathbf{R}$. *Thus you should not use the Riesz representation theorem as a tool in your solution.*

Solution.

(a) Let $u, v, w \in V$ be given and note that

$$\varphi_{v+w}(u) = \langle u, v + w \rangle = \langle u, v \rangle + \langle u, w \rangle = \varphi_v(u) + \varphi_w(u).$$

Similarly, for any $u, v \in V$ and any $\lambda \in \mathbf{R}$,

$$\varphi_{\lambda v}(u) = \langle u, \lambda v \rangle = \lambda \langle u, v \rangle = \lambda \varphi_v(u).$$

Thus $v \mapsto \varphi_v$ is linear. Suppose that $v \in V$ is such that $\varphi_v = 0$, i.e. $\langle u, v \rangle = 0$ for every $u \in U$. It follows that $v \in V^\perp$ and hence, by 6.48(c), $v = 0$. Thus $v \mapsto \varphi_v$ is injective.

(b) By 3.111, 3.65, and part (a), the map $v \mapsto \varphi_v$ must be an isomorphism.

Exercise 6.C.14. Suppose that $e_1, ..., e_n$ is an orthonormal basis of V. Explain why the dual basis (see 3.112) of $e_1, ..., e_n$ is $e_1, ..., e_n$ under the identification of V' with V provided by the Riesz representation theorem (6.58).

Solution. Let $\varphi_1, ..., \varphi_n$ be the dual basis of $e_1, ..., e_n$. Using the notation of 6.58, for any $j, k \in \{1, ..., n\}$ we have

$$\varphi_{e_k}(e_j) = \langle e_j, e_k \rangle = \begin{cases} 1 & \text{if } j = k, \\ 0 & \text{if } j \neq k. \end{cases}$$

Thus $\varphi_{e_k} = \varphi_k$ for each $k \in \{1, ..., n\}$, i.e. we may identify e_k with φ_k.

Exercise 6.C.15. In $\mathbf{R}^4$, let

$$U = \operatorname{span}((1, 1, 0, 0), (1, 1, 1, 2)).$$

Find $u \in U$ such that $\|u - (1, 2, 3, 4)\|$ is as small as possible.

Solution. Let $u_1 = (1, 1, 0, 0)$ and $u_2 = (1, 1, 1, 2)$, so that u_1, u_2 is a basis of U. Performing the Gram-Schmidt procedure on the list u_1, u_2 yields the list

$$e_1 = \frac{1}{\sqrt{2}}(1, 1, 0, 0), \quad e_2 = \frac{1}{\sqrt{5}}(0, 0, 1, 2),$$

which is an orthonormal basis of U. Let $v = (1, 2, 3, 4)$. According to 6.61, to minimize $\|u - v\|$ we should take $u = P_U v$. This can be calculated using 6.57(i):

$$P_U v = \langle v, e_1 \rangle e_1 + \langle v, e_2 \rangle e_2 = \left(\frac{3}{2}, \frac{3}{2}, \frac{11}{5}, \frac{22}{5} \right).$$

Exercise 6.C.16. Suppose $C[-1, 1]$ is the vector space of continuous real-valued functions on the interval $[-1, 1]$ with inner product given by

$$\langle f, g \rangle = \int_{-1}^{1} fg$$

for all $f, g \in C[-1, 1]$. Let U be the subspace of $C[-1, 1]$ defined by

$$U = \{ f \in C[-1, 1] : f(0) = 0 \}.$$

(a) Show that $U^\perp = \{0\}$.

(b) Show that 6.49 and 6.52 do not hold without the finite-dimensional hypothesis.

Solution.

(a) Certainly $0 \in U^\perp$. Suppose that $g \in U^\perp$. Let $f : [-1, 1] \to \mathbf{R}$ be given by $f(x) = x^2 g(x)$ and note that $f \in U$, so that

$$0 = \langle f, g \rangle = \int_{-1}^{1} [xg(x)]^2 \, dx.$$

Because the integrand $[xg(x)]^2$ is continuous and non-negative we must have $xg(x) = 0$ for every $x \in [-1, 1]$, which implies $g(x) = 0$ for all non-zero $x \in [-1, 1]$. The continuity of g gives us $g(0) = 0$ also and thus $g = 0$. We may conclude that $U^\perp = \{0\}$.

(b) From part (a) we have $U \oplus U^\perp = U \neq C[-1, 1]$ and thus 6.49 does not hold. Part (a) and 6.48(b) give us

$$\left(U^\perp\right)^\perp = \{0\}^\perp = C[-1, 1] \neq U,$$

so that 6.52 does not hold.

Exercise 6.C.17. Find $p \in \mathcal{P}_3(\mathbf{R})$ such that $p(0) = 0$, $p'(0) = 0$, and

$$\int_0^1 |2 + 3x - p(x)|^2 \, dx$$

is as small as possible.

Solution. Equip $\mathcal{P}_3(\mathbf{R})$ with the inner product

$$\langle p, q\rangle = \int_0^1 p(x)q(x)\,\mathrm{d}x$$

and let $U = \{p \in \mathcal{P}_3(\mathbf{R}) : p(0) = p'(0) = 0\}$. It is straightforward to verify that U is a subspace of $\mathcal{P}_3(\mathbf{R})$ and that x^2, x^3 is a basis of U. Performing the Gram-Schmidt procedure on this basis yields the orthonormal basis e_1, e_2 of U, where

$$e_1 = \sqrt{5}x^2, \quad e_2 = 6\sqrt{7}\big(x^3 - \tfrac{5}{6}x^2\big).$$

Let $q(x) = 2 + 3x$. According to 6.61, to minimize $\|q - p\|^2 = \int_0^1 |2 + 3x - p(x)|^2\,\mathrm{d}x$ we should take $p = P_U q$. This can be calculated using 6.57(i):

$$P_U q = \langle q, e_1\rangle e_1 + \langle q, e_2\rangle e_2 = 24x^2 - \frac{203}{10}x^3.$$

Exercise 6.C.18. Find $p \in \mathcal{P}_5(\mathbf{R})$ that makes $\int_{-\pi}^{\pi} |\sin x - p(x)|^2\,dx$ as small as possible.

> *The polynomial 6.65 is an excellent approximation to the answer to this exercise, but here you are asked to find the exact solution, which involves powers of π. A computer that can perform symbolic integration should help.*

Solution. Equip $C[-\pi, \pi]$ with the inner product

$$\langle p, q\rangle = \int_{-\pi}^{\pi} p(x)q(x)\,\mathrm{d}x$$

and let $U = \mathcal{P}_5(\mathbf{R})$. Performing the Gram-Schmidt procedure on the basis $1, x, x^2, x^3, x^4, x^5$ of U gives us the orthonormal basis

$$e_1 = \frac{1}{\sqrt{2\pi}}, \quad e_2 = \sqrt{\frac{3}{2\pi^3}}x, \quad e_3 = -\frac{1}{2}\sqrt{\frac{5}{2\pi^5}}(\pi^2 - 3x^2),$$

$$e_4 = -\frac{1}{2}\sqrt{\frac{7}{2\pi^7}}(3\pi^2 x - 5x^3), \quad e_5 = \frac{3}{8\sqrt{2\pi^9}}(3\pi^4 - 30\pi^2x^2 + 35x^4),$$

$$e_6 = -\frac{1}{8}\sqrt{\frac{11}{2\pi^{11}}}(15\pi^4 x - 70\pi^2x^3 + 63x^5).$$

According to 6.61, to minimize $\|\sin x - p\|^2 = \int_{-\pi}^{\pi} |\sin x - p(x)|^2\,\mathrm{d}x$ we should take $p = P_U(\sin x)$. This can be calculated using 6.57(i):

$$P_U(\sin x) = \frac{105(1465 - 153\pi^2 + \pi^4)}{8\pi^6}x - \frac{315(1155 - 125\pi^2 + \pi^4)}{4\pi^8}x^3 + \frac{693(945 - 105\pi^2 + \pi^4)}{8\pi^{10}}x^5.$$

Exercise 6.C.19. Suppose V is finite-dimensional and $P \in \mathcal{L}(V)$ is an orthogonal projection of V onto some subspace of V. Prove that $P^\dagger = P$.

Solution. Suppose U is the subspace of V such that $P = P_U$. Using 6.57(e) and 6.52, observe that

$$(\operatorname{null} P)^\perp = (\operatorname{null} P_U)^\perp = (U^\perp)^\perp = U.$$

Thus $P|_{(\operatorname{null} P)^\perp} = P|_U$. Since $P = P_U$, 6.57(b) shows that $P|_{(\operatorname{null} P)^\perp}$ is simply the identity operator on $(\operatorname{null} P)^\perp = U$. Combining this with 6.57(d), we have

$$P^\dagger = \left(P|_{(\operatorname{null} P)^\perp}\right)^{-1} P_{\operatorname{range} P} = P_U = P.$$

Exercise 6.C.20. Suppose V is finite-dimensional and $T \in \mathcal{L}(V, W)$. Show that

$$\operatorname{null} T^\dagger = (\operatorname{range} T)^\perp \quad \text{and} \quad \operatorname{range} T^\dagger = (\operatorname{null} T)^\perp.$$

Solution. Because $(T|_{(\operatorname{null} T)^\perp})^{-1}$ is injective (by 6.67), we have

$$\operatorname{null} T^\dagger = \operatorname{null} P_{\operatorname{range} T} = (\operatorname{range} T)^\perp,$$

where we have used 6.57(e) for the last equality.

6.57(d) shows that $P_{\operatorname{range} T}$ is a surjection from W onto $\operatorname{range} T$, and 6.67 shows that $(T|_{(\operatorname{null} T)^\perp})^{-1}$ is a surjection from $\operatorname{range} T$ onto $(\operatorname{null} T)^\perp$. Thus $\operatorname{range} T^\dagger = (\operatorname{null} T)^\perp$.

$$W \xrightarrow{P_{\operatorname{range} T}} \operatorname{range} T \xrightarrow{(T|_{(\operatorname{null} T)^\perp})^{-1}} (\operatorname{null} T)^\perp$$

Exercise 6.C.21. Suppose $T \in \mathcal{L}(\mathbf{F}^3, \mathbf{F}^2)$ is defined by

$$T(a, b, c) = (a + b + c, 2b + 3c).$$

(a) For $(x, y) \in \mathbf{F}^2$, find a formula for $T^\dagger(x, y)$.

(b) Verify that the equation $TT^\dagger = P_{\operatorname{range} T}$ from 6.69(b) holds with the formula for $T^\dagger$ obtained in (a).

(c) Verify that the equation $T^\dagger T = P_{(\operatorname{null} T)^\perp}$ from 6.69(c) holds with the formula for $T^\dagger$ obtained in (a).

Solution.

(a) We proceed as in example 6.71. Note that T is surjective, so that $P_{\operatorname{range} T}$ is the identity operator on W. Note further that

$$\operatorname{null} T = \{(a, b, c) \in \mathbf{F}^3 : a + b + c = 0, 2b + 3c = 0\} = \operatorname{span}((1, -3, 2)).$$

For $(x, y) \in \mathbf{F}^2$, it follows that

$$T^\dagger(x,y) = (T|_{(\operatorname{null} T)^\perp})^{-1} P_{\operatorname{range} T}(x,y) = (T|_{(\operatorname{null} T)^\perp})^{-1}(x,y).$$

If $(T|_{(\operatorname{null} T)^\perp})^{-1}(x,y) = (a,b,c) \in \mathbf{F}^3$ then (a,b,c) must satisfy $T(a,b,c) = (x,y)$ and $\langle (a,b,c), (1,-3,2)\rangle = 0$. In other words, (a,b,c) must satisfy the following equations

$$a+b+c = x,$$

$$2b+3c = y,$$

$$a-3b+2c = 0.$$

Solving this system of equations yields the solution

$$a = \tfrac{1}{14}(13x-5y), \quad b = \tfrac{1}{14}(3x+y), \quad c = \tfrac{1}{7}(-x+2y).$$

Thus

$$T^\dagger(x,y) = \tfrac{1}{14}(13x-5y, 3x+y, -2x+4y).$$

(b) As noted in part (a), $P_{\operatorname{range} T}$ is the identity operator on W. Observe that

$$TT^\dagger(x,y) = \tfrac{1}{14}T(13x-5y, 3x+y, -2x+4y) = \tfrac{1}{14}(14x, 14y) = (x,y).$$

Thus $TT^\dagger = P_{\operatorname{range} T}$.

(c) As noted in part (a), $\operatorname{null} T = \operatorname{span}((1,-3,2))$. Thus

$$(\operatorname{null} T)^\perp = \{(a,b,c) \in \mathbf{F}^3 : a-3b+2c = 0\}.$$

It is straightforward to verify that $(1,1,1), (0,2,3)$ is a basis of $(\operatorname{null} T)^\perp$. Performing the Gram-Schmidt procedure on this basis gives us the orthonormal basis

$$\frac{1}{\sqrt{3}}(1,1,1), \quad \frac{1}{\sqrt{42}}(-5,1,4).$$

A tedious calculation using the formula for $T^\dagger$ found in part (a) and the formula for $P_{(\operatorname{null} T)^\perp}$ given by 6.57(i) shows that

$$T^\dagger T(a,b,c) = P_{(\operatorname{null} T)^\perp}(a,b,c) = \tfrac{1}{14}(13a+3b-2c, 3a+5b+6c, -2a+6b+10c).$$

Exercise 6.C.22. Suppose V is finite-dimensional and $T \in \mathcal{L}(V,W)$. Prove that

$$TT^\dagger T = T \quad \text{and} \quad T^\dagger T T^\dagger = T^\dagger.$$

Both formulas above clearly hold if T is invertible because in that case we can replace $T^\dagger$ with T^{-1}.

Solution. By 6.69(b) we have $TT^\dagger = P_{\operatorname{range} T}$. It follows that $TT^\dagger T = P_{\operatorname{range} T} T = T$, since $P_{\operatorname{range} T}$ is the identity operator on $\operatorname{range} T$ by 6.57(b).

By 6.69(c) we have $T^\dagger T = P_{(\operatorname{null} T)^\perp}$ and thus $T^\dagger T T^\dagger = P_{(\operatorname{null} T)^\perp} T^\dagger = T^\dagger$, since $T^\dagger$ maps into $(\operatorname{null} T)^\perp$ and $P_{(\operatorname{null} T)^\perp}$ is the identity operator on $(\operatorname{null} T)^\perp$.

Exercise 6.C.23. Suppose V and W are finite-dimensional and $T \in \mathcal{L}(V, W)$. Prove that

$$\left(T^\dagger\right)^\dagger = T.$$

The equation above is analogous to the equation $\left(T^{-1}\right)^{-1} = T$ *that holds if* T *is invertible.*

Solution. By Exercise 6.C.20 and 6.52 we have

$$\left(T^\dagger\right)^\dagger = \left(T^\dagger|_{(\operatorname{null} T^\dagger)^\perp}\right)^{-1} P_{\operatorname{range} T^\dagger} = \left(T^\dagger|_{\left((\operatorname{range} T)^\perp\right)^\perp}\right)^{-1} P_{(\operatorname{null} T)^\perp} = (T^\dagger|_{\operatorname{range} T})^{-1} P_{(\operatorname{null} T)^\perp}.$$

Because $P_{\operatorname{range} T}$ is the identity operator on $\operatorname{range} T$ we have $T^\dagger|_{\operatorname{range} T} = (T|_{(\operatorname{null} T)^\perp})^{-1}$. Thus

$$\left(T^\dagger\right)^\dagger = \left((T|_{(\operatorname{null} T)^\perp})^{-1}\right)^{-1} P_{(\operatorname{null} T)^\perp} = T|_{(\operatorname{null} T)^\perp}\, P_{(\operatorname{null} T)^\perp} = T P_{(\operatorname{null} T)^\perp}.$$

For any $v \in V$ we have $v = u + w$, where $u \in \operatorname{null} T$ and $w \in (\operatorname{null} T)^\perp$. It follows that

$$\left(T^\dagger\right)^\dagger(v) = T P_{(\operatorname{null} T)^\perp}(v) = Tw = Tv.$$

Thus $\left(T^\dagger\right)^\dagger = T$.

Chapter 7. Operators on Inner Product Spaces

7.A. Self-Adjoint and Normal Operators

Exercise 7.A.1. Suppose n is a positive integer. Define $T \in \mathcal{L}(\mathbf{F}^n)$ by

$$T(z_1, ..., z_n) = (0, z_1, ..., z_{n-1}).$$

Find a formula for $T^*(z_1, ..., z_n)$.

Solution. Observe that

$$\begin{aligned} \langle (w_1, ..., w_n), T^*(z_1, ..., z_n) \rangle &= \langle T(w_1, ..., w_n), (z_1, ..., z_n) \rangle \\ &= \langle (0, w_1, ..., w_{n-1}), (z_1, ..., z_n) \rangle \\ &= w_1 z_2 + \cdots + w_{n-1} z_n \\ &= \langle (w_1, ..., w_n), (z_2, ..., z_n, 0) \rangle. \end{aligned}$$

Thus $T^*(z_1, ..., z_n) = (z_2, ..., z_n, 0)$.

Exercise 7.A.2. Suppose $T \in \mathcal{L}(V, W)$. Prove that

$$T = 0 \quad \Leftrightarrow \quad T^* = 0 \quad \Leftrightarrow \quad T^*T = 0 \quad \Leftrightarrow \quad TT^* = 0.$$

Solution. Suppose $T = 0$, fix $w \in W$, and observe that

$$\langle v, T^*w \rangle = \langle Tv, w \rangle = \langle 0, w \rangle = 0$$

for any $v \in V$. Thus $T^*w \in V^\perp$. It follows from 6.48(c) that $T^*w = 0$ and hence that $T^* = 0$. Combining this with 7.5(c) shows that $T = 0$ if and only if $T^* = 0$.

That $T = 0$ implies $T^*T = 0$ is clear. Suppose that $T^*T = 0$, let $v \in V$ be given, and observe that

$$0 = \langle 0, v \rangle = \langle T^*Tv, v \rangle = \langle Tv, Tv \rangle = \|Tv\|^2 \quad \Rightarrow \quad Tv = 0.$$

Thus $T = 0$, so that $T = 0$ if and only if $T^*T = 0$. Replacing T with T^* in this result and using 7.5(c) shows that $T^* = 0$ if and only if $TT^* = 0$.

Exercise 7.A.3. Suppose $T \in \mathcal{L}(V)$ and $\lambda \in \mathbf{F}$. Prove that

$$\lambda \text{ is an eigenvalue of } T \quad \Leftrightarrow \quad \overline{\lambda} \text{ is an eigenvalue of } T^*.$$

Solution. Observe that

$$\begin{aligned}
\lambda \text{ is an eigenvalue of } T &\iff T-\lambda I \text{ is not surjective} && (5.7(c))\\
&\iff \operatorname{range}(T-\lambda I) \neq V\\
&\iff (\operatorname{range}(T-\lambda I))^{\perp} \neq \{0\} && (7.48(c))\\
&\iff \operatorname{null}(T-\lambda I)^{*} \neq \{0\} && (7.6(a))\\
&\iff \operatorname{null}(T^{*}-\overline{\lambda} I) \neq \{0\} && (7.5(a),\ (b),\ (e))\\
&\iff T^{*}-\overline{\lambda} I \text{ is not injective}\\
&\iff \overline{\lambda} \text{ is an eigenvalue of } T^{*}. && (5.7(b))
\end{aligned}$$

Exercise 7.A.4. Suppose $T \in \mathcal{L}(V)$ and U is a subspace of V. Prove that

$$U \text{ is invariant under } T \quad \Leftrightarrow \quad U^{\perp} \text{ is invariant under } T^{*}.$$

Solution. Suppose that U is invariant under T and let $v \in U^{\perp}$ be given. Observe that

$$\langle u, T^{*}v\rangle = \langle Tu, v\rangle = 0$$

for any $u \in U$, where the last equality follows since $Tu \in U$ and $v \in U^{\perp}$. Thus $T^{*}v \in U^{\perp}$ and it follows that $U^{\perp}$ is invariant under T^{*}.

Now suppose that $U^{\perp}$ is invariant under T^{*}. The previous paragraph shows that $(U^{\perp})^{\perp}$ is invariant under $(T^{*})^{*}$, which by 6.52 and 7.5(c) is exactly the statement that U is invariant under T.

Exercise 7.A.5. Suppose $T \in \mathcal{L}(V, W)$. Suppose $e_1, \ldots, e_n$ is an orthonormal basis of V and $f_1, \ldots, f_m$ is an orthonormal basis of W. Prove that

$$\|Te_1\|^2 + \cdots + \|Te_n\|^2 = \|T^{*}f_1\|^2 + \cdots + \|T^{*}f_m\|^2.$$

The numbers $\|Te_1\|^2, \ldots, \|Te_n\|^2$ in the equation above depend on the orthonormal basis $e_1, \ldots, e_n$, but the right side of the equation does not depend on $e_1, \ldots, e_n$. Thus the equation above shows that the sum on the left side does not depend on which orthonormal basis $e_1, \ldots, e_n$ is used.

Solution. Using 6.30(b), observe that

$$\begin{aligned}
\sum_{j=1}^{n} \|Te_j\|^2 = \sum_{j=1}^{n}\sum_{k=1}^{n} \left|\langle Te_j, f_k\rangle\right|^2 = \sum_{j=1}^{n}\sum_{k=1}^{n} \left|\langle e_j, T^{*}f_k\rangle\right|^2 \\
= \sum_{k=1}^{n}\sum_{j=1}^{n} \left|\langle T^{*}f_k, e_j\rangle\right|^2 = \sum_{k=1}^{n} \|T^{*}f_k\|^2.
\end{aligned}$$

Exercise 7.A.6. Suppose $T \in \mathcal{L}(V, W)$. Prove that

(a) T is injective $\Leftrightarrow$ T^* is surjective;

(b) T is surjective $\Leftrightarrow$ T^* is injective.

Solution.

(a) Observe that

$$\begin{aligned} T \text{ is injective} &\Leftrightarrow \operatorname{null} T = \{0\} \\ &\Leftrightarrow (\operatorname{range} T^*)^\perp = \{0\} && (7.6(c)) \\ &\Leftrightarrow \operatorname{range} T^* = V && (7.48(c)) \\ &\Leftrightarrow T^* \text{ is surjective.} \end{aligned}$$

(b) Part (a) shows that T^* is injective if and only if $(T^*)^*$ is surjective, which by 7.5(c) is exactly the statement that T^* is injective if and only if T is surjective.

Exercise 7.A.7. Prove that if $T \in \mathcal{L}(V, W)$, then

(a) $\dim \operatorname{null} T^* = \dim \operatorname{null} T + \dim W - \dim V$;

(b) $\dim \operatorname{range} T^* = \dim \operatorname{range} T$.

Solution.

(a) We have

$$\begin{aligned} \dim \operatorname{null} T^* = \dim\, (\operatorname{range} T)^\perp = \dim W - \dim \operatorname{range} T \\ = \dim \operatorname{null} T + \dim W - \dim V, \end{aligned}$$

where the first equality is 7.6(a), the second equality is 6.51, and the last equality follows from the fundamental theorem of linear maps (3.21).

(b) We have

$$\dim \operatorname{range} T^* = \dim\, (\operatorname{null} T)^\perp = \dim V - \dim \operatorname{null} T = \dim \operatorname{range} T,$$

where the first equality is 7.6(b), the second equality is 6.51, and the last equality follows from the fundamental theorem of linear maps (3.21).

Exercise 7.A.8. Suppose A is an m-by-n matrix with entries in $\mathbf{F}$. Use (b) in Exercise 7 to prove that the row rank of A equals the column rank of A.

This exercise asks for yet another alternative proof of a result that was previously proved in 3.57 and 3.133.

Solution. For a column vector v with entries in $\mathbf{F}$, let $\overline{v}$ denote the column vector obtained by taking the complex conjugate of each entry of v, e.g.

$$v = \begin{pmatrix} 1-i \\ \pi + 9i \\ \sqrt{2} \end{pmatrix} \quad \Rightarrow \quad \overline{v} = \begin{pmatrix} 1+i \\ \pi - 9i \\ \sqrt{2} \end{pmatrix}.$$

For a matrix M with entries in $\mathbf{F}$, let $\overline{M}$ be the matrix obtained by taking the complex conjugate of each entry of M. Suppose that $v_1, ..., v_\ell$ is a basis of the span of the columns of M. We claim that the list $\overline{v_1}, ..., \overline{v_\ell}$ is linearly independent. Indeed, for scalars $a_1, ..., a_\ell \in \mathbf{F}$, observe that

$$\begin{aligned} a_1\overline{v_1} + \cdots + a_\ell\overline{v_\ell} = 0 \quad &\Rightarrow \quad \overline{a_1}v_1 + \cdots + \overline{a_\ell}v_\ell = 0 \\ &\Rightarrow \quad \overline{a_1} = \cdots = \overline{a_\ell} = 0 \quad \Rightarrow \quad a_1 = \cdots = a_\ell = 0. \end{aligned}$$

It follows that the column rank of $\overline{M}$ is greater than or equal to the column rank of M. By replacing M with $\overline{M}$ in this result, we see that the column rank of $\overline{M}$ must equal the column rank of M.

Let $T \in \mathcal{L}(\mathbf{F}^n, \mathbf{F}^m)$ be such that the matrix of T with respect to the standard orthonormal bases of $\mathbf{F}^n$ and $\mathbf{F}^m$ is A; 7.9 shows that the matrix of T^* with respect to the standard orthonormal bases of $\mathbf{F}^m$ and $\mathbf{F}^n$ is A^*. Using Exercise 7.A.7 and our previous discussion, it follows that

$$\begin{aligned} \text{column rank of } A &= \dim \operatorname{range} T = \dim \operatorname{range} T^* \\ &= \text{column rank of } A^* = \text{column rank of } A^{\mathrm{t}} = \text{row rank of } A. \end{aligned}$$

Exercise 7.A.9. Prove that the product of two self-adjoint operators on V is self-adjoint if and only if the two operators commute.

Solution. Suppose $S, T \in \mathcal{L}(V)$ are self-adjoint and observe that, by 7.5(d),

$$(ST)^* = T^*S^* = TS.$$

Thus $(ST)^* = ST$ if and only if $TS = ST$. That is, ST is self-adjoint if and only if S and T commute.

Exercise 7.A.10. Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$. Prove that T is self-adjoint if and only if

$$\langle Tv, v \rangle = \langle T^*v, v \rangle$$

for all $v \in V$.

Solution. Note that

$$\langle Tv, v \rangle = \langle T^*v, v \rangle \text{ for all } v \in V \quad \Leftrightarrow \quad \langle (T - T^*)v, v \rangle = 0 \text{ for all } v \in V.$$

The desired equivalence then follows from 7.13.

Exercise 7.A.11. Define an operator $S : \mathbf{F}^2 \to \mathbf{F}^2$ by $S(w, z) = (-z, w)$.

(a) Find a formula for S^*.

(b) Show that S is normal but not self-adjoint.

(c) Find all eigenvalues of S.

If $\mathbf{F} = \mathbf{R}$, *then* S *is the operator on* $\mathbf{R}^2$ *of counterclockwise rotation by* $90°$.

Solution.

(a) Observe that

$$\langle S(w, z), (x, y)\rangle = \langle(-z, w), (x, y)\rangle = -zx + wy = \langle(w, z), (y, -x)\rangle.$$

Thus $S^*(x, y) = (y, -x)$.

(b) Certainly $S^* \neq S$, but notice that $S^* = S^{-1}$. It follows that $SS^* = I = S^*S$, so that S is normal.

(c) As shown in Example 5.9, S has no eigenvalues if $\mathbf{F} = \mathbf{R}$ and S has $\pm i$ as eigenvalues if $\mathbf{F} = \mathbf{C}$.

Exercise 7.A.12. An operator $B \in \mathcal{L}(V)$ is called *skew* if

$$B^* = -B.$$

Suppose that $T \in \mathcal{L}(V)$. Prove that T is normal if and only if there exist commuting operators A and B such that A is self-adjoint, B is a skew operator, and $T = A + B$.

Solution. Suppose there exist such operators A and B. Observe that

$$\begin{aligned} TT^* - T^*T &= (A + B)(A^* + B^*) - (A^* + B^*)(A + B) \\ &= (A + B)(A - B) - (A - B)(A + B) = A^2 - B^2 - A^2 + B^2 = 0, \end{aligned}$$

where we have used 7.5(a) for the first equality, that A is self-adjoint and B is skew for the second equality, and that A and B commute for the third equality. Thus T is normal.

Suppose that T is normal and define

$$A = \frac{T + T^*}{2} \quad \text{and} \quad B = \frac{T - T^*}{2}.$$

Certainly $A + B = T$, and 7.5(a), (b), and (c) show that A is self-adjoint and B is skew. Observe that

$$(T + T^*)(T - T^*) - (T - T^*)(T + T^*) = T^2 - (T^*)^2 - T^2 + (T^*)^2 = 0,$$

where we have used that T and T^* commute for the first equality. It follows that $4AB = 4BA$ and hence that A and B commute.

Exercise 7.A.13. Suppose $\mathbf{F} = \mathbf{R}$. Define $\mathcal{A} \in \mathcal{L}(\mathcal{L}(V))$ by $\mathcal{A}T = T^*$ for all $T \in \mathcal{L}(V)$.

(a) Find all eigenvalues of $\mathcal{A}$.

(b) Find the minimal polynomial of $\mathcal{A}$.

Solution.

(a) We are looking for $T \neq 0$ and $\lambda \in \mathbf{R}$ such that $T^* = \lambda T$. Taking the adjoint of both sides of this equation and using 7.5 shows that $T = \lambda T^*$ and thus $T^* = \lambda^2 T^*$. Exercise 7.A.2 shows that $T^* \neq 0$ since $T \neq 0$ and thus $\lambda^2 = 1$, so that ± 1 are the only possible eigenvalues of $\mathcal{A}$. These are indeed eigenvalues of $\mathcal{A}$, since

$$I^* = I \quad \text{and} \quad (-I)^* = -I^* = -I,$$

where we have used 7.5.

(b) By part (a) and 5.27(a), the minimal polynomial of $\mathcal{A}$ has two distinct zeros and hence must have degree at least two. Using 7.5(c), observe that

$$(\mathcal{A}^2 - I)(T) = (T^*)^* - T = 0$$

for any $T \in \mathcal{L}(V)$. Thus $z^2 - 1$ is the minimal polynomial of $\mathcal{A}$.

Exercise 7.A.14. Define an inner product on $\mathcal{P}_2(\mathbf{R})$ by $\langle p, q \rangle = \int_0^1 pq$. Define an operator $T \in \mathcal{L}(\mathcal{P}_2(\mathbf{R}))$ by

$$T(ax^2 + bx + c) = bx.$$

(a) Show that with this inner product, the operator T is not self-adjoint.

(b) The matrix of T with respect to the basis $1, x, x^2$ is

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

This matrix equals its conjugate transpose, even though T is not self-adjoint. Explain why this is not a contradiction.

Solution.

(a) Let $p, q \in \mathcal{P}_2(\mathbf{R})$ be given by $p = 2x$ and $q = 1$, so that $Tp = p$ and $Tq = 0$. Observe that

$$\langle Tp, q \rangle = \int_0^1 2x \, \mathrm{d}x = 1 \neq 0 = \langle p, Tq \rangle.$$

Thus T is not self-adjoint.

(b) The result in 7.9 requires that the basis of $\mathcal{P}_2(\mathbf{R})$ is orthonormal, but $1, x, x^2$ is not an orthonormal basis:

$$\langle 1, x \rangle = \int_0^1 x \, \mathrm{d}x = \tfrac{1}{2} \neq 0.$$

Exercise 7.A.15. Suppose $T \in \mathcal{L}(V)$ is invertible. Prove that

(a) T is self-adjoint $\quad\Leftrightarrow\quad$ T^{-1} is self-adjoint;

(b) T is normal $\quad\Leftrightarrow\quad$ T^{-1} is normal.

Solution.

(a) Suppose that T is self-adjoint. Using 7.5(f), observe that

$$\left(T^{-1}\right)^* = \left(T^*\right)^{-1} = T^{-1}.$$

Thus T^{-1} is self-adjoint. Replacing T with T^{-1} in the previous result and using that $\left(T^{-1}\right)^{-1} = T$ gives us the desired equivalence.

(b) Suppose that T is normal. Using 7.5(f), observe that

$$T^{-1}\left(T^{-1}\right)^* = T^{-1}\left(T^*\right)^{-1} = \left(T^*T\right)^{-1} = \left(TT^*\right)^{-1} = \left(T^*\right)^{-1}T^{-1} = \left(T^{-1}\right)^*T^{-1}.$$

Thus T^{-1} is normal. Replacing T with T^{-1} in the previous result and using that $\left(T^{-1}\right)^{-1} = T$ gives us the desired equivalence.

Exercise 7.A.16. Suppose $\mathbf{F} = \mathbf{R}$.

(a) Show that the set of self-adjoint operators on V is a subspace of $\mathcal{L}(V)$.

(b) What is the dimension of the subspace of $\mathcal{L}(V)$ in (a) [in terms of $\dim V$]?

Solution.

(a) The zero operator is self-adjoint by Exercise 7.A.2; closure under operator addition and closure under scalar multiplication follow from 7.5(a) and 7.5(b).

(b) Suppose $\dim V = n$. By 3.71 and 7.9, it will suffice to find the dimension of the subspace $\mathcal{E}$ of $\mathbf{R}^{n,n}$ consisting of those matrices A such that $A = A^{\mathrm{t}}$. For $k \in \{1, \ldots, n\}$ let $E_{k,k}$ be the matrix with a 1 in the k^{th} diagonal entry and zeros elsewhere, and for $j, k \in \{1, \ldots, n\}$ with $j < k$ let $E_{j,k}$ be the matrix with a 1 in the j^{th} row and k^{th} column, a 1 in the k^{th} row and j^{th} column, and zeros elsewhere. Let $\mathcal{B}$ be the list consisting of the matrices $E_{j,k}$ with $j \leq k$. It is straightforward to verify that $\mathcal{B}$ is linearly independent. Since any $A \in \mathcal{E}$ satisfies $A = A^{\mathrm{t}}$, the entry in row j and column k of A must equal the entry in row k and column j of A. It follows that $\mathcal{B}$ spans $\mathcal{E}$ and hence that $\mathcal{B}$ is a basis of $\mathcal{E}$. A simple counting argument shows that $\mathcal{B}$ has length $n(n+1)/2$ and thus

$$\dim \mathcal{E} = \frac{n(n+1)}{2}.$$

Exercise 7.A.17. Suppose $\mathbf{F} = \mathbf{C}$. Show that the set of self-adjoint operators on V is not a subspace of $\mathcal{L}(V)$.

Solution. 7.5(e) shows that the identity operator I is self-adjoint. Let $v \in V$ be non-zero and observe that

$$(iI)(v) = iv \neq -iv = (\bar{i}I)(v) = (iI)^*(v),$$

where we have used 7.5(b). It follows that iI is not self-adjoint, hence that the set of self-adjoint operators on V is not closed under scalar multiplication, and hence that this set is not a subspace of V.

Exercise 7.A.18. Suppose $\dim V \geq 2$. Show that the set of normal operators on V is not a subspace of $\mathcal{L}(V)$.

Solution. Let $e_1, e_2, \ldots, e_n$ be an orthonormal basis of V and let $S, T \in \mathcal{L}(V)$ be the operators whose matrices with respect to this basis are

$$A = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ -1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

Note that S is self-adjoint and hence normal. Note further that T satisfies $T^* = -T$, so that $TT^* = T^*T = -T^2$; it follows that T is also normal. However, some calculations reveal that

$$(A+B)(A+B)^* = \begin{pmatrix} 2 & -1 & 0 & \cdots & 0 \\ -1 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix} \neq (A+B)^*(A+B) = \begin{pmatrix} 2 & 1 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

Thus $S + T$ is not normal. It follows that the set of normal operators on V is not closed under addition and hence cannot be a subspace of $\mathcal{L}(V)$.

Exercise 7.A.19. Suppose $T \in \mathcal{L}(V)$ and $\|T^*v\| \leq \|Tv\|$ for every $v \in V$. Prove that T is normal.

This exercise fails on infinite-dimensional inner product spaces, leading to what are called hyponormal operators, which have a well-developed theory.

Solution. Let $e_1, \ldots, e_n$ be an orthonormal basis of V. It follows from Exercise 7.A.5 that

$$\|Te_1\|^2 + \cdots + \|Te_n\|^2 = \|T^*e_1\|^2 + \cdots + \|T^*e_n\|^2. \tag{$*$}$$

By assumption we have $\|T^*e_k\| \leq \|Te_k\|$ for each $k \in \{1, ..., n\}$. In fact, each of these inequalities must be an equality, otherwise the right-hand side of $(*)$ would be strictly less than the left-hand side. Because $e_1, ..., e_n$ was an arbitrary orthonormal basis of V, we have now shown that $\|T^*e_1\| = \|Te_1\|$ for any $e_1 \in V$ such that $\|e_1\| = 1$. Thus for any non-zero $v \in V$ we have

$$\left\|T^*\left(\frac{v}{\|v\|}\right)\right\| = \left\|T\left(\frac{v}{\|v\|}\right)\right\| \quad \Rightarrow \quad \|T^*v\| = \|Tv\|.$$

7.20 allows us to conclude that T is normal.

Exercise 7.A.20. Suppose $P \in \mathcal{L}(V)$ is such that $P^2 = P$. Prove that the following are equivalent.

(a) P is self-adjoint.

(b) P is normal.

(c) There is a subspace U of V such that $P = P_U$.

Solution. Certainly (a) implies (b).

Suppose (b) holds and note that

$$\operatorname{null} P = \operatorname{null} P^* = (\operatorname{range} P)^\perp,$$

where the first equality is 7.21(a) and the second equality is 7.6(a). It follows from Exercise 6.C.9 that (c) holds.

Suppose that (c) holds, let $v = u_1 + x_1$ and $w = u_2 + x_2$ be given, where $u_1, u_2 \in U$ and $x_1, x_2 \in U^\perp$, and observe that

$$\langle P_U v, w\rangle = \langle u_1, u_2 + x_2\rangle = \langle u_1, u_2\rangle = \langle u_1 + x_1, u_2\rangle = \langle v, P_U w\rangle.$$

Thus $P_U = P$ is self-adjoint, i.e. (a) holds.

Exercise 7.A.21. Suppose $D : \mathcal{P}_8(\mathbf{R}) \to \mathcal{P}_8(\mathbf{R})$ is the differentiation operator defined by $Dp = p'$. Prove that there does not exist an inner product on $\mathcal{P}_8(\mathbf{R})$ that makes D a normal operator.

Solution. If $T \in \mathcal{L}(V)$ is normal then $\operatorname{null} T^2 = \operatorname{null} T$ (we will prove a stronger result in Exercise 7.A.27), since if $v \in \operatorname{null} T^2$ then $v \in \operatorname{null} T^*T$ by 7.20 and thus

$$\langle T^*Tv, v\rangle = 0 \quad \Leftrightarrow \quad \langle Tv, Tv\rangle = 0 \quad \Leftrightarrow \quad Tv = 0.$$

Notice that $\operatorname{null} D^2 \neq \operatorname{null} D$ since $D^2(x) = 0$ but $D(x) = 1$. It follows from our previous discussion that there does not exist an inner product on $\mathcal{P}_8(\mathbf{R})$ that makes D a normal operator.

Exercise 7.A.22. Give an example of an operator $T \in \mathcal{L}(\mathbf{R}^3)$ such that T is normal but not self-adjoint.

Solution. Let $T \in \mathcal{L}(\mathbf{R}^3)$ be the operator whose matrix with respect to the standard orthonormal basis of $\mathbf{R}^3$ is

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

It follows from 7.9 that $T^* = -T$ and hence that $TT^* = T^*T = -T^2$. Thus T is normal but not self-adjoint (since $T \neq 0$).

Exercise 7.A.23. Suppose T is a normal operator on V. Suppose also that $v, w \in V$ satisfy the equations

$$\|v\| = \|w\| = 2, \quad Tv = 3v, \quad Tw = 4w.$$

Show that $\|T(v+w)\| = 10$.

Solution. Because v and w are eigenvectors of T corresponding to distinct eigenvalues, they must be orthogonal by 7.22. The Pythagorean theorem then implies that

$$\|T(v+w)\|^2 = \|3v+4w\|^2 = \|3v\|^2 + \|4w\|^2 = 36 + 64 = 100.$$

Thus $\|T(v+w)\| = 10$.

Exercise 7.A.24. Suppose $T \in \mathcal{L}(V)$ and

$$a_0 + a_1 z + a_2 z^2 + \cdots + a_{m-1} z^{m-1} + z^m$$

is the minimal polynomial of T. Prove that the minimal polynomial of T^* is

$$\overline{a_0} + \overline{a_1} z + \overline{a_2} z^2 + \cdots + \overline{a_{m-1}} z^{m-1} + z^m.$$

This exercise shows that the minimal polynomial of T^* equals the minimal polynomial of T if $\mathbf{F} = \mathbf{R}$.

Solution. For $p \in \mathcal{P}(\mathbf{F})$, let $\overline{p} \in \mathcal{P}(\mathbf{F})$ be the polynomial whose coefficients are the complex conjugates of the coefficients of p; notice that $\deg \overline{p} = \deg p$. Letting $p \in \mathcal{P}(\mathbf{F})$ be the minimal polynomial of T, i.e.

$$p(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_{m-1} z^{m-1} + z^m,$$

our aim is to show that $\overline{p}$ is the minimal polynomial of T^*. Notice that, by 7.5 and Exercise 7.A.2,

$$p(T) = 0 \quad \Rightarrow \quad [p(T)]^* = 0 \quad \Rightarrow \quad \overline{p}(T^*) = 0.$$

Suppose $s \in \mathcal{P}(\mathbf{F})$ satisfies $\deg s < \deg p$. It follows that $\deg \overline{s} < \deg p$ and hence that

$$\overline{s}(T) \neq 0 \quad \Rightarrow \quad [\overline{s}(T)]^* \neq 0 \quad \Rightarrow \quad s(T^*) \neq 0.$$

Thus the minimal polynomial of T^* must have degree at least $\deg p$. Since $\overline{p}$ is monic, $\deg \overline{p} = \deg p$, and $\overline{p}(T^*) = 0$, we may conclude that $\overline{p}$ is the minimal polynomial of T^*.

Exercise 7.A.25. Suppose $T \in \mathcal{L}(V)$. Prove that T is diagonalizable if and only if T^* is diagonalizable.

Solution. Let $p \in \mathcal{P}(\mathbf{F})$ be the minimal polynomial of T; using the notation of Exercise 7.A.24, $\overline{p}$ is the minimal polynomial of T^*. For $\alpha \in \mathbf{F}$ notice that

$$p(\alpha) = 0 \quad \Leftrightarrow \quad \overline{p(\alpha)} = 0 \quad \Leftrightarrow \quad \overline{p}(\overline{\alpha}) = 0.$$

It follows that if p is of the form $p(z) = (z - \alpha_1) \cdots (z - \alpha_m)$ for some distinct $\alpha_1, ..., \alpha_m \in \mathbf{F}$ then $\overline{p}$ is of the form $\overline{p}(z) = (z - \overline{\alpha_1}) \cdots (z - \overline{\alpha_m})$, where $\overline{\alpha_1}, ..., \overline{\alpha_m}$ are distinct. By 5.62 this is exactly the statement that T^* is diagonalizable if T is diagonalizable. Replacing T with T^* in this implication and using that $(T^*)^* = T$ gives us the desired equivalence.

Exercise 7.A.26. Fix $u, x \in V$. Define $T \in \mathcal{L}(V)$ by $Tv = \langle v, u \rangle x$ for every $v \in V$.

(a) Prove that if V is a real vector space, then T is self-adjoint if and only if the list u, x is linearly dependent.

(b) Prove that T is normal if and only if the list u, x is linearly dependent.

Solution. Note that example 7.3 gives us the formula

$$T^* v = \langle v, x \rangle u.$$

(a) Suppose that u, x is linearly dependent, say $x = \lambda u$ for some $\lambda \in \mathbf{R}$, and observe that

$$Tv = \langle v, u \rangle x = \lambda \langle v, u \rangle u = \langle v, \lambda u \rangle u = \langle v, x \rangle u = T^* v$$

for any $v \in V$. Thus T is self-adjoint.

Now suppose that T is self-adjoint. If $u = 0$ then we are done, so suppose that $u \neq 0$. Since T is self-adjoint we must have

$$Tv = \langle v, u \rangle x = \langle v, x \rangle u = T^* v$$

for every $v \in V$. In particular,

$$\langle u, u \rangle x = \langle u, x \rangle u \quad \Rightarrow \quad x = \frac{\langle u, x \rangle}{\langle u, u \rangle} u,$$

demonstrating that u, x is linearly dependent.

(b) Note that

$$(TT^* - T^*T)v = \langle v, x \rangle \langle u, u \rangle x - \langle v, u \rangle \langle x, x \rangle u$$

for any $v \in V$. Suppose that u, x is linearly dependent, say $x = \lambda u$ for some $\lambda \in \mathbf{F}$, and observe that

$$\begin{aligned}(TT^* - T^*T)v &= \langle v, x\rangle\langle u, u\rangle x - \langle v, u\rangle\langle x, x\rangle u \\ &= |\lambda|^2\langle v, u\rangle\langle u, u\rangle u - |\lambda|^2\langle v, u\rangle\langle u, u\rangle u = 0\end{aligned}$$

for every $v \in V$. Thus T is normal.

Conversely, suppose that T is normal. If $u = 0$ then we are done, so suppose that $u \neq 0$ and observe that

$$(TT^* - T^*T)x = \langle x, x\rangle\langle u, u\rangle x - \langle x, u\rangle\langle x, x\rangle u = 0 \quad \Rightarrow \quad x = \frac{\langle x, u\rangle}{\langle u, u\rangle}u,$$

demonstrating that u, x is linearly dependent.

Exercise 7.A.27. Suppose $T \in \mathcal{L}(V)$ is normal. Prove that

$$\operatorname{null} T^k = \operatorname{null} T \quad \text{and} \quad \operatorname{range} T^k = \operatorname{range} T$$

for every positive integer k.

Solution. We will use induction to prove that $\operatorname{null} T^k = \operatorname{null} T$ for every positive integer k. The base case $k = 1$ is clear, so suppose that the result holds for some positive integer k. Certainly $\operatorname{null} T^k \subseteq \operatorname{null} T^{k+1}$, so suppose that $v \in \operatorname{null} T^{k+1}$. It follows from 7.20 that $v \in \operatorname{null} T^*T^k$ and hence that

$$\left\langle T^*T^k v, T^{k-1}v\right\rangle = 0 \quad \Leftrightarrow \quad \left\langle T^k v, T^k v\right\rangle = 0 \quad \Leftrightarrow \quad T^k v = 0.$$

Thus $\operatorname{null} T^{k+1} = \operatorname{null} T^k = \operatorname{null} T$, where the last equality is our induction hypothesis. This completes the induction step and the proof.

If T is normal then T^* is normal, and 7.5(d) shows that $\left(T^k\right)^* = (T^*)^k$ for every positive integer k. It then follows from the previous result that

$$\begin{aligned}\operatorname{range} T^k &= \left(\operatorname{null}\left(T^k\right)^*\right)^\perp && \text{(7.6(d))} \\ &= \left(\operatorname{null}\left(T^*\right)^k\right)^\perp \\ &= (\operatorname{null} T^*)^\perp && (T^* \text{ is normal}) \\ &= \operatorname{range} T && \text{(7.6(d))}\end{aligned}$$

for any positive integer k.

Exercise 7.A.28. Suppose $T \in \mathcal{L}(V)$ is normal. Prove that if $\lambda \in \mathbf{F}$, then the minimal polynomial of T is not a polynomial multiple of $(x - \lambda)^2$.

Solution. We will prove the contrapositive. Suppose $T \in \mathcal{L}(V)$ has minimal polynomial $p \in \mathcal{P}(\mathbf{F})$ of the form $p(x) = (x-\lambda)^2 q(x)$ for some $\lambda \in \mathbf{F}$ and some $q \in \mathcal{P}(\mathbf{F})$. Because the polynomial $(x-\lambda)q(x)$ has degree strictly less than p, there must exist some $v \in V$ such that $q(T)v \notin \operatorname{null}(T-\lambda I)$. Since $p(T) = 0$ we have $q(T)v \in \operatorname{null}(T-\lambda I)^2$ and thus $\operatorname{null}(T-\lambda I)^2 \neq \operatorname{null}(T-\lambda I)$. It follows from Exercise 7.A.27 that $T - \lambda I$ is not normal. The contrapositive of 7.21(d) allows us to conclude that T is not normal.

Exercise 7.A.29. Prove or give a counterexample: If $T \in \mathcal{L}(V)$ and there is an orthonormal basis $e_1, \ldots, e_n$ of V such that $\|Te_k\| = \|T^* e_k\|$ for each $k = 1, \ldots, n$, then T is normal.

Solution. This is false. Let T be the operator on $\mathbf{F}^2$ whose matrix with respect to the standard orthonormal basis e_1, e_2 of $\mathbf{F}^2$ is

$$\begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix};$$

As we showed in Exercise 7.A.18, T is not normal. However,

$$\|Te_1\| = \|T^* e_1\| = \sqrt{2} \quad \text{and} \quad \|Te_2\| = \|T^* e_2\| = 1.$$

Exercise 7.A.30. Suppose that $T \in \mathcal{L}(\mathbf{F}^3)$ is normal and $T(1,1,1) = (2,2,2)$. Suppose $(z_1, z_2, z_3) \in \operatorname{null} T$. Prove that $z_1 + z_2 + z_3 = 0$.

Solution. If $u := (z_1, z_2, z_3) = 0$ then we are done, so suppose that $u \neq 0$. It follows that u is an eigenvector of T corresponding to the eigenvalue 0. Note that $v := (1,1,1)$ is an eigenvector of T corresponding to the eigenvalue 2. Since these are eigenvectors of a normal operator corresponding to distinct eigenvalues, they must be orthogonal by 7.22. That is,

$$\langle u, v \rangle = z_1 + z_2 + z_3 = 0.$$

Exercise 7.A.31. Fix a positive integer n. In the inner product space of continuous real-valued functions on $[-\pi, \pi]$ with inner product $\langle f, g \rangle = \int_{-\pi}^{\pi} fg$, let

$$V = \operatorname{span}(1, \cos x, \cos 2x, \ldots, \cos nx, \sin x, \sin 2x, \ldots, \sin nx).$$

(a) Define $D \in \mathcal{L}(V)$ by $Df = f'$. Show that $D^* = -D$. Conclude that D is normal but not self-adjoint.

(b) Define $T \in \mathcal{L}(V)$ by $Tf = f''$. Show that T is self-adjoint.

Solution.

(a) For each $k \in \{1, \ldots, n\}$, let

$$v = \frac{1}{\sqrt{2\pi}}, \quad e_k = \frac{\cos kx}{\sqrt{\pi}}, \quad \text{and} \quad f_k = \frac{\sin kx}{\sqrt{\pi}}.$$

If we let $\mathcal{B} = v, e_1, ..., e_n, f_1, ..., f_n$, then $\mathcal{B}$ is an orthonormal basis of V, as shown in Exercise 6.B.4. Observe that $Dv = 0, De_k = -kf_k$, and $Df_k = ke_k$ for each $k \in \{1, ..., n\}$. It follows from 7.9 that

$$D^*v = 0 = -Dv, \quad D^*e_k = kf_k = -De_k, \quad \text{and} \quad D^*f_k = -ke_k = -Df_k$$

for each $k \in \{1, ..., n\}$. Thus $D^* = -D$, so that D^* is normal ($D^*D = DD^* = -D^2$) but not self-adjoint (since $V \neq 0$).

(b) Notice that $T = D^2$. It follows from 7.5 that

$$T^* = \left(D^2\right)^* = \left(D^*\right)^2 = (-D)^2 = D^2 = T.$$

Thus T is self-adjoint.

Exercise 7.A.32. Suppose $T : V \to W$ is a linear map. Show that under the standard identification of V with V' (see 6.58) and the corresponding identification of W with W', the adjoint map $T^* : W \to V$ corresponds to the dual map $T' : W' \to V'$. More precisely, show that

$$T'(\varphi_w) = \varphi_{T^*w}$$

for all $w \in W$, where φ_w and φ_{T^*w} are defined as in 6.58.

Solution. For any $v \in V$ and $w \in W$, observe that

$$[T'(\varphi_w)]v = \varphi_w(Tv) = \langle Tv, w\rangle = \langle v, T^*w\rangle = \varphi_{T^*w}(v).$$

7.B. Spectral Theorem

Exercise 7.B.1. Prove that a normal operator on a complex inner product space is self-adjoint if and only if all its eigenvalues are real.

This exercise strengthens the analogy (for normal operators) between self-adjoint operators and real numbers.

Solution. Suppose V is a complex inner product space and $T \in \mathcal{L}(V)$ is normal. If T is self-adjoint then 7.12 shows that every eigenvalue of T is real. Suppose that every eigenvalue of T is real. The complex spectral theorem (7.31) implies that there is an orthonormal basis of V with respect to which the matrix $\mathcal{M}(T)$ is diagonal. Because each eigenvalue of T is real, the diagonal entries of $\mathcal{M}(T)$ must be real. It follows that $\mathcal{M}(T)$ equals its own conjugate transpose and hence that T is self-adjoint.

Exercise 7.B.2. Suppose $\mathbf{F} = \mathbf{C}$. Suppose $T \in \mathcal{L}(V)$ is normal and has only one eigenvalue. Prove that T is a scalar multiple of the identity operator.

Solution. Suppose that $\lambda \in \mathbf{C}$ is the sole eigenvalue of T. The complex spectral theorem (7.31) implies that there is an orthonormal basis of V with respect to which $\mathcal{M}(T)$ is diagonal. Because λ is the only eigenvalue of T, each diagonal entry of $\mathcal{M}(T)$ must be equal to λ. Thus $T = \lambda I$.

Exercise 7.B.3. Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$ is normal. Prove that the set of eigenvalues of T is contained in $\{0, 1\}$ if and only if there is a subspace U of V such that $T = P_U$.

Solution. If there exists a subspace U of V such that $T = P_U$ then $T^2 = T$ and it follows from Exercise 5.A.8 that the set of eigenvalues of T is contained in $\{0, 1\}$.

Suppose that the set of eigenvalues of T is contained in $\{0, 1\}$. The complex spectral theorem (7.31) implies that there is an orthonormal basis $\mathcal{B}$ of V consisting of eigenvectors of T. Each basis vector in $\mathcal{B}$ must correspond either to the eigenvalue 0 or the eigenvalue 1. Let $u_1, \ldots, u_m$ be those basis vectors in $\mathcal{B}$ corresponding to the eigenvalue 1 and let $v_1, \ldots, v_n$ be those basis vectors in $\mathcal{B}$ corresponding to the eigenvalue 0; either of these lists may be empty. Let $U = \operatorname{span}(u_1, \ldots, u_m)$. Because $\mathcal{B}$ is an orthonormal basis of V, it follows that $U^\perp = \operatorname{span}(v_1, \ldots, v_n)$ and hence that

$$P_U u_k = u_k = T u_k \quad \text{and} \quad P_U v_k = 0 = T v_k.$$

Thus $T = P_U$.

Exercise 7.B.4. Prove that a normal operator on a complex inner product space is skew (meaning that it equals the negative of its adjoint) if and only if all its eigenvalues are purely imaginary (meaning that they have real part equal to 0).

Solution. Suppose V is a complex inner product space and $T \in \mathcal{L}(V)$ is normal. By the complex spectral theorem (7.31) there is an orthonormal basis of V with respect to which $\mathcal{M}(T)$ is of the form

$$\begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix},$$

where $\lambda_1, \dots, \lambda_n \in \mathbf{C}$ are the eigenvalues of T. Observe that

$$T \text{ is skew} \quad \Leftrightarrow \quad \mathcal{M}(T) = -[\mathcal{M}(T)]^* \quad \Leftrightarrow \quad \text{each } \lambda_k = -\overline{\lambda_k}$$
$$\Leftrightarrow \quad \text{each } \lambda_k \text{ is purely imaginary} \quad \Leftrightarrow \quad \text{each eigenvalue of } T \text{ is purely imaginary.}$$

Exercise 7.B.5. Prove or give a counterexample: If $T \in \mathcal{L}(\mathbf{C}^3)$ is a diagonalizable operator, then T is normal (with respect to the usual inner product).

Solution. This is false. Consider the basis

$$v_1 = (1,0,0), \quad v_2 = (0,1,0), \quad v_3 = (1,0,1)$$

of $\mathbf{C}^3$ and define $T \in \mathcal{L}(\mathbf{C}^3)$ by

$$Tv_1 = v_1, \quad Tv_2 = v_2, \quad Tv_3 = 2v_3.$$

Observe that T is diagonalizable since $\mathbf{C}^3$ has a basis v_1, v_2, v_3 consisting of eigenvectors of T. Observe further that v_1, v_3 are eigenvectors of T corresponding to distinct eigenvalues, and that v_1 and v_3 are not orthogonal: $\langle v_1, v_3 \rangle = 1$. It follows from 7.22 that T is not normal.

Exercise 7.B.6. Suppose V is a complex inner product space and $T \in \mathcal{L}(V)$ is a normal operator such that $T^9 = T^8$. Prove that T is self-adjoint and $T^2 = T$.

Solution. The complex spectral theorem (7.31) implies that there is an orthonormal basis $e_1, \dots, e_n$ of V consisting of eigenvectors of T, so that $Te_k = \lambda_k e_k$ for each $k \in \{1, \dots, n\}$, where $\lambda_1, \dots, \lambda_n$ are the eigenvalues of T. By assumption we have

$$T^9 e_k = T^8 e_k \quad \Leftrightarrow \quad \lambda_k^9 e_k = \lambda_k^8 e_k \quad \Leftrightarrow \quad \lambda_k^9 = \lambda_k^8 \quad \Leftrightarrow \quad \lambda_k \in \{0, 1\}.$$

It follows from Exercise 7.B.3 that $T = P_U$ for some subspace U of V, so that $T^2 = T$, and Exercise 7.B.1 (or Exercise 7.A.20) shows that T is self-adjoint.

Exercise 7.B.7. Give an example of an operator T on a complex vector space such that $T^9 = T^8$ but $T^2 \neq T$.

Solution. Let T be the operator on $\mathbf{C}^2$ whose matrix with respect to the standard basis of $\mathbf{C}^2$ is

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

It follows that $T^9 = T^8 = T^2 = 0$ but $T \neq 0$.

Exercise 7.B.8. Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$. Prove that T is normal if and only if every eigenvector of T is also an eigenvector of T^*.

Solution. If T is normal then every eigenvector of T is an eigenvector of T^* by 7.21(e).

Suppose that every eigenvector of T is also an eigenvector of T^*. By Schur's theorem (6.38) there is an orthonormal basis $e_1, \dots, e_n$ of V with respect to which $A := \mathcal{M}(T)$ is upper-triangular. It follows that $Te_1 = A_{1,1}e_1$, so that e_1 is an eigenvector of T. Our assumption implies that e_1 is also an eigenvector of T^*, say $T^*e_1 = \mu_1 e_1$. On the other hand, by 7.9,

$$T^*e_1 = \overline{A_{1,1}}e_1 + \overline{A_{1,2}}e_2 + \cdots + \overline{A_{1,n}}e_n.$$

It follows from unique representation that $A_{1,2} = \cdots = A_{1,n} = 0$. Thus $Te_2 = A_{2,2}e_2$, so that e_2 is an eigenvector of T. Our assumption implies that e_2 is also an eigenvector of T^*, say $T^*e_2 = \mu_2 e_2$. On the other hand, by 7.9,

$$T^*e_2 = \overline{A_{2,2}}e_2 + \overline{A_{2,3}}e_3 + \cdots + \overline{A_{2,n}}e_n.$$

It follows from unique representation that $A_{2,3} = \cdots = A_{2,n} = 0$. Continuing in this manner, we see that A is a diagonal matrix. The complex spectral theorem allows us to conclude that T is normal.

Exercise 7.B.9. Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$. Prove that T is normal if and only if there exists a polynomial $p \in \mathcal{P}(\mathbf{C})$ such that $T^* = p(T)$.

Solution. If there exists such a polynomial p then T commutes with $T^* = p(T)$ by 5.17(b). Thus T is normal.

Suppose that T is normal. By the complex spectral theorem (7.31), there is an orthonormal basis $e_1, \dots, e_n$ of V such that each e_k is an eigenvector of T. Let $\lambda_1, \dots, \lambda_m$ be the distinct eigenvalues of T. Exercise 4.7 shows that there is a polynomial $p \in \mathcal{P}(\mathbf{C})$ satisfying $p(\lambda_j) = \overline{\lambda_j}$ for each $j \in \{1, \dots, m\}$. For any $k \in \{1, \dots, n\}$ we have $Te_k = \lambda_j e_k$ for some $j \in \{1, \dots, m\}$. It follows that

$$p(T)e_k = p(\lambda_j)e_k = \overline{\lambda_j}e_k = T^*e_k,$$

where we have used 7.21(e) for the last equality. Thus $p(T) = T^*$.

Exercise 7.B.10. Suppose V is a complex inner product space. Prove that every normal operator on V has a square root.

*An operator $S \in \mathcal{L}(V)$ is called a **square root** of $T \in \mathcal{L}(V)$ if $S^2 = T$. We will discuss more about square roots of operators in Sections 7C and 8C.*

Solution. Let $T \in \mathcal{L}(V)$ be normal. The complex spectral theorem (7.31) implies that there is an orthonormal basis $e_1, \ldots, e_n$ of V such that $Te_k = \lambda_k e_k$ for each $k \in \{1, \ldots, n\}$, where $\lambda_1, \ldots, \lambda_n \in \mathbf{C}$ are the eigenvalues of T. Because any complex number has a square root, for each k there exists some $\mu_k \in \mathbf{C}$ such that $\mu_k^2 = \lambda_k$. Define $S \in \mathcal{L}(V)$ by $Se_k = \mu_k e_k$ and observe that

$$S^2 e_k = \mu_k^2 e_k = \lambda_k e_k = Te_k.$$

Thus $S^2 = T$.

Exercise 7.B.11. Prove that every self-adjoint operator on V has a cube root.

*An operator $S \in \mathcal{L}(V)$ is called a **cube root** of $T \in \mathcal{L}(V)$ if $S^3 = T$.*

Solution. Let $T \in \mathcal{L}(V)$ be self-adjoint. The relevant spectral theorem (7.29 or 7.31) implies that there is an orthonormal basis $e_1, \ldots, e_n$ of V such that $Te_k = \lambda_k e_k$ for each $k \in \{1, \ldots, n\}$, where $\lambda_1, \ldots, \lambda_n \in \mathbf{F}$ are the eigenvalues of T. Note that each λ_k must be real by 7.12. Because any real number has a cube root, for each k there exists some $\mu_k \in \mathbf{R}$ such that $\mu_k^3 = \lambda_k$. Define $S \in \mathcal{L}(V)$ by $Se_k = \mu_k e_k$ and observe that

$$S^3 e_k = \mu_k^3 e_k = \lambda_k e_k = Te_k.$$

Thus $S^3 = T$.

Exercise 7.B.12. Suppose V is a complex vector space and $T \in \mathcal{L}(V)$ is normal. Prove that if S is an operator on V that commutes with T, then S commutes with T^*.

The result in this exercise is called Fuglede's theorem.

Solution. By Exercise 7.B.9 there is a polynomial $p = \sum_{k=0}^{m} a_k z^k \in \mathcal{P}(\mathbf{C})$ such that $T^* = p(T)$. Since $ST = TS$, a straightforward induction argument shows that $ST^k = T^k S$ for any non-negative integer k. It follows that

$$ST^* = Sp(T) = S\left(\sum_{k=0}^{m} a_k T^k\right) = \sum_{k=0}^{m} a_k ST^k = \sum_{k=0}^{m} a_k T^k S = \left(\sum_{k=0}^{m} a_k T^k\right) S = p(T)S = T^*S.$$

Exercise 7.B.13. Without using the complex spectral theorem, use the version of Schur's theorem that applies to two commuting operators (take $\mathcal{E} = \{T, T^*\}$ in Exercise 20 in Section 6B) to give a different proof that if $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$ is normal, then T has a diagonal matrix with respect to some orthonormal basis of V.

Solution. Since T and T^* commute, Exercise 6.B.20 implies that there is an orthonormal basis $e_1, ..., e_n$ of V such that $\mathcal{M}(T)$ and $\mathcal{M}(T^*)$ are upper-triangular. Because $\mathcal{M}(T^*)$ is the conjugate transpose of $\mathcal{M}(T)$ (by 7.9), it must be that $\mathcal{M}(T)$ is also lower-triangular. Thus $\mathcal{M}(T)$ is diagonal.

Exercise 7.B.14. Suppose $\mathbf{F} = \mathbf{R}$ and $T \in \mathcal{L}(V)$. Prove that T is self-adjoint if and only if all pairs of eigenvectors corresponding to distinct eigenvalues of T are orthogonal and $V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)$, where $\lambda_1, ..., \lambda_m$ denote the distinct eigenvalues of T.

Solution. If T is self-adjoint then 7.22 shows that all pairs of eigenvectors corresponding to distinct eigenvalues of T are orthogonal, and the real spectral theorem (7.29) shows that T is diagonalizable, which by 5.55 is equivalent to

$$V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T),$$

where $\lambda_1, ..., \lambda_m$ are the distinct eigenvalues of T.

Now suppose that all pairs of eigenvectors corresponding to distinct eigenvalues of T are orthogonal and

$$V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T),$$

where $\lambda_1, ..., \lambda_m$ are the distinct eigenvalues of T. Choose an orthonormal basis for each eigenspace $E(\lambda_k, T)$. Our hypotheses ensure that the list obtained by concatenating these orthonormal bases is an orthonormal basis of V consisting of eigenvectors of T. It follows from the real spectral theorem (7.29) that T is self-adjoint.

Exercise 7.B.15. Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$. Prove that T is normal if and only if all pairs of eigenvectors corresponding to distinct eigenvalues of T are orthogonal and $V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)$, where $\lambda_1, ..., \lambda_m$ denote the distinct eigenvalues of T.

Solution. If T is normal then 7.22 shows that all pairs of eigenvectors corresponding to distinct eigenvalues of T are orthogonal, and the complex spectral theorem (7.31) shows that T is diagonalizable, which by 5.55 is equivalent to

$$V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T),$$

where $\lambda_1, ..., \lambda_m$ are the distinct eigenvalues of T.

Now suppose that all pairs of eigenvectors corresponding to distinct eigenvalues of T are orthogonal and

$$V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T),$$

where $\lambda_1, ..., \lambda_m$ are the distinct eigenvalues of T. Choose an orthonormal basis for each eigenspace $E(\lambda_k, T)$. Our hypotheses ensure that the list obtained by concatenating these orthonormal bases is an orthonormal basis of V consisting of eigenvectors of T. It follows from the complex spectral theorem (7.31) that T is normal.

Exercise 7.B.16. Suppose $\mathbf{F} = \mathbf{C}$ and $\mathcal{E} \subseteq \mathcal{L}(V)$. Prove that there is an orthonormal basis of V with respect to which every element of $\mathcal{E}$ has a diagonal matrix if and only if S and T are commuting normal operators for all $S, T \in \mathcal{E}$.

This exercise extends the complex spectral theorem to the context of a collection of commuting normal operators.

Solution. If there exists such an orthonormal basis of V then each $T \in \mathcal{E}$ is normal by the complex spectral theorem (7.31) and each pair $S, T \in \mathcal{E}$ commutes by 5.74 (since diagonal matrices always commute).

Suppose S and T are commuting normal operators for all $S, T \in \mathcal{E}$. By Exercise 6.B.20, there is an orthonormal basis of V with respect to which the matrix of each $T \in \mathcal{E}$ is upper-triangular. Because T is normal this matrix must actually be diagonal, as the proof of the complex spectral theorem (7.31) shows.

Exercise 7.B.17. Suppose $\mathbf{F} = \mathbf{R}$ and $\mathcal{E} \subseteq \mathcal{L}(V)$. Prove that there is an orthonormal basis of V with respect to which every element of $\mathcal{E}$ has a diagonal matrix if and only if S and T are commuting self-adjoint operators for all $S, T \in \mathcal{E}$.

This exercise extends the real spectral theorem to the context of a collection of commuting self-adjoint operators.

Solution. If there exists such an orthonormal basis of V then each $T \in \mathcal{E}$ is self-adjoint by the real spectral theorem (7.29) and each pair $S, T \in \mathcal{E}$ commutes by 5.74 (since diagonal matrices always commute).

Suppose S and T are commuting self-adjoint operators for all $S, T \in \mathcal{E}$. The real spectral theorem (7.29) shows that each $T \in \mathcal{E}$ is diagonalizable and thus by Exercise 5.E.2 there is a basis $v_1, ..., v_n$ of V with respect to which the matrix of each $T \in \mathcal{E}$ is diagonal. Perform the Gram-Schmidt procedure on $v_1, ..., v_n$ to obtain an orthonormal basis $e_1, ..., e_n$ of V such that

$$\operatorname{span}(e_1, ..., e_k) = \operatorname{span}(v_1, ..., v_k)$$

for each $k \in \{1, ..., n\}$. Let $T \in \mathcal{E}$ be given and note that $\operatorname{span}(v_1, ..., v_k) = \operatorname{span}(e_1, ..., e_k)$ is invariant under T for each $k \in \{1, ..., n\}$ since $\mathcal{M}(T, (v_1, ..., v_n))$ is diagonal. Thus, by 5.39,

$\mathcal{M}(T, (e_1, ..., e_n))$ is upper-triangular. Because T is self-adjoint this matrix must actually be diagonal, as the proof of the real spectral theorem (7.29) shows.

Exercise 7.B.18. Give an example of a real inner product space V, an operator $T \in \mathcal{L}(V)$, and real numbers $b^2 < 4c$ such that

$$T^2 + bT + cI$$

is not invertible.

> *This exercise shows that the hypothesis that T is self-adjoint cannot be deleted in 7.26, even for real vector spaces.*

Solution. Let $V = \mathbf{R}^2$ with the usual inner product and let $T \in \mathcal{L}(\mathbf{R}^2)$ be a counterclockwise rotation about the origin by 90°, so that the matrix of T with respect to the standard orthonormal basis of $\mathbf{R}^2$ is

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

Observe that $T^2 + I$ (i.e. taking $b = 0$ and $c = 1$) is zero and hence not invertible.

Exercise 7.B.19. Suppose $T \in \mathcal{L}(V)$ is self-adjoint and U is a subspace of V that is invariant under T.

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

(b) Prove that $T|_U \in \mathcal{L}(U)$ is self-adjoint.

(c) Prove that $T|_{U^\perp} \in \mathcal{L}(U^\perp)$ is self-adjoint.

Solution.

(a) This is immediate from Exercise 7.A.4 and the fact that $T^* = T$.

(b) Because T is self-adjoint we have $\langle Tx, y\rangle = \langle x, Ty\rangle$ for all $x, y \in V$. In particular this equality holds for all $x, y \in U$. It follows that $T|_U$ is self-adjoint.

(c) This follows by replacing U with $U^\perp$ in part (b), which is valid by part (a).

Exercise 7.B.20. Suppose $T \in \mathcal{L}(V)$ is normal and U is a subspace of V that is invariant under T.

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

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

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

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

This exercise can be used to give yet another proof of the complex spectral theorem (use induction on $\dim V$ *and the result that* T *has an eigenvector).*

Solution.

(a) Let $e_1, ..., e_m$ be an orthonormal basis of U and let $e_{m+1}, ..., e_n$ be an orthonormal basis of $U^\perp$, so that $e_1, ..., e_n$ is an orthonormal basis of V. Because U is invariant under T, the matrix of T with respect to $e_1, ..., e_n$ is of the form

$$\begin{pmatrix} A_{1,1} & \cdots & A_{1,m} & A_{1,m+1} & \cdots & A_{1,n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ A_{m,1} & \cdots & A_{m,m} & A_{m,m+1} & \cdots & A_{m,n} \\ 0 & \cdots & 0 & A_{m+1,m+1} & \cdots & A_{m+1,n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & A_{n,m+1} & \cdots & A_{n,n} \end{pmatrix}.$$

It follows from 7.9 and the matrix above that

$$\sum_{k=1}^{m} \|Te_k\|^2 = \sum_{k=1}^{m}\sum_{j=1}^{m} |A_{j,k}|^2 \quad \text{and} \quad \sum_{k=1}^{m} \|T^*e_k\|^2 = \sum_{k=1}^{m}\sum_{j=1}^{n} |A_{k,j}|^2 = \sum_{k=1}^{n}\sum_{j=1}^{m} |A_{j,k}|^2,$$

where we have swapped the indices j and k for the last equality. For each $k \in \{1, ..., m\}$ we have $\|Te_k\|^2 = \|T^*e_k\|^2$ and thus

$$0 = \sum_{k=1}^{m} \|T^*e_k\|^2 - \sum_{k=1}^{m} \|Te_k\|^2 = \sum_{k=1}^{n}\sum_{j=1}^{m} |A_{j,k}|^2 - \sum_{k=1}^{m}\sum_{j=1}^{m} |A_{j,k}|^2 = \sum_{k=m+1}^{n}\sum_{j=1}^{m} |A_{j,k}|^2.$$

It follows that $A_{j,k} = 0$ for each $j \in \{1, ..., m\}$ and each $k \in \{m+1, ..., n\}$, so that the matrix of T with respect to $e_1, ..., e_n$ is of the form

$$\begin{pmatrix} A_{1,1} & \cdots & A_{1,m} & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ A_{m,1} & \cdots & A_{m,m} & 0 & \cdots & 0 \\ 0 & \cdots & 0 & A_{m+1,m+1} & \cdots & A_{m+1,n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & A_{n,m+1} & \cdots & A_{n,n} \end{pmatrix}.$$

Thus $U^\perp$ is invariant under T.

(b) With respect to the orthonormal basis $e_1, ..., e_n$ of V from part (a), 7.9 shows that the matrix of T^* is

$$\begin{pmatrix} \overline{A_{1,1}} & \cdots & \overline{A_{m,1}} & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \overline{A_{1,m}} & \cdots & \overline{A_{m,m}} & 0 & \cdots & 0 \\ 0 & \cdots & 0 & \overline{A_{m+1,m+1}} & \cdots & \overline{A_{n,m+1}} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \overline{A_{m+1,n}} & \cdots & \overline{A_{n,n}} \end{pmatrix}.$$

Thus U is invariant under T^*.

(c) With respect to the orthonormal basis $e_1, ..., e_m$ of U from part (a) we have

$$\mathcal{M}(T|_U) = \begin{pmatrix} A_{1,1} & \cdots & A_{1,m} \\ \vdots & \ddots & \vdots \\ A_{m,1} & \cdots & A_{m,m} \end{pmatrix} \quad \Rightarrow \quad \mathcal{M}((T|_U)^*) = \begin{pmatrix} \overline{A_{1,1}} & \cdots & \overline{A_{m,1}} \\ \vdots & \ddots & \vdots \\ \overline{A_{1,m}} & \cdots & \overline{A_{m,m}} \end{pmatrix}.$$

Part (b) shows that

$$\mathcal{M}((T^*)|_U) = \begin{pmatrix} \overline{A_{1,1}} & \cdots & \overline{A_{m,1}} \\ \vdots & \ddots & \vdots \\ \overline{A_{1,m}} & \cdots & \overline{A_{m,m}} \end{pmatrix} = \mathcal{M}((T|_U)^*).$$

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

(d) Using part (c), notice that

$$(T|_U)^* T|_U = (T^*)|_U \; T|_U = (T^*T)|_U = (TT^*)|_U = T|_U \; (T^*)|_U = T|_U \; (T|_U)^*.$$

Thus $T|_U$ is normal. Replacing U with $U^\perp$ in this result, which is valid by part (a), shows that $T|_{U^\perp}$ is normal.

Exercise 7.B.21. Suppose that T is a self-adjoint operator on a finite-dimensional inner product space and that 2 and 3 are the only eigenvalues of T. Prove that

$$T^2 - 5T + 6I = 0.$$

Solution. Exercise 7.B.14 if $\mathbf{F} = \mathbf{R}$, or Exercise 7.B.15 if $\mathbf{F} = \mathbf{C}$, shows that

$$V = E(2, T) \oplus E(3, T).$$

Thus any $v \in V$ is of the form $v = x + y$, where $x \in E(2, T)$ and $y \in E(3, T)$. It follows that

$$(T^2 - 5T + 6I)v = (T - 3I)(T - 2I)x + (T - 2I)(T - 3I)y = 0.$$

Exercise 7.B.22. Give an example of an operator $T \in \mathcal{L}(\mathbf{C}^3)$ such that 2 and 3 are the only eigenvalues of T and $T^2 - 5T + 6I \neq 0$.

Solution. Let e_1, e_2, e_3 be the standard basis of $\mathbf{C}^3$ and define $T \in \mathcal{L}(\mathbf{C}^3)$ by

$$Te_1 = 2e_1, \quad Te_2 = e_1 + 2e_2, \quad Te_3 = 3e_3,$$

so that the matrix of T with respect to e_1, e_2, e_3 is

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

Since this matrix is upper-triangular, we see that 2 and 3 are the only eigenvalues of T. However, notice that

$$(T^2 - 5T + 6I)e_2 = -e_1 \neq 0.$$

Exercise 7.B.23. Suppose $T \in \mathcal{L}(V)$ is self-adjoint, $\lambda \in \mathbf{F}$, and $\varepsilon > 0$. Suppose there exists $v \in V$ such that $\|v\| = 1$ and

$$\|Tv - \lambda v\| < \varepsilon.$$

Prove that T has an eigenvalue λ' such that $|\lambda - \lambda'| < \varepsilon$.

This exercise shows that for a self-adjoint operator, a number that is close to satisfying an equation that would make it an eigenvalue is close to an eigenvalue.

Solution. The relevant spectral theorem (7.29 or 7.31) implies that there is an orthonormal basis $e_1, ..., e_n$ of V consisting of eigenvectors of T, so that $Te_k = \lambda_k e_k$ for each $k \in \{1, ..., n\}$, where $\lambda_1, ..., \lambda_n \in \mathbf{F}$ are the eigenvalues of T. It follows from 6.30 and that

$$\|Tv - \lambda v\|^2 = |\lambda_1 - \lambda|^2 |\langle v, e_1 \rangle|^2 + \cdots + |\lambda_n - \lambda|^2 |\langle v, e_n \rangle|^2.$$

Let λ' be the eigenvalue in $\{\lambda_1, ..., \lambda_n\}$ which minimizes $|\lambda_k - \lambda|$ and observe that

$$\begin{aligned} |\lambda' - \lambda|^2 &= |\lambda' - \lambda|^2 \|v\|^2 \\ &= |\lambda' - \lambda|^2 |\langle v, e_1 \rangle|^2 + \cdots + |\lambda' - \lambda|^2 |\langle v, e_n \rangle|^2 \\ &\leq |\lambda_1 - \lambda|^2 |\langle v, e_1 \rangle|^2 + \cdots + |\lambda_n - \lambda|^2 |\langle v, e_n \rangle|^2 \\ &= \|Tv - \lambda v\|^2 \\ &< \varepsilon^2. \end{aligned}$$

Thus $|\lambda' - \lambda| < \varepsilon$.

Exercise 7.B.24. Suppose U is a finite-dimensional vector space and $T \in \mathcal{L}(U)$.

(a) Suppose $\mathbf{F} = \mathbf{R}$. Prove that T is diagonalizable if and only if there is a basis of U such that the matrix of T with respect to this basis equals its transpose.

(b) Suppose $\mathbf{F} = \mathbf{C}$. Prove that T is diagonalizable if and only if there is a basis of U such that the matrix of T with respect to this basis commutes with its conjugate transpose.

This exercise adds another equivalence to the list of conditions equivalent to diagonalizability in 5.55.

Solution.

(a) If T is diagonalizable then there is a basis $u_1, ..., u_n$ of U such that

$$A := \mathcal{M}(T, (u_1, ..., u_n))$$

is diagonal. It follows that $A = A^{\mathrm{t}}$.

Now suppose that there is a basis $u_1, ..., u_n$ of U such that $A = A^{\mathrm{t}}$, where

$$A = \mathcal{M}(T, (u_1, ..., u_n)).$$

For $u = a_1 u_1 + \cdots + a_n u_n$ and $v = b_1 u_1 + \cdots + b_n u_n$ in U, define

$$\langle u, v \rangle = a_1 b_1 + \cdots + a_n b_n;$$

it is straightforward to verify that this defines an inner product on U. Notice that $u_1, ..., u_n$ is an orthonormal basis of U with respect to this inner product; it follows that T is self-adjoint with respect to this inner product and so we may apply the real spectral theorem (7.29) to obtain a basis of U consisting of eigenvectors of T. Thus T is diagonalizable.

(b) If T is diagonalizable then there is a basis $u_1, ..., u_n$ of U such that

$$A := \mathcal{M}(T, (u_1, ..., u_n))$$

is diagonal. It follows that A^* is also diagonal and hence that $AA^* = A^*A$, since diagonal matrices always commute.

Now suppose that there is a basis $u_1, ..., u_n$ of U such that $AA^* = A^*A$, where

$$A = \mathcal{M}(T, (u_1, ..., u_n)).$$

For $u = a_1 u_1 + \cdots + a_n u_n$ and $v = b_1 u_1 + \cdots + b_n u_n$ in U, define

$$\langle u, v \rangle = a_1 \overline{b_1} + \cdots + a_n \overline{b_n};$$

it is straightforward to verify that this defines an inner product on U. Notice that $u_1, ..., u_n$ is an orthonormal basis of U with respect to this inner product; it follows that T is normal with respect to this inner product and so we may apply the complex spectral theorem (7.31) to obtain a basis of U consisting of eigenvectors of T. Thus T is diagonalizable.

Exercise 7.B.25. Suppose that $T \in \mathcal{L}(V)$ and there is an orthonormal basis $e_1, ..., e_n$ of V consisting of eigenvectors of T, with corresponding eigenvalues $\lambda_1, ..., \lambda_n$. Show that if $k \in \{1, ..., n\}$, then the pseudoinverse $T^\dagger$ satisfies the equation

$$T^\dagger e_k = \begin{cases} \frac{1}{\lambda_k} e_k & \text{if } \lambda_k \neq 0, \\ 0 & \text{if } \lambda_k = 0. \end{cases}$$

Solution. By the relevant spectral theorem (7.29 or 7.31), T is either self-adjoint if $\mathbf{F} = \mathbf{R}$ or normal if $\mathbf{F} = \mathbf{C}$. In either case, 7.21 shows that

$$\operatorname{null} T = \operatorname{null} T^* \quad \text{and} \quad \operatorname{range} T = \operatorname{range} T^*.$$

Let $k \in \{1, ..., n\}$ be given. If $\lambda_k = 0$ then

$$e_k \in \operatorname{null} T = \operatorname{null} T^* = (\operatorname{range} T)^\perp = \operatorname{null} P_{\operatorname{range} T},$$

where we have used 6.57(e) and 7.6(a). It follows that

$$T^\dagger e_k = (T|_{(\operatorname{null} T)^\perp})^{-1} P_{\operatorname{range} T} e_k = (T|_{(\operatorname{null} T)^\perp})^{-1}(0) = 0.$$

If $\lambda_k \neq 0$ then observe that

$$\lambda_k^{-1} e_k = T(\lambda_k^{-2} e_k) \quad \Rightarrow \quad \lambda_k^{-1} e_k \in \operatorname{range} T = \operatorname{range} T^* = (\operatorname{null} T)^\perp,$$

where we have used 7.6(b). Because the restriction of T to $(\operatorname{null} T)^\perp$ is an isomorphism between $(\operatorname{null} T)^\perp$ and $\operatorname{range} T$ (see 6.67), it follows that $(T|_{(\operatorname{null} T)^\perp})^{-1} T(\lambda_k^{-1} e_k) = \lambda_k^{-1} e_k$ and hence that

$$\begin{aligned} T^\dagger e_k = (T|_{(\operatorname{null} T)^\perp})^{-1} P_{\operatorname{range} T} e_k &= (T|_{(\operatorname{null} T)^\perp})^{-1} P_{\operatorname{range} T} T(\lambda_k^{-1} e_k) \\ &= (T|_{(\operatorname{null} T)^\perp})^{-1} T(\lambda_k^{-1} e_k) = \lambda_k^{-1} e_k. \end{aligned}$$

7.C. Positive Operators

Exercise 7.C.1. Suppose $T \in \mathcal{L}(V)$. Prove that if both T and $-T$ are positive operators, then $T = 0$.

Solution. Let $v \in V$ be given. Observe that $\langle Tv, v\rangle \geq 0$ since T is positive and

$$\langle -Tv, v\rangle \geq 0 \quad \Leftrightarrow \quad -\langle Tv, v\rangle \geq 0 \quad \Leftrightarrow \quad \langle Tv, v\rangle \leq 0$$

since $-T$ is positive. Thus $\langle Tv, v\rangle = 0$ and it follows from 7.43 that $Tv = 0$. Hence $T = 0$.

Exercise 7.C.2. Suppose $T \in \mathcal{L}(\mathbf{F}^4)$ is the operator whose matrix (with respect to the standard basis) is

$$\begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{pmatrix}.$$

Show that T is an invertible positive operator.

Solution. Note that the matrix in question equals its conjugate transpose; since the standard basis of $\mathbf{F}^4$ is orthonormal, it follows that T is self-adjoint. Some calculations reveal that T has 4 distinct eigenvalues:

$$\frac{3 \pm \sqrt{5}}{2}, \quad \frac{5 \pm \sqrt{5}}{2}.$$

Notice that each eigenvalue is strictly positive. It follows from 5.7 and 7.38(b) that T is an invertible positive operator.

Exercise 7.C.3. Suppose n is a positive integer and $T \in \mathcal{L}(\mathbf{F}^n)$ is the operator whose matrix (with respect to the standard basis) consists of all 1's. Show that T is a positive operator.

Solution. The matrix of T with respect to the standard basis of $\mathbf{F}^n$ (which is orthonormal) equals its conjugate transpose; it follows that T is self-adjoint. As shown in Exercise 5.B.3 (a), the eigenvalues of T are contained in $\{0, 1\}$. Thus, by 7.38(b), T is positive.

Exercise 7.C.4. Suppose n is a positive integer with $n > 1$. Show that there exists an n-by-n matrix A such that all of the entries of A are positive numbers and $A = A^*$, but the operator on $\mathbf{F}^n$ whose matrix (with respect to the standard basis) equals A is not a positive operator.

Solution. Let

$$A = \begin{pmatrix} 1 & 2 & \cdots & 2 \\ 2 & 1 & \cdots & 2 \\ \vdots & \vdots & \ddots & \vdots \\ 2 & 2 & \cdots & 1 \end{pmatrix}.$$

Observe that each entry of A is positive and $A = A^*$. Let $T \in \mathcal{L}(\mathbf{F}^n)$ be the operator on $\mathbf{F}^n$ whose matrix with respect to the standard basis equals A, i.e.

$$T(x_1, x_2, ..., x_n) = (x_1 + 2x_2 + \cdots + 2x_n, 2x_1 + x_2 + \cdots + 2x_n, ..., 2x_1 + 2x_2 + \cdots + x_n).$$

Notice that

$$T(-1, 1, 0, ..., 0) = (1, -1, 0, ..., 0).$$

It follows that -1 is an eigenvalue of T. Thus, by 7.38(b), T is not a positive operator.

Exercise 7.C.5. Suppose $T \in \mathcal{L}(V)$ is self-adjoint. Prove that T is a positive operator if and only if for every orthonormal basis $e_1, ..., e_n$ of V, all entries on the diagonal of $\mathcal{M}(T, (e_1, ..., e_n))$ are nonnegative numbers.

Solution. Suppose that T is a positive operator and let $e_1, ..., e_n$ be an orthonormal basis of V. For $k \in \{1, ..., n\}$, 6.30(a) shows that the k^{th} diagonal entry of $\mathcal{M}(T, (e_1, ..., e_n))$ is equal to $\langle Te_k, e_k \rangle$; this must be non-negative by the positivity of T.

Suppose that for every orthonormal basis $e_1, ..., e_n$ of V, all entries on the diagonal of $\mathcal{M}(T, (e_1, ..., e_n))$ are nonnegative numbers. Because T is self-adjoint, the relevant spectral theorem (7.29 or 7.31) implies the existence of an orthonormal basis $e_1, ..., e_n$ of V with respect to which the matrix of T is diagonal. By assumption each diagonal entry of this matrix is non-negative and it then follows from 7.38(c) that T is a positive operator.

Exercise 7.C.6. Prove that the sum of two positive operators on V is a positive operator.

Solution. Suppose that $S, T \in \mathcal{L}(V)$ are positive operators and note that $S + T$ is self-adjoint by 7.5(a). Let $v \in V$ be given and observe that

$$\langle (S + T)v, v \rangle = \langle Sv + Tv, v \rangle = \langle Sv, v \rangle + \langle Tv, v \rangle \geq 0.$$

Thus $S + T$ is positive.

Exercise 7.C.7. Suppose $S \in \mathcal{L}(V)$ is an invertible positive operator and $T \in \mathcal{L}(V)$ is a positive operator. Prove that $S + T$ is invertible.

Solution. Let us prove the following lemma (see 7.61 also).

Lemma L.11. If $T \in \mathcal{L}(V)$ is a positive operator then T is invertible if and only if $\langle Tv, v\rangle > 0$ for every non-zero $v \in V$.

Proof. Suppose that T is invertible and let $v \in V$ be non-zero. It follows that Tv is non-zero and thus, by the contrapositive of 7.43, we have $\langle Tv, v\rangle > 0$. If T is not invertible then there exists some non-zero $v \in V$ such that $Tv = 0$, which gives us $\langle Tv, v\rangle = 0$. □

Let $v \in V$ be non-zero. Lemma L.11 shows that $\langle Sv, v\rangle > 0$ and it follows from the positivity of T that

$$\langle (S+T)v, v\rangle = \langle Sv, v\rangle + \langle Tv, v\rangle > 0.$$

Another application of Lemma L.11 allows us to conclude that $S + T$ is invertible.

Exercise 7.C.8. Suppose $T \in \mathcal{L}(V)$. Prove that T is a positive operator if and only if the pseudoinverse $T^\dagger$ is a positive operator.

Solution. Suppose that T is a positive operator. By 7.38(c) there is an orthonormal basis $e_1, \ldots, e_n$ of V such that $Te_k = \lambda_k e_k$ with $\lambda_k \geq 0$. For each $k \in \{1, \ldots, n\}$ let

$$\mu_k = \begin{cases} \lambda_k^{-1} & \text{if } \lambda_k \neq 0, \\ 0 & \text{if } \lambda_k = 0. \end{cases}$$

It follows from Exercise 7.B.25 that $T^\dagger e_k = \mu_k e_k$ for each $k \in \{1, \ldots, n\}$. Because each μ_k is non-negative, 7.38(c) implies that $T^\dagger$ is a positive operator.

Replacing T with $T^\dagger$ in the preceding result and using that $\left(T^\dagger\right)^\dagger = T$ (see Exercise 6.C.23) gives us the converse statement.

Exercise 7.C.9. Suppose $T \in \mathcal{L}(V)$ is a positive operator and $S \in \mathcal{L}(W, V)$. Prove that S^*TS is a positive operator on W.

Solution. For any $w \in W$ observe that

$$\langle S^*TSw, w\rangle = \langle T(Sw), Sw\rangle \geq 0,$$

where we have used the positivity of T.

Exercise 7.C.10. Suppose T is a positive operator on V. Suppose $v, w \in V$ are such that

$$Tv = w \quad \text{and} \quad Tw = v.$$

Prove that $v = w$.

Solution. Notice that

$$\|v-w\|^2 = \langle v-w, v-w\rangle = \langle Tw - Tv, v-w\rangle = -\langle T(v-w), v-w\rangle \leq 0,$$

where we have used that T is a positive operator for the final inequality. It follows that $\|v-w\|^2 = 0$, which is the case if and only if $v = w$.

Exercise 7.C.11. Suppose T is a positive operator on V and U is a subspace of V invariant under T. Prove that $T|_U \in \mathcal{L}(U)$ is a positive operator on U.

Solution. Exercise 7.B.19 (b) shows that $T|_U$ is self-adjoint, and for any $u \in U$ we have

$$\langle T|_U(u), u\rangle = \langle Tu, u\rangle \geq 0,$$

where we have used that T is a positive operator. Thus $T|_U$ is a positive operator.

Exercise 7.C.12. Suppose $T \in \mathcal{L}(V)$ is a positive operator. Prove that T^k is a positive operator for every positive integer k.

Solution. By 7.38(c) there is an orthonormal basis $e_1, \ldots, e_n$ of V such that $Te_j = \lambda_j e_j$ with $\lambda_j \geq 0$ for each $j \in \{1, \ldots, n\}$. Let k be a positive integer and observe that $T^k e_j = \lambda_j^k e_j$ for each $j \in \{1, \ldots, n\}$. Because each λ_j^k is non-negative, 7.38(c) allows us to conclude that T^k is a positive operator.

Exercise 7.C.13. Suppose $T \in \mathcal{L}(V)$ is self-adjoint and $\alpha \in \mathbf{R}$.

(a) Prove that $T - \alpha I$ is a positive operator if and only if α is less than or equal to every eigenvalue of T.

(b) Prove that $\alpha I - T$ is a positive operator if and only if α is greater than or equal to every eigenvalue of T.

Solution.

(a) The relevant spectral theorem (7.29 or 7.31) implies the existence of an orthonormal basis $e_1, \ldots, e_n$ such that $Te_k = \lambda_k e_k$ for some eigenvalues $\lambda_1, \ldots, \lambda_n$; note that each λ_k is real by 7.12. It follows that the matrix of $T - \alpha I$ with respect to $e_1, \ldots, e_n$ is

$$\begin{pmatrix} \lambda_1 - \alpha & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n - \alpha \end{pmatrix}.$$

If α is less than or equal to every eigenvalue of T then each diagonal entry of this matrix is non-negative and it follows from 7.38(c) that $T - \alpha I$ is a positive operator; if α is greater than some eigenvalue of T then at least one diagonal entry of this matrix is negative and it follows from Exercise 7.C.5 that $T - \alpha I$ is not a positive operator.

(b) It is straightforward to modify the argument in part (a) to prove part (b).

Exercise 7.C.14. Suppose T is a positive operator on V and $v_1, \ldots, v_m \in V$. Prove that

$$\sum_{j=1}^{m}\sum_{k=1}^{m}\langle Tv_k, v_j\rangle \geq 0.$$

Solution. Using the positivity of T, observe that

$$\sum_{j=1}^{m}\sum_{k=1}^{m}\langle Tv_k, v_j\rangle = \langle Tv_1 + \cdots + Tv_m, v_1 + \cdots + v_m\rangle = \langle T(v_1 + \cdots + v_m), v_1 + \cdots + v_m\rangle \geq 0.$$

Exercise 7.C.15. Suppose $T \in \mathcal{L}(V)$ is self-adjoint. Prove that there exist positive operators $A, B \in \mathcal{L}(V)$ such that

$$T = A - B \quad \text{and} \quad \sqrt{T^*T} = A + B \quad \text{and} \quad AB = BA = 0.$$

Solution. The relevant spectral theorem (7.29 or 7.31) implies the existence of an orthonormal basis $e_1, \ldots, e_n$ such that $Te_k = \lambda_k e_k$ for some eigenvalues $\lambda_1, \ldots, \lambda_n$; note that each λ_k is real by 7.12. Note further that the operator $R \in \mathcal{L}(V)$ given by $Re_k = |\lambda_k| e_k$ is positive by 7.38(c) and satisfies $R^2 e_k = \lambda_k^2 e_k = T^2 e_k$; it follows from the uniqueness in 7.39 that $R = \sqrt{T^2} = \sqrt{T^*T}$.

For each $k \in \{1, \ldots, n\}$ define

$$\alpha_k = \begin{cases} \lambda_k & \text{if } \lambda_k \geq 0, \\ 0 & \text{if } \lambda_k < 0, \end{cases} \quad \text{and} \quad \beta_k = \begin{cases} -\lambda_k & \text{if } \lambda_k \leq 0, \\ 0 & \text{if } \lambda_k > 0. \end{cases}$$

Notice that, for each $k \in \{1, \ldots, n\}$,

$$\alpha_k, \beta_k \geq 0, \quad \alpha_k - \beta_k = \lambda_k, \quad \alpha_k + \beta_k = |\lambda_k|, \quad \text{and} \quad \alpha_k \beta_k = 0.$$

Define $A, B \in \mathcal{L}(V)$ by $Ae_k = \alpha_k e_k$ and $Be_k = \beta_k e_k$ and note that A and B are positive operators by 7.38(c). Furthermore,

$$(A - B)e_k = (\alpha_k - \beta_k)e_k = \lambda_k e_k = Te_k, \quad (A + B)e_k = (\alpha_k + \beta_k)e_k = |\lambda_k| e_k = \sqrt{T^*T} e_k,$$
$$ABe_k = BAe_k = \alpha_k \beta_k e_k = 0.$$

It follows that

$$T = A - B, \quad \sqrt{T^*T} = A + B \quad \text{and} \quad AB = BA = 0.$$

Exercise 7.C.16. Suppose T is a positive operator on V. Prove that

$$\operatorname{null} \sqrt{T} = \operatorname{null} T \quad \text{and} \quad \operatorname{range} \sqrt{T} = \operatorname{range} T.$$

Solution. Note that $\sqrt{T}$ is positive, hence self-adjoint, hence normal; it follows from Exercise 7.A.27 that

$$\operatorname{null} \sqrt{T} = \operatorname{null} \left(\sqrt{T}\right)^2 = \operatorname{null} T \quad \text{and} \quad \operatorname{range} \sqrt{T} = \operatorname{range} \left(\sqrt{T}\right)^2 = \operatorname{range} T.$$

Exercise 7.C.17. Suppose that $T \in \mathcal{L}(V)$ is a positive operator. Prove that there exists a polynomial p with real coefficients such that $\sqrt{T} = p(T)$.

Solution. By 7.38(c) there is an orthonormal basis $e_1, ..., e_n$ of V such that $Te_k = \lambda_k e_k$ with $\lambda_k \geq 0$, where $\lambda_1, ..., \lambda_n$ are the eigenvalues of T; note that $\sqrt{T} e_k = \sqrt{\lambda_k} e_k$. Exercise 4.7 shows that there is a polynomial $p \in \mathcal{P}(\mathbf{R})$ satisfying $p(\lambda_k) = \sqrt{\lambda_k}$ for each $k \in \{1, ..., n\}$ and it follows that

$$p(T)e_k = p(\lambda_k)e_k = \sqrt{\lambda_k} e_k = \sqrt{T} e_k.$$

Thus $\sqrt{T} = p(T)$.

Exercise 7.C.18. Suppose S and T are positive operators on V. Prove that ST is a positive operator if and only if S and T commute.

Solution. If S and T do not commute then, by Exercise 7.A.9, ST is not self-adjoint and hence not a positive operator.

If S and T commute then Exercise 7.A.9 shows that ST is self-adjoint and Exercise 7.B.16/ Exercise 7.B.17 shows that there is an orthonormal basis $e_1, ..., e_n$ of V consisting of eigenvectors of both S and T, say $Se_k = \mu_k e_k$ and $Te_k = \lambda_k e_k$; each μ_k and each λ_k is a non-negative real number by 7.38(b). It follows that $STe_k = \mu_k \lambda_k e_k$, so that each e_k is an eigenvector of ST with a corresponding non-negative real eigenvalue. Thus, by 7.38(c), ST is a positive operator.

Exercise 7.C.19. Show that the identity operator on $\mathbf{F}^2$ has infinitely many self-adjoint square roots.

Solution. For any $t \in (0, 1)$, let R_t be the operator on $\mathbf{F}^2$ whose matrix with respect to the standard orthonormal basis of $\mathbf{F}^2$ is

$$\begin{pmatrix} \sqrt{1-t^2} & t \\ t & -\sqrt{1-t^2} \end{pmatrix};$$

note that each $t \in (0, 1)$ gives a distinct operator R_t. Because the matrix above equals its conjugate transpose, R_t is self-adjoint. A calculation shows that $R_t^2 = I$.

Exercise 7.C.20. Suppose $T \in \mathcal{L}(V)$ and $e_1, ..., e_n$ is an orthonormal basis of V. Prove that T is a positive operator if and only if there exist $v_1, ..., v_n \in V$ such that

$$\langle Te_k, e_j \rangle = \langle v_k, v_j \rangle$$

for all $j, k = 1, ..., n$.

The numbers $\{\langle Te_k, e_j \rangle\}_{j,k=1,...,n}$ are the entries in the matrix of T with respect to the orthonormal basis $e_1, ..., e_n$.

Solution. Suppose that T is a positive operator and for each $k \in \{1, ..., n\}$ let $v_k = \sqrt{T} e_k$. It follows that

$$\langle Te_k, e_j \rangle = \langle \sqrt{T} e_k, \sqrt{T} e_j \rangle = \langle v_k, v_j \rangle$$

for all $j, k \in \{1, ..., n\}$.

Now suppose that there exist $v_1, ..., v_n \in V$ such that

$$\langle Te_k, e_j \rangle = \langle v_k, v_j \rangle$$

for all $j, k \in \{1, ..., n\}$. Define $R \in \mathcal{L}(V)$ by $Re_k = v_k$, so that

$$\langle Te_k, e_j \rangle = \langle Re_k, Re_j \rangle = \langle R^* Re_k, e_j \rangle$$

for all $j, k \in \{1, ..., n\}$. It follows that

$$\mathcal{M}(T, (e_1, ..., e_n)) = \mathcal{M}(R^* R, (e_1, ..., e_n))$$

and hence that $T = R^* R$. We may use 7.38(f) to conclude that T is a positive operator.

Exercise 7.C.21. Suppose n is a positive integer. The n-by-n *Hilbert matrix* is the n-by-n matrix whose entry in row j, column k is $\frac{1}{j+k-1}$. Suppose $T \in \mathcal{L}(V)$ is an operator whose matrix with respect to some orthonormal basis of V is the n-by-n Hilbert matrix. Prove that T is a positive invertible operator.

Example: The 4-by-4 Hilbert matrix is

$$\begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} & \frac{1}{6} \\ \frac{1}{4} & \frac{1}{5} & \frac{1}{6} & \frac{1}{7} \end{pmatrix}.$$

Solution. (This solution uses some integration of complex-valued functions of a real variable.) Suppose that $e_1, ..., e_n$ is the orthonormal basis of V with respect to which the matrix of T is the n-by-n Hilbert matrix. For any $v = x_1 e_1 + \cdots + x_n e_n \in V$, observe that

$$
\begin{aligned}
\langle Tv, v\rangle &= \left\langle \sum_{k=1}^{n} x_k T e_k, \sum_{j=1}^{n} x_j e_j \right\rangle \\
&= \sum_{j,k=1}^{n} \overline{x_j} x_k \langle T e_k, e_j \rangle \\
&= \sum_{j,k=1}^{n} \frac{\overline{x_j} x_k}{j+k-1} \\
&= \sum_{j,k=1}^{n} \overline{x_j} x_k \int_0^1 t^{j+k-2}\,\mathrm{d}t \\
&= \int_0^1 \sum_{j,k=1}^{n} \overline{x_j} x_k t^{j+k-2}\,\mathrm{d}t \\
&= \int_0^1 \left(\sum_{j=1}^{n} \overline{x_j} t^{j-1} \right) \left(\sum_{k=1}^{n} x_k t^{k-1} \right) \mathrm{d}t \\
&= \int_0^1 \left(\overline{\sum_{j=1}^{n} x_j t^{j-1}} \right) \left(\sum_{k=1}^{n} x_k t^{k-1} \right) \mathrm{d}t \\
&= \int_0^1 \left| \sum_{k=1}^{n} x_k t^{k-1} \right|^2 \mathrm{d}t.
\end{aligned}
$$

Observe that the integrand above is a non-negative and continuous function $[0,1] \to \mathbf{R}$. It follows that the integral is non-negative and vanishes if and only if

$$
\begin{aligned}
\left| \sum_{k=1}^{n} x_k t^{k-1} \right|^2 = 0 \text{ for all } t \in [0,1] \quad &\Leftrightarrow \quad \sum_{k=1}^{n} x_k t^{k-1} = 0 \text{ for all } t \in [0,1] \\
&\Leftrightarrow \quad x_k = 0 \text{ for all } k \in \{1, \dots, n\} \quad \Leftrightarrow \quad v = 0.
\end{aligned}
$$

Thus $\langle Tv, v\rangle \geq 0$ for every $v \in V$ and $\langle Tv, v\rangle = 0$ if and only if $v = 0$; it follows from Lemma L.11 that T is an invertible positive operator.

Exercise 7.C.22. Suppose $T \in \mathcal{L}(V)$ is a positive operator and $u \in V$ is such that $\|u\| = 1$ and $\|Tu\| \geq \|Tv\|$ for all $v \in V$ with $\|v\| = 1$. Show that u is an eigenvector of T corresponding to the largest eigenvalue of T.

Solution. Suppose $0 \leq \lambda_1 < \cdots < \lambda_n$ are the distinct eigenvalues of T; each λ_k is a non-negative real number by 7.38(b). Exercise 7.B.14/Exercise 7.B.15 shows that

$$
V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_n, T)
$$

and that all pairs of eigenvectors of T corresponding to distinct eigenvalues are orthogonal. Suppose that $u = v_1 + \cdots + v_n$ where each $v_k \in E(\lambda_k, T)$. It follows from the Pythagorean theorem that $1 = \|u\|^2 = \|v_1\|^2 + \cdots + \|v_n\|^2$ and hence that

$$\|Tu\|^2 = \lambda_1^2\|v_1\|^2 + \cdots + \lambda_n^2\|v_n\|^2 \leq \lambda_n^2\left(\|v_1\|^2 + \cdots + \|v_n\|^2\right) = \lambda_n^2.$$

Letting w be a unit eigenvector corresponding to the eigenvalue λ_n, our hypothesis implies that $\lambda_n^2 = \|Tw\|^2 \leq \|Tu\|^2$. Thus $\|Tu\|^2 = \lambda_n^2$, which gives us

$$\|Tu\|^2 = \lambda_n^2 \quad \Leftrightarrow \quad \lambda_1^2\|v_1\|^2 + \cdots + \lambda_n^2\left(\|v_n\|^2 - 1\right) = 0.$$

Note that if $\lambda_n = 0$ then T must be the zero operator and the desired result is clear. If $\lambda_n \neq 0$ then the equation above shows that $\|v_n\|^2 = 1$; combining this with $1 = \|v_1\|^2 + \cdots + \|v_n\|^2$ shows that $v_1 = \cdots = v_{n-1} = 0$ and hence that $u = v_n \in E(\lambda_n, T)$, as desired.

Exercise 7.C.23. For $T \in \mathcal{L}(V)$ and $u, v \in V$, define $\langle u, v\rangle_T$ by $\langle u, v\rangle_T = \langle Tu, v\rangle$.

(a) Suppose $T \in \mathcal{L}(V)$. Prove that $\langle \cdot, \cdot\rangle_T$ is an inner product on V if and only if T is an invertible positive operator (with respect to the original inner product $\langle \cdot, \cdot\rangle$).

(b) Prove that every inner product on V is of the form $\langle \cdot, \cdot\rangle_T$ for some positive invertible operator $T \in \mathcal{L}(V)$.

Solution.

(a) Suppose that T is an invertible positive operator. We must verify each property of definition 6.2.

Positivity. For any $v \in V$ we have $\langle v, v\rangle_T = \langle Tv, v\rangle \geq 0$ by the positivity of T.

Definiteness. We have $\langle 0, 0\rangle_T = \langle T(0), 0\rangle = 0$, and for $v \neq 0$ we have

$$\langle v, v\rangle_T = \langle Tv, v\rangle > 0$$

by Lemma L.11.

Additivity in first slot. For $u, v, w \in V$ we have

$$\langle u + v, w\rangle_T = \langle T(u + v), w\rangle = \langle Tu + Tv, w\rangle = \langle Tu, w\rangle + \langle Tv, w\rangle = \langle u, w\rangle_T + \langle v, w\rangle_T.$$

Homoegeneity in first slot. For $u, v \in V$ and $\lambda \in \mathbf{F}$ we have

$$\langle \lambda u, v\rangle_T = \langle T(\lambda u), v\rangle = \langle \lambda Tu, v\rangle = \lambda\langle Tu, v\rangle = \lambda\langle u, v\rangle_T.$$

Conjugate symmetry. For $u, v \in V$ we have

$$\langle u, v\rangle_T = \langle Tu, v\rangle = \overline{\langle v, Tu\rangle} = \overline{\langle Tv, u\rangle} = \overline{\langle v, u\rangle_T},$$

where we have used that T is self-adjoint for the third equality.

Thus $\langle \cdot, \cdot\rangle_T$ is an inner product on V.

Now suppose that $\langle \cdot, \cdot\rangle_T$ is an inner product on V and let $u, v \in V$ be given. Observe that

$$\langle Tu, v\rangle = \langle u, v\rangle_T = \overline{\langle v, u\rangle_T} = \overline{\langle Tv, u\rangle} = \langle u, Tv\rangle.$$

Thus T is self-adjoint. Furthermore, for any $v \in V$,

$$\langle Tv, v\rangle = \langle v, v\rangle_T \geq 0.$$

Thus T is a positive operator. Finally, for $v \neq 0$,

$$\langle Tv, v\rangle = \langle v, v\rangle_T > 0.$$

It follows from Lemma L.11 that T is invertible.

(b) Let $\langle\cdot,\cdot\rangle_1$ be the original inner product on V and let $\langle\cdot,\cdot\rangle_2$ be an arbitrary inner product on V; we need to show that there exists an invertible operator $T \in \mathcal{L}(V)$ which is positive with respect to $\langle\cdot,\cdot\rangle_1$ such that $\langle u, v\rangle_2 = \langle Tu, v\rangle_1$ for every $u, v \in V$.

Let $e_1, \ldots, e_n$ be an orthonormal basis of V with respect to $\langle\cdot,\cdot\rangle_1$ and let $f_1, \ldots, f_n$ be an orthonormal basis of V with respect to $\langle\cdot,\cdot\rangle_2$. Define $R \in \mathcal{L}(V)$ by $Rf_k = e_k$ and note that R is invertible since R maps a basis to a basis. Now define $T \in \mathcal{L}(V)$ by $T = R^*R$, where R^* is the adjoint of R with respect to $\langle\cdot,\cdot\rangle_1$. Observe that T is invertible by 7.5(f) and Exercise 3.D.2, and T is positive with respect to $\langle\cdot,\cdot\rangle_1$ by 7.38(f). Furthermore, for any $u = x_1 f_1 + \cdots + x_n f_n$ and $v = y_1 f_1 + \cdots + y_n f_n$ in V, observe that

$$\langle u, v\rangle_2 = x_1\overline{y_1} + \cdots + x_n\overline{y_n} = \langle Ru, Rv\rangle_1 = \langle R^*Ru, v\rangle_1 = \langle Tu, v\rangle_1.$$

Exercise 7.C.24. Suppose S and T are positive operators on V. Prove that

$$\operatorname{null}(S+T) = \operatorname{null} S \cap \operatorname{null} T.$$

Solution. Exercise 7.C.6 shows that $S+T$ is a positive operator. Observe that

$$\begin{aligned} Sv = 0 \text{ and } Tv = 0 \quad &\Leftrightarrow \quad \langle Sv, v\rangle = 0 \text{ and } \langle Tv, v\rangle = 0 \\ &\Leftrightarrow \quad \langle Sv, v\rangle + \langle Tv, v\rangle = 0 \quad \Leftrightarrow \quad \langle (S+T)v, v\rangle = 0 \quad \Leftrightarrow \quad (S+T)v = 0, \end{aligned}$$

where we have used 7.43 several times. Thus $\operatorname{null}(S+T) = \operatorname{null} S \cap \operatorname{null} T$.

Exercise 7.C.25. Let T be the second derivative operator in Exercise 31(b) in Section 7A. Show that $-T$ is a positive operator.

Solution. Define the orthonormal basis $\mathcal{B} = v, e_1, \ldots, e_n, f_1, \ldots, f_n$ as in Exercise 7.A.31 and observe that

$$-Tv = 0, \quad -Te_k = k^2 e_k, \quad -Tf_k = k^2 f_k.$$

It follows that the matrix of $-T$ with respect to $\mathcal{B}$ is diagonal with non-negative diagonal entries; 7.38(c) allows us to conclude that $-T$ is a positive operator.

7.D. Isometries, Unitary Operators, and Matrix Factorization

Exercise 7.D.1. Suppose $\dim V \geq 2$ and $S \in \mathcal{L}(V, W)$. Prove that S is an isometry if and only if Se_1, Se_2 is an orthonormal list in W for every orthonormal list e_1, e_2 of length two in V.

Solution. Suppose S is an isometry and suppose e_1, e_2 is an orthonormal list in V. Let $U = \operatorname{span}(e_1, e_2)$ and note that e_1, e_2 is an orthonormal basis of U. Note further that $S|_U$ is an isometry; it follows from 7.49(d) that Se_1, Se_2 is an orthonormal list in W.

Now suppose that Se_1, Se_2 is an orthonormal list in W for every orthonormal list e_1, e_2 in V. Let $e_1, e_2, ..., e_n$ be an orthonormal basis of V and let $j < k$ in $\{1, ..., n\}$ be given. Since e_j, e_k is an orthonormal list in V, our hypothesis guarantees that Se_j, Se_k is an orthonormal list in V. Thus $\|Se_j\| = \|Se_k\| = 1$ and $\langle Se_j, Se_k \rangle = 0$. Because $j < k$ in $\{1, ..., n\}$ were arbitrary, it follows that $Se_1, ..., Se_n$ is an orthonormal list in W. We may use 7.49(d) to conclude that S is an isometry.

Exercise 7.D.2. Suppose $T \in \mathcal{L}(V, W)$. Prove that T is a scalar multiple of an isometry if and only if T preserves orthogonality.

The phrase "T preserves orthogonality" means that $\langle Tu, Tv \rangle = 0$ for all $u, v \in V$ such that $\langle u, v \rangle = 0$.

Solution. Suppose that $T = \lambda S$ for some $\lambda \in \mathbf{F}$ and some isometry $S \in \mathcal{L}(V, W)$. For any $u, v \in V$ such that $\langle u, v \rangle = 0$, observe that

$$\langle Tu, Tv \rangle = \langle \lambda Su, \lambda Sv \rangle = |\lambda|^2 \langle Su, Sv \rangle = |\lambda|^2 \langle u, v \rangle = 0,$$

where we have used 7.49(c) for the third equality. Thus T preserves orthogonality.

Now suppose that T preserves orthogonality. Let $e_1, ..., e_n$ be an orthonormal basis of V and let $k \in \{1, ..., n\}$ be given. Using the identity $\langle u + v, u - v \rangle = \|u\|^2 - \|v\|^2$, observe that

$$\langle e_1 + e_k, e_1 - e_k \rangle = \|e_1\| - \|e_k\| = 0 \quad \Rightarrow \quad \langle Te_1 + Te_k, Te_1 - Te_k \rangle = \|Te_1\| - \|Te_k\| = 0.$$

Thus, letting $\lambda = \|Te_1\|$, we have $\lambda = \|Te_k\|$ for each $k \in \{1, ..., n\}$. If $\lambda = 0$ then $T = 0I$, so that T is a scalar multiple of the identity operator, which is certainly an isometry. If $\lambda \neq 0$ then let $S = \lambda^{-1}T$. Observe that, for any distinct $j, k \in \{1, ..., n\}$,

$$\langle e_j, e_k \rangle = 0 \quad \Rightarrow \quad \langle Te_j, Te_k \rangle = 0 \quad \Leftrightarrow \quad \langle \lambda Se_j, \lambda Se_k \rangle = 0 \quad \Leftrightarrow \quad |\lambda|^2 \langle Se_j, Se_k \rangle = 0.$$

Since $\lambda \neq 0$, this last equation implies that $\langle Se_j, Se_k \rangle = 0$. Furthermore, using that λ is a non-negative real number,

$$\|Se_k\| = \lambda^{-1} \|Te_k\| = 1.$$

Thus $Se_1, ..., Se_n$ is an orthonormal list in W. It follows from 7.49(d) that S is an isometry and hence that $T = \lambda S$ is a scalar multiple of an isometry.

Exercise 7.D.3.

(a) Show that the product of two unitary operators on V is a unitary operator.

(b) Show that the inverse of a unitary operator on V is a unitary operator.

This exercise shows that the set of unitary operators on V is a group, where the group operation is the usual product of two operators.

Solution.

(a) Suppose $S, T \in \mathcal{L}(V)$ are unitary operators. For any $v \in V$, observe that

$$\|STv\| = \|Sv\| = \|v\|.$$

Thus ST is a unitary operator.

(b) Suppose that $S \in \mathcal{L}(V)$ is a unitary operator. For any $v \in V$ we have $v = Su$ for some $u \in V$. It follows that

$$\|S^{-1}v\| = \|u\| = \|Su\| = \|v\|.$$

Thus S^{-1} is a unitary operator.

Exercise 7.D.4. Suppose $\mathbf{F} = \mathbf{C}$ and $A, B \in \mathcal{L}(V)$ are self-adjoint. Show that $A + iB$ is unitary if and only if $AB = BA$ and $A^2 + B^2 = I$.

Solution. Suppose $A + iB$ is unitary. It follows from 7.53(f) that $(A + iB)^* = A - iB$ is also unitary. Thus, for any $v \in V$,

$$\|Av\|^2 + \|Bv\|^2 + i(\langle(AB - BA)v, v\rangle) = \|Av + iBv\|^2 = \|v\|^2,$$
$$\|Av\|^2 + \|Bv\|^2 - i(\langle(AB - BA)v, v\rangle) = \|Av - iBv\|^2 = \|v\|^2.$$

Subtracting the latter of these equations from the former, we see that

$$\langle(AB - BA)v, v\rangle = 0$$

for every $v \in V$. It follows from 7.13 that $AB = BA$. We can now use 7.53(c) to see that

$$I = (A + iB)^*(A + iB) = (A - iB)(A + iB) = A^2 + B^2 + i(AB - BA) = A^2 + B^2.$$

If $AB = BA$ and $A^2 + B^2 = I$ then

$$(A + iB)^*(A - iB) = (A - iB)(A + iB) = A^2 + B^2 + i(AB - BA) = I.$$

Thus, by 7.49(b), $A + iB$ is unitary.

Exercise 7.D.5. Suppose $S \in \mathcal{L}(V)$. Prove that the following are equivalent.

(a) S is a self-adjoint unitary operator.

(b) $S = 2P - I$ for some orthogonal projection P on V.

(c) There exists a subspace U of V such that $Su = u$ for every $u \in U$ and $Sw = -w$ for every $w \in U^\perp$.

Solution. Suppose that (a) holds, i.e. suppose that S is a self-adjoint unitary operator. Let $P = \frac{1}{2}(S + I)$ and notice that P is self-adjoint and that $S = 2P - I$. Notice further that, by 7.53(b), $S^2 = I$; it follows that

$$P^2 = \frac{S^2 + 2S + I}{4} = \frac{S + I}{2} = P.$$

We may now invoke Exercise 7.A.20 to see that P is an orthogonal projection. Thus (b) holds.

Suppose that (b) holds, so that there is some subspace U of V such that $S = 2P_U - I$. For any $u \in U$ and any $w \in U^\perp$ it follows that

$$Su = 2P_U u - u = u \quad \text{and} \quad Sw = 2P_U w - w = -w,$$

where we have used 6.57. Thus (c) holds.

Suppose that (c) holds, i.e. suppose there exists a subspace U of V such that $Su = u$ for every $u \in U$ and $Sw = -w$ for every $w \in U^\perp$. Let $v_1 = u_1 + w_1$ and $v_2 = u_2 + w_2$ in V be given, where $u_1, u_2 \in U$ and $w_1, w_2 \in U^\perp$. Observe that

$$\langle Sv_1, v_2 \rangle = \langle u_1 - w_1, u_2 + w_2 \rangle = \langle u_1, u_2 \rangle - \langle w_1, w_2 \rangle = \langle u_1 + w_1, u_2 - w_2 \rangle = \langle Sv_1, v_2 \rangle.$$

Thus S is self-adjoint. Furthermore,

$$S^2 v_1 = S(S(u_1 + w_1)) = S(u_1 - w_1) = u_1 + w_1 = v_1.$$

Thus S is its own inverse. Combining this with the fact that S is self-adjoint and 7.53(c), we see that S is a self-adjoint unitary operator, i.e. (a) holds.

Exercise 7.D.6. Suppose T_1, T_2 are both normal operators on $\mathbf{F}^3$ with $2, 5, 7$ as eigenvalues. Prove that there exists a unitary operator $S \in \mathcal{L}(\mathbf{F}^3)$ such that $T_1 = S^* T_2 S$.

Solution. Let $\lambda_1 = 2, \lambda_2 = 5, \lambda_3 = 7$, and let $e_1, e_2, e_3, f_1, f_2, f_3 \in V$ be such that

$$T_1 e_k = \lambda_k e_k \quad \text{and} \quad T_2 f_k = \lambda_k f_k.$$

Without loss of generality, we may assume that each e_k and each f_k is a unit vector. Combining this with 7.22, we see that each list e_1, e_2, e_3 and f_1, f_2, f_3 is an orthonormal basis of $\mathbf{F}^3$. Define $S \in \mathcal{L}(V)$ by $Se_k = f_k$ and note that S is a unitary operator by 7.53(d). Note further that

$$S^* T_2 S e_k = S^* T_2 f_k = \lambda_k S^* f_k = \lambda_k e_k = T_1 e_k,$$

where we have used that $S^* = S^{-1}$, which holds by 7.53(c). Thus $T_1 = S^*T_2S$.

Exercise 7.D.7. Give an example of two self-adjoint operators $T_1, T_2 \in \mathcal{L}(\mathbf{F}^4)$ such that the eigenvalues of both operators are $2, 5, 7$ but there does not exist a unitary operator $S \in \mathcal{L}(\mathbf{F}^4)$ such that $T_1 = S^*T_2S$. Be sure to explain why there is no unitary operator with the required property.

Solution. Let T_1 and T_2 be the operators on $\mathbf{F}^4$ whose matrices with respect to the standard basis of $\mathbf{F}^4$ are

$$\begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 7 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 7 \end{pmatrix}.$$

Since these matrices are diagonal and the standard basis of $\mathbf{F}^4$ is orthonormal, we see that T_1 and T_2 are self-adjoint and that their eigenvalues are precisely $2, 5, 7$. If there was an isometry $S \in \mathcal{L}(\mathbf{F}^4)$ such that $T_1 = S^*T_2S$ then since $S^*S = I$ (by 7.42), we would have

$$T_1 - 2I = S^*T_2S - 2I = S^*(T_2 - 2I)S.$$

It would then follow from Exercise 3.D.8 that $\dim \operatorname{null}(T_1 - 2I) = \dim \operatorname{null}(T_2 - 2I)$. However, from the matrices of T_1 and T_2 above we can see that

$$\dim \operatorname{null}(T_1 - 2I) = 2 \neq 1 = \dim \operatorname{null}(T_2 - 2I).$$

Exercise 7.D.8. Prove or give a counterexample: If $S \in \mathcal{L}(V)$ and there exists an orthonormal basis $e_1, \ldots, e_n$ of V such that $\|Se_k\| = 1$ for each e_k, then S is a unitary operator.

Solution. This is false. Let e_1, e_2 be the standard orthonormal basis of $\mathbf{R}^2$ and let $S \in \mathcal{L}(\mathbf{R}^2)$ be given by $Se_1 = Se_2 = e_1$. Observe that $\|Se_1\| = \|Se_2\| = \|e_1\| = 1$, but S is not a unitary operator because S is not injective.

Exercise 7.D.9. Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$. Suppose every eigenvalue of T has absolute value 1 and $\|Tv\| \leq \|v\|$ for every $v \in V$. Prove that T is a unitary operator.

Solution. By Schur's theorem (6.38) there is an orthonormal basis $e_1, \ldots, e_n$ of V with respect to which $A := \mathcal{M}(T, (e_1, \ldots, e_n))$ is upper-triangular. Note that $|A_{j,j}| = 1$ for each $j \in \{1, \ldots, n\}$ since each diagonal entry of A is an eigenvalue of T. For any $k \in \{2, \ldots, n\}$ it follows that

$$\sum_{j=1}^{k} |A_{j,k}|^2 = 1 + \sum_{j=1}^{k-1} |A_{j,k}|^2 = \|Te_k\|^2 \leq \|e_k\|^2 = 1.$$

Thus $\sum_{j=1}^{k-1} |A_{j,k}|^2 = 0$, which gives us $A_{1,k} = \cdots = A_{k-1,k} = 0$. Hence A is a diagonal matrix, i.e. $e_1, \dots, e_n$ is an orthonormal basis of V consisting of eigenvectors of T. By assumption each eigenvalue of T has absolute value 1 and thus, by 7.55, T is a unitary operator.

Exercise 7.D.10. Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$ is a self-adjoint operator such that $\|Tv\| \leq \|v\|$ for all $v \in V$.

(a) Show that $I - T^2$ is a positive operator.

(b) Show that $T + i\sqrt{I - T^2}$ is a unitary operator.

Solution.

(a) For any $v \in V$ observe that

$$\begin{aligned} \langle (I - T^2)v, v \rangle = \langle v - T^2 v, v \rangle = \langle v, v \rangle - \langle T^2 v, v \rangle \\ = \langle v, v \rangle - \langle Tv, Tv \rangle = \|v\|^2 - \|Tv\|^2 \geq 0. \end{aligned}$$

Thus $I - T^2$ is a positive operator.

(b) Let us prove the following lemma.

Lemma L.12. If $S \in \mathcal{L}(V)$ is positive and $T \in \mathcal{L}(V)$ is such that $ST = TS$, then $\sqrt{S}T = T\sqrt{S}$.

Proof. By Exercise 7.C.17 there is a polynomial $p \in \mathcal{P}(\mathbf{R})$ such that $\sqrt{S} = p(S)$. Because S and T commute, we can argue as in Exercise 7.B.12 to see that $\sqrt{S}$ and T commute. □

A straightforward calculation (or 5.17) shows that T and $I - T^2$ commute. It follows from Lemma L.12 that T and $\sqrt{I - T^2}$ commute. Observe that

$$T^2 + \left(\sqrt{I - T^2}\right)^2 = I.$$

Thus, by Exercise 7.D.4, $T + i\sqrt{I - T^2}$ is a unitary operator.

Exercise 7.D.11. Suppose $S \in \mathcal{L}(V)$. Prove that S is a unitary operator if and only if

$$\{Sv : v \in V \text{ and } \|v\| \leq 1\} = \{v \in V : \|v\| \leq 1\}.$$

Solution. Let $X = \{Sv : v \in V \text{ and } \|v\| \leq 1\}$ and $Y = \{v \in V : \|v\| \leq 1\}$, so that our goal is to show that S is a unitary operator if and only if $X = Y$.

Suppose that S is a unitary operator. If $Sv \in X$ for some $v \in V$ such that $\|v\| \leq 1$, then observe that $\|Sv\| = \|v\| \leq 1$; it follows that $Sv \in Y$ and hence that $X \subseteq Y$. Now suppose $v \in Y$, so that $\|v\| \leq 1$, and note that S is invertible and that S^{-1} is an isometry by 7.53. It follows that

$$v = S(S^{-1}v) \quad \text{and} \quad \|S^{-1}v\| = \|v\| \leq 1.$$

Thus $v \in X$, so that $Y \subseteq X$. We may conclude that $X = Y$.

Now suppose that S is not unitary. If S is not invertible then S is not surjective and thus there exists some necessarily non-zero $w \notin \operatorname{range} S$, which implies $\|w\|^{-1}w \notin \operatorname{range} S$. Thus $\|w\|^{-1}w \in Y$ but $\|w\|^{-1}w \notin X$, so that $X \neq Y$. Suppose that S is invertible. Because S is not unitary, there must exist some $v \in V$ such that $\|Sv\| \neq \|v\|$; note that v must be non-zero. By replacing v with $\|v\|^{-1}v$ if necessary, we may assume that $\|v\| = 1$. Consider the following cases.

Case 1. If $\|Sv\| > 1$ then $Sv \in X$ but $Sv \notin Y$. Thus $X \neq Y$.

Case 2. If $\|Sv\| < 1$ then note that $\|Sv\| \neq 0$ since $v \neq 0$ and S is injective. Let $u = \|Sv\|^{-1}Sv$, so that $u \in Y$. We claim that $u \notin X$. Indeed, if $u \in X$ then $u = Sw$ for some $w \in V$ such that $\|w\| \leq 1$. It follows from the injectivity of S that

$$\frac{Sv}{\|Sv\|} = Sw \quad \Rightarrow \quad v = \|Sv\|w \quad \Rightarrow \quad \|v\| = \|Sv\|\|w\| < 1,$$

contradicting that $\|v\| = 1$. Thus $u \notin X$, so that $X \neq Y$.

In any case, we have shown that if S is not unitary then $X \neq Y$. We may conclude that S is a unitary operator if and only if $X = Y$.

Exercise 7.D.12. Prove or give a counterexample: If $S \in \mathcal{L}(V)$ is invertible and $\|S^{-1}v\| = \|Sv\|$ for every $v \in V$, then S is unitary.

Solution. This is false. For a counterexample, consider the operator $S \in \mathcal{L}(\mathbf{C}^2)$ whose matrix with respect to the standard basis of $\mathbf{C}^2$ is

$$\begin{pmatrix} i & \sqrt{2} \\ \sqrt{2} & -i \end{pmatrix}.$$

A calculation shows that S is its own inverse, so that $\|S^{-1}v\| = \|Sv\|$ for every $v \in V$. However, observe that

$$\|S(1,0)\| = \|(i, \sqrt{2})\| = \sqrt{3} \neq 1 = \|(1,0)\|.$$

Thus S is not a unitary operator.

Exercise 7.D.13. Explain why the columns of a square matrix of complex numbers form an orthonormal list in $\mathbf{C}^n$ if and only if the rows of the matrix form an orthonormal list in $\mathbf{C}^n$.

Solution. Suppose $A \in \mathbf{C}^{n,n}$ and let $S \in \mathcal{L}(\mathbf{C}^n)$ be such that A is the matrix of S with respect to the standard orthonormal basis of $\mathbf{C}^n$. Observe that the columns of A form an orthonormal list in $\mathbf{C}^n$

$$\Leftrightarrow \quad S \text{ is unitary} \quad \Leftrightarrow \quad \text{the rows of } A \text{ form an orthonormal list in } \mathbf{C}^n;$$

the first equivalence is 7.49(e) and the second equivalence is 7.53(e).

Exercise 7.D.14. Suppose $v \in V$ with $\|v\| = 1$ and $b \in \mathbf{F}$. Also suppose $\dim V \geq 2$. Prove that there exists a unitary operator $S \in \mathcal{L}(V)$ such that $\langle Sv, v\rangle = b$ if and only if $|b| \leq 1$.

Solution. If there exists such a unitary operator then, using the Cauchy-Schwarz inequality,

$$|b| = |\langle Sv, v\rangle| \leq \|Sv\|\|v\| = \|v\|^2 \leq 1.$$

Now suppose that $|b| \leq 1$. Let $e_1 = v$ and extend this to an orthonormal basis $e_1, e_2, ..., e_n$ of V. Define $S \in \mathcal{L}(V)$ by

$$Se_1 = be_1 + \sqrt{1 - |b|^2}e_2, \quad Se_2 = -\sqrt{1 - |b|^2}e_1 + be_2, \quad Se_k = e_k \text{ for } k > 2.$$

Given that $e_1, e_2, ..., e_n$ is an orthonormal basis of V, some straightforward calculations show that $Se_1, Se_2, ..., Se_n$ is also an orthonormal basis of V. Thus S is a unitary operator by 7.53(d). Furthermore,

$$\langle Sv, v\rangle = \left\langle be_1 + \sqrt{1 - |b|^2}e_2, e_1\right\rangle = b\|e\|^2 = b.$$

Exercise 7.D.15. Suppose T is a unitary operator on V such that $T - I$ is invertible.

(a) Prove that $(T + I)(T - I)^{-1}$ is a skew operator (meaning that it equals the negative of its adjoint).

(b) Prove that if $\mathbf{F} = \mathbf{C}$, then $i(T + I)(T - I)^{-1}$ is a self-adjoint operator.

The function $z \mapsto i(z + 1)(z - 1)^{-1}$ maps the unit circle in $\mathbf{C}$ (except for the point 1) to $\mathbf{R}$. Thus (b) illustrates the analogy between the unitary operators and the unit circle in $\mathbf{C}$, along with the analogy between the self-adjoint operators and $\mathbf{R}$.

Solution.

(a) Using that $T^*T = TT^* = I$, which holds by 7.53(b), observe that

$$\begin{aligned} \left[(T + I)(T - I)^{-1}\right]^* &= R^*R(T - T^*), \\ -(T + I)(T - I)^{-1} &= (T - T^*)R^*R, \end{aligned} \tag{1}$$

where $R = (T - I)^{-1}$. Observe further that

$$(R^*R)^{-1}T = (2I - T - T^*)T = T(2I - T - T^*) = T(R^*R)^{-1}.$$

Since $T^{-1} = T^*$ (by 7.53(c)), taking the inverse of both sides of the equation $(R^*R)^{-1}T = T(R^*R)^{-1}$ shows that $(R^*R)T^* = T^*(R^*R)$. Hence R^*R commutes with

T^*, and we can similarly show that R^*R commutes with T. Thus R^*R commutes with $T - T^*$. It follows from the expressions in (1) that

$$\left[(T+I)(T-I)^{-1}\right]^* = -(T+I)(T-I)^{-1}.$$

(b) Using part (a) and 7.5(b), we have

$$\left[i(T+I)(T-I)^{-1}\right]^* = (-i)\left[-(T+I)(T-I)^{-1}\right] = i(T+I)(T-I)^{-1}.$$

Exercise 7.D.16. Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$ is self-adjoint. Prove that

$$(T+iI)(T-iI)^{-1}$$

is a unitary operator and 1 is not an eigenvalue of this operator.

Solution. Let $Q = (T+iI)(T-iI)^{-1}$. The complex spectral theorem (7.31) guarantees the existence of an orthonormal basis $e_1, ..., e_n$ of V such that $Te_k = \lambda_k e_k$ for some eigenvalues $\lambda_1, ..., \lambda_n$, which must be real since T is self-adjoint. For each $k \in \{1, ..., n\}$ a routine calculation shows that

$$Qe_k = \frac{\lambda_k + i}{\lambda_k - i} e_k \quad \text{and} \quad \left|\frac{\lambda_k + i}{\lambda_k - i}\right| = 1.$$

Thus, by 7.55, Q is a unitary operator. Now observe that for any $v \in V$,

$$Qv = v \quad \Rightarrow \quad Q^*v = v \quad \Rightarrow \quad (T-iI)v = (T+iI)v \quad \Rightarrow \quad 2iv = 0 \quad \Rightarrow \quad v = 0.$$

Thus 1 is not an eigenvalue of Q.

Exercise 7.D.17. Explain why the characterization of unitary matrices given by 7.57 holds.

Solution. Let $Q \in \mathbf{F}^{n,n}$ be a matrix and let $S \in \mathcal{L}(\mathbf{F}^n)$ be such that Q is the matrix of S with respect to the standard orthonormal basis of $\mathbf{F}^n$. By definition, Q is a unitary matrix if and only if 7.49(e) holds (we are taking both orthonormal bases $e_1, ..., e_n$ and $f_1, ..., f_m$ in the statement of 7.49 to be the standard orthonormal basis of $\mathbf{F}^n$). Observe that

$$7.57(\text{a}) \quad \Leftrightarrow \quad 7.49(\text{e}) \quad \Leftrightarrow \quad 7.49(\text{a}) \quad \Leftrightarrow \quad 7.53(\text{a}) \quad \Leftrightarrow \quad 7.53(\text{e}) \quad \Leftrightarrow \quad 7.57(\text{b}).$$

Thus 7.57(a) and 7.57(b) are equivalent. After identifying elements of $\mathbf{F}^n$ with column vectors, i.e. n-by-1 matrices, note that $\|Qv\| = \|Sv\|$ for any $v \in \mathbf{F}^n$. As noted before, Q is a unitary matrix if and only if S is a unitary operator. The equivalence of 7.57(a) and 7.57(c) is now immediate from the definition of a unitary operator. Finally, by 7.9, the matrix of S^* is Q^*. The equivalence of 7.57(a) and 7.57(d) then follows from the equivalence of 7.53(a) and 7.53(b).

Exercise 7.D.18. A square matrix A is called *symmetric* if it equals its transpose. Prove that if A is a symmetric matrix with real entries, then there exists a unitary matrix Q with real entries such that Q^*AQ is a diagonal matrix.

Solution. Suppose $A \in \mathbf{R}^{n,n}$ and let $e_1, ..., e_n$ be the standard orthonormal basis of $\mathbf{R}^n$. Let $T \in \mathcal{L}(\mathbf{R}^n)$ be such that $\mathcal{M}(T, (e_1, ..., e_n)) = A$. By assumption $A = A^{\mathrm{t}}$ and thus T is self-adjoint. It follows from the real spectral theorem (7.29) that there exists an orthonormal basis $f_1, ..., f_n$ of $\mathbf{R}^n$ such that $D := \mathcal{M}(T, (f_1, ..., f_n))$ is diagonal. Let

$$Q = \mathcal{M}(I, (f_1, ..., f_n), (e_1, ..., e_n)).$$

Certainly the identity operator on $\mathbf{R}^n$ is an isometry. It follows from the equivalence of 7.49(a) and 7.49(e) that Q is a unitary matrix. On one hand, 7.57 shows that $Q^{-1} = Q^*$. On the other hand, 3.82 shows that

$$Q^{-1} = \mathcal{M}(T, (e_1, ..., e_n), (f_1, ..., f_n)).$$

Thus, using the change-of-basis formula (3.84), we have that $Q^*AQ = D$ is diagonal.

Exercise 7.D.19. Suppose n is a positive integer. For this exercise, we adopt the notation that a typical element z of $\mathbf{C}^n$ is denoted by $z = (z_0, z_1, ..., z_{n-1})$. Define linear functionals $\omega_0, \omega_1, ..., \omega_{n-1}$ on $\mathbf{C}^n$ by

$$\omega_j(z_0, z_1, ..., z_{n-1}) = \frac{1}{\sqrt{n}} \sum_{m=0}^{n-1} z_m e^{-2\pi i j m/n}.$$

The *discrete Fourier transform* is the operator $\mathcal{F} : \mathbf{C}^n \to \mathbf{C}^n$ defined by

$$\mathcal{F}z = (\omega_0(z), \omega_1(z), ..., \omega_{n-1}(z)).$$

(a) Show that $\mathcal{F}$ is a unitary operator on $\mathbf{C}^n$.

(b) Show that if $(z_0, ..., z_{n-1}) \in \mathbf{C}^n$ and z_n is defined to equal z_0, then

$$\mathcal{F}^{-1}(z_0, z_1, ..., z_{n-1}) = \mathcal{F}(z_n, z_{n-1}, ..., z_1).$$

(c) Show that $\mathcal{F}^4 = I$.

The discrete Fourier transform has many important applications in data analysis. The usual Fourier transform involves expressions of the form $\int_{-\infty}^{\infty} f(x)e^{-2\pi i t x}\,dx$ for complex-valued integrable functions f defined on $\mathbf{R}$.

Solution.

(a) For this exercise, let us count from 0 to $n-1$ instead of from 1 to n for columns of matrices etc. Let $e_0, ..., e_{n-1}$ be the standard orthonormal basis of $\mathbf{C}^n$ and observe that

$$\omega_j e_k = \frac{\alpha^{jk}}{\sqrt{n}}, \quad \text{where} \quad \alpha = e^{-2\pi i/n}.$$

It follows that

$$Q := \mathcal{M}(\mathcal{F}, (e_0, ..., e_{n-1})) = \frac{1}{\sqrt{n}} \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \alpha & \alpha^2 & \cdots & \alpha^{n-1} \\ 1 & \alpha^2 & \alpha^4 & \cdots & \alpha^{2(n-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \alpha^{n-1} & \alpha^{2(n-1)} & \cdots & \alpha^{(n-1)(n-1)} \end{pmatrix}.$$

Thus the Euclidean inner product of the j^{th} and k^{th} columns of Q is

$$\frac{1}{n}\sum_{m=0}^{n-1}\left(\alpha^j\overline{\alpha}^k\right)^m = \frac{1}{n}\sum_{m=0}^{n-1}\left(e^{2\pi i(k-j)/n}\right)^m.$$

If $j = k$ then $e^{2\pi i(k-j)/n} = 1$, so that

$$\frac{1}{n}\sum_{m=0}^{n-1}\left(e^{2\pi i(k-j)/n}\right)^m = \frac{1}{n}\sum_{m=0}^{n-1} 1 = 1,$$

and if $j \neq k$ then $e^{2\pi i(k-j)/n} \neq 1$ and the geometric series formula gives us

$$\frac{1}{n}\sum_{m=0}^{n-1}\left(e^{2\pi i(k-j)/n}\right)^m = \frac{1-\left[e^{2\pi i(k-j)/n}\right]^n}{n\left(1-e^{2\pi i(k-j)/n}\right)} = \frac{1-e^{2\pi i(k-j)}}{n\left(1-e^{2\pi i(k-j)/n}\right)} = 0,$$

where we have used that $e^{2\pi i(k-j)} = 1$ since $k - j$ is an integer. We have now shown that the columns of Q form an orthonormal list in $\mathbf{C}^n$ with respect to the Euclidean inner product. It follows from 7.49 that $\mathcal{F}$ is a unitary operator.

(b) Define $\mathcal{E} \in \mathcal{L}(\mathbf{C}^n)$ by $\mathcal{E}(z_0, z_1, ..., z_{n-1}) = \mathcal{F}(z_0, z_{n-1}, ..., z_1)$ and observe that

$$\mathcal{E}e_0 = \mathcal{F}e_0 \quad \text{and} \quad \mathcal{E}e_k = \mathcal{F}e_{n-k} \text{ for } k \in \{1, ..., n-1\}.$$

Thus, letting $X = \mathcal{M}(\mathcal{E}, (e_0, ..., e_{n-1}))$, we have

$$X_{j,k} = \begin{cases} Q_{j,k} & \text{if } k = 0, \\ Q_{j,n-k} & \text{if } k \in \{1, ..., n-1\}. \end{cases}$$

For $k \in \{1, ..., n-1\}$, observe that

$$\overline{\alpha}^{kj} = e^{2\pi ijk/n} = e^{2\pi ijk/n}e^{-2\pi ij} = e^{-2\pi ij(n-k)/n} = \alpha^{j(n-k)}.$$

It follows that

$$Q^*_{j,k} = \overline{Q_{k,j}} = \frac{\overline{\alpha}^{kj}}{\sqrt{n}} = \frac{\alpha^{j(n-k)}}{\sqrt{n}} = Q_{j,n-k}.$$

By inspection of Q, it is also clear that $Q^*_{j,0} = Q_{j,0}$. Thus $Q^*_{j,k} = X_{j,k}$ for all j, k and it follows that $\mathcal{E} = \mathcal{F}^* = \mathcal{F}^{-1}$.

(c) The formula in part (b) shows that

$$\mathcal{F}^2(z_0, z_1, ..., z_{n-1}) = (z_0, z_{n-1}, ..., z_1) \quad \Rightarrow \quad \mathcal{F}^4(z_0, z_1, ..., z_{n-1}) = (z_0, z_1, ..., z_{n-1}).$$

Exercise 7.D.20. Suppose A is a square matrix with linearly independent columns. Prove that there exist unique matrices R and Q such that R is lower triangular with only positive numbers on its diagonal, Q is unitary, and $A = RQ$.

Solution. If A has linearly independent columns then Exercise 7.A.7 (b) shows that A^* also has linearly independent columns. It follows from 7.58 that there exists a unitary matrix P and an upper triangular matrix U with only positive numbers on its diagonal such that $A^* = PU$. It follows that $A = RQ$, where $R = U^*$ and $Q = P^*$. Observe that R is lower triangular with only positive numbers on its diagonal and that Q is unitary by 7.57.

7.E. Singular Value Decomposition

Exercise 7.E.1. Suppose $T \in \mathcal{L}(V, W)$. Show that $T = 0$ if and only if all singular values of T are 0.

Solution. Let N be the number of positive singular values of T. It will suffice to show that $T = 0$ if and only if $N = 0$. Indeed, 7.68(b) shows that

$$N = 0 \quad \Leftrightarrow \quad \dim \operatorname{range} T = 0 \quad \Leftrightarrow \quad T = 0.$$

Exercise 7.E.2. Suppose $T \in \mathcal{L}(V, W)$ and $s > 0$. Prove that s is a singular value of T if and only if there exist nonzero vectors $v \in V$ and $w \in W$ such that

$$Tv = sw \quad \text{and} \quad T^* w = sv.$$

*The vectors v, w satisfying both equations above are called a **Schmidt pair**. Erhard Schmidt introduced the concept of singular values in 1907.*

Solution. Suppose that s is a singular value of T. Thus, letting $s_1, \ldots, s_m$ be the positive singular values of T, we have $s = s_k$ for some $k \in \{1, \ldots, m\}$. Let $e_1, \ldots, e_m$ and $f_1, \ldots, f_m$ be the orthonormal lists obtained from the SVD (7.70). It follows from 7.70 and 7.75 that

$$Te_k = s_k f_k \quad \text{and} \quad T^* f_k = s_k e_k.$$

Thus we can take $v = e_k$ and $w = f_k$.

Conversely, suppose there exist non-zero vectors $v \in V$ and $w \in W$ such that

$$Tv = sw \quad \text{and} \quad T^* w = sv.$$

It follows that $T^* T v = s^2 v$, so that s^2 is an eigenvalue of $T^* T$. Thus $\sqrt{s^2} = s$ is a singular value of T, where we have used that s is non-negative.

Exercise 7.E.3. Give an example of $T \in \mathcal{L}(\mathbf{C}^2)$ such that 0 is the only eigenvalue of T and the singular values of T are $5, 0$.

Solution. Let $T \in \mathcal{L}(\mathbf{C}^2)$ be the operator whose matrix with respect to the standard basis of $\mathbf{C}^2$ is

$$\begin{pmatrix} 0 & 5 \\ 0 & 0 \end{pmatrix}$$

and note that 0 is the only eigenvalue of T. Note further that the matrix of $T^* T$ is

$$\begin{pmatrix} 0 & 0 \\ 5 & 0 \end{pmatrix} \begin{pmatrix} 0 & 5 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 25 \end{pmatrix}.$$

Thus the singular values of T are $5, 0$.

Exercise 7.E.4. Suppose that $T \in \mathcal{L}(V, W)$, s_1 is the largest singular value of T, and s_n is the smallest singular value of T. Prove that

$$\{\|Tv\| : v \in V \text{ and } \|v\| = 1\} = [s_n, s_1].$$

Solution. We consider several cases.

Case 1. If $s_1 = s_n = 0$ then all singular values of T equal 0 and Exercise 7.E.1 shows that $T = 0$. Thus

$$\{\|Tv\| : v \in V \text{ and } \|v\| = 1\} = \{0\} = [0, 0] = [s_n, s_1].$$

Case 2. If $s_1 = s_n > 0$ then note that all singular values of $s_1^{-1}T$ equal 1. It follows from 7.69 that $s_1^{-1}T$ is an isometry and hence that

$$\{\|Tv\| : v \in V \text{ and } \|v\| = 1\} = \{s_1\} = [s_1, s_1] = [s_n, s_1].$$

Case 3. Suppose that $s_1 > s_n$, which implies $s_1 > 0$. Let $s_1, \ldots, s_m$ be the positive singular values of T and let $e_1, \ldots, e_m$ and $f_1, \ldots, f_m$ be the orthonormal lists obtained from the SVD (7.70). For any $v \in V$, 7.70 and 6.24 show that

$$Tv = s_1\langle v, e_1\rangle f_1 + \cdots + s_m\langle v, e_m\rangle f_m \quad \Rightarrow \quad \|Tv\|^2 = s_1^2|\langle v, e_1\rangle|^2 + \cdots + s_m^2|\langle v, e_m\rangle|^2.$$

For any $v \in V$ such that $\|v\| = 1$, it follows from Bessel's inequality (6.26) that

$$\|Tv\|^2 \leq s_1^2\left(|\langle v, e_1\rangle|^2 + \cdots + |\langle v, e_m\rangle|^2\right) \leq s_1^2\|v\|^2 = s_1^2 \quad \Rightarrow \quad \|Tv\| \leq s_1.$$

Thus $\{\|Tv\| : v \in V \text{ and } \|v\| = 1\} \subseteq [0, s_1]$. To complete the exercise, we now consider two subcases.

Case 3.1. If $s_n > 0$ then it must be that $m = n = \dim V$, so that $e_1, \ldots, e_n$ is an orthonormal basis of V. It follows that

$$\|Tv\|^2 \geq s_n^2\left(|\langle v, e_1\rangle|^2 + \cdots + |\langle v, e_n\rangle|^2\right) = s_n^2\|v\|^2 = s_n^2 \quad \Rightarrow \quad \|Tv\| \geq s_n.$$

Thus $\{\|Tv\| : v \in V \text{ and } \|v\| = 1\} \subseteq [s_n, s_1]$. For $s \in [s_n, s_1]$, let

$$v = \sqrt{\frac{s^2 - s_n^2}{s_1^2 - s_n^2}}e_1 + \sqrt{\frac{s_1^2 - s^2}{s_1^2 - s_n^2}}e_n.$$

A calculation shows that $\|v\| = 1$ and $\|Tv\| = s$. Thus $[s_n, s_1] \subseteq \{\|Tv\| : v \in V \text{ and } \|v\| = 1\}$ and we may conclude that

$$\{\|Tv\| : v \in V \text{ and } \|v\| = 1\} = [s_n, s_1].$$

Case 3.2. If $s_n = 0$ then it must be that $m < n = \dim V$. Extend the orthonormal list $e_1, \ldots, e_m$ to an orthonormal basis $e_1, \ldots, e_n$ of V. As noted in the discussion after the proof of 7.70, it follows that $e_n \in \operatorname{null} T$. For $s \in [0, s_1]$, let

$$v = \frac{s}{s_1} e_1 + \sqrt{1 - \frac{s^2}{s_1^2}} e_n.$$

A calculation shows that $\|v\| = 1$ and $\|Tv\| = s$. Thus $[0, s_1] \subseteq \{\|Tv\| : v \in V \text{ and } \|v\| = 1\}$ and we may conclude that

$$\{\|Tv\| : v \in V \text{ and } \|v\| = 1\} = [0, s_1] = [s_n, s_1].$$

Exercise 7.E.5. Suppose $T \in \mathcal{L}(\mathbf{C}^2)$ is defined by $T(x, y) = (-4y, x)$. Find the singular values of T.

Solution. The matrix of T with respect to the standard basis of $\mathbf{C}^2$ is

$$A = \begin{pmatrix} 0 & -4 \\ 1 & 0 \end{pmatrix} \quad \Rightarrow \quad A^* A = \begin{pmatrix} 0 & 1 \\ -4 & 0 \end{pmatrix} \begin{pmatrix} 0 & -4 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 16 \end{pmatrix}.$$

Thus the singular values of T are $4, 1$.

Exercise 7.E.6. Find the singular values of the differentiation operator $D \in \mathcal{L}(\mathcal{P}_2(\mathbf{R}))$ defined by $Dp = p'$, where the inner product on $\mathcal{P}_2(\mathbf{R})$ is as in Example 6.34.

Solution. As shown in Example 6.34, the list

$$e_1 = \sqrt{\tfrac{1}{2}}, \quad e_2 = \sqrt{\tfrac{3}{2}} x, \quad e_3 = \sqrt{\tfrac{45}{8}} (x^2 - \tfrac{1}{3})$$

is an orthonormal basis of $\mathcal{P}_2(\mathbf{R})$ with respect to the inner product given in Example 6.34. Observe that

$$De_1 = 0, \quad De_2 = \sqrt{\tfrac{3}{2}} = \sqrt{3} e_1, \quad De_3 = \sqrt{\tfrac{45}{8}} (2x) = \sqrt{15} e_2.$$

Thus the matrix of D with respect to e_1, e_2, e_3 is

$$A = \begin{pmatrix} 0 & \sqrt{3} & 0 \\ 0 & 0 & \sqrt{15} \\ 0 & 0 & 0 \end{pmatrix} \quad \Rightarrow \quad A^* A = \begin{pmatrix} 0 & 0 & 0 \\ \sqrt{3} & 0 & 0 \\ 0 & \sqrt{15} & 0 \end{pmatrix} \begin{pmatrix} 0 & \sqrt{3} & 0 \\ 0 & 0 & \sqrt{15} \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 15 \end{pmatrix}.$$

It follows that the singular values of D are $\sqrt{15}, \sqrt{3}, 0$.

Exercise 7.E.7. Suppose that $T \in \mathcal{L}(V)$ is self-adjoint or that $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$ is normal. Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of T, each included in this list as many times as the dimension of the corresponding eigenspace. Show that the singular values of T are $|\lambda_1|, \ldots, |\lambda_n|$, after these numbers have been sorted into decreasing order.

Solution. By Exercise 7.B.14/Exercise 7.B.15, we know that there is a decomposition of V into a direct sum of mutually orthogonal eigenspaces of T. Thus $n = \dim V$ and there exists an orthonormal basis $e_1, ..., e_n$ of V such that $Te_k = \lambda_k e_k$ for each $k \in \{1, ..., n\}$. It follows from 7.21(e) that

$$T^* e_k = \overline{\lambda_k} e_k \quad \Rightarrow \quad T^* T e_k = |\lambda_k|^2 e_k.$$

Thus each vector in the list $|\lambda_1|, ..., |\lambda_n|$ is a singular value of T. As noted in the table on p. 272, the length of the list of singular values of T is exactly $\dim V = n$. Thus, after being sorted into decreasing order, $|\lambda_1|, ..., |\lambda_n|$ are the singular values of T.

Exercise 7.E.8. Suppose $T \in \mathcal{L}(V, W)$. Suppose $s_1 \geq s_2 \geq \cdots \geq s_m > 0$ and $e_1, ..., e_m$ is an orthonormal list in V and $f_1, ..., f_m$ is an orthonormal list in W such that

$$Tv = s_1 \langle v, e_1 \rangle f_1 + \cdots + s_m \langle v, e_m \rangle f_m$$

for every $v \in V$.

(a) Prove that $f_1, ..., f_m$ is an orthonormal basis of $\operatorname{range} T$.

(b) Prove that $e_1, ..., e_m$ is an orthonormal basis of $(\operatorname{null} T)^\perp$.

(c) Prove that $s_1, ..., s_m$ are the positive singular values of T.

(d) Prove that if $k \in \{1, ..., m\}$, then e_k is an eigenvector of T^*T with corresponding eigenvalue s_k^2.

(e) Prove that

$$TT^* w = s_1^2 \langle w, f_1 \rangle f_1 + \cdots + s_m^2 \langle w, f_m \rangle f_m$$

for all $w \in W$.

Solution.

(a) Because the equation

$$Tv = s_1 \langle v, e_1 \rangle f_1 + \cdots + s_m \langle v, e_m \rangle f_m$$

holds for all $v \in V$, it is clear that $f_1, ..., f_m$ spans $\operatorname{range} T$. By assumption $f_1, ..., f_m$ is orthonormal and hence linearly independent (6.25). Thus $f_1, ..., f_m$ is an orthonormal basis of $\operatorname{range} T$.

(b) If $v \in \operatorname{null} T$ then

$$0 = Tv = s_1 \langle v, e_1 \rangle f_1 + \cdots + s_m \langle v, e_m \rangle f_m \quad \Rightarrow \quad \langle v, e_1 \rangle = \cdots = \langle v, e_m \rangle = 0,$$

where we have used that $f_1, ..., f_m$ is linearly independent and that each s_k is strictly positive. Thus $\{e_1, ..., e_m\} \subseteq (\operatorname{null} T)^\perp$. By assumption $e_1, ..., e_m$ is orthonormal. Furthermore, by 6.67 we have $\dim (\operatorname{null} T)^\perp = \dim \operatorname{range} T = m$. It follows from 6.28 that $e_1, ..., e_m$ is an orthonormal basis of $(\operatorname{null} T)^\perp$.

(c) Extend the orthonormal list $e_1, ..., e_m$ to an orthonormal basis $e_1, ..., e_{\dim V}$ of V and extend the orthonormal list $f_1, ..., f_m$ to an orthonormal basis $f_1, ..., f_{\dim V}$ of W. As in the discussion after the proof of 7.70, we have

$$Te_k = \begin{cases} s_k f_k & \text{if } k \in \{1, \ldots, m\}, \\ 0 & \text{otherwise.} \end{cases}$$

For any $j \in \{1, \ldots, \dim W\}$, it follows that

$$T^* f_j = \sum_{k=1}^{\dim V} \langle T^* f_j, e_k \rangle e_k = \sum_{k=1}^{\dim V} \langle f_j, Te_k \rangle e_k = \begin{cases} s_j e_j & \text{if } j \in \{1, \ldots, m\}, \\ 0 & \text{otherwise.} \end{cases}$$

Thus

$$T^* T e_k = \begin{cases} s_k^2 e_k & \text{if } k \in \{1, \ldots, m\}, \\ 0 & \text{otherwise.} \end{cases}$$

From this expression and the fact that $s_1 \geq s_2 \geq \cdots \geq s_m > 0$, we see that the list $s_1^2, \ldots, s_m^2$ consists of the eigenvalues of T^*T listed in decreasing order with each eigenvalue appearing as many times as the dimension of the corresponding eigenspace of T^*T. It follows that the singular values of T are precisely $s_1, \ldots, s_m$.

(d) We proved this in part (c).

(e) Because the equation

$$Tv = s_1 \langle v, e_1 \rangle f_1 + \cdots + s_m \langle v, e_m \rangle f_m$$

holds for all $v \in V$, the calculation in the proof of 7.75 shows that equation 7.77 holds for all $w \in W$. Thus, for any $w \in W$,

$$TT^* w = T\left(\sum_{k=1}^{m} s_k \langle w, f_k \rangle e_k \right) = \sum_{j=1}^{m} s_j \left\langle \sum_{k=1}^{m} s_k \langle w, f_k \rangle e_k, e_j \right\rangle f_j = \sum_{j=1}^{m} s_j^2 \langle w, f_j \rangle f_j.$$

Exercise 7.E.9. Suppose $T \in \mathcal{L}(V, W)$. Show that T and T^* have the same positive singular values.

Solution. Let $s_1, \ldots, s_m$ be the positive singular values of T^*. By the SVD (7.70), there exists an orthonormal list $f_1, \ldots, f_m$ in W and an orthonormal list $e_1, \ldots, e_m$ in V such that

$$T^* w = s_1 \langle w, f_1 \rangle e_1 + \cdots + s_m \langle w, f_m \rangle e_m$$

for every $w \in W$. It follows from 7.75 that

$$Tv = (T^*)^* v = s_1 \langle v, e_1 \rangle f_1 + \cdots + s_m \langle v, e_m \rangle f_m$$

for every $v \in V$. Thus, by Exercise 7.E.8, the positive singular values of T are $s_1, \ldots, s_m$. We have now shown that if $s_1, \ldots, s_m$ are the positive singular values of T^* then $s_1, \ldots, s_m$ are the positive singular values of T. Replacing T with T^* in this result and using that $(T^*)^* = T$ gives us the desired equivalence.

Exercise 7.E.10. Suppose $T \in \mathcal{L}(V, W)$ has singular values $s_1, \ldots, s_n$. Prove that if T is an invertible linear map, then T^{-1} has singular values

$$\frac{1}{s_n}, \ldots, \frac{1}{s_1}.$$

Solution. Because T is invertible, 7.68 shows that each singular value of T must be positive. By the SVD (7.70), there exists an orthonormal list $e_1, \ldots, e_n$ in V and an orthonormal list $f_1, \ldots, f_n$ in W such that

$$Tv = s_1 \langle v, e_1 \rangle f_1 + \cdots + s_n \langle v, e_n \rangle f_n$$

for every $v \in V$. Note that $T^{-1} = T^\dagger$ by 6.69(a). It follows from 7.75 that

$$T^{-1} w = \frac{\langle w, f_n \rangle}{s_n} e_n + \cdots + \frac{\langle w, f_1 \rangle}{s_1} e_1$$

for every $w \in W$. Since

$$s_1 \geq \cdots \geq s_n > 0 \quad \Rightarrow \quad \frac{1}{s_n} \geq \cdots \geq \frac{1}{s_1} > 0,$$

Exercise 7.E.8 shows that $s_n^{-1}, \ldots, s_1^{-1}$ are the singular values of T^{-1} (we are taking the orthonormal lists required by Exercise 7.E.8 to be $f_n, \ldots, f_1$ and $e_n, \ldots, e_1$).

Exercise 7.E.11. Suppose that $T \in \mathcal{L}(V, W)$ and $v_1, \ldots, v_n$ is an orthonormal basis of V. Let $s_1, \ldots, s_n$ denote the singular values of T.

(a) Prove that $\|Tv_1\|^2 + \cdots + \|Tv_n\|^2 = s_1^2 + \cdots + s_n^2$.

(b) Prove that if $W = V$ and T is a positive operator, then

$$\langle Tv_1, v_1 \rangle + \cdots + \langle Tv_n, v_n \rangle = s_1 + \cdots + s_n.$$

See the comment after Exercise 5 in Section 7A.

Solution.

(a) Let $s_1, \ldots, s_m$ be the positive singular values of T. As discussed after the proof of 7.70, there exists an orthonormal basis $e_1, \ldots, e_n$ of V and an orthonormal basis $f_1, \ldots, f_{\dim W}$ such that

$$Te_k = \begin{cases} s_k f_k & \text{if } k \in \{1, \ldots, m\}, \\ 0 & \text{otherwise.} \end{cases}$$

It follows from Exercise 7.A.5 that

$$\sum_{k=1}^{n} \|Tv_k\|^2 = \sum_{k=1}^{n} \|Te_k\|^2 = \sum_{k=1}^{m} \|s_k f_k\|^2 = \sum_{k=1}^{m} s_k^2 = \sum_{k=1}^{n} s_k^2,$$

where we have used that $s_k = 0$ if $k > m$ for the last equality.

(b) Because T is a positive operator, T must be self-adjoint and each of its eigenvalues must be non-negative. Recalling that singular values are non-negative, Exercise 7.E.7 shows that $s_1, \ldots, s_n$ must be the eigenvalues of T. It follows from 7.39 that $\sqrt{s_1}, \ldots, \sqrt{s_n}$ are the eigenvalues of $\sqrt{T}$ and another application of Exercise 7.E.7 shows that the singular values of $\sqrt{T}$ are $\sqrt{s_1}, \ldots, \sqrt{s_n}$. Thus, by part (a),

$$\sum_{k=1}^{n} \langle Tv_n, v_n \rangle = \sum_{k=1}^{n} \|\sqrt{T} v_n\|^2 = s_1 + \cdots + s_n.$$

Exercise 7.E.12.

(a) Give an example of a finite-dimensional vector space and an operator T on it such that the singular values of T^2 do not equal the squares of the singular values of T.

(b) Suppose $T \in \mathcal{L}(V)$ is normal. Prove that the singular values of T^2 equal the squares of the singular values of T.

Solution.

(a) Consider the operator $T \in \mathcal{L}(\mathbf{F}^2)$ whose matrix with respect to the standard basis of $\mathbf{F}^2$ is

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

A calculation shows that the singular values of T are $1, 0$. However, since $T^2 = 0$, the singular values of T^2 are $0, 0$.

(b) Suppose that $s_1, \ldots, s_n$ are the singular values of T, so that there is an orthonormal basis $e_1, \ldots, e_n$ of V such that $T^*Te_k = s_k^2 e_k$. Using that T is normal, observe that

$$(T^2)^* T^2 e_k = (T^*T)^2 e_k = s_k^4 e_k.$$

Thus the singular values of T^2 are $s_1^2, \ldots, s_n^2$.

Exercise 7.E.13. Suppose $T_1, T_2 \in \mathcal{L}(V)$. Prove that T_1 and T_2 have the same singular values if and only if there exist unitary operators $S_1, S_2 \in \mathcal{L}(V)$ such that $T_1 = S_1 T_2 S_2$.

Solution. Suppose there exist such unitary operators and observe that

$$\begin{aligned} T_1^* = S_2^* T_2^* S_1^* = S_2^{-1} T_2^* S_1^{-1} \quad &\Rightarrow \quad T_1^* T_1 = S_2^{-1} T_2^* T_2 S_2 \\ &\Rightarrow \quad T_1^* T_1 - \lambda I = S_2^{-1} (T_2^* T_2 - \lambda I) S_2 \text{ for all } \lambda \in \mathbf{F} \\ &\Rightarrow \quad \dim E(\lambda, T_1^* T_1) = \dim E(\lambda, T_2^* T_2) \text{ for all } \lambda \in \mathbf{F}. \end{aligned}$$

Thus T_1 and T_2 have the same singular values.

Now suppose that T_1 and T_2 have the same singular values $s_1, \dots, s_m, \dots, s_n$, where $s_1, \dots, s_m$ are positive and $s_k = 0$ if $k > m$ (it may be the case that $m = n$). By the SVD (7.70), there exist orthonormal bases

$$e_1, \dots, e_n, \quad f_1, \dots, f_n, \quad g_1, \dots, g_n, \quad h_1, \dots, h_n$$

of V such that

$$T_1 e_k = \begin{cases} s_k f_k & \text{if } k \in \{1, \dots, m\}, \\ 0 & \text{otherwise,} \end{cases} \quad \text{and} \quad T_2 g_k = \begin{cases} s_k h_k & \text{if } k \in \{1, \dots, m\}, \\ 0 & \text{otherwise.} \end{cases}$$

Define $S_1, S_2 \in \mathcal{L}(V)$ by $S_1 h_k = f_k$ and $S_2 e_k = g_k$ and note that S_1 and S_2 are unitary operators by 7.53(d). Furthermore,

$$k \in \{1, \dots, m\} \quad \Rightarrow \quad S_1 T_2 S_2 e_k = S_1 T_2 g_k = s_k S_1 h_k = s_k f_k = T_1 e_k,$$

$$k > m \quad \Rightarrow \quad S_1 T_2 S_2 e_k = S_1 T_2 g_k = 0 = T_1 e_k.$$

Thus $T_1 = S_1 T_2 S_2$.

Exercise 7.E.14. Suppose $T \in \mathcal{L}(V, W)$. Let s_n denote the smallest singular value of T. Prove that $s_n \|v\| \leq \|Tv\|$ for every $v \in V$.

Solution. If $v = 0$ then the inequality is clear, so suppose that $v \neq 0$. It follows from Exercise 7.E.4 that

$$s_n \leq \left\| T\left(\frac{v}{\|v\|}\right) \right\| \quad \Rightarrow \quad s_n \|v\| \leq \|Tv\|.$$

Exercise 7.E.15. Suppose $T \in \mathcal{L}(V)$ and $s_1 \geq \cdots \geq s_n$ are the singular values of T. Prove that if λ is an eigenvalue of T, then $s_1 \geq |\lambda| \geq s_n$.

Solution. Let $v \in V$ be such that $Tv = \lambda v$ and $\|v\| = 1$. It follows from Exercise 7.E.4 that

$$|\lambda| = \|\lambda v\| = \|Tv\| \in [s_n, s_1].$$

Exercise 7.E.16. Suppose $T \in \mathcal{L}(V, W)$. Prove that $(T^*)^\dagger = \left(T^\dagger\right)^*$.

Compare the result in this exercise to the analogous result for invertible linear maps [see 7.5(f)].

Solution. Let $s_1, \dots, s_m$ be the positive singular values of T. The SVD (7.70) and 7.75 imply the existence of an orthonormal list $e_1, \dots, e_m$ in V and an orthonormal list $f_1, \dots, f_m$ in W such that

$$T^* w = s_1 \langle w, f_1 \rangle e_1 + \cdots + s_m \langle w, f_m \rangle e_m$$

for every $w \in W$. Since $s_1, \ldots, s_m$ are also the positive singular values of T^* (see Exercise 7.E.9), another application of 7.75 shows that

$$(T^*)^\dagger v = \frac{\langle v, e_1 \rangle}{s_1} f_1 + \cdots + \frac{\langle v, e_m \rangle}{s_m} f_m$$

for every $v \in V$.

Also by 7.75, we have

$$T^\dagger w = \frac{\langle w, f_m \rangle}{s_m} e_m + \cdots + \frac{\langle w, f_1 \rangle}{s_1} e_1$$

for every $w \in W$. It follows from Exercise 7.E.8 that $s_m^{-1}, \ldots, s_m^{-1}$ are the positive singular values of $T^\dagger$ and we can apply 7.75 once more to see that

$$(T^\dagger)^* v = \frac{\langle v, e_m \rangle}{s_m} f_m + \cdots + \frac{\langle v, e_1 \rangle}{s_1} f_1$$

for every $v \in V$. Thus $(T^*)^\dagger = (T^\dagger)^*$.

Exercise 7.E.17. Suppose $T \in \mathcal{L}(V)$. Prove that T is self-adjoint if and only if $T^\dagger$ is self-adjoint.

Solution. Observe that

$$T = T^* \quad \Leftrightarrow \quad T^\dagger = (T^*)^\dagger \quad \Leftrightarrow \quad T^\dagger = (T^\dagger)^*,$$

where we have used Exercise 6.C.23 for the first equivalence and Exercise 7.E.16 for the second equivalence.

7.F. Consequences of Singular Value Decomposition

Exercise 7.F.1. Prove that if $S, T \in \mathcal{L}(V, W)$, then $|\|S\| - \|T\|| \leq \|S - T\|$.

The inequality above is called the ***reverse triangle inequality****.*

Solution. The proof is essentially the same as the proof given in Exercise 4.2.

Exercise 7.F.2. Suppose that $T \in \mathcal{L}(V)$ is self-adjoint or that $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$ is normal. Prove that

$$\|T\| = \max\{|\lambda| : \lambda \text{ is an eigenvalue of } T\}.$$

Solution. Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of T, included in this list as many times as the dimension of the corresponding eigenspace. As shown in Exercise 7.E.7, the singular values of T are $|\lambda_1|, \ldots, |\lambda_n|$ (sorted in decreasing order). It follows from 7.88(a) that

$$\|T\| = \max\{|\lambda| : \lambda \text{ is an eigenvalue of } T\}.$$

Exercise 7.F.3. Suppose $T \in \mathcal{L}(V, W)$ and $v \in V$. Prove that

$$\|Tv\| = \|T\|\|v\| \quad \Leftrightarrow \quad T^*Tv = \|T\|^2 v.$$

Solution. Suppose that $T^*Tv = \|T\|^2 v$. Using 7.88(c) and 7.91, it follows that

$$\|T\|^2\|v\| = \|T^*Tv\| \leq \|T^*\|\|Tv\| = \|T\|\|Tv\| \quad \Rightarrow \quad \|T\|\|v\| \leq \|Tv\|.$$

Since $\|Tv\| \leq \|T\|\|v\|$ (by 7.88(c)), we may conclude that $\|Tv\| = \|T\|\|v\|$.

Suppose that $\|Tv\| = \|T\|\|v\|$ and observe that

$$\begin{aligned}
\left\|T^*Tv - \|T\|^2 v\right\|^2 &= \langle T^*Tv - \|T\|^2 v, T^*Tv - \|T\|^2 v\rangle \\
&= \|T^*Tv\|^2 + \|T\|^4\|v\|^2 - 2\,\mathrm{Re}\langle T^*Tv, \|T\|^2 v\rangle \\
&\leq \|T^*\|^2\|Tv\|^2 + \|T\|^2\|Tv\|^2 - 2\|T\|^2\|Tv\|^2 && (7.88(c)) \\
&= 0. && (7.91)
\end{aligned}$$

Thus $T^*Tv = \|T\|^2 v$.

Exercise 7.F.4. Suppose $T \in \mathcal{L}(V, W), v \in V$, and $\|Tv\| = \|T\|\|v\|$. Prove that if $u \in V$ and $\langle u, v\rangle = 0$, then $\langle Tu, Tv\rangle = 0$.

Solution. By Exercise 7.F.4 we must have $T^*Tv = \|T\|^2 v$. It follows that

$$\langle Tu, Tv \rangle = \langle u, T^*Tv \rangle = \langle u, \|T\|^2 v \rangle = \|T\|^2 \langle u, v \rangle = 0.$$

Exercise 7.F.5. Suppose U is a finite-dimensional inner product space, $T \in \mathcal{L}(V, U)$, and $S \in \mathcal{L}(U, W)$. Prove that

$$\|ST\| \leq \|S\|\|T\|.$$

Solution. By 7.88(c) we have, for any $v \in V$,

$$\|STv\| \leq \|S\|\|Tv\| \leq \|S\|\|T\|\|v\|.$$

It follows from the minimality of $\|ST\|$ that $\|ST\| \leq \|S\|\|T\|$.

Exercise 7.F.6. Prove or give a counterexample: If $S, T \in \mathcal{L}(V)$, then $\|ST\| = \|TS\|$.

Solution. This is false. For a counterexample, consider the operators $S, T \in \mathcal{L}(\mathbf{R}^2)$ whose matrices with respect to the standard basis of $\mathbf{R}^2$ are

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

A routine calculation shows that $ST = 0$ whereas $TS \neq 0$. It follows from 7.87(b) that

$$\|ST\| = 0 \neq \|TS\|.$$

Exercise 7.F.7. Show that defining $d(S, T) = \|S - T\|$ for $S, T \in \mathcal{L}(V, W)$ makes d a metric on $\mathcal{L}(V, W)$.

This exercise is intended for readers who are familiar with metric spaces.

Solution. Certainly d is non-negative and $d(T, T) = 0$ for any $T \in \mathcal{L}(V, W)$. Furthermore, $d(S, T) = 0$ implies $S = T$ by 7.87(b), and we have $d(S, T) = d(T, S)$ for any $S, T \in \mathcal{L}(V, W)$ since $\|Sv - Tv\| = \|Tv - Sv\|$ for any $v \in V$. Finally, 7.87(d) shows that

$$d(R, T) \leq d(R, S) + d(S, T)$$

for any $R, S, T \in \mathcal{L}(V, W)$.

Exercise 7.F.8.

(a) Prove that if $T \in \mathcal{L}(V)$ and $\|I - T\| < 1$, then T is invertible.

(b) Suppose that $S \in \mathcal{L}(V)$ is invertible. Prove that if $T \in \mathcal{L}(V)$ and

$$\|S - T\| < 1/\|S^{-1}\|,$$

then T is invertible.

This exercise shows that the set of invertible operators in $\mathcal{L}(V)$ is an open subset of $\mathcal{L}(V)$, using the metric defined in Exercise 7.

Solution.

(a) We will prove the contrapositive statement. Suppose that T is not invertible, so that there is some non-zero $v \in V$ such that $Tv = 0$. It follows that

$$\|I - T\|\|v\| \geq \|(I - T)v\| = \|v\| \quad \Rightarrow \quad \|I - T\| \geq 1.$$

(b) The proof is a generalization of the proof in part (a). We will again prove the contrapositive statement. Suppose that T is not invertible, so that there is some non-zero $v \in V$ such that $Tv = 0$. Observe that

$$\|v\| = \|S^{-1}Sv\| \leq \|S^{-1}\|\|Sv\| = \|S^{-1}\|\|(S - T)v\| \leq \|S^{-1}\|\|S - T\|\|v\|$$

$$\Rightarrow \quad \|S - T\| \geq \frac{1}{\|S^{-1}\|}.$$

Exercise 7.F.9. Suppose $T \in \mathcal{L}(V)$. Prove that for every $\varepsilon > 0$, there exists an invertible operator $S \in \mathcal{L}(V)$ such that $0 < \|T - S\| < \varepsilon$.

Solution. Because T can have only finitely many eigenvalues, we can choose some $\delta \in (0, \varepsilon)$ such that δ is not an eigenvalue of T. Letting $S = T - \delta I$, it follows that S is invertible and that $S \neq T$. Thus

$$0 < \|T - S\| = \|\delta I\| = \delta < \varepsilon.$$

Exercise 7.F.10. Suppose $\dim V > 1$ and $T \in \mathcal{L}(V)$ is not invertible. Prove that for every $\varepsilon > 0$, there exists $S \in \mathcal{L}(V)$ such that $0 < \|T - S\| < \varepsilon$ and S is not invertible.

Solution. Since T is not invertible, there exists some $e_1 \in \operatorname{null} T$ such that $\|e_1\| = 1$. Extend this to an orthonormal basis $e_1, \dots, e_n$ of V and note that $n \geq 2$. Define $S \in \mathcal{L}(V)$ by $Se_1 = 0$ and $Se_k = Te_k - \frac{\varepsilon}{2}e_k$ for $k \geq 2$. Notice that S is not invertible and that $S \neq T$. Notice further that, for any $v \in V$,

$$\|(T - S)v\|^2 = \left\|\frac{\varepsilon}{2}(\langle v, e_2\rangle e_2 + \dots + \langle v, e_n\rangle e_n)\right\|^2$$

$$= \frac{\varepsilon^2}{4}\left(|\langle v, e_2\rangle|^2 + \dots + |\langle v, e_n\rangle|^2\right) \leq \frac{\varepsilon^2}{4}\|v\|^2.$$

Thus $\|(T - S)v\| \leq \frac{\varepsilon}{2}\|v\|$ for any $v \in V$. It follows that $\|T - S\| \leq \frac{\varepsilon}{2} < \varepsilon$ and hence that $0 < \|T - S\| < \varepsilon$, since $S \neq T$.

Exercise 7.F.11. Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$. Prove that for every $\varepsilon > 0$ there exists a diagonalizable operator $S \in \mathcal{L}(V)$ such that $0 < \|T - S\| < \varepsilon$.

Solution. We will prove that S may be chosen to have $\dim V$ distinct eigenvalues, which implies the desired result by 5.58.

By Schur's theorem (6.38), there is an orthonormal basis $e_1, ..., e_n$ of V with respect to which the matrix of T is upper-triangular, say

$$\mathcal{M}(T, (e_1, ..., e_n)) = \begin{pmatrix} \lambda_1 & * & \cdots & * \\ 0 & \lambda_2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

Let $D \in \mathcal{L}(V)$ be given by $De_k = ke_k$, so that

$$\mathcal{M}(D, (e_1, ..., e_n)) = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & n \end{pmatrix}.$$

It is straightforward to verify that $\|D\| = n$. For $\delta \in \mathbf{R}$ and $j, k \in \{1, ..., n\}$ such that $j < k$, notice that

$$\lambda_j + j\delta = \lambda_k + k\delta \quad \Leftrightarrow \quad \delta = \frac{\lambda_j - \lambda_k}{k - j}.$$

Since there are only finitely many such choices of j and k, we may choose a $\delta \in \left(0, \frac{\varepsilon}{n}\right)$ such that $\lambda_j + j\delta \neq \lambda_k + k\delta$ for each $j, k \in \{1, ..., n\}$ satisfying $j < k$. It follows that the diagonal entries of the upper-triangular matrix

$$\mathcal{M}(T + \delta D, (e_1, ..., e_n)) = \begin{pmatrix} \lambda_1 + \delta & * & \cdots & * \\ 0 & \lambda_2 + 2\delta & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n + n\delta \end{pmatrix},$$

and hence the n eigenvalues of the operator $S := T + \delta D$, are distinct. Furthermore,

$$0 < \delta \quad \Rightarrow \quad S \neq T \quad \Rightarrow \quad 0 < \|T - S\|,$$
$$\delta < \frac{\varepsilon}{n} \quad \Rightarrow \quad \|T - S\| = \delta\|D\| < \varepsilon.$$

Exercise 7.F.12. Suppose $T \in \mathcal{L}(V)$ is a positive operator. Show that $\|\sqrt{T}\| = \sqrt{\|T\|}$.

Solution. Let s be the largest singular value of $\sqrt{T}$, i.e. $\|\sqrt{T}\| = s$. Since $\sqrt{T}$ is positive and hence normal, Exercise 7.E.12 (b) shows that the singular values of T equal the squares of the singular values of $\sqrt{T}$. Given that singular values are non-negative, it follows that the largest singular value of T is s^2, i.e. $\|T\| = s^2$. Thus $\|\sqrt{T}\| = s = \sqrt{\|T\|}$.

Exercise 7.F.13. Suppose $S, T \in \mathcal{L}(V)$ are positive operators. Show that

$$\|S - T\| \leq \max\{\|S\|, \|T\|\} \leq \|S + T\|.$$

Solution. Let us prove a couple of useful lemmas.

Lemma L.13. If $A \in \mathcal{L}(V)$ is a self-adjoint operator then $\|A\|I - A$ is a positive operator.

Proof. First note that $\|A\|I - A$ is a real-linear combination of self-adjoint operators and hence is itself self-adjoint (see 7.5). Now suppose that $\lambda \in \mathbf{R}$ is an eigenvalue of $\|A\|I - A$, say $\|A\|v - Av = \lambda v$ for some $v \in V$. It follows that $\lambda - \|A\|$ is an eigenvalue of A and hence, by Exercise 7.F.2,

$$|\lambda - \|A\|| \leq \|A\| \quad \Rightarrow \quad \lambda \geq 0.$$

Thus $\|A\|I - A$ is a positive operator by 7.38(b). □

Lemma L.14. If A and $B - A$ are positive operators then $\|A\| \leq \|B\|$.

Proof. Note that $B = (B - A) + A$ is a positive operator by Exercise 7.C.6. It follows from Lemma L.13 that $\|B\|I - B$ is a positive operator and hence that

$$\|B\|I - A = (\|B\|I - B) + (B - A)$$

is a positive operator. Suppose $\lambda \geq 0$ is an eigenvalue of A, say $Av = \lambda v$ for some $v \in V$. Observe that

$$(\|B\|I - A)v = (\|B\| - \lambda)v,$$

so that $\|B\| - \lambda$ is an eigenvalue of the positive operator $\|B\|I - A$. It follows that $\|B\| - \lambda$ is non-negative and hence that $\lambda \leq \|B\|$. Since this was true for any eigenvalue of A, and each such eigenvalue is non-negative, Exercise 7.F.2 shows that

$$\|A\| = \max\{|\lambda| : \lambda \text{ is an eigenvalue of } A\} \leq \|B\|,$$

as desired. □

Returning to the exercise, suppose $\lambda \in \mathbf{R}$ is an eigenvalue of the self-adjoint operator $S - T$, say $(S - T)v = \lambda v$ for some $v \in V$. It follows that

$$(\|S\|I - (S - T))v = (\|S\| - \lambda)v$$

and hence that $\|S\| - \lambda$ is an eigenvalue of $\|S\|I - (S - T)$. Notice that $\|S\|I - (S - T)$ is a positive operator by Lemma L.13 and Exercise 7.C.6, so that its eigenvalues are non-negative. Thus $\lambda \leq \|S\|$ and a similar argument with the operator $S - T + \|T\|I$ shows that $-\|T\| \leq \lambda$. It follows that $|\lambda| \leq \max\{\|S\|, \|T\|\}$ and hence, by Exercise 7.F.2,

$$\|S - T\| = \max\{|\lambda| : \lambda \text{ is an eigenvalue of } S - T\} \leq \max\{\|S\|, \|T\|\}.$$

Applying Lemma L.14 twice, first with $A = T$ and $B = S + T$ and then with $A = S$ and $B = S + T$, shows that

$$\|S\| \le \|S+T\| \quad \text{and} \quad \|T\| \le \|S+T\| \quad \Rightarrow \quad \max\{\|S\|, \|T\|\} \le \|S+T\|.$$

Exercise 7.F.14. Suppose U and W are subspaces of V such that $\|P_U - P_W\| < 1$. Prove that $\dim U = \dim W$.

Solution. Using the identities $P_U = I - P_{U^\perp}$ and $P_W = I - P_{W^\perp}$ (see Exercise 6.C.5) and Exercise 7.F.8 (a), we see that the operators $P_{U^\perp} + P_W$ and $P_U + P_{W^\perp}$ are invertible. It follows that

$$U \cap W^\perp = (\operatorname{null} P_{U^\perp}) \cap (\operatorname{null} P_W) \subseteq \operatorname{null}(P_{U^\perp} + P_W) = \{0\} \quad \Rightarrow \quad U \cap W^\perp = \{0\},$$

$$V = \operatorname{range}(P_U + P_{W^\perp}) \subseteq (\operatorname{range} P_U) + (\operatorname{range} P_{W^\perp}) = U + W^\perp \quad \Rightarrow \quad U + W^\perp = V.$$

Thus, using 6.51,

$$\dim V = \dim(U + W^\perp) = \dim U + \dim W^\perp - \dim(U \cap W^\perp) = \dim V + \dim U - \dim W$$

$$\Rightarrow \quad \dim U = \dim W.$$

Exercise 7.F.15. Define $T \in \mathcal{L}(\mathbf{F}^3)$ by

$$T(z_1, z_2, z_3) = (z_3, 2z_1, 3z_2).$$

Find (explicitly) a unitary operator $S \in \mathcal{L}(\mathbf{F}^3)$ such that $T = S\sqrt{T^*T}$.

Solution. Let e_1, e_2, e_3 be the standard orthonormal basis of $\mathbf{F}^3$. A routine calculation shows that the singular value decomposition of T is

$$Tv = 3\langle v, e_2\rangle e_3 + 2\langle v, e_1\rangle e_2 + \langle v, e_3\rangle e_1.$$

As the proof of 7.93 shows, if we take $S \in \mathcal{L}(V)$ to be the unitary operator defined by

$$Se_2 = e_3, \quad Se_1 = e_2, \quad \text{and} \quad Se_3 = e_1,$$

then $T = S\sqrt{T^*T}$.

Exercise 7.F.16. Suppose $S \in \mathcal{L}(V)$ is a positive invertible operator. Prove that there exists $\delta > 0$ such that T is a positive operator for every self-adjoint operator $T \in \mathcal{L}(V)$ with $\|S - T\| < \delta$.

Solution. Let μ be the least singular value of S and note that $\mu > 0$ since S is invertible. Note further that $\sqrt{\mu} > 0$ is the least singular value of $\sqrt{S}$ by Exercise 7.E.12 (b). Thus, by Exercise 7.E.14,

$$\langle Sv, v\rangle = \left\langle S^{1/2}v, S^{1/2}v\right\rangle = \left\|S^{1/2}v\right\|^2 \ge \mu\|v\|^2 \tag{1}$$

for every $v \in V$. Suppose $T \in \mathcal{L}(V)$ is a self-adjoint operator satisfying $\|S - T\| < \mu$. For any $v \in V$, the Cauchy-Schwarz inequality gives us

$$\langle (S-T)v, v\rangle \le \|(S-T)v\|\|v\| \le \|S-T\|\|v\|^2 < \mu\|v\|^2$$
$$\Rightarrow \quad \mu\|v\|^2 + \langle (T-S)v, v\rangle > 0. \qquad (2)$$

Combining inequalities (1) and (2) we obtain, for any $v \in V$,

$$\langle Tv, v\rangle = \langle Sv, v\rangle + \langle (T-S)v, v\rangle \ge \mu\|v\|^2 + \langle (T-S)v, v\rangle > 0.$$

Thus T is a positive operator.

Exercise 7.F.17. Prove that if $u \in V$ and φ_u is the linear functional on V defined by the equation $\varphi_u(v) = \langle v, u\rangle$, then $\|\varphi_u\| = \|u\|$.

Here we are thinking of the scalar field $\mathbf{F}$ *as an inner product space with* $\langle \alpha, \beta\rangle = \alpha\overline{\beta}$ *for all* $\alpha, \beta \in \mathbf{F}$. *Thus* $\|\varphi_u\|$ *means the norm of* φ_u *as a linear map from* V *to* $\mathbf{F}$.

Solution. The Cauchy-Schwarz inequality shows that

$$|\varphi_u(v)| = |\langle v, u\rangle| \le \|u\|\|v\|$$

for any $v \in V$. It follows that $\|\varphi_u\| \le \|u\|$. Since $|\varphi_u(u)| = |\langle u, u\rangle| = \|u\|$, we may conclude that $\|\varphi_u\| = \|u\|$.

Exercise 7.F.18. Suppose $e_1, \dots, e_n$ is an orthonormal basis of V and $T \in \mathcal{L}(V, W)$.

(a) Prove that $\max\{\|Te_1\|, \dots, \|Te_n\|\} \le \|T\| \le \left(\|Te_1\|^2 + \cdots + \|Te_n\|^2\right)^{1/2}$.

(b) Prove that $\|T\| = \left(\|Te_1\|^2 + \cdots + \|Te_n\|^2\right)^{1/2}$ if and only if $\dim \operatorname{range} T \le 1$.

Here $e_1, \dots, e_n$ *is an arbitrary orthonormal basis of* V, *not necessarily connected with a singular value decomposition of* T. *If* $s_1, \dots, s_n$ *is the list of singular values of* T, *then the right side of the inequality above equals* $(s_1^2 + \cdots + s_n^2)^{1/2}$, *as was shown in* *Exercise 11(a) in Section 7E*.

Solution.

(a) For any $k \in \{1, \dots, n\}$ we have $\|Te_k\| \le \|T\|\|e_k\| = \|T\|$ and thus

$$\max\{\|Te_1\|, \dots, \|Te_n\|\} \le \|T\|.$$

Let $s_1, \dots, s_n$ be the singular values of T and suppose that $s_1, \dots, s_m$ are the positive singular values of T. Suppose further that

$$Tv = s_1\langle v, f_1\rangle g_1 + \cdots + s_m\langle v, f_m\rangle g_m$$

is a singular value decomposition of T. For any $v \in V$, observe that

$$\begin{aligned}\|Tv\|^2 &= s_1^2|\langle v, f_1\rangle|^2 + \cdots + s_m^2|\langle v, f_m\rangle|^2 \\ &\le (s_1^2 + \cdots + s_m^2)\|v\|^2 \\ &\le (s_1^2 + \cdots + s_n^2)\|v\|^2.\end{aligned}$$

Thus $\|Tv\| \leq (s_1^2 + \cdots + s_n^2)^{1/2}\|v\|$ for every $v \in V$, which implies that

$$\|T\| \leq (s_1^2 + \cdots + s_n^2)^{1/2} = \left(\|Te_1\|^2 + \cdots + \|Te_n\|^2\right)^{1/2},$$

where we have used Exercise 7.E.11 (a) for the last equality.

(b) Let $s_1, \ldots, s_n$ be the singular values of T. By Exercise 7.E.11 (a) and 7.68(b), it will suffice to show that $\|T\| = (s_1^2 + \cdots + s_n^2)^{1/2}$ if and only if T has at most one positive singular value.

If T has at most one positive singular value then $s_k = 0$ for $k \geq 2$. It follows that

$$\|T\| = s_1 = (s_1^2)^{1/2} = (s_1^2 + \cdots + s_n^2)^{1/2}.$$

If T has at least two positive singular values then $s_2 > 0$ and it follows that

$$\|T\| = s_1 = (s_1^2)^{1/2} < (s_1^2 + s_2^2 + \cdots + s_n^2)^{1/2}.$$

Exercise 7.F.19. Prove that if $T \in \mathcal{L}(V, W)$, then $\|T^*T\| = \|T\|^2$.

This formula for $\|T^*T\|$ leads to the important subject of C^*-algebras.

Solution. For any $v \in V$, observe that

$$\|(T^*T)^{1/2}\, v\|^2 = \langle (T^*T)^{1/2}\, v, (T^*T)^{1/2}\, v\rangle = \langle T^*Tv, v\rangle = \langle Tv, Tv\rangle = \|Tv\|^2 \leq \|T\|^2\|v\|^2.$$

It follows that $\|(T^*T)^{1/2}\| \leq \|T\|$. By the polar decomposition (7.93), there exists a unitary operator $S \in \mathcal{L}(V)$ such that $T = S(T^*T)^{1/2}$. Note that

$$\|S\| = \max\{\|Sv\| : v \in V \text{ and } \|V\| = 1\} = \max\{1\} = 1.$$

It then follows from Exercise 7.F.5 that

$$\|T\| = \|S(T^*T)^{1/2}\| \leq \|S\|\|(T^*T)^{1/2}\| = \|(T^*T)^{1/2}\|.$$

Thus $\|(T^*T)^{1/2}\| = \|T\|$, which is equivalent to $\|T^*T\| = \|T\|^2$ by Exercise 7.F.12.

Exercise 7.F.20. Suppose $T \in \mathcal{L}(V)$ is normal. Prove that $\|T^k\| = \|T\|^k$ for every positive integer k.

Solution. Let k be a positive integer and suppose $s_1, \ldots, s_n$ are the singular values of T, so that $\|T\| = s_1$. As in Exercise 7.E.12 (b), let $e_1, \ldots, e_n$ be an orthonormal basis of V such that $T^*Te_j = s_j^2 e_j$ for each $j \in \{1, \ldots, n\}$. Using the normality of T, it follows that

$$(T^k)^*T^k e_j = (T^*T)^k e_j = s_j^{2k} e_j$$

for each $j \in \{1, \ldots, n\}$. Thus the singular values of T^k are $s_1^k, \ldots, s_n^k$; these are still in decreasing order since the function $x \mapsto x^k$ is strictly increasing on $[0, \infty)$. It follows that

$$\|T^k\| = s_1^k = \|T\|^k.$$

Exercise 7.F.21. Suppose $\dim V > 1$ and $\dim W > 1$. Prove that the norm on $\mathcal{L}(V, W)$ does not come from an inner product. In other words, prove that there does not exist an inner product on $\mathcal{L}(V, W)$ such that

$$\max\{\|Tv\| : v \in V \text{ and } \|v\| \le 1\} = \sqrt{\langle T, T\rangle}$$

for all $T \in \mathcal{L}(V, W)$.

Solution. Let $v_1, ..., v_m$ be an orthonormal basis of V and let $w_1, ..., w_n$ be an orthonormal basis of W; note that $m, n \ge 2$. Let $S, T \in \mathcal{L}(V, W)$ be the linear maps given by

$$Sv_1 = w_1, \quad Sv_2 = w_2, \quad \text{and} \quad Sv_k = 0 \text{ if } k > 2,$$

$$Tv_1 = -w_1, \quad Tv_2 = w_2, \quad \text{and} \quad Tv_k = 0 \text{ if } k > 2.$$

It is straightforward to verify that

$$\|S + T\| = \|S - T\| = 2 \quad \text{and} \quad \|S\| = \|T\| = 1$$

$$\Rightarrow \quad \|S+T\|^2 + \|S-T\|^2 = 8 \ne 4 = 2\left(\|S\|^2 + \|T\|^2\right).$$

Thus the norm on $\mathcal{L}(V, W)$ does not satisfy the parallelogram equality (6.21) and hence does not come from an inner product.

Exercise 7.F.22. Suppose $T \in \mathcal{L}(V, W)$. Let $n = \dim V$ and let $s_1 \ge \cdots \ge s_n$ denote the singular values of T. Prove that if $1 \le k \le n$, then

$$\min\{\|T|_U\| : U \text{ is a subspace of } V \text{ with} \dim U = k\} = s_{n-k+1}.$$

Solution. For $k \in \{1, ..., n\}$, let $E_k = \{\|T|_U\| : U \text{ is a subspace of } V \text{ with} \dim U = k\}$. If T has no positive singular values, i.e. $T = 0$, then

$$\min E_k = \min\{0\} = 0 = s_{n-k+1}$$

for any $k \in \{1, ..., n\}$. Furthermore, since the only subspace of V with dimension n is V itself, we have

$$\min E_n = \min\{\|T\|\} = \|T\| = s_1.$$

We may therefore assume that T has at least one singular value and that $1 \le k < n$.

Suppose that $s_1 \ge \cdots \ge s_m$ are the positive singular values of T, where $m \ge 1$, and let

$$Tv = s_1\langle v, e_1\rangle f_1 + \cdots + s_m\langle v, e_m\rangle f_m$$

be a singular value decomposition of T. Extend the orthonormal list $e_1, ..., e_m$ to an orthonormal basis $e_1, ..., e_n$ of V, so that

$$Te_j = \begin{cases} s_j f_j & \text{if } 1 \le j \le m, \\ 0 & \text{if } m < j \le n. \end{cases}$$

For $1 \le k < n$ let $X = \operatorname{span}(e_{n-k+1}, \dots, e_n)$ and note that $\dim X = k$. We consider two cases.

Case 1. If $1 \le k \le n-m$, so that $m+1 \le n-k+1$, then $s_{n-k+1} = 0$. Furthermore, T vanishes on each of the basis vectors $e_{n-k+1}, \dots, e_n$, so that $T|_X = 0$. Since E_k is bounded below by zero, it follows that $\min E_k = 0 = s_{n-k+1}$.

Case 2. Suppose that $n-m < k < n$, so that $1 \le n-k < m$, and let U be any subspace of V satisfying $\dim U = k$. If $v \in V$ is such that $\|v\| \le 1$ then note that $P_U v \in U$ and that $\|P_U v\| \le \|v\| \le 1$. It follows that $\|TP_U v\| \le \|T|_U\|$, since

$$\|T|_U\| = \max\{\|Tu\| : u \in U \text{ and } \|u\| \le 1\}.$$

Thus $\|TP_U\| \le \|T|_U\|$. Next, observe that

$$\dim \operatorname{range}(TP_{U^\perp}) \le \dim \operatorname{range} P_{U^\perp} = \dim U^\perp = \dim V - \dim U = n-k.$$

It follows from 7.92 that $\|T - TP_{U^\perp}\| \ge s_{n-k+1}$. Combining these inequalities with the identity $P_U = I - P_{U^\perp}$ from Exercise 6.C.5, we have

$$\|T|_U\| \ge \|TP_U\| = \|T - TP_{U^\perp}\| \ge s_{n-k+1}.$$

Thus s_{n-k+1} is a lower bound of E_k. Now observe that, for any $x \in X$,

$$\|(T|_X)x\|^2 = s_{n-k+1}^2 |\langle x, e_{n-k+1}\rangle|^2 + \cdots + s_m^2 |\langle x, e_m\rangle|^2 \le s_{n-k+1}^2 \|x\|^2.$$

It follows that $\|T|_X\| \le s_{n-k+1}$; in fact, this is an equality since $\|e_{n-k+1}\| = 1$ and

$$(T|_X)(e_{n-k+1}) = s_{n-k+1} f_{n-k+1} \quad \Rightarrow \quad \|(T|_X)(e_{n-k+1})\| = s_{n-k+1}.$$

Thus the lower bound s_{n-k+1} belongs to E_k, i.e. $\min E_k = s_{n-k+1}$.

Exercise 7.F.23. Suppose $T \in \mathcal{L}(V, W)$. Show that T is uniformly continuous with respect to the metrics on V and W that arise from the norms on those spaces (see Exercise 23 in Section 6B).

Solution. Let $\varepsilon > 0$ be given and let $\delta = \varepsilon(1 + \|T\|)^{-1}$. If $u, v \in V$ are such that $\|u - v\| < \delta$, then observe that

$$\|Tu - Tv\| = \|T(u-v)\| \le \|T\|\|u - v\| < \|T\|\delta < \varepsilon.$$

Thus T is uniformly continuous.

Exercise 7.F.24. Suppose $T \in \mathcal{L}(V)$ is invertible. Prove that

$$\|T^{-1}\| = \|T\|^{-1} \quad \Leftrightarrow \quad \frac{T}{\|T\|} \text{ is a unitary operator.}$$

Solution. Let $s_1 \ge \cdots \ge s_n$ be the singular values of T, so that $\|T\| = s_1$; note that each singular value is strictly positive since T is invertible. Note further that $s_n^{-1} \ge \cdots \ge s_1^{-1}$ are the singular values of T^{-1} by Exercise 7.E.10, which gives us $\|T^{-1}\| = s_n^{-1}$. Thus we wish to prove that

$$s_n^{-1} = s_1^{-1} \quad \Leftrightarrow \quad s_1^{-1}T \text{ is a unitary operator.}$$

Indeed,

$$\begin{aligned} s_n^{-1} = s_1^{-1} \quad &\Leftrightarrow \quad s_1 = s_n \\ &\Leftrightarrow \quad \text{each singular value of } T \text{ equals } s_1 \\ &\Leftrightarrow \quad \text{each singular value of } s_1^{-1}T \text{ equals } 1 \\ &\Leftrightarrow \quad s_1^{-1}T \text{ is a unitary operator.} \end{aligned} \tag{7.69}$$

Exercise 7.F.25. Fix $u, x \in V$ with $u \neq 0$. Define $T \in \mathcal{L}(V)$ by $Tv = \langle v, u\rangle x$ for every $v \in V$. Prove that

$$\sqrt{T^*T}v = \frac{\|x\|}{\|u\|}\langle v, u\rangle u$$

for every $v \in V$.

Solution. Let $R \in \mathcal{L}(V)$ be given by

$$Rv = \frac{\|x\|}{\|u\|}\langle v, u\rangle u.$$

Our aim is to show that $R = \sqrt{T^*T}$. Using the formula $T^*v = \langle v, x\rangle u$, shown in example 7.3, a routine calculation shows that

$$R^2 v = \|x\|^2 \langle v, u\rangle u = T^*Tv.$$

Furthermore, for any $v \in V$,

$$\langle Rv, v\rangle = \frac{\|x\|}{\|u\|}|\langle v, u\rangle|^2 \geq 0.$$

Thus R is a positive square root of T^*T. It follows from uniqueness (7.39) that $R = \sqrt{T^*T}$.

Exercise 7.F.26. Suppose $T \in \mathcal{L}(V)$. Prove that T is invertible if and only if there exists a unique unitary operator $S \in \mathcal{L}(V)$ such that $T = S\sqrt{T^*T}$.

Solution. Suppose that T is invertible. The polar decomposition (7.93) provides us with a unitary operator $S \in \mathcal{L}(V)$ such that $T = S\sqrt{T^*T}$. To see that S is unique, note that $\sqrt{T^*T}$ is invertible, since

$$\operatorname{null}\sqrt{T^*T} = \operatorname{null} T^*T = \operatorname{null} T = \{0\},$$

where the first equality follows from Exercise 7.C.16 and the second equality follows from 7.64(b). Thus S is uniquely determined by the formula $S = T\left(\sqrt{T^*T}\right)^{-1}$.

Now suppose that T is not invertible. Let $s_1, \ldots, s_m$ be the positive singular values of T, let

$$Tv = s_1\langle v, e_1\rangle f_1 + \cdots + s_m\langle v, e_m\rangle f_m$$

be a singular value decomposition of T, and extend $e_1, \ldots, e_m$ and $f_1, \ldots, f_m$ to orthonormal bases $e_1, \ldots, e_n$ and $f_1, \ldots, f_n$ of V. Because T is not invertible, it must be the case that m is strictly less than n. Define $R, S \in \mathcal{L}(V)$ by

$$Re_k = Se_k = f_k \text{ for } k < n, \quad Re_n = f_n, \quad \text{and} \quad Se_n = -f_n.$$

As in the proof of the polar decomposition (7.93), R and S are both unitary operators and both satisfy $T = R\sqrt{T^*T} = S\sqrt{T^*T}$, however $R \neq S$.

Exercise 7.F.27. Suppose $T \in \mathcal{L}(V)$ and $s_1, \ldots, s_n$ are the singular values of T. Let $e_1, \ldots, e_n$ and $f_1, \ldots, f_n$ be orthonormal bases of V such that

$$Tv = s_1\langle v, e_1\rangle f_1 + \cdots + s_n\langle v, e_n\rangle f_n$$

for all $v \in V$. Define $S \in \mathcal{L}(V)$ by

$$Sv = \langle v, e_1\rangle f_1 + \cdots + \langle v, e_n\rangle f_n.$$

(a) Show that S is unitary and $\|T - S\| = \max\{|s_1 - 1|, \ldots, |s_n - 1|\}$.

(b) Show that if $E \in \mathcal{L}(V)$ is unitary, then $\|T - E\| \geq \|T - S\|$.

This exercise finds a unitary operator S that is as close as possible (among the unitary operators) to a given operator T.

Solution.

(a) S was shown to be unitary in the proof of the polar decomposition (7.93). Suppose that $\max\{|s_1 - 1|, \ldots, |s_n - 1|\} = |s_j - 1|$, where $j \in \{1, \ldots, n\}$. For any $v \in V$, observe that

$$\|(T - S)v\|^2 = |s_1 - 1|^2|\langle v, e_1\rangle|^2 + \cdots + |s_n - 1|^2|\langle v, e_n\rangle|^2 \leq |s_j - 1|^2\|v\|^2.$$

Thus $\|T - S\| \leq |s_j - 1|$. Since $\|(T - S)e_j\| = |s_j - 1|$, we must have $\|T - S\| = |s_j - 1|$.

(b) Let $k \in \{1, \ldots, n\}$ be given and note that, since E is unitary, we must have $\|Ee_k\| = \|e_k\| = 1$. It follows from the reverse triangle inequality (see Exercise 6.A.20) that

$$\|T - E\| \geq \|(T - E)e_k\| \geq |\|Te_k\| - \|Ee_k\|| = |\|s_k f_k\| - 1| = |s_k - 1|.$$

Thus $\|T - E\| \geq \max\{|s_1 - 1|, \ldots, |s_n - 1|\} = \|T - S\|$.

Exercise 7.F.28. Suppose $T \in \mathcal{L}(V)$. Prove that there exists a unitary operator $S \in \mathcal{L}(V)$ such that $T = \sqrt{TT^*}S$.

Solution. Let $s_1, ..., s_m$ be the positive singular values of T, let

$$Tv = s_1\langle v, e_1\rangle f_1 + \cdots + s_m\langle v, e_m\rangle f_m$$

be a singular value decomposition of T, and extend $e_1, ..., e_m$ and $f_1, ..., f_m$ to orthonormal bases $e_1, ..., e_n$ and $f_1, ..., f_n$ of V. Define the unitary operator $S \in \mathcal{L}(V)$ as in the proof of the polar decomposition (7.93). Using the formula for TT^* from Exercise 7.E.8 (e), we see that

$$\sqrt{TT^*}v = s_1\langle v, f_1\rangle f_1 + \cdots + s_m\langle v, f_m\rangle f_m$$

for every $v \in V$. For any $k \in \{1, ..., m\}$ and any $v \in V$, notice that $\langle Sv, f_k\rangle = \langle v, e_k\rangle$. It follows that

$$\sqrt{TT^*}Sv = s_1\langle Sv, f_1\rangle f_1 + \cdots + s_m\langle Sv, f_m\rangle f_m = s_1\langle v, e_1\rangle f_1 + \cdots + s_m\langle v, e_m\rangle f_m = Tv.$$

Thus S is a unitary operator satisfying $T = \sqrt{TT^*}S$.

Exercise 7.F.29. Suppose $T \in \mathcal{L}(V)$.

(a) Use the polar decomposition to show that there exists a unitary operator $S \in \mathcal{L}(V)$ such that $TT^* = ST^*TS^*$.

(b) Show how (a) implies that T and T^* have the same singular values.

Solution.

(a) The polar decomposition (7.93) gives us a unitary operator $S \in \mathcal{L}(V)$ satisfying $T = S\sqrt{T^*T}$; as we showed in Exercise 7.F.28, S also satisfies $T = \sqrt{TT^*}S$. It follows that

$$\sqrt{TT^*}S = S\sqrt{T^*T} \quad \Rightarrow \quad \sqrt{TT^*} = S\sqrt{T^*T}S^*$$
$$\Rightarrow \quad TT^* = S\sqrt{T^*T}S^*S\sqrt{T^*T}S^* = ST^*TS^*.$$

(b) The identity $TT^* = ST^*TS^*$ and Exercise 5.A.13 show that the list of eigenvalues of T^*T is the same as the list of eigenvalues of TT^*. It follows that the list of singular values of T is the same as the list of singular values of T^*.

Exercise 7.F.30. Suppose $T \in \mathcal{L}(V)$, $S \in \mathcal{L}(V)$ is a unitary operator, and $R \in \mathcal{L}(V)$ is a positive operator such that $T = SR$. Prove that $R = \sqrt{T^*T}$.

This exercise shows that if we write T as the product of a unitary operator and a positive operator (as in the polar decomposition 7.93), then the positive operator equals $\sqrt{T^*T}$.

Solution. Observe that

$$T^*T = (SR)^*SR = R^*S^*SR = R^2,$$

where we have used that R is self-adjoint and that S is unitary. Thus R is a positive square root of T^*T; by uniqueness (7.39) we must have $R = \sqrt{T^*T}$.

Exercise 7.F.31. Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$ is normal. Prove that there exists a unitary operator $S \in \mathcal{L}(V)$ such that $T = S\sqrt{T^*T}$ and such that S and $\sqrt{T^*T}$ both have diagonal matrices with respect to the same orthonormal basis of V.

Solution. As the polar decomposition (7.93) and Exercise 7.F.28 show, there exists a unitary operator $S \in \mathcal{L}(V)$ such that

$$T = S\sqrt{T^*T} = \sqrt{TT^*}S = \sqrt{T^*T}S,$$

where we have used that T is normal for the last equality. Thus S and $\sqrt{T^*T}$ commute, which by Exercise 7.B.16 is equivalent to the existence of an orthonormal basis of V with respect to which S and $\sqrt{T^*T}$ both have diagonal matrices.

Exercise 7.F.32. Suppose that $T \in \mathcal{L}(V, W)$ and $T \neq 0$. Let $s_1, \ldots, s_m$ denote the positive singular values of T. Show that there exists an orthonormal basis $e_1, \ldots, e_m$ of $(\operatorname{null} T)^\perp$ such that

$$T\left(E\left(\frac{e_1}{s_1}, \ldots, \frac{e_m}{s_m}\right)\right)$$

equals the ball in $\operatorname{range} T$ of radius 1 centered at 0.

Solution. Let B be the ball in $\operatorname{range} T$ of radius 1 centered at 0 and let

$$Tv = s_1\langle v, e_1\rangle f_1 + \cdots + s_m\langle v, e_m\rangle f_m$$

be a singular value decomposition of T; as we showed in Exercise 7.E.8 (b), $e_1, \ldots, e_m$ is an orthonormal basis of $(\operatorname{null} T)^\perp$. Note that

$$Tv \in B \quad \Leftrightarrow \quad \|Tv\|^2 = s_1^2|\langle v, e_1\rangle|^2 + \cdots + s_m^2|\langle v, e_m\rangle|^2 \leq 1.$$

Note further that

$$\begin{aligned} Tv \in T\left(E\left(\frac{e_1}{s_1}, \ldots, \frac{e_m}{s_m}\right)\right) \quad &\Leftrightarrow \quad v \in E\left(\frac{e_1}{s_1}, \ldots, \frac{e_m}{s_m}\right) \\ &\Leftrightarrow \quad s_1^2|\langle v, e_1\rangle|^2 + \cdots + s_m^2|\langle v, e_m\rangle|^2 \leq 1. \end{aligned}$$

Thus $T(E(s_1^{-1}e_1, \ldots, s_m^{-1}e_m)) = B$.

Chapter 8. Operators on Complex Vector Spaces

8.A. Generalized Eigenvectors and Nilpotent Operators

Exercise 8.A.1. Suppose $T \in \mathcal{L}(V)$. Prove that if $\dim \operatorname{null} T^4 = 8$ and $\dim \operatorname{null} T^6 = 9$, then $\dim \operatorname{null} T^m = 9$ for all integers $m \geq 5$.

Solution. By 8.1 we have $\dim \operatorname{null} T^5 \in \{8, 9\}$. Notice that

$$\dim \operatorname{null} T^5 = 8 \quad \Rightarrow \quad \operatorname{null} T^4 = \operatorname{null} T^5 \quad \Rightarrow \quad \operatorname{null} T^5 = \operatorname{null} T^6 \quad \Rightarrow \quad \dim \operatorname{null} T^6 = 8,$$

where we have used 8.2 for the second implication. Since $\dim \operatorname{null} T^6 = 9$ it must then be the case that $\dim \operatorname{null} T^5 = 9$, whence $\operatorname{null} T^5 = \operatorname{null} T^6$. It follows from 8.2 that, for all integers $m \geq 5$,

$$\operatorname{null} T^m = \operatorname{null} T^5 \quad \Rightarrow \quad \dim \operatorname{null} T^m = \dim \operatorname{null} T^5 = 9.$$

Exercise 8.A.2. Suppose $T \in \mathcal{L}(V)$, m is a positive integer, $v \in V$, and $T^{m-1}v \neq 0$ but $T^m v = 0$. Prove that $v, Tv, T^2v, \ldots, T^{m-1}v$ is linearly independent.

The result in this exercise is used in the proof of 8.45.

Solution. Suppose $a_0, \ldots, a_{m-1}$ are scalars such that

$$a_0 v + a_1 Tv + \cdots + a_{m-1} T^{m-1} v = 0.$$

Apply T^{m-1} to both sides of this equation to obtain $a_0 T^{m-1} v = 0$. Since $T^{m-1}v \neq 0$, it must be the case that $a_0 = 0$. Thus we have the equation

$$a_1 Tv + \cdots + a_{m-1} T^{m-1} v = 0.$$

Now apply T^{m-2} to both sides of this equation to obtain $a_1 T^{m-1} v = 0$, which implies $a_1 = 0$. By continuing in this manner, we see that each of the scalars $a_0, \ldots, a_{m-1}$ is zero. Thus $v, Tv, \ldots, T^{m-1}v$ is linearly independent.

Exercise 8.A.3. Suppose $T \in \mathcal{L}(V)$. Prove that

$$V = \operatorname{null} T \oplus \operatorname{range} T \quad \Leftrightarrow \quad \operatorname{null} T^2 = \operatorname{null} T.$$

Solution. By Exercise 5.D.4 it will suffice to show that

$$\operatorname{null} T \cap \operatorname{range} T = \{0\} \quad \Leftrightarrow \quad \operatorname{null} T^2 = \operatorname{null} T.$$

First suppose that $\operatorname{null} T \cap \operatorname{range} T = \{0\}$. To show that $\operatorname{null} T^2 = \operatorname{null} T$, it will suffice to show that $\operatorname{null} T^2 \subseteq \operatorname{null} T$ (by 8.1). Suppose therefore that $v \in \operatorname{null} T^2$, so that $T^2 v = 0$, and notice that $Tv \in \operatorname{null} T \cap \operatorname{range} T = \{0\}$. Hence $Tv = 0$, i.e. $v \in \operatorname{null} T$. Thus $\operatorname{null} T^2 \subseteq \operatorname{null} T$, as desired.

Now suppose that $\operatorname{null} T^2 = \operatorname{null} T$ and suppose that $v \in \operatorname{null} T \cap \operatorname{range} T$, so that $Tv = 0$ and $v = Tu$ for some $u \in V$. It follows that

$$T^2 u = Tv = 0 \quad \Rightarrow \quad u \in \operatorname{null} T^2 \quad \Rightarrow \quad u \in \operatorname{null} T \quad \Rightarrow \quad Tu = v = 0.$$

Thus $\operatorname{null} T \cap \operatorname{range} T = \{0\}$, as desired.

Exercise 8.A.4. Suppose $T \in \mathcal{L}(V)$, $\lambda \in \mathbf{F}$, and m is a positive integer such that the minimal polynomial of T is a polynomial multiple of $(z - \lambda)^m$. Prove that

$$\dim \operatorname{null} (T - \lambda I)^m \geq m.$$

Solution. Let p be the minimal polynomial of T; by assumption we have $p(z) = (z - \lambda)^m q(z)$ for some polynomial q. Consider the chain of inclusions provided by 8.1:

$$\{0\} \subseteq \operatorname{null} (T - \lambda I)^1 \subseteq \cdots \subseteq \operatorname{null} (T - \lambda I)^m.$$

Notice that it will suffice to prove that each of these m inclusions is strict, since this will imply that the dimension increases by at least 1 at each inclusion. Suppose therefore, by way of contradiction, that there is some inclusion in this chain which is not strict, i.e. there exists some $k \in \{0, \ldots, m - 1\}$ such that

$$\operatorname{null} (T - \lambda I)^k = \operatorname{null} (T - \lambda I)^{k+1} \quad \Rightarrow \quad \operatorname{null} (T - \lambda I)^k = \operatorname{null} (T - \lambda I)^m \quad \text{(by 8.2)}.$$

Define $r(z) = (z - \lambda)^k q(z)$ and note that $\deg r < \deg p$. For any $v \in V$ observe that, since $0 = p(T)v = (T - \lambda I)^m q(T)v$,

$$q(T)v \in \operatorname{null} (T - \lambda I)^m \quad \Rightarrow \quad q(T)v \in \operatorname{null} (T - \lambda I)^k \quad \Rightarrow \quad r(T)v = 0.$$

Thus r is a polynomial of lesser degree than p which annihilates T, contradicting that p is the minimal polynomial of T.

Exercise 8.A.5. Suppose $T \in \mathcal{L}(V)$ and m is a positive integer. Prove that

$$\dim \operatorname{null} T^m \leq m \dim \operatorname{null} T.$$

Hint: Exercise 21 in Section 3B may be useful.

Solution. Exercise 3.B.22 shows that

$$\begin{aligned} \dim \operatorname{null} T^m \leq \dim \operatorname{null} T^{m-1} + \dim \operatorname{null} T &\leq \dim \operatorname{null} T^{m-2} + 2 \dim \operatorname{null} T \\ &\leq \cdots \leq m \dim \operatorname{null} T. \end{aligned}$$

Exercise 8.A.6. Suppose $T \in \mathcal{L}(V)$. Show that

$$V = \operatorname{range} T^0 \supseteq \operatorname{range} T^1 \supseteq \cdots \supseteq \operatorname{range} T^k \supseteq \operatorname{range} T^{k+1} \supseteq \cdots.$$

Solution. Suppose k is a non-negative integer and $w \in \operatorname{range} T^{k+1}$, so that there is some $v \in V$ such that

$$w = T^{k+1}v = T^k(Tv) \in \operatorname{range} T^k.$$

Thus $\operatorname{range} T^{k+1} \subseteq \operatorname{range} T^k$.

Exercise 8.A.7. Suppose $T \in \mathcal{L}(V)$ and m is a nonnegative integer such that

$$\operatorname{range} T^m = \operatorname{range} T^{m+1}.$$

Prove that $\operatorname{range} T^k = \operatorname{range} T^m$ for all $k > m$.

Solution. It will suffice to prove that $\operatorname{range} T^{m+\ell} = \operatorname{range} T^{m+\ell+1}$ for all positive integers ℓ. By Exercise 8.A.6, this is equivalent to showing that $\operatorname{range} T^{m+\ell} \subseteq \operatorname{range} T^{m+\ell+1}$ for all positive integers ℓ. Suppose therefore that ℓ is a positive integer and that $w \in \operatorname{range} T^{m+\ell}$, so that $w = T^{m+\ell}v$ for some $v \in V$. It follows that

$$T^m v \in \operatorname{range} T^m \quad \Rightarrow \quad T^m v \in \operatorname{range} T^{m+1} \quad \Rightarrow \quad T^m v = T^{m+1}u \text{ for some } u \in V.$$

Thus $w = T^{m+\ell+1}u \in \operatorname{range} T^{m+\ell+1}$. Hence $\operatorname{range} T^{m+\ell} \subseteq \operatorname{range} T^{m+\ell+1}$, as desired.

Exercise 8.A.8. Suppose $T \in \mathcal{L}(V)$. Prove that

$$\operatorname{range} T^{\dim V} = \operatorname{range} T^{\dim V+1} = \operatorname{range} T^{\dim V+2} = \cdots.$$

Solution. The proof is similar to 8.3. By Exercise 8.A.7, we need only prove that

$$\operatorname{range} T^{\dim V} = \operatorname{range} T^{\dim V+1}.$$

Seeking a contradiction, suppose that this is not true. It then follows from Exercise 8.A.6 and Exercise 8.A.7 that

$$V = \operatorname{range} T^0 \supsetneq \operatorname{range} T^1 \supsetneq \cdots \supsetneq \operatorname{range} T^{\dim V} \supsetneq \operatorname{range} T^{\dim V+1}.$$

At each of the strict inclusions in this chain, the dimension decreases by at least 1. Thus

$$\dim \operatorname{range} T^{\dim V+1} \leq -1,$$

which is a contradiction since the dimension of a vector space must be a non-negative integer.

Exercise 8.A.9. Suppose $T \in \mathcal{L}(V)$ and m is a nonnegative integer. Prove that

$$\operatorname{null} T^m = \operatorname{null} T^{m+1} \quad \Leftrightarrow \quad \operatorname{range} T^m = \operatorname{range} T^{m+1}.$$

Solution. Since $\operatorname{null} T^m$ is a subspace of $\operatorname{null} T^{m+1}$ and $\operatorname{range} T^{m+1}$ is a subspace of $\operatorname{range} T^m$ (see 8.1 and Exercise 8.A.6), it will suffice to show that

$$\dim \operatorname{null} T^m = \dim \operatorname{null} T^{m+1} \quad \Leftrightarrow \quad \dim \operatorname{range} T^m = \dim \operatorname{range} T^{m+1}.$$

Indeed, by the fundamental theorem of linear maps (3.21), we have

$$\dim \operatorname{null} T^{m+1} - \dim \operatorname{null} T^m = \dim \operatorname{range} T^m - \dim \operatorname{range} T^{m+1}.$$

Exercise 8.A.10. Define $T \in \mathcal{L}(\mathbf{C}^2)$ by $T(w,z) = (z,0)$. Find all generalized eigenvectors of T.

Solution. Observe that the matrix of T with respect to the standard basis of $\mathbf{C}^2$ is

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

It follows that the only eigenvalue of T is 0 and that $T^2 = 0$, so that every non-zero $v \in \mathbf{C}^2$ is a generalized eigenvector of T corresponding to the eigenvalue 0.

Exercise 8.A.11. Suppose that $T \in \mathcal{L}(V)$. Prove that there is a basis of V consisting of generalized eigenvectors of T if and only if the minimal polynomial of T equals $(z-\lambda_1)\cdots(z-\lambda_m)$ for some $\lambda_1, \dots, \lambda_m \in \mathbf{F}$.

Assume $\mathbf{F} = \mathbf{R}$ because the case $\mathbf{F} = \mathbf{C}$ follows from 5.27(b) and 8.9.

This exercise states that the condition for there to be a basis of V consisting of generalized eigenvectors of T is the same as the condition for there to be a basis with respect to which T has an upper-triangular matrix (see 5.44).

Caution: *If T has an upper-triangular matrix with respect to a basis $v_1, \dots, v_n$ of V, then v_1 is an eigenvector of T but it is not necessarily true that $v_2, \dots, v_n$ are generalized eigenvectors of T.*

Solution. Suppose that there is a basis $v_1, \dots, v_m$ of V consisting of generalized eigenvectors of T, i.e. there exist scalars $\lambda_1, \dots, \lambda_m \in \mathbf{F}$ and positive integers $k_1, \dots, k_m$ such that

$$(T - \lambda_n I)^{k_n} v_n = 0$$

for each $n \in \{1, \dots, m\}$. Let p be the polynomial given by

$$p(z) = (z-\lambda_1)^{k_1} \cdots (z-\lambda_m)^{k_m}$$

and notice that p is a product of linear factors. Notice further that $p(T)$ annihilates each of the basis vectors $v_1, \dots, v_m$, so that $p(T) = 0$. It follows from 5.29 that p is a polynomial multiple of the minimal polynomial of T, from which it follows that the minimal polynomial of T is a product of linear factors.

Conversely, if the minimal polynomial of T splits into linear factors, then notice that T has an eigenvalue. The proof of 8.9 then shows that there is a basis of V consisting of generalized eigenvectors of T.

Exercise 8.A.12. Suppose $T \in \mathcal{L}(V)$ is such that every vector in V is a generalized eigenvector of T. Prove that there exists $\lambda \in \mathbf{F}$ such that $T - \lambda I$ is nilpotent.

Solution. First let us prove the following lemma.

Lemma L.15. If $T \in \mathcal{L}(V)$ is such that u, v, and $u + v$ are generalized eigenvectors of T then u and v correspond to the same eigenvalue.

Proof. Suppose that u, v, and $u + v$ correspond to the eigenvalues α, β, and γ respectively. Since the list $u, v, u + v$ is linearly dependent, 8.12 implies that the eigenvalues α, β, and γ cannot be distinct, i.e. at least two of these eigenvalues are equal. If $\alpha = \beta$ then we are done, so suppose that $\beta = \gamma$; the case where $\alpha = \gamma$ is handled similarly. Observe that

$$\beta = \gamma \quad \Rightarrow \quad u + v, v \in \operatorname{null}(T - \beta I)^{\dim V} \quad \Rightarrow \quad u \in \operatorname{null}(T - \beta I)^{\dim V}.$$

Thus u corresponds to both α and β; it follows from 8.11 that $\alpha = \beta$. $\square$

Returning to the exercise, note that the desired result is clear if $V = \{0\}$. Suppose therefore that $n := \dim V \geq 1$ and fix a non-zero $u \in V$. By assumption u is a generalized eigenvector of T and thus corresponds to some eigenvalue $\lambda \in \mathbf{F}$, i.e. $(T - \lambda I)^n u = 0$. Let $v \in V$ be non-zero. If $u + v = 0$ then

$$(T - \lambda I)^n v = -(T - \lambda I)^n u = 0,$$

and if $u + v \neq 0$ then, by assumption, v and $u + v$ are generalized eigenvectors of T. It follows from Lemma L.15 that u and v correspond to the same eigenvalue, so that $(T - \lambda I)^n v = 0$. Thus $(T - \lambda I)^n = 0$, i.e. $T - \lambda I$ is nilpotent.

Exercise 8.A.13. Suppose $S, T \in \mathcal{L}(V)$ and ST is nilpotent. Prove that TS is nilpotent.

Solution. There is an integer k such that $(ST)^k = 0$, which implies

$$(TS)^{k+1} = T(ST)^k S = 0.$$

Thus TS is nilpotent.

Exercise 8.A.14. Suppose $T \in \mathcal{L}(V)$ is nilpotent and $T \neq 0$. Prove T is not diagonalizable.

Solution. 8.18 shows that the minimal polynomial of T is z^m for some positive integer m; since $T \neq 0$ we must have $m \geq 2$. It follows from 5.62 that T is not diagonalizable.

Exercise 8.A.15. Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$. Prove that T is diagonalizable if and only if every generalized eigenvector of T is an eigenvector of T.

For $\mathbf{F} = \mathbf{C}$, this exercise adds another equivalence to the list of conditions for diagonalizability in 5.55.

Solution. If every generalized eigenvector of T is an eigenvector of T, then the basis of V consisting of generalized eigenvectors of T provided by 8.9 is in fact a basis of V consisting of eigenvectors of T. That is, T is diagonalizable.

Now suppose that there is some $v \in V$ such that v is a generalized eigenvector of T but v is not an eigenvector of T, i.e. there exists $\lambda \in \mathbf{F}$ (we need not require $\mathbf{F} = \mathbf{C}$ for this implication) such that

$$(T - \lambda I)^m v = 0 \quad \text{and} \quad (T - \lambda I)v \neq 0$$

for some integer $m \geq 2$. Let p be the minimal polynomial of T and let p_v be the unique monic polynomial of smallest degree satisfying $p_v(T)v = 0$ (see Exercise 5.C.7.) A small modification of 5.29 shows that p_v must divide $(z - \lambda)^m$, so that $p_v = (z - \lambda)^k$ for some non-negative integer k. Since $v = (T - \lambda I)^0 v$ and $(T - \lambda I)^1 v$ are both non-zero, it must be the case that $k \geq 2$. It then follows from Exercise 5.C.7 (b) that the minimal polynomial of T has $(z - \lambda)^k$ as a factor, where $k \geq 2$. We may use 5.62 to conclude that T is not diagonalizable.

Exercise 8.A.16.

(a) Give an example of nilpotent operators S, T on the same vector space such that neither $S + T$ nor ST is nilpotent.

(b) Suppose $S, T \in \mathcal{L}(V)$ are nilpotent and $ST = TS$. Prove that $S + T$ and ST are nilpotent.

Solution.

(a) Let $S, T \in \mathcal{L}(\mathbf{F}^2)$ be the operators whose matrices with respect to the standard basis of $\mathbf{F}^2$ are

$$\mathcal{M}(S) = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad \mathcal{M}(T) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Some straightforward calculations show that $S^2 = T^2 = 0$, so that S and T are nilpotent, and that $(S + T)^2 \neq 0$ and $(ST)^2 \neq 0$, so that neither $S + T$ nor ST is nilpotent (by the contrapositive of 8.16).

(b) By 8.16 we have $S^n = T^n = 0$, where $n = \dim V$. Using that S and T commute, we may write $(ST)^n = S^n T^n = 0$. Thus ST is nilpotent. Furthermore, since S and T commute, we may apply the binomial theorem:

$$(S+T)^{2n} = \sum_{k=0}^{2n} \binom{2n}{k} S^{2n-k} T^k.$$

Notice that $S^{2n-k} = 0$ for $0 \leq k \leq n$ and $T^k = 0$ for $n \leq k \leq 2n$. It follows that each term in the sum above is zero, so that $(S+T)^{2n} = 0$. Thus $S+T$ is nilpotent.

Exercise 8.A.17. Suppose $T \in \mathcal{L}(V)$ is nilpotent and m is a positive integer such that $T^m = 0$.

(a) Prove that $I - T$ is invertible and that $(I-T)^{-1} = I + T + \cdots + T^{m-1}$.

(b) Explain how you would guess the formula above.

Solution.

(a) A calculation shows that

$$(I-T)(I + T + \cdots + T^{m-1}) = I - T^m = I.$$

(b) We might guess this formula by analogy with the formula

$$(1-z)(1 + z + \cdots + z^{m-1}) = 1 - z^m$$

for $z \in \mathbf{F}$.

Exercise 8.A.18. Suppose $T \in \mathcal{L}(V)$ is nilpotent. Prove that $T^{1+\dim \operatorname{range} T} = 0$.

If $\dim \operatorname{range} T < \dim V - 1$*, then this exercise improves 8.16.*

Solution. Let $\dim \operatorname{range} T = k$ and, seeking a contradiction, suppose that $T^{k+1} \neq 0$. By 8.16 we have $T^{\dim V} = 0$ and thus we can let m be the smallest integer such that $T^m = 0$ and $T^{m-1} \neq 0$; note that $k + 2 \leq m \leq n$. Let $v \in V$ be such that $T^{m-1}v \neq 0$ and $T^m v = 0$. It then follows from Exercise 8.A.2 that the list

$$v, Tv, \ldots, T^{m-1}v$$

is linearly independent, from which it follows that $k \geq m - 1 \geq k + 1$—a contradiction.

Exercise 8.A.19. Suppose $T \in \mathcal{L}(V)$ is not nilpotent. Show that

$$V = \operatorname{null} T^{\dim V - 1} \oplus \operatorname{range} T^{\dim V - 1}.$$

For operators that are not nilpotent, this exercise improves 8.4.

Solution. Observe that

$$T \text{ is not nilpotent} \quad \Rightarrow \quad \operatorname{null} T^{\dim V - 1} = \operatorname{null} T^{\dim V}$$

$$\Rightarrow \quad \operatorname{null} T^{2(\dim V-1)} = \operatorname{null} T^{\dim V-1} \quad \Rightarrow \quad V = \operatorname{null} T^{\dim V-1} \oplus \operatorname{range} T^{\dim V-1},$$

where we have used Exercise 8.A.21 for the first implication, 8.3 for the second implication, and Exercise 8.A.3 for the third implication.

Exercise 8.A.20. Suppose V is an inner product space and $T \in \mathcal{L}(V)$ is normal and nilpotent. Prove that $T = 0$.

Solution. Observe that

$$T^{\dim V} = 0 \quad \Leftrightarrow \quad \operatorname{null} T^{\dim V} = V \quad \Leftrightarrow \quad \operatorname{null} T = V \quad \Leftrightarrow \quad T = 0,$$

where we have used Exercise 7.A.27 for the second equivalence.

Exercise 8.A.21. Suppose $T \in \mathcal{L}(V)$ is such that $\operatorname{null} T^{\dim V-1} \neq \operatorname{null} T^{\dim V}$. Prove that T is nilpotent and that $\dim \operatorname{null} T^k = k$ for every integer k with $0 \leq k \leq \dim V$.

Solution. Consider the chain of inclusions provided by 8.1:

$$\{0\} = \operatorname{null} T^0 \subseteq \operatorname{null} T^1 \subseteq \cdots \subseteq \operatorname{null} T^{\dim V-1} \subseteq \operatorname{null} T^{\dim V}.$$

Since $\operatorname{null} T^{\dim V-1} \neq \operatorname{null} T^{\dim V}$, 8.2 shows that each of these inclusions must be strict. Note that if the dimension increased by more than 1 at some inclusion in this chain then we would have $\operatorname{null} T^{\dim V} > \dim V$, which cannot happen. Hence it must be the case that the dimension increases by exactly 1 at each inclusion in this chain, whence $\dim \operatorname{null} T^k = k$ for every integer $k \in \{0, \ldots, \dim V\}$. It follows from this that $T^{\dim V} = 0$, so that T is nilpotent.

Exercise 8.A.22. Suppose $T \in \mathcal{L}(\mathbf{C}^5)$ is such that $\operatorname{range} T^4 \neq \operatorname{range} T^5$. Prove that T is nilpotent.

Solution. Exercise 8.A.9 shows that $\operatorname{null} T^4 \neq \operatorname{null} T^5$. It then follows from Exercise 8.A.21 that T is nilpotent.

Exercise 8.A.23. Give an example of an operator T on a finite-dimensional real vector space such that 0 is the only eigenvalue of T but T is not nilpotent.

This exercise shows that the implication (b) $\Rightarrow$ (a) in 8.17 does not hold without the hypothesis that $\mathbf{F} = \mathbf{C}$.

Solution. Let $T \in \mathcal{L}(\mathbf{R}^3)$ be the operator whose matrix with respect to the standard basis of $\mathbf{R}^3$ is

$$\mathcal{M}(T) = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

A straightforward calculation reveals that 0 is the only eigenvalue of T and that $T^3 \neq 0$; it follows from 8.16 that T is not nilpotent.

Exercise 8.A.24. For each item in Example 8.15, find a basis of the domain vector space such that the matrix of the nilpotent operator with respect to that basis has the upper-triangular form promised by 8.18(c).

Solution.

(a) If e_1, e_2, e_3, e_4 is the standard basis of $\mathbf{F}^4$ then observe that the matrix of T with respect to the basis e_3, e_4, e_1, e_2 is

$$\begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

(b) Let $v_1 = \left(1, \frac{1}{3}, \frac{2}{3}\right), v_2 = \left(\frac{1}{6}, \frac{1}{6}, 0\right)$, and $v_3 = \left(0, \frac{1}{54}, 0\right)$. Routine calculations show that v_1, v_2, v_3 is a basis of $\mathbf{F}^3$ and that the matrix of the nilpotent operator in part (b) with respect to this basis is

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

(c) Observe that the matrix of the differentiation operator on $\mathcal{P}_m(\mathbf{R})$ with respect to the basis $1, x, \ldots, x^m$ is

$$\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 2 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & m \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}.$$

Exercise 8.A.25. Suppose that V is an inner product space and $T \in \mathcal{L}(V)$ is nilpotent. Show that there is an orthonormal basis of V with respect to which the matrix of T has the upper-triangular form promised by 8.18(c).

Solution. Combining 8.18(b) with 6.37, we see that T has an upper-triangular matrix with respect to some orthonormal basis of V, and combining 5.41 with 8.17(a) shows that the diagonal entries of this matrix are zero.

8.B. Generalized Eigenspace Decomposition

Exercise 8.B.1. Define $T \in \mathcal{L}(\mathbf{C}^2)$ by $T(w,z) = (-z,w)$. Find the generalized eigenspaces corresponding to the distinct eigenvalues of T.

Solution. As example 5.9 shows, T has the two distinct eigenvalues $\pm i$ with corresponding eigenspaces

$$E(-i,T) = \operatorname{span}((1,i)) \quad \text{and} \quad E(i,T) = \operatorname{span}((1,-i)).$$

Since $\dim \mathbf{C}^2 = 2$, 8.12 shows that there can be no other generalized eigenvectors linearly independent from the two above. Thus

$$G(-i,T) = E(-i,T) = \operatorname{span}((1,i)) \quad \text{and} \quad G(i,T) = E(i,T) = \operatorname{span}((1,-i)).$$

Exercise 8.B.2. Suppose $T \in \mathcal{L}(V)$ is invertible. Prove that $G(\lambda,T) = G\big(\frac{1}{\lambda}, T^{-1}\big)$ for every $\lambda \in \mathbf{F}$ with $\lambda \neq 0$.

Solution. By 8.20 it will suffice to show that

$$\operatorname{null}(T - \lambda I)^n = \operatorname{null}(T^{-1} - \lambda^{-1} I)^n,$$

where $n = \dim V$. Suppose therefore that $v \in \operatorname{null}(T - \lambda I)^n$, i.e.

$$0 = (T - \lambda I)^n v = \sum_{k=0}^{n} (-1)^k \lambda^k \binom{n}{k} T^{n-k} v.$$

Applying the operator $(-1)^{-n}\lambda^{-n}T^{-n}$ to both sides of this equation and using that $(-1)^{k-n} = (-1)^{n-k}$ for any $k \in \{0, ..., n\}$, we find that

$$0 = \sum_{k=0}^{n} (-1)^{n-k} (\lambda^{-1})^{n-k} \binom{n}{k} T^{-k} v = (T^{-1} - \lambda^{-1} I)^n v.$$

Thus $v \in \operatorname{null}(T^{-1} - \lambda^{-1} I)^n$ and it follows that $\operatorname{null}(T - \lambda I)^n \subseteq \operatorname{null}(T^{-1} - \lambda^{-1} I)^n$. Replacing T with T^{-1} and λ with λ^{-1} in this inclusion gives us the desired result.

Exercise 8.B.3. Suppose $T \in \mathcal{L}(V)$. Suppose $S \in \mathcal{L}(V)$ is invertible. Prove that T and $S^{-1}TS$ have the same eigenvalues with the same multiplicities.

Solution. We showed in Exercise 5.A.13 (a) that T and $S^{-1}TS$ have the same eigenvalues. Suppose that $\lambda \in \mathbf{F}$ is an eigenvalue of T and $S^{-1}TS$. Using the identity $p(T) = p(S^{-1}TS)$ for a polynomial p, observe that

$$S^{-1}(T - \lambda I)^n S = (S^{-1}TS - \lambda I)^n,$$

where $\dim V = n$. It now follows from Exercise 3.D.8 that

$$\dim \operatorname{null} (T - \lambda I)^n = \dim \operatorname{null} (S^{-1}TS - \lambda I)^n,$$

i.e. the multiplicity of λ as an eigenvalue of T equals the multiplicity of λ as an eigenvalue of $S^{-1}TS$.

Exercise 8.B.4. Suppose $\dim V \geq 2$ and $T \in \mathcal{L}(V)$ is such that

$$\operatorname{null} T^{\dim V - 2} \neq \operatorname{null} T^{\dim V - 1}.$$

Prove that T has at most two distinct eigenvalues.

Solution. Let $\dim V = n \geq 2$ and note that our hypothesis implies $\dim \operatorname{null} T^{n-1} \geq n - 1$. There are then two cases.

Case 1. If $\dim \operatorname{null} T^{n-1} = n$ then $T^{n-1} = 0$ and it follows from 8.17(a) that 0 is the only eigenvalue of T.

Case 2. Suppose that $\dim \operatorname{null} T^{n-1} = n - 1$, so that we can find $n - 1$ linearly independent generalized eigenvectors of T corresponding to the eigenvalue 0. It follows from 8.12 that T can have at most one non-zero eigenvalue (otherwise we would have at least $n + 1$ linearly independent vectors in a vector space of dimension n).

In either case, T has at most two distinct eigenvalues.

Exercise 8.B.5. Suppose $T \in \mathcal{L}(V)$ and 3 and 8 are eigenvalues of T. Let $n = \dim V$. Prove that $V = (\operatorname{null} T^{n-2}) \oplus (\operatorname{range} T^{n-2})$.

Solution. If 0 is not an eigenvalue of T, i.e. T is injective, then $\operatorname{null} T^{n-2} = \{0\}$ and the desired result follows from Exercise 5.D.4.

If 0 is an eigenvalue of T then T has at least three distinct eigenvalues and it follows from Exercise 8.B.4 and 8.2 that

$$\operatorname{null} T^{n-2} = \operatorname{null} T^{n-1} = \operatorname{null} T^{2(n-2)}.$$

Thus, by Exercise 8.A.3, $V = (\operatorname{null} T^{n-2}) \oplus (\operatorname{range} T^{n-2})$.

Exercise 8.B.6. Suppose $T \in \mathcal{L}(V)$ and λ is an eigenvalue of T. Explain why the exponent of $z - \lambda$ in the factorization of the minimal polynomial of T is the smallest positive integer m such that $(T - \lambda I)^m|_{G(\lambda, T)} = 0$.

Solution. Let p be the minimal polynomial of T and let m be the exponent of $z - \lambda$ in the factorization of p, i.e. $p(z) = (z - \lambda)^m q(z)$ for some polynomial q with $q(\lambda) \neq 0$.

We claim that $\operatorname{null} (T - \lambda I)^m = \operatorname{null} (T - \lambda I)^{m+1}$. By 8.1 it will suffice to prove the inclusion $\operatorname{null} (T - \lambda I)^{m+1} \subseteq \operatorname{null} (T - \lambda I)^m$, so suppose that $v \in \operatorname{null} (T - \lambda I)^{m+1}$, define $w = (T - \lambda I)^m v$, and notice that $Tw = \lambda w$. Now observe that

$$0 = p(T)v = q(T)(T - \lambda I)^m v = q(T)w = q(\lambda)w,$$

where we have used that $Tw = \lambda w$ for the last equality. Since $q(\lambda) \neq 0$ it must be the case that $w = 0$, i.e. $v \in \operatorname{null}(T - \lambda I)^m$. Thus $\operatorname{null}(T - \lambda I)^{m+1} \subseteq \operatorname{null}(T - \lambda I)^m$, as desired.

It now follows from 8.2 that $\operatorname{null}(T - \lambda I)^m = \operatorname{null}(T - \lambda I)^{\dim V} = G(\lambda, T)$, from which it is clear that $(T - \lambda I)^m|_{G(\lambda,T)} = 0$. To see that m is minimal, let k be any positive integer such that $(T - \lambda I)^k|_{G(\lambda,T)} = 0$ and let $s(z) = (z - \lambda)^k q(z)$. The equation $(T - \lambda I)^k|_{G(\lambda,T)} = 0$ implies that

$$\operatorname{null}(T - \lambda I)^k = G(\lambda, T) = \operatorname{null}(T - \lambda I)^m.$$

Let $v \in V$ be given. Since $0 = p(T)v = (T - \lambda I)^m q(T)v$, we either have $q(T)v = 0$, in which case $s(T)v = 0$, or

$$q(T)v \in \operatorname{null}(T - \lambda I)^m = \operatorname{null}(T - \lambda I)^k,$$

in which case $s(T)v = 0$ also. Thus $s(T) = 0$ and the minimality of the degree of p implies that $\deg p \leq \deg s$, from which it follows that $m \leq k$.

Exercise 8.B.7. Suppose $T \in \mathcal{L}(V)$ and λ is an eigenvalue of T with multiplicity d. Prove that $G(\lambda, T) = \operatorname{null}(T - \lambda I)^d$.

If $d < \dim V$, then this exercise improves 8.20.

Solution. Let p be the minimal polynomial of T and let m be the positive integer such that $p(z) = (z - \lambda)^m q(z)$ with $q(\lambda) \neq 0$. As we showed in Exercise 8.B.6,

$$G(\lambda, T) = \operatorname{null}(T - \lambda I)^m = \operatorname{null}(T - \lambda I)^{m+1}.$$

It then follows from Exercise 8.A.4 and 8.2 that

$$d = \dim G(\lambda, T) = \dim \operatorname{null}(T - \lambda I)^m \geq m$$
$$\Rightarrow \quad \operatorname{null}(T - \lambda I)^d = \operatorname{null}(T - \lambda I)^m = G(\lambda, T).$$

Exercise 8.B.8. Suppose $T \in \mathcal{L}(V)$ and $\lambda_1, \ldots, \lambda_m$ are the distinct eigenvalues of T. Prove that

$$V = G(\lambda_1, T) \oplus \cdots \oplus G(\lambda_m, T)$$

if and only if the minimal polynomial of T equals $(z - \lambda_1)^{k_1} \cdots (z - \lambda_m)^{k_m}$ for some positive integers $k_1, \ldots, k_m$.

The case $\mathbf{F} = \mathbf{C}$ follows immediately from 5.27(b) and the generalized eigenspace decomposition (8.22); thus this exercise is interesting only when $\mathbf{F} = \mathbf{R}$.

Solution. Suppose that

$$V = G(\lambda_1, T) \oplus \cdots \oplus G(\lambda_m, T).$$

For each $i \in \{1, ..., m\}$ let $d_i = \dim G(\lambda_i, T)$ and let $q(z) = (z - \lambda_1)^{d_1} \cdots (z - \lambda_m)^{d_m}$. For any $v \in V$ we have $v = v_1 + \cdots + v_m$, where each $v_i \in G(\lambda_i, T)$. Fix $i \in \{1, ..., m\}$ and note that $G(\lambda_i, T) = \operatorname{null} (T - \lambda_i I)^{d_i}$ by Exercise 8.B.7; it follows that

$$q(T)v_i = \left(\prod_{j \neq i} (T - \lambda_j I)^{d_j} \right) (T - \lambda_i I)^{d_i} v_i = 0.$$

Thus $q(T)v = 0$, whence $q(T) = 0$. It now follows from 5.29 that the minimal polynomial of T is a factor of q and hence must be of the form $(z - \lambda_1)^{k_1} \cdots (z - \lambda_m)^{k_m}$ for some non-negative integers $k_1, ..., k_m$. In fact each k_i must be positive since each λ_i is an eigenvalue of T and the eigenvalues of T are precisely the zeros of the minimal polynomial (by 5.27(a)).

Now suppose that the minimal polynomial of T equals $(z - \lambda_1)^{k_1} \cdots (z - \lambda_m)^{k_m}$ for some positive integers $k_1, ..., k_m$. The proof of 8.22(c) shows that the sum

$$G(\lambda_1, T) \oplus \cdots \oplus G(\lambda_m, T)$$

is direct and Exercise 8.A.11 shows that any vector in V can be expressed as a linear combination of generalized eigenvectors of T. Thus

$$V = G(\lambda_1, T) \oplus \cdots \oplus G(\lambda_m, T).$$

Exercise 8.B.9. Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$. Prove that there exist $D, N \in \mathcal{L}(V)$ such that $T = D + N$, the operator D is diagonalizable, N is nilpotent, and $DN = ND$.

Solution. Let $\lambda_1, ..., \lambda_m$ be the distinct eigenvalues of T and let

$$V = G(\lambda_1, T) \oplus \cdots \oplus G(\lambda_m, T)$$

be the generalized eigenspace decomposition of V. For any $v \in V$ we have $v = v_1 + \cdots + v_m$, where each $v_k \in G(\lambda_k, T)$. Define $D \in \mathcal{L}(V)$ by $Dv_k = \lambda_k v_k$, so that $D|_{G(\lambda_k, T)} = \lambda_k I$, and let $N = T - D$. Certainly D is diagonalizable. Furthermore, for any $k \in \{1, ..., m\}$,

$$N|_{G(\lambda_k, T)} = (T - D)|_{G(\lambda_k, T)} = (T - \lambda_k I)|_{G(\lambda_k, T)}$$

is nilpotent by 8.22(b). It follows that

$$N^{\dim V} v = N^{\dim V} v_1 + \cdots + N^{\dim V} v_m = 0$$

for any $v = v_1 + \cdots + v_m \in V = G(\lambda_1, T) \oplus \cdots \oplus G(\lambda_m, T)$. Thus N is nilpotent. Now, since $DN = DT - D^2$ and $ND = TD - D^2$, to show that D and N commute it will suffice to show that D and T commute. Indeed, for any $k \in \{1, ..., m\}$ and $v_k \in G(\lambda_k, T)$,

$$TDv_k = \lambda_k T v_k = DTv_k,$$

where we have used that $D|_{G(\lambda_k, T)} = \lambda_k I$ and that $G(\lambda_k, T)$ is invariant under T (by 8.22(a)). It follows that $TDv = DTv$ for any $v = v_1 + \cdots + v_m \in V = G(\lambda_1, T) \oplus \cdots \oplus G(\lambda_m, T)$. Thus T and D commute.

Exercise 8.B.10. Suppose V is a complex inner product space, $e_1, \ldots, e_n$ is an orthonormal basis of T, and $T \in \mathcal{L}(V)$. Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of T, each included as many times as its multiplicity. Prove that

$$|\lambda_1|^2 + \cdots + |\lambda_n|^2 \leq \|Te_1\|^2 + \cdots + \|Te_n\|^2.$$

See the comment after Exercise 5 in Section 7A.

Solution. Let $\lambda \in \mathbf{C}$ be an eigenvalue of T, let $d = \dim G(\lambda, T)$, and consider the restriction operator $R := T|_{G(\lambda,T)}$. Certainly λ is an eigenvalue of R and R has no other eigenvalues. By Schur's theorem (6.38) there is an orthonormal basis $f_1, \ldots, f_d$ of $G(\lambda, T)$ with respect to which the matrix $A := \mathcal{M}(R)$ is upper-triangular. For any $k \in \{1, \ldots, d\}$ it then follows that $|\lambda|^2 = \|\lambda f_k\|^2 \leq \|Tf_k\|^2$, whence

$$d|\lambda|^2 \leq \|Tf_1\|^2 + \cdots + \|Tf_d\|^2.$$

Summing this inequality over the finitely many distinct eigenvalues of T shows that

$$|\lambda_1|^2 + \cdots + |\lambda_n|^2 \leq \|Tg_1\|^2 + \cdots + \|Tg_n\|^2,$$

where $g_1, \ldots, g_n$ is the orthonormal basis of V obtained by combining the orthonormal bases of the generalized eigenspaces of T found in the previous discussion (the generalized eigenspace decomposition (8.22) shows that this provides an orthonormal basis for all of V). As we showed in Exercise 7.A.5, the quantity $\|Te_1\|^2 + \cdots + \|Te_n\|^2$ does not depend on which orthonormal basis of V is used and thus

$$|\lambda_1|^2 + \cdots + |\lambda_n|^2 \leq \|Te_1\|^2 + \cdots + \|Te_n\|^2.$$

Exercise 8.B.11. Give an example of an operator on $\mathbf{C}^4$ whose characteristic polynomial equals $(z-7)^2(z-8)^2$.

Solution. Let $T \in \mathcal{L}(\mathbf{C}^4)$ be the operator whose matrix with respect to the standard basis e_1, e_2, e_3, e_4 of $\mathbf{C}^4$ is

$$\begin{pmatrix} 7 & 0 & 0 & 0 \\ 0 & 7 & 0 & 0 \\ 0 & 0 & 8 & 0 \\ 0 & 0 & 0 & 8 \end{pmatrix}.$$

A straightforward calculation shows that 7 and 8 are the only eigenvalues of T and that

$$E(7,T) = G(7,T) = \operatorname{span}(e_1, e_2) \quad \text{and} \quad E(8,T) = G(8,T) = \operatorname{span}(e_3, e_4).$$

Thus the multiplicities of the eigenvalues 7 and 8 both equal 2, from which it follows that the characteristic polynomial of T is $(z-7)^2(z-8)^2$.

Exercise 8.B.12. Give an example of an operator on $\mathbf{C}^4$ whose characteristic polynomial equals $(z-1)(z-5)^3$ and whose minimal polynomial equals $(z-1)(z-5)^2$.

Solution. Let $T \in \mathcal{L}(\mathbf{C}^4)$ be the operator whose matrix with respect to the standard basis e_1, e_2, e_3, e_4 of $\mathbf{C}^4$ is

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 5 & 1 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 5 \end{pmatrix}.$$

A straightforward calculation shows that 1 and 5 are the only eigenvalues of T and that

$$E(1,T) = G(1,T) = \operatorname{span}(e_1) \quad \text{and} \quad G(5,T) = \operatorname{span}(e_2, e_3, e_4).$$

Thus the multiplicity of the eigenvalue 1 equals 1 and the multiplicity of the eigenvalue 5 equals 3, from which it follows that the characteristic polynomial of T is $(z-1)(z-5)^3$. Another calculation shows that $(T-I)(T-5I) \neq 0$ and that $(T-I)(T-5I)^2 = 0$. Thus the minimal polynomial of T is $(z-1)(z-5)^2$.

Exercise 8.B.13. Give an example of an operator on $\mathbf{C}^4$ whose characteristic and minimal polynomials both equal $z(z-1)^2(z-3)$.

Solution. Let $T \in \mathcal{L}(\mathbf{C}^4)$ be the operator whose matrix with respect to the standard basis e_1, e_2, e_3, e_4 of $\mathbf{C}^4$ is

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

A straightforward calculation shows that $0, 1$, and 3 are the only eigenvalues of T and that

$$E(0,T) = G(0,T) = \operatorname{span}(e_1), \quad E(3,T) = G(3,T) = \operatorname{span}(e_4),$$
$$\text{and} \quad G(1,T) = \operatorname{span}(e_2, e_3).$$

Thus the multiplicities of the eigenvalues 0 and 3 both equal 1 and the multiplicity of the eigenvalue 1 equals 2, from which it follows that the characteristic polynomial of T is $z(z-1)^2(z-3)$. Another calculation shows that

$$T(T-I)(T-3I) \neq 0 \quad \text{and} \quad T(T-I)^2(T-3I) = 0.$$

Thus the minimal polynomial of T is $z(z-1)^2(z-3)$.

Exercise 8.B.14. Give an example of an operator on $\mathbf{C}^4$ whose characteristic polynomial equals $z(z-1)^2(z-3)$ and whose minimal polynomial equals $z(z-1)(z-3)$.

Solution. Let $T \in \mathcal{L}(\mathbf{C}^4)$ be the operator whose matrix with respect to the standard basis e_1, e_2, e_3, e_4 of $\mathbf{C}^4$ is

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

A straightforward calculation shows that $0, 1$, and 3 are the only eigenvalues of T and that

$$E(0,T) = G(0,T) = \operatorname{span}(e_1), \quad E(3,T) = G(3,T) = \operatorname{span}(e_4),$$
$$\text{and} \quad E(1,T) = G(1,T) = \operatorname{span}(e_2, e_3).$$

Thus the multiplicities of the eigenvalues 0 and 3 both equal 1 and the multiplicity of the eigenvalue 1 equals 2, from which it follows that the characteristic polynomial of T is $z(z-1)^2(z-3)$. Another calculation shows that $T(T-I)(T-3I) = 0$. Thus the minimal polynomial of T is $z(z-1)(z-3)$.

Exercise 8.B.15. Let T be the operator on $\mathbf{C}^4$ defined by

$$T(z_1, z_2, z_3, z_4) = (0, z_1, z_2, z_3).$$

Find the characteristic polynomial and the minimal polynomial of T.

Solution. Let e_1, e_2, e_3, e_4 be the standard basis of $\mathbf{C}^4$. The matrix of T with respect to the basis e_4, e_3, e_2, e_1 is

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Thus 0 is the only eigenvalue of T and it then follows from 8.28 that the characteristic polynomial of T is z^4. Since $T^3 \neq 0$ it must be the case that the minimal polynomial of T is also z^4.

Exercise 8.B.16. Let T be the operator on $\mathbf{C}^6$ defined by

$$T(z_1, z_2, z_3, z_4, z_5, z_6) = (0, z_1, z_2, 0, z_4, 0).$$

Find the characteristic polynomial and the minimal polynomial of T.

Solution. Let $e_1, e_2, e_3, e_4, e_5, e_6$ be the standard basis of $\mathbf{C}^4$. The matrix of T with respect to the basis $e_3, e_2, e_1, e_5, e_4, e_6$ is

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Thus 0 is the only eigenvalue of T and it then follows from 8.28 that the characteristic polynomial of T is z^6. A straightforward calculation shows that $T^2 \neq 0$ and $T^3 = 0$. Thus the minimal polynomial of T is z^3.

Exercise 8.B.17. Suppose $\mathbf{F} = \mathbf{C}$ and $P \in \mathcal{L}(V)$ is such that $P^2 = P$. Prove that the characteristic polynomial of P is $z^m(z-1)^n$, where $m = \dim \operatorname{null} P$ and $n = \dim \operatorname{range} P$.

Solution. Exercise 5.A.8 shows that the only possible eigenvalues of P are 0 and 1. It follows from 8.28 that the characteristic polynomial of P is of the form $z^\ell(z-1)^k$, where

$$\ell = \dim G(0, P) \quad \text{and} \quad k = \dim V - \ell.$$

Since $\operatorname{null} P^2 = \operatorname{null} P$, 8.2 shows that

$$G(0, P) = \operatorname{null} P^{\dim V} = \operatorname{null} P \quad \Rightarrow \quad \ell = \dim G(0, P) = \dim \operatorname{null} P = m.$$

It now follows from the fundamental theorem of linear maps that

$$k = \dim V - \ell = \dim V - \dim \operatorname{null} P = \dim \operatorname{range} T = n.$$

Exercise 8.B.18. Suppose $T \in \mathcal{L}(V)$ and λ is an eigenvalue of T. Explain why the following four numbers equal each other.

(a) The exponent of $z - \lambda$ in the factorization of the minimal polynomial of T.

(b) The smallest positive integer m such that $(T - \lambda I)^m|_{G(\lambda,T)} = 0$.

(c) The smallest positive integer m such that

$$\operatorname{null}(T - \lambda I)^m = \operatorname{null}(T - \lambda I)^{m+1}.$$

(d) The smallest positive integer m such that

$$\operatorname{range}(T - \lambda I)^m = \operatorname{range}(T - \lambda I)^{m+1}.$$

Solution. We showed that (a) and (b) are equal in Exercise 8.B.6, and a very small modification of that argument shows that (a) and (c) are also equal. Finally, the fact that (c) and (d) are equal follows from Exercise 8.A.9.

Exercise 8.B.19. Suppose $\mathbf{F} = \mathbf{C}$ and $S \in \mathcal{L}(V)$ is a unitary operator. Prove that the constant term in the characteristic polynomial of S has absolute value 1.

Solution. Suppose $(z-\lambda_1)^{d_1}\cdots(z-\lambda_m)^{d_m}$ is the characteristic polynomial of S and note that the constant term of this polynomial is $\pm\lambda_1^{d_1}\cdots\lambda_m^{d_m}$. Note further that $|\lambda_k|=1$ for each $k\in\{1,\ldots,m\}$ by 7.54. It follows that

$$|\pm\lambda_1^{d_1}\cdots\lambda_m^{d_m}|=|\lambda_1|^{d_1}\cdots|\lambda_m|^{d_m}=1.$$

Exercise 8.B.20. Suppose that $\mathbf{F}=\mathbf{C}$ and $V_1,\ldots,V_m$ are nonzero subspaces of V such that

$$V=V_1\oplus\cdots\oplus V_m.$$

Suppose $T\in\mathcal{L}(V)$ and each V_k is invariant under T. For each k, let p_k denote the characteristic polynomial of $T|_{V_k}$. Prove that the characteristic polynomial of T equals $p_1\cdots p_m$.

Solution. It will suffice to prove the case where $m=2$; a straightforward induction argument will then prove the general statement. Suppose therefore that $V=U\oplus W$, where U and W are non-zero subspaces of V invariant under T.

Let $E(T)$ be the collection of eigenvalues of T and define $E(T|_U)$ and $E(T|_W)$ similarly. We claim that $E(T)=E(T|_U)\cup E(T|_W)$. Certainly any eigenvalue of $T|_U$ or of $T|_W$ must also be an eigenvalue of T, so that $E(T|_U)\cup E(T|_W)\subseteq E(T)$. For the reverse inclusion, suppose that $\lambda\in\mathbf{C}$ is an eigenvalue of T, so that $Tv=\lambda v$ for some non-zero $v\in V$. Let $u\in U$ and $w\in W$ be such that $v=u+w$ and observe that

$$Tu-\lambda u=\lambda w-Tw.$$

Since U and W are invariant under T we have $Tu-\lambda u\in U$ and $\lambda w-Tw\in W$, whence

$$Tu-\lambda u=\lambda w-Tw\in U\cap W=\{0\}.$$

Thus $Tu=\lambda u$ and $Tw=\lambda w$. Note that at least one of u and w must be non-zero, since v is non-zero. It follows that $\lambda\in E(T|_U)\cup E(T|_W)$ and we may conclude that

$$E(T)=E(T|_U)\cup E(T|_W),$$

as claimed.

Next we claim that $G(\lambda,T)=G(\lambda,T|_U)\oplus G(\lambda,T|_W)$ for any $\lambda\in\mathbf{C}$. First, note that

$$G(\lambda,T|_U)\subseteq U\quad\text{and}\quad G(\lambda,T|_W)\subseteq W.$$

Since $U\cap W=\{0\}$, it follows that $G(\lambda,T|_U)\cap G(\lambda,T|_W)=\{0\}$. Thus the sum

$$G(\lambda,T|_U)\oplus G(\lambda,T|_W)$$

is indeed direct.

Now suppose that

$$u + w \in G(\lambda, T|_U) \oplus G(\lambda, T|_W),$$

$$\text{i.e.} \quad u \in U \text{ and } (T|_U - \lambda I)^{\dim U} u = 0, \quad w \in W \text{ and } (T|_W - \lambda I)^{\dim W} w = 0.$$

It follows that $(T - \lambda I)^{\dim V} u = (T - \lambda I)^{\dim V} w = 0$ and thus $u + w \in G(\lambda, T)$. Hence

$$G(\lambda, T|_U) \oplus G(\lambda, T|_W) \subseteq G(\lambda, T).$$

Now suppose that $v \in G(\lambda, T)$, i.e. $(T - \lambda I)^{\dim V} v = 0$, and let $u \in U$ and $w \in W$ be such that $v = u + w$. Observe that

$$(T - \lambda I)^{\dim V}(u + w) = 0 \quad \Leftrightarrow \quad (T - \lambda I)^{\dim V} u = -(T - \lambda I)^{\dim V} w.$$

Since U and W are invariant under T we have $(T - \lambda I)^{\dim V} u \in U$ and $(T - \lambda I)^{\dim V} w \in W$, whence

$$(T - \lambda I)^{\dim V} u = -(T - \lambda I)^{\dim V} w \in U \cap W = \{0\}.$$

Thus

$$(T - \lambda I)^{\dim V} u = 0 \quad \Rightarrow \quad (T|_U - \lambda I)^{\dim V} u = 0 \quad \Rightarrow \quad (T|_U - \lambda I)^{\dim U} u = 0,$$

where we have used 8.3 for the last implication. It follows that $u \in G(\lambda, T|_U)$ and we can similarly show that $w \in G(\lambda, T|_W)$. Hence $v = u + w \in G(\lambda, T|_U) \oplus G(\lambda, T|_W)$ and we may conclude that

$$G(\lambda, T) = G(\lambda, T|_U) \oplus G(\lambda, T|_W),$$

as claimed.

We have now proved the following:

(i) $E(T) = E(T|_U) \cup E(T|_W)$;

(ii) $G(\lambda, T) = G(\lambda, T|_U) \oplus G(\lambda, T|_W)$ for any $\lambda \in \mathbf{C}$.

Let $\lambda_1, \ldots, \lambda_n$ denote the distinct eigenvalues of T, let

$$m_i = \dim G(\lambda_i, T), \quad k_i = \dim G(\lambda_i, T|_U), \quad \text{and} \quad \ell_i = \dim G(\lambda_i, T|_W),$$

and let r, p, and q be the characteristic polynomials of $T, T|_U$, and $T|_W$ respectively; by definition we have $r(z) = (z - \lambda_1)^{m_1} \cdots (z - \lambda_n)^{m_n}$. It follows from (i) that

$$p(z) = (z - \lambda_1)^{k_1} \cdots (z - \lambda_n)^{k_n} \quad \text{and} \quad q(z) = (z - \lambda_1)^{\ell_1} \cdots (z - \lambda_n)^{\ell_n},$$

and it follows from (ii) that $k_i + \ell_i = m_i$ for each $i \in \{1, \ldots, n\}$. Thus

$$p(z)q(z) = (z - \lambda_1)^{k_1 + \ell_1} \cdots (z - \lambda_n)^{k_n + \ell_n} = (z - \lambda_1)^{m_1} \cdots (z - \lambda_n)^{m_n} = r(z),$$

as desired.

Exercise 8.B.21. Suppose $p, q \in \mathcal{P}(\mathbf{C})$ are monic polynomials with the same zeros and q is a polynomial multiple of p. Prove that there exists $T \in \mathcal{L}(\mathbf{C}^{\deg q})$ such that the characteristic polynomial of T is q and the minimal polynomial of T is p.

This exercise implies that every monic polynomial is the characteristic polynomial of some operator.

Solution. Let $\lambda_1, \ldots, \lambda_m$ be the distinct zeros of p and q and suppose that

$$p(z) = (z - \lambda_1)^{k_1} \cdots (z - \lambda_m)^{k_m} \quad \text{and} \quad q(z) = (z - \lambda_1)^{\ell_1} \cdots (z - \lambda_m)^{\ell_m}$$

for some positive integers $k_1, \ldots, k_m, \ell_1, \ldots, \ell_m$; note that $k_i \leq \ell_i$ for each $i \in \{1, \ldots, m\}$ since q is a polynomial multiple of p. Fix $i \in \{1, \ldots, m\}$. If $k_i = 1$ then let A_i be the $\ell_i \times \ell_i$ diagonal matrix with diagonal entries equal to λ_i. If $k_i > 1$ and $k_i = \ell_i$ then let A_i be the $\ell_i \times \ell_i$ matrix with diagonal entries equal to λ_i and entries directly above the diagonal equal to 1, i.e.

$$A_i = \begin{pmatrix} \lambda_i & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_i & 1 & \cdots & 0 & 0 \\ 0 & 0 & \lambda_i & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_i & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda_i \end{pmatrix}.$$

If $1 < k_i < \ell_i$ then let B_i be the $(\ell_i - k_i) \times (\ell_i - k_i)$ diagonal matrix with diagonal entries equal to λ_i, let C_i be the $k_i \times k_i$ matrix with diagonal entries equal to λ_i and entries directly above the diagonal equal to 1, and let A_i be the block diagonal matrix

$$\begin{pmatrix} B_i & 0 \\ 0 & C_i \end{pmatrix}.$$

Now let $T \in \mathcal{L}(\mathbf{C}^n)$, where $n = \deg q$, be the operator whose matrix with respect to the standard basis is the block diagonal matrix

$$A := \begin{pmatrix} A_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & A_m \end{pmatrix}.$$

It follows that the distinct eigenvalues of T are $\lambda_1, \ldots, \lambda_m$, since these are precisely the distinct diagonal elements of A. Fix $i \in \{1, \ldots, m\}$. For any positive integer d, a calculation (see Exercise 8.B.22) shows that

$$(A-\lambda_i I_n)^d = \begin{pmatrix} \left(A_1-\lambda_i I_{\ell_1}\right)^d & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & \left(A_i-\lambda_i I_{\ell_i}\right)^d & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & \left(A_m-\lambda_i I_{\ell_m}\right)^n \end{pmatrix},$$

where I_n is the $n \times n$ identity matrix and I_{ℓ_j} is the $\ell_j \times \ell_j$ identity matrix. For $j \neq i$, note that the upper-triangular matrix $A_j - \lambda_i I_{\ell_j}$ has non-zero diagonal entries since $\lambda_j \neq \lambda_i$; it follows that the upper-triangular matrix $\left(A_j - \lambda_i I_{\ell_j}\right)^d$ also has non-zero diagonal entries and hence is injective. Another calculation shows that $(A_i - \lambda_i I_{\ell_i})^{\ell_i}$ is the $\ell_i \times \ell_i$ zero matrix. It follows from this discussion that

$$\dim G(\lambda_i, T) = \dim \operatorname{null} (A - \lambda_i I_n)^n = \dim \operatorname{null} (A_i - \lambda_i I_{\ell_i})^n = \ell_i.$$

Hence q is the characteristic polynomial of T. It also follows from this discussion that the matrix of $(T - \lambda_i I)|_{G(\lambda_i, T)}$ with respect to the standard basis is $A_i - \lambda_i I_{\ell_i}$. A final calculation shows that k_i is the least integer k such that $(A_i - \lambda_i I_{\ell_i})^k = 0$ and thus, by Exercise 8.B.18, k_i must be the exponent of $z - \lambda_i$ in the factorization of the minimal polynomial of T. Hence p is the minimal polynomial of T.

Exercise 8.B.22. Suppose A and B are block diagonal matrices of the form

$$A = \begin{pmatrix} A_1 & & 0 \\ & \ddots & \\ 0 & & A_m \end{pmatrix}, \quad B = \begin{pmatrix} B_1 & & 0 \\ & \ddots & \\ 0 & & B_m \end{pmatrix},$$

where A_k and B_k are square matrices of the same size for each $k = 1, \dots, m$. Show that AB is a block diagonal matrix of the form

$$AB = \begin{pmatrix} A_1 B_1 & & 0 \\ & \ddots & \\ 0 & & A_m B_m \end{pmatrix}.$$

Solution. For each $k \in \{1, \dots, m\}$ suppose that A_k and B_k are $\ell_k \times \ell_k$ matrices and let $n = \ell_1 + \cdots + \ell_m$, so that A and B are $n \times n$ matrices. Let $S, T \in \mathcal{L}(\mathbf{F}^n)$ be the operators whose matrices with respect to the standard basis $e_1, \dots, e_n$ are $\mathcal{M}(S) = A$ and $\mathcal{M}(T) = B$, so that $\mathcal{M}(ST) = AB$. Let E_1 be the list $e_1, \dots, e_{\ell_1}$ and, if $m \geq 2$, for each $k \in \{2, \dots, m\}$ let E_k be the list

$$e_{\ell_1 + \cdots + \ell_{k-1} + 1}, \dots, e_{\ell_1 + \cdots + \ell_k}.$$

Now let $U_k = \operatorname{span} E_k$ for each $k \in \{1, \dots, m\}$ and note that E_k is a basis of U_k. Note further that each U_k is invariant under both S and T and that the matrices of $S|_{U_k}$ and $T|_{U_k}$ with respect to E_k are A_k and B_k respectively. It follows that each U_k is invariant under ST and that the matrix of $(ST)|_{U_k}$ with respect to the basis E_k is $A_k B_k$. Thus the k^{th} block on the diagonal of AB equals A_kB_k, and all entries of AB off the "block diagonal" must be zero (otherwise U_k would not be invariant under ST).

Exercise 8.B.23. Suppose $\mathbf{F} = \mathbf{R}, T \in \mathcal{L}(V)$, and $\lambda \in \mathbf{C}$.

(a) Show that $u + iv \in G(\lambda, T_{\mathbf{C}})$ if and only if $u - iv \in G(\overline{\lambda}, T_{\mathbf{C}})$.

(b) Show that the multiplicity of λ as an eigenvalue of $T_{\mathbf{C}}$ equals the multiplicity of $\overline{\lambda}$ as an eigenvalue of $T_{\mathbf{C}}$.

(c) Use (b) and the result about the sum of multiplicities (8.25) to show that if $\dim V$ is an odd number, then $T_{\mathbf{C}}$ has a real eigenvalue.

(d) Use (c) and the result about real eigenvalues of $T_{\mathbf{C}}$ (Exercise 17 in Section 5A) to show that if $\dim V$ is an odd number, then T has an eigenvalue (thus giving an alternative proof of 5.34).

See Exercise 33 in Section 3B for the definition of the complexification $T_{\mathbf{C}}$.

Solution.

(a) Analogously to complex conjugation, for $u, v \in V$ let us define $\overline{u + iv} = u - iv$. This operation has the properties we would expect of it, such as

$$\overline{(u + iv) + (x + iy)} = \overline{u + iv} + \overline{x + iy} \quad \text{and} \quad \overline{\alpha(u + iv)} = \overline{\alpha}\,\overline{u + iv} \text{ for } \alpha \in \mathbf{C}.$$

Furthermore, observe that

$$\overline{T_{\mathbf{C}}(u + iv)} = \overline{Tu + iTv} = Tu - iTv = T_{\mathbf{C}}(u - iv) = T_{\mathbf{C}}\big(\overline{u + iv}\big).$$

It follows from these algebraic properties that, for a non-negative integer m,

$$\overline{(T_{\mathbf{C}} - \lambda I)^m(u + iv)} = \big(T_{\mathbf{C}} - \overline{\lambda} I\big)^m(u - iv).$$

Combining this identity with the obvious equation $\overline{0} = 0$ (where 0 is the zero vector in $V_{\mathbf{C}}$) shows that $u + iv \in G(\lambda, T_{\mathbf{C}})$ if and only if $u - iv \in G(\overline{\lambda}, T_{\mathbf{C}})$.

(b) Let $u_1 + iv_1, \ldots, u_m + iv_m$ be a basis of $G(\lambda, T_{\mathbf{C}})$; we claim that $u_1 - iv_1, \ldots, u_1 - iv_m$ is a basis of $G(\overline{\lambda}, T_{\mathbf{C}})$. Part (a) shows that each $u_k - iv_k$ indeed belongs to $G(\overline{\lambda}, T_{\mathbf{C}})$. Suppose that $\alpha_1, \ldots, \alpha_m \in \mathbf{C}$ are such that

$$\alpha_1(u_1 - iv_1) + \cdots + \alpha_m(u_m - iv_m) = 0.$$

Taking the complex conjugate of both sides (see part (a)) shows that

$$\overline{\alpha_1}(u_1 + iv_1) + \cdots + \overline{\alpha_m}(u_m + iv_m) = 0.$$

The linear independence of the list $u_1 + iv_1, \ldots, u_m + iv_m$ now shows that

$$\overline{\alpha_1} = \cdots = \overline{\alpha_m} = 0 \quad \Rightarrow \quad \alpha_1 = \cdots = \alpha_m = 0.$$

Thus the list $u_1 - iv_1, ..., u_1 - iv_m$ is linearly independent. Now let $u + iv \in G(\overline{\lambda}, T_{\mathbf{C}})$ be given. By part (a) we have $u - iv \in G(\lambda, T)$ and thus there exist scalars $\alpha_1, ..., \alpha_m \in \mathbf{C}$ such that

$$u - iv = \alpha_1(u_1 + iv_1) + \cdots + \alpha_m(u_m + iv_m).$$

Taking the complex conjugate of both sides shows that

$$u + iv = \overline{\alpha_1}(u - iv_1) + \cdots + \overline{\alpha_m}(u_m - iv_m).$$

Thus $u_1 - iv_1, ..., u_m - iv_m$ spans $G(\overline{\lambda}, T_{\mathbf{C}})$. Hence $u_1 - iv_1, ..., u_m - iv_m$ is a basis of $G(\overline{\lambda}, T_{\mathbf{C}})$, as claimed. It follows that $\dim G(\lambda, T_{\mathbf{C}}) = \dim G(\overline{\lambda}, T_{\mathbf{C}})$, i.e. the multiplicity of λ as an eigenvalue of $T_{\mathbf{C}}$ equals the multiplicity of $\overline{\lambda}$ as an eigenvalue of $T_{\mathbf{C}}$.

(c) We will prove the contrapositive. Suppose that $T_{\mathbf{C}}$ has no real eigenvalues. Since non-real eigenvalues of $T_{\mathbf{C}}$ come in pairs (by part (a)) and both eigenvalues of this pair have the same multiplicity (by part (b)), the sum of the multiplicities of all the eigenvalues of $T_{\mathbf{C}}$ must be an even number. Thus, by 8.25, $\dim V$ is an even number.

(d) This is immediate from part (c) and Exercise 5.A.17.

8.C. Consequences of Generalized Eigenspace Decomposition

Exercise 8.C.1. Suppose $T \in \mathcal{L}(\mathbf{C}^3)$ is the operator defined by

$$T(z_1, z_2, z_3) = (z_2, z_3, 0).$$

Prove that T does not have a square root.

Solution. Notice that T is nilpotent. Thus, if $S \in \mathcal{L}(\mathbf{C}^3)$ satisfies $S^2 = T$, then S must also be nilpotent. It follows from 8.16 that $S^3 = 0$. However, note that

$$0 \neq T^2(0,0,1) = S^4(0,0,1).$$

Thus there can be no $S \in \mathcal{L}(\mathbf{C}^3)$ satisfying $S^2 = T$.

Exercise 8.C.2. Define $T \in \mathcal{L}(\mathbf{F}^5)$ by $T(x_1, x_2, x_3, x_4, x_5) = (2x_2, 3x_3, -x_4, 4x_5, 0)$.

(a) Show that T is nilpotent.

(b) Find a square root of $I + T$.

Solution.

(a) Notice that the matrix of T with respect to the standard basis of $\mathbf{F}^5$ is

$$\begin{pmatrix} 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

It follows from 8.18 that T is nilpotent.

(b) Note that $T^4 \neq 0$ and $T^5 = 0$. Following the strategy outlined in the proof of 8.39, we should attempt to solve the equation

$$\left(I + a_1 T + a_2 T^2 + a_3 T^3 + a_4 T^4\right)^2 = I + T$$

for the coefficients a_1, a_2, a_3, a_4. After calculating, we find that

$$a_1 = \frac{1}{2}, \quad a_2 = -\frac{1}{8}, \quad a_3 = \frac{1}{16}, \quad a_4 = -\frac{5}{128}.$$

Thus the operator $\frac{1}{128}\left(128I + 64T - 16T^2 + 8T^3 - 5T^4\right)$ is a square root of $I + T$ (a matrix calculation, by hand or otherwise, verifies this).

Exercise 8.C.3. Suppose V is a complex vector space. Prove that every invertible operator on V has a cube root.

Solution. The proof is almost identical to the proof of 8.41, replacing "square root" with "cube root" where applicable. The crux of the argument is the existence of cube roots for operators of the form $I + T$, where T is nilpotent. As in the proof of 8.39, proving this existence amounts to showing that we can always solve the equation

$$\left(I + a_1 T + a_2 T^2 + \cdots + a_{m-1} T^{m-1}\right)^3 = I + T$$

for $a_1, a_2, ..., a_{m-1}$, where m is some positive integer. By multiplying out the left-hand side, we notice that the coefficient of T^k is always a degree 1 polynomial in a_k with constant term involving sums and products of $a_1, ..., a_{k-1}$. Thus, having found $a_1, ..., a_{k-1}$, we can always solve for a_k.

Exercise 8.C.4. Suppose V is a real vector space. Prove that the operator $-I$ on V has a square root if and only if $\dim V$ is an even number.

Solution. Suppose that $\dim V$ is an even number and let $v_1, ..., v_{2n}$ be a basis of V. Define $R \in \mathcal{L}(V)$ by

$$Re_{2k-1} = -e_{2k} \quad \text{and} \quad Re_{2k} = e_{2k-1}$$

for each $k \in \{1, ..., n\}$. It follows that

$$R^2 e_{2k-1} = -e_{2k-1} \quad \text{and} \quad R^2 e_{2k} = -e_{2k}$$

for each $k \in \{1, ..., n\}$, so that $R^2 = -I$, i.e. R is a square root of $-I$.

Now suppose that $\dim V$ is an odd number and let $T \in \mathcal{L}(V)$ be given. By 5.34, there exists an eigenvalue $\lambda \in \mathbf{R}$ of T, say $Tv = \lambda v$ for some non-zero $v \in V$. It follows that $T^2 v = \lambda^2 v$ and hence that $T^2 v \neq -v$, since $\lambda^2 \geq 0$. Thus no operator $T \in \mathcal{L}(V)$ satisfies $T^2 = -I$.

Exercise 8.C.5. Suppose $T \in \mathcal{L}(\mathbf{C}^2)$ is the operator defined by

$$T(w, z) = (-w - z, 9w + 5z).$$

Find a Jordan basis for T.

Solution. The matrix of T with respect to the basis $(1, -3), (1, -2)$ is

$$\begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}.$$

Thus $(1, -3), (1, -2)$ is a Jordan basis for T.

Exercise 8.C.6. Find a basis of $\mathcal{P}_4(\mathbf{R})$ that is a Jordan basis for the differentiation operator D on $\mathcal{P}_4(\mathbf{R})$ defined by $Dp = p'$.

Solution. The matrix of D with respect to the basis $1, x, \frac{1}{2}x^2, \frac{1}{3}x^3, \frac{1}{4}x^4$ is

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Thus $1, x, \frac{1}{2}x^2, \frac{1}{3}x^3, \frac{1}{4}x^4$ is a Jordan basis for D.

Exercise 8.C.7. Suppose $T \in \mathcal{L}(V)$ is nilpotent and $v_1, ..., v_n$ is a Jordan basis for T. Prove that the minimal polynomial of T is z^{m+1}, where m is the length of the longest consecutive string of 1's that appears on the line directly above the diagonal in the matrix of T with respect to $v_1, ..., v_n$.

Solution. The matrix of T with respect to $v_1, ..., v_n$ is a block diagonal matrix of the form

$$A = \begin{pmatrix} A_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & A_p \end{pmatrix},$$

where each A_k is a $d_k \times d_k$ matrix of the form

$$A_k = \begin{pmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{pmatrix};$$

the diagonal entries are zero as T, being nilpotent, has only zero as an eigenvalue. A calculation shows that $A_k^{d_k-1} \neq 0$ and $A_k^{d_k} = 0$. Note that the length of the string of 1's appearing on the line directly above the diagonal of A_k is exactly $d_k - 1$. It follows that $d_k \leq m+1$ for each $k \in \{1, ..., n\}$, so that $A_k^{m+1} = 0$, and $d_\ell = m+1$ for some $\ell \in \{1, ..., n\}$, so that $A_\ell^m \neq 0$. Thus, by Exercise 8.A.22,

$$A^{m+1} = \begin{pmatrix} A_1^{m+1} & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & A_\ell^{m+1} & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & A_p^{m+1} \end{pmatrix} = \begin{pmatrix} 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 \end{pmatrix} = 0,$$

$$A^m = \begin{pmatrix} A_1^m & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & A_\ell^m & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & A_p^m \end{pmatrix} \neq 0.$$

Hence $T^{m+1} = 0$ and $T^m \neq 0$. Furthermore, zero is the only eigenvalue of T and hence the only root of the minimal polynomial of T. We may conclude that the minimal polynomial of T is z^{m+1}.

Exercise 8.C.8. Suppose $T \in \mathcal{L}(V)$ and $v_1, ..., v_n$ is a basis of V that is a Jordan basis for T. Describe the matrix of T^2 with respect to this basis.

Solution. The matrix of T with respect to $v_1, ..., v_n$ is a block diagonal matrix of the form

$$A = \begin{pmatrix} A_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & A_p \end{pmatrix}, \quad \text{where } A_k = \begin{pmatrix} \lambda_k & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_k \end{pmatrix}.$$

A calculation shows that

$$A_k^2 = \begin{pmatrix} \lambda_k^2 & 2\lambda_k & 1 & & 0 \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & 1 \\ & & & \ddots & 2\lambda_k \\ 0 & & & & \lambda_k^2 \end{pmatrix}.$$

Thus, by Exercise 8.A.22, the matrix of T^2 with respect to $v_1, ..., v_n$ is

$$A^2 = \begin{pmatrix} A_1^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & A_p^2 \end{pmatrix}.$$

Exercise 8.C.9. Suppose $T \in \mathcal{L}(V)$ is nilpotent. Explain why there exist $v_1, ..., v_n \in V$ and non-negative integers $m_1, ..., m_n$ such that (a) and (b) below both hold.

(a) $T^{m_1}v_1, ..., Tv_1, v_1, ..., T^{m_n}v_n, ..., Tv_n, v_n$ is a basis of V.

(b) $T^{m_1+1}v_1 = \cdots = T^{m_n+1}v_n = 0$.

Solution. By 8.45 there exists a Jordan basis for T, i.e. a basis with respect to which the matrix of T is of the form

$$A = \begin{pmatrix} A_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & A_n \end{pmatrix}, \quad \text{where } A_k = \begin{pmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{pmatrix};$$

the diagonal entries of each A_k are zero as T, being nilpotent, has only zero as an eigenvalue. Fix $k \in \{1, ..., n\}$ and suppose that A_k is an $(m_k + 1) \times (m_k + 1)$ matrix, where m_k is some non-negative integer. Suppose that the sub-list of the Jordan basis corresponding to the A_k block on the block diagonal of A is $u_{m_k+1}, u_{m_k}, ..., u_3, u_2, u_1$. The form of A_k shows that

$$Tu_1 = u_2, \quad T^2u_1 = Tu_2 = u_3, \quad ..., \quad T^{m_k}u_1 = u_{m_k+1}, \quad T^{m_k+1}u_1 = 0.$$

Thus we can take $v_k = u_k$ for each $k \in \{1, ..., n\}$ and (a) and (b) will both hold.

Exercise 8.C.10. Suppose $T \in \mathcal{L}(V)$ and $v_1, \ldots, v_n$ is a basis of V that is a Jordan basis for T. Describe the matrix of T with respect to the basis $v_n, \ldots, v_1$ obtained by reversing the order of the v's.

Solution. The matrix of T with respect to $v_1, \ldots, v_n$ is a block diagonal matrix of the form

$$A = \begin{pmatrix} A_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & A_p \end{pmatrix}, \quad \text{where } A_k = \begin{pmatrix} \lambda_k & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_k \end{pmatrix}.$$

Reversing the order of the sub-list of $v_1, \ldots, v_n$ corresponding to A_k has the effect of transposing A_k, and reversing the order of the entire basis $v_1, \ldots, v_n$ has the effect of reversing the order of the blocks on the block diagonal of A. Thus the matrix of T with respect to the basis $v_n, \ldots, v_1$ is

$$\begin{pmatrix} A_p^{\mathrm{t}} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & A_1^{\mathrm{t}} \end{pmatrix}.$$

Exercise 8.C.11. Suppose $T \in \mathcal{L}(V)$. Explain why every vector in each Jordan basis for T is a generalized eigenvector of T.

Solution. Suppose that $v_1, \ldots, v_n$ is a Jordan basis for T, i.e. the matrix of T with respect to $v_1, \ldots, v_n$ is of the form

$$A = \begin{pmatrix} A_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & A_p \end{pmatrix}, \quad \text{where } A_k = \begin{pmatrix} \lambda_k & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_k \end{pmatrix}.$$

Let v_m be a vector in the Jordan basis $v_1, \ldots, v_n$. The form of A above shows that there is a $k \in \{1, \ldots, p\}$ such that either $(T - \lambda_k I)v_m = 0$ or

$$Tv_m = \lambda_k v_m + v_{m-1} \quad \Rightarrow \quad (T - \lambda_k I)v_m = v_{m-1}.$$

Similarly, either $(T - \lambda_k I)v_{m-1} = 0$, which gives us $(T - \lambda_k I)^2 v_m = 0$, or

$$Tv_{m-1} = \lambda_k v_{m-1} + v_{m-2} \quad \Rightarrow \quad (T - \lambda_k I)v_{m-1} = (T - \lambda_k I)^2 v_m = v_{m-2}.$$

Continuing in this fashion, we find a positive integer ℓ, no greater than the size of A_k, such that $(T - \lambda_k I)^\ell v_m = 0$. Thus v_m is a generalized eigenvector of T.

Exercise 8.C.12. Suppose $T \in \mathcal{L}(V)$ is diagonalizable. Show that $\mathcal{M}(T)$ is a diagonal matrix with respect to every Jordan basis for T.

Solution. Let $v_1, ..., v_n$ be a Jordan basis for T. It follows from Exercise 8.C.11 that $v_1, ..., v_n$ is a basis of V consisting of generalized eigenvectors of T. Because T is diagonalizable, every generalized eigenvector of T is an eigenvector of T (as we showed in Exercise 8.A.15; note that the proof of this implication does not use the hypothesis of Exercise 8.A.15 that $\mathbf{F} = \mathbf{C}$). Thus $v_1, ..., v_n$ is a basis of V consisting of eigenvectors of T and hence the matrix of T with respect to this basis is diagonal.

Exercise 8.C.13. Suppose $T \in \mathcal{L}(V)$ is nilpotent. Prove that if $v_1, ..., v_n$ are vectors in V and $m_1, ..., m_n$ are nonnegative integers such that

$$T^{m_1}v_1, ..., Tv_1, v_1, ..., T^{m_n}v_n, ..., Tv_n, v_n \text{ is a basis of } V$$

and

$$T^{m_1+1}v_1 = \cdots = T^{m_n+1}v_n = 0,$$

then $T^{m_1}v_1, ..., T^{m_n}v_n$ is a basis of $\operatorname{null} T$.

This exercise shows that $n = \dim \operatorname{null} T$. Thus the positive integer n that appears above depends only on T and not on the specific Jordan basis chosen for T.

Solution. The linear independence of the basis

$$T^{m_1}v_1, ..., Tv_1, v_1, ..., T^{m_n}v_n, ..., Tv_n, v_n$$

gives us the linear independence of the list $T^{m_1}v_1, ..., T^{m_n}v_n$. The condition

$$T^{m_1+1}v_1 = \cdots = T^{m_n+1}v_n = 0$$

shows that each of the vectors $T^{m_1}v_1, ..., T^{m_n}v_n$ belongs to $\operatorname{null} T$. Suppose that $v \in \operatorname{null} T$. There are scalars $a_{j,k}$ such that

$$v = \sum_{j=1}^{n} \sum_{k=0}^{m_j} a_{j,k} T^k v_j \quad \Rightarrow \quad 0 = Tv = \sum_{j=1}^{n} \sum_{k=0}^{m_j} a_{j,k} T^{k+1} v_j = \sum_{\substack{j=1 \\ m_j \neq 0}}^{n} \sum_{k=0}^{m_j - 1} a_{j,k} T^{k+1} v_j,$$

where we have used that $T^{m_1+1}v_1 = \cdots = T^{m_n+1}v_n = 0$ for the last equality. It follows from the linear independence of the basis

$$T^{m_1}v_1, ..., Tv_1, v_1, ..., T^{m_n}v_n, ..., Tv_n, v_n$$

that $a_{j,k} = 0$ for all $j \in \{1, ..., n\}$ such that $m_j \geq 1$ and all $k \in \{0, ..., m_j - 1\}$. Thus

$$v = a_{1,m_1} T^{m_1} v_1 + \cdots + a_{n,m_n} T^{m_n} v_n.$$

Hence $T^{m_1}v_1, ..., T^{m_n}v_n$ spans $\operatorname{null} T$ and we may conclude that this list is a basis of $\operatorname{null} T$.

Exercise 8.C.14. Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$. Prove that there does not exist a direct sum decomposition of V into two nonzero subspaces invariant under T if and only if the minimal polynomial of T is of the form $(z - \lambda)^{\dim V}$ for some $\lambda \in \mathbf{C}$.

Solution. First suppose that there exist non-zero subspaces U and W which are invariant under T and satisfy $V = U \oplus W$. If the union of the set of eigenvalues of $T|_U$ and the set of eigenvalues of $T|_W$ contains at least two complex numbers then T has at least two eigenvalues. It follows that the minimal polynomial of T has at least two roots and hence is not of the form $(z-\lambda)^{\dim V}$ for any $\lambda \in \mathbf{C}$. Suppose therefore that $T|_U$ and $T|_W$ both have a single eigenvalue $\mu \in \mathbf{C}$. As we argued in Exercise 8.B.20, it follows that μ is the only eigenvalue of T and hence that the minimal polynomial of T is of the form $(z-\mu)^\ell$ for some positive integer $\ell \leq \dim V$. Let $m = \max\{\dim U, \dim W\}$ and notice that $m < \dim V$ since U and W are proper subspaces of V. Notice further that $V = G(\mu, T), U = G(\mu, T|_U)$, and $W = G(\mu, T|_W)$ by 8.22(c). It follows from 8.22(b) and 8.16 that

$$(T|_U - \mu I)^m = (T|_W - \mu I)^m = 0.$$

Let $v = u + w \in V = U \oplus W$ be given and observe that

$$(T-\mu I)^m|_{G(\mu,T)}\, v = (T-\mu I)^m(u+w) = (T|_U - \mu I)^m u + (T|_W - \mu I)^m w = 0.$$

Thus $(T-\mu I)^m|_{G(\mu,T)} = 0$. Hence, by Exercise 8.B.6, we must have $\ell \leq m < \dim V$. It follows that the minimal polynomial of T is not of the form $(z-\lambda)^{\dim V}$ for any $\lambda \in \mathbf{C}$.

Now suppose that the minimal polynomial of T is not of the form $(z-\lambda)^{\dim V}$ for any $\lambda \in \mathbf{C}$ and let $\lambda_1, \ldots, \lambda_m$ be the distinct eigenvalues of T. If $m \geq 2$ then, by 8.22(c),

$$V = G(\lambda_1, T) \oplus [G(\lambda_2, T) \oplus \cdots \oplus G(\lambda_m, T)]$$

is a direct sum decomposition of V into two non-zero subspaces of V invariant under T. If $m = 1$ then the minimal polynomial of T must be of the form $(z-\lambda_1)^\ell$ where ℓ is a positive integer satisfying $\ell < \dim V$. By 8.46 there exists a Jordan basis $v_1, \ldots, v_{\dim V}$ for T, so that the matrix of T with respect to this basis is of the form

$$A = \begin{pmatrix} A_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & A_p \end{pmatrix}, \quad \text{where } A_k = \begin{pmatrix} \lambda_1 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_1 \end{pmatrix}.$$

Note that if $p = 1$, so that $A = A_1$ is a $(\dim V) \times (\dim V)$ matrix, then $(A - \lambda_1 I)^\ell \neq 0$, contradicting that $(z - \lambda_1)^\ell$ is the minimal polynomial of T. Thus $p \geq 2$. Suppose that A_1 is a $d \times d$ matrix; since $p \geq 2$ we must have $d < \dim V$. Let

$$U = \operatorname{span}(v_1, \ldots, v_d) \quad \text{and} \quad W = \operatorname{span}(v_{d+1}, \ldots, v_{\dim V}).$$

Then U and W are non-zero, invariant under T (since A is block diagonal), and decompose V as the direct sum $V = U \oplus W$.

8.D. Trace: A Connection Between Matrices and Operators

Exercise 8.D.1. Suppose V is an inner product space and $v, w \in V$. Define an operator $T \in \mathcal{L}(V)$ by $Tu = \langle u, v\rangle w$. Find a formula for $\operatorname{tr} T$.

Solution. Let $e_1, \dots, e_n$ be an orthonormal basis of V and suppose that

$$v = a_1 e_1 + \cdots + a_n e_n \quad \text{and} \quad w = b_1 e_1 + \cdots + b_n e_n.$$

Observe that $\langle Te_k, e_k\rangle = \langle \langle e_k, v\rangle w, e_k\rangle = \overline{a_k} b_k$ for each $k \in \{1, \dots, n\}$. It follows from 8.55 that

$$\operatorname{tr} T = \sum_{k=1}^{n} \overline{a_k} b_k = \langle w, v\rangle.$$

Exercise 8.D.2. Suppose $P \in \mathcal{L}(V)$ satisfies $P^2 = P$. Prove that

$$\operatorname{tr} P = \dim \operatorname{range} P.$$

Solution. Since $P(P - 1) = 0$, the minimal polynomial of P is either z, $z - 1$, or $z(z - 1)$. In any case, the minimal polynomial of P splits into distinct linear factors and hence P is diagonalizable by 5.62. Thus there is a basis of V with respect to which the matrix $A := \mathcal{M}(T)$ is diagonal. Since the only possible eigenvalues of P are 0 or 1, each diagonal entry of A is either 0 or 1. It follows that $\operatorname{tr} P$ is the number of diagonal entries of A equal to 1; denoting this number by m, it is clear from the form of A that $\dim \operatorname{range} P = \operatorname{rank} A = m$.

Exercise 8.D.3. Suppose $T \in \mathcal{L}(V)$ and $T^5 = T$. Prove that the real and imaginary parts of $\operatorname{tr} T$ are both integers.

Solution. First suppose that $\mathbf{F} = \mathbf{C}$. The equation $T^5 = T$ is equivalent to

$$T(T - I)(T + I)(T - iI)(T + iI) = 0.$$

Thus the eigenvalues of T are contained in the set $\{0, \pm 1, \pm i\}$. It then follows from 8.52 that the real and imaginary parts of $\operatorname{tr} T$ are both integers.

Before we proceed, let us prove the following lemma.

Lemma L.16. Let V be a real vector space and suppose $T \in \mathcal{L}(V)$. Then $\operatorname{tr} T_{\mathbf{C}} = \operatorname{tr} T$.

Proof. Let $v_1, \dots, v_n$ be a basis of V. As we showed in Exercise 2.B.11, $v_1, \dots, v_n$ is also a basis of $V_{\mathbf{C}}$. Observe that $T_{\mathbf{C}} v_k = Tv_k$ for each $k \in \{1, \dots, n\}$. Thus

$$\mathcal{M}(T_{\mathbf{C}}, (v_1, \dots, v_n)) = \mathcal{M}(T, (v_1, \dots, v_n)).$$

It is now immediate that $\operatorname{tr} T_{\mathbf{C}} = \operatorname{tr} T$. $\square$

Now suppose that $\mathbf{F} = \mathbf{R}$. Since $T^5 = T$ we also have $T_{\mathbf{C}}^5 = T_{\mathbf{C}}$ and it follows from our previous discussion that the real and imaginary parts of $\operatorname{tr} T_{\mathbf{C}}$ are integers. Thus, by Lemma L.16, $\operatorname{tr} T$ is an integer.

Exercise 8.D.4. Suppose V is an inner product space and $T \in \mathcal{L}(V)$. Prove that

$$\operatorname{tr} T^* = \overline{\operatorname{tr} T}.$$

Solution. Let $e_1, \dots, e_n$ be an orthonormal basis of V and let A be the matrix of T with respect to $e_1, \dots, e_n$. It follows from 7.9 that the matrix of T^* with respect to $e_1, \dots, e_n$ is A^*. Since the diagonal entries of A^* are the complex conjugates of the diagonal entries of A, we obtain the equation $\operatorname{tr} T^* = \overline{\operatorname{tr} T}$.

Exercise 8.D.5. Suppose V is an inner product space. Suppose $T \in \mathcal{L}(V)$ is a positive operator and $\operatorname{tr} T = 0$. Prove that $T = 0$.

Solution. By 7.38(c) there is an orthonormal basis $e_1, \dots, e_n$ of V with respect to which the matrix of T is diagonal with only non-negative numbers on the diagonal. Since $\operatorname{tr} T = 0$ the sum of these non-negative diagonal entries is zero, which can be the case only if each diagonal entry is zero. Thus the matrix of T with respect to $e_1, \dots, e_n$ is the zero matrix. Hence $T = 0$.

Exercise 8.D.6. Suppose V is an inner product space and $P, Q \in \mathcal{L}(V)$ are orthogonal projections. Prove that $\operatorname{tr}(PQ) \geq 0$.

Solution. Suppose that P and Q are orthogonal projections onto subspaces U and W respectively. Let $e_1, \dots, e_m$ be an orthonormal basis of W and let $f_1, \dots, f_n$ be an orthonormal basis of $W^\perp$, so that $e_1, \dots, e_m, f_1, \dots, f_n$ is an orthonormal basis of V. It follows from 8.55 that

$$\begin{aligned}\operatorname{tr}(PQ) &= \langle PQe_1, e_1\rangle + \cdots + \langle PQe_m, e_m\rangle + \langle PQf_1, f_1\rangle + \cdots + \langle PQf_n, f_n\rangle \\ &= \langle Pe_1, e_1\rangle + \cdots + \langle Pe_m, e_m\rangle,\end{aligned}$$

where we have used that $Q|_W = I$ and $Q|_{W^\perp} = 0$. As noted in 7.35(b), orthogonal projections are positive operators. It follows that

$$\operatorname{tr}(PQ) = \langle Pe_1, e_1\rangle + \cdots + \langle Pe_m, e_m\rangle \geq 0.$$

Exercise 8.D.7. Suppose $T \in \mathcal{L}(\mathbf{C}^3)$ is the operator whose matrix is

$$\begin{pmatrix} 51 & -12 & -21 \\ 60 & -40 & -28 \\ 57 & -68 & 1 \end{pmatrix}.$$

Someone tells you (accurately) that -48 and 24 are eigenvalues of T. Without using a computer or writing anything down, find the third eigenvalue of T.

Solution. On one hand, the trace of the matrix above is $51 - 40 + 1 = 12$. On the other hand, by 8.52, the trace of this matrix is the sum of the eigenvalues of T. Letting x be the third eigenvalue of T, it follows that

$$12 = -48 + 24 + x \quad \Rightarrow \quad x = 36.$$

Exercise 8.D.8. Prove or give a counterexample: If $S, T \in \mathcal{L}(V)$, then $\operatorname{tr}(ST) = (\operatorname{tr} S)(\operatorname{tr} T)$.

Solution. This is false. For a counterexample, let S and T be the operators on $\mathbf{F}^2$ whose matrices with respect to the standard basis are

$$\mathcal{M}(S) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad \mathcal{M}(T) = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \quad \Rightarrow \quad \mathcal{M}(ST) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Then $\operatorname{tr} S = \operatorname{tr} T = 1$ but $\operatorname{tr}(ST) = 0$. Thus $\operatorname{tr}(ST) \neq (\operatorname{tr} S)(\operatorname{tr} T)$.

Exercise 8.D.9. Suppose $T \in \mathcal{L}(V)$ is such that $\operatorname{tr}(ST) = 0$ for all $S \in \mathcal{L}(V)$. Prove that $T = 0$.

Solution. We will prove the contrapositive. Suppose that $T \neq 0$, so that there exists some $v_1 \in V$ such that $Tv_1 \neq 0$. Extend v_1 to a basis $v_1, \ldots, v_n$ of V. Suppose that

$$Tv_1 = A_{1,1}v_1 + \cdots + A_{n,1}v_n.$$

Since $Tv_1 \neq 0$ there must exist some $i \in \{1, \ldots, n\}$ such that $A_{i,1} \neq 0$. Define $S \in \mathcal{L}(V)$ by $Sv_i = v_1$ and $Sv_k = 0$ for $k \neq i$. Now observe that the matrix of ST with respect to $v_1, \ldots, v_n$ is

$$\begin{pmatrix} A_{i,1} & A_{i,2} & \cdots & A_{i,n} \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}.$$

Thus $\operatorname{tr}(ST) = A_{i,1} \neq 0$.

Exercise 8.D.10. Prove that the trace is the only linear functional $\tau : \mathcal{L}(V) \to \mathbf{F}$ such that

$$\tau(ST) = \tau(TS)$$

for all $S, T \in \mathcal{L}(V)$ and $\tau(I) = \dim V$.

Hint: Suppose that $v_1, \ldots, v_n$ is a basis of V. For $j, k \in \{1, \ldots, n\}$, define $P_{j,k} \in \mathcal{L}(V)$ by $P_{j,k}(a_1 v_1 + \cdots + a_n v_n) = a_k v_j$. Prove that

$$\tau(P_{j,k}) = \begin{cases} 1 & \text{if } j = k, \\ 0 & \text{if } j \neq k. \end{cases}$$

Then for $T \in \mathcal{L}(V)$, use the equation $T = \sum_{k=1}^{n} \sum_{j=1}^{n} \mathcal{M}(T)_{j,k} P_{j,k}$ to show that $\tau(T) = \operatorname{tr} T$.

Solution. Let $v_1, \ldots, v_n$ be a basis of V and for $j, k \in \{1, \ldots, n\}$ define $P_{j,k} \in \mathcal{L}(V)$ as in the hint. For any $j, k \in \{1, \ldots, n\}$ observe that

$$P_{j,k} P_{k,j}(a_1 v_1 + \cdots + a_n v_n) = P_{j,k}(a_j v_k) = a_j v_j = P_{j,j}(a_1 v_1 + \cdots + a_n v_n).$$

Thus $P_{j,k} P_{k,j} = P_{j,j}$. It follows that

$$\tau(P_{j,j}) = \tau(P_{j,k} P_{k,j}) = \tau(P_{k,j} P_{j,k}) = \tau(P_{k,k}),$$

which implies

$$n = \tau(I) = \tau(P_{1,1} + \cdots + P_{n,n}) = \tau(P_{1,1}) + \cdots + \tau(P_{n,n}) = n\tau(P_{1,1}).$$

Thus $\tau(P_{1,1}) = \cdots = \tau(P_{n,n}) = 1$. Now observe that, for $j \neq k$,

$$P_{k,k} P_{j,k}(a_1 v_1 + \cdots + a_n v_n) = P_{k,k}(a_k v_j) = 0,$$

$$\text{and} \quad P_{j,k} P_{k,k}(a_1 v_1 + \cdots + a_n v_n) = P_{j,k}(a_k v_k) = a_k v_j = P_{j,k}(a_1 v_1 + \cdots + a_n v_n).$$

Thus $P_{k,k} P_{j,k} = 0$ and $P_{j,k} P_{k,k} = P_{j,k}$. It follows that

$$\tau(P_{j,k}) = \tau(P_{j,k} P_{k,k}) = \tau(P_{k,k} P_{j,k}) = \tau(0) = 0.$$

We have now shown that

$$\tau(P_{j,k}) = \begin{cases} 1 & \text{if } j = k, \\ 0 & \text{if } j \neq k. \end{cases}$$

For any $T \in \mathcal{L}(V)$, it follows that

$$\tau(T) = \tau\left(\sum_{k=1}^{n} \sum_{j=1}^{n} \mathcal{M}(T)_{j,k} P_{j,k} \right) = \sum_{k=1}^{n} \sum_{j=1}^{n} \mathcal{M}(T)_{j,k} \tau(P_{j,k}) = \sum_{k=1}^{n} \mathcal{M}(T)_{k,k} = \operatorname{tr} T.$$

Exercise 8.D.11. Suppose V and W are inner product spaces and $T \in \mathcal{L}(V, W)$. Prove that if $e_1, ..., e_n$ is an orthonormal basis of V and $f_1, ..., f_m$ is an orthonormal basis of W, then

$$\operatorname{tr}(T^*T) = \sum_{k=1}^{n} \sum_{j=1}^{m} \left|\langle Te_k, f_j \rangle\right|^2.$$

The numbers $\langle Te_k, f_j \rangle$ are the entries of the matrix of T with respect to the orthonormal bases $e_1, ..., e_n$ and $f_1, ..., f_m$. These numbers depend on the bases, but $\operatorname{tr}(T^*T)$ does not depend on a choice of bases. Thus this exercise shows that the sum of the squares of the absolute values of the matrix entries does not depend on which orthonormal bases are used.

Solution. The matrix of T with respect to $e_1, ..., e_n$ and $f_1, ..., f_m$ is

$$\begin{pmatrix} \langle Te_1, f_1 \rangle & \langle Te_2, f_1 \rangle & \cdots & \langle Te_n, f_1 \rangle \\ \langle Te_1, f_2 \rangle & \langle Te_2, f_2 \rangle & \cdots & \langle Te_n, f_2 \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle Te_1, f_m \rangle & \langle Te_2, f_m \rangle & \cdots & \langle Te_n, f_m \rangle \end{pmatrix}$$

and thus by 7.9 the matrix of T^* with respect to $f_1, ..., f_m$ and $e_1, ..., e_n$ is

$$\begin{pmatrix} \overline{\langle Te_1, f_1 \rangle} & \overline{\langle Te_1, f_2 \rangle} & \cdots & \overline{\langle Te_1, f_m \rangle} \\ \overline{\langle Te_2, f_1 \rangle} & \overline{\langle Te_2, f_2 \rangle} & \cdots & \overline{\langle Te_2, f_m \rangle} \\ \vdots & \vdots & \ddots & \vdots \\ \overline{\langle Te_n, f_1 \rangle} & \overline{\langle Te_n, f_2 \rangle} & \cdots & \overline{\langle Te_n, f_m \rangle} \end{pmatrix}.$$

It follows that the k^{th} diagonal entry of the matrix of T^*T with respect to $e_1, ..., e_n$ is

$$\sum_{j=1}^{m} \left|\langle Te_k, f_j \rangle\right|^2,$$

from which we obtain the desired formula.

Exercise 8.D.12. Suppose V and W are finite-dimensional inner product spaces.

(a) Prove that $\langle S, T\rangle = \operatorname{tr}(T^*S)$ defines an inner product on $\mathcal{L}(V, W)$.

(b) Suppose $e_1, \dots, e_n$ is an orthonormal basis of V and $f_1, \dots, f_m$ is an orthonormal basis of W. Show that the inner product on $\mathcal{L}(V, W)$ from (a) is the same as the standard inner product on $\mathbf{F}^{mn}$, where we identify each element of $\mathcal{L}(V, W)$ with its matrix (with respect to the bases just mentioned) and then with an element of $\mathbf{F}^{mn}$.

*Caution: The norm of a linear map $T \in \mathcal{L}(V, W)$ as defined by 7.86 is not the same as the norm that comes from the inner product in (a) above. Unless explicitly stated otherwise, always assume that $\|T\|$ refers to the norm as defined by 7.86. The norm that comes from the inner product in (a) is called the **Frobenius norm** or the **Hilbert-Schmidt norm**.*

Solution.

(a) We shall verify each property in definition 6.2.

Positivity. For any $T \in \mathcal{L}(V)$, Exercise 8.D.11 shows that $\langle T, T\rangle = \operatorname{tr}(T^*T)$ is non-negative.

Definiteness. Certainly $\langle 0, 0\rangle = 0$. Suppose that $T \in \mathcal{L}(V)$ is such that $\langle T, T\rangle = 0$, i.e. $\operatorname{tr}(T^*T) = 0$. Since T^*T is a positive operator, Exercise 8.D.5 shows that $T^*T = 0$ and Exercise 7.A.2 then gives us $T = 0$.

Additivity/homogeneity in first slot. These properties follow from the linearity of the trace (see 8.56).

Conjugate symmetry. Let $S, T \in \mathcal{L}(V)$ be given and observe that

$$\langle S, T\rangle = \operatorname{tr}(T^*S) = \operatorname{tr}\big((S^*T)^*\big) = \overline{\operatorname{tr}(S^*T)} = \overline{\langle T, S\rangle},$$

where we have used Exercise 8.D.4 for the third equality.

(b) As we showed in Exercise 8.D.11, the matrix of T^* with respect to $f_1, \dots, f_m$ and $e_1, \dots, e_n$ is

$$\begin{pmatrix} \overline{\langle Te_1, f_1\rangle} & \overline{\langle Te_1, f_2\rangle} & \cdots & \overline{\langle Te_1, f_m\rangle} \\ \overline{\langle Te_2, f_1\rangle} & \overline{\langle Te_2, f_2\rangle} & \cdots & \overline{\langle Te_2, f_m\rangle} \\ \vdots & \vdots & \ddots & \vdots \\ \overline{\langle Te_n, f_1\rangle} & \overline{\langle Te_n, f_2\rangle} & \cdots & \overline{\langle Te_n, f_m\rangle} \end{pmatrix}.$$

The matrix of S with respect to $e_1, \dots, e_n$ and $f_1, \dots, f_m$ is

$$\begin{pmatrix} \langle Se_1, f_1 \rangle & \langle Se_2, f_1 \rangle & \cdots & \langle Se_n, f_1 \rangle \\ \langle Se_1, f_2 \rangle & \langle Se_2, f_2 \rangle & \cdots & \langle Se_n, f_2 \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle Se_1, f_m \rangle & \langle Se_2, f_m \rangle & \cdots & \langle Se_n, f_m \rangle \end{pmatrix}.$$

It follows that the k^{th} diagonal entry of the matrix of T^*S with respect to $e_1, ..., e_n$ is $\sum_{j=1}^{m} \langle Se_k, f_j \rangle \overline{\langle Te_k, f_j \rangle}$, from which we obtain the desired formula

$$\langle S, T \rangle = \operatorname{tr}(T^*S) = \sum_{k=1}^{n} \sum_{j=1}^{m} \langle Se_k, f_j \rangle \overline{\langle Te_k, f_j \rangle}.$$

Exercise 8.D.13. Find $S, T \in \mathcal{L}(\mathcal{P}(\mathbf{F}))$ such that $ST - TS = I$.

Hint: Make an appropriate modification of the operators in Example 3.9.

This exercise shows that additional hypotheses are needed on S and T to extend 8.57 to the setting of infinite-dimensional vector spaces.

Solution. Let $S \in \mathcal{L}(\mathcal{P}(\mathbf{F}))$ be the differentiation operator, i.e. $Sp = p'$, and let $T \in \mathcal{L}(\mathcal{P}(\mathbf{F}))$ be the multiplication by x operator, i.e. $(Tp)(x) = xp(x)$. Observe that

$$((ST - TS)p)(x) = [p(x) + xp'(x)] - xp'(x) = p(x).$$

Thus $ST - TS = I$.

Chapter 9. Multilinear Algebra and Determinants

9.A. Bilinear Forms and Quadratic Forms

Exercise 9.A.1. Prove that if β is a bilinear form on $\mathbf{F}$, then there exists $c \in \mathbf{F}$ such that

$$\beta(x, y) = cxy$$

for all $x, y \in \mathbf{F}$.

Solution. Using the linearity of β in each slot, we have

$$\beta(x, y) = \beta(x \cdot 1, y \cdot 1) = xy\beta(1, 1)$$

for all $x, y \in \mathbf{F}$. Thus we can take $c = \beta(1, 1)$.

Exercise 9.A.2. Let $n = \dim V$. Suppose β is a bilinear form on V. Prove that there exist $\varphi_1, ..., \varphi_n, \tau_1, ..., \tau_n \in V'$ such that

$$\beta(u, v) = \varphi_1(u) \cdot \tau_1(v) + \cdots + \varphi_n(u) \cdot \tau_n(v)$$

for all $u, v \in V$.

This exercise shows that if $n = \dim V$, then every bilinear form on V is of the form given by the last bullet point of Example 9.2.

Solution. Let $e_1, ..., e_n$ be a basis of V and for $j \in \{1, ..., n\}$ define $\varphi_j, \tau_j \in V'$ by

$$\varphi_j\left(\sum_{k=1}^{n} a_k e_k\right) = a_j \quad \text{and} \quad \tau_j\left(\sum_{k=1}^{n} b_k e_k\right) = \sum_{k=1}^{n} b_k \beta(e_j, e_k).$$

For any $u = \sum_{k=1}^{n} a_k e_k, v = \sum_{k=1}^{n} b_k e_k \in V$, observe that

$$\sum_{j=1}^{n} \varphi_j(u)\tau_j(v) = \sum_{j=1}^{n}\sum_{k=1}^{n} a_j b_k \beta(e_j, e_k) = \beta\left(\sum_{j=1}^{n} a_j e_j, \sum_{k=1}^{n} b_k e_k\right) = \beta(u, v).$$

Thus $\varphi_1, ..., \varphi_n, \tau_1, ..., \tau_n$ are the desired linear functionals.

Exercise 9.A.3. Suppose $\beta : V \times V \to \mathbf{F}$ is a bilinear form on V and also is a linear functional on $V \times V$. Prove that $\beta = 0$.

Solution. Let $u, v \in V$ be given. On one hand, since β is a linear functional on $V \times V$ we must have $2\beta(u, v) = \beta(2u, 2v)$. On the other hand, since β is a bilinear form on V we must have $2\beta(u, v) = \beta(u, 2v)$. Thus

$$0 = \beta(2u, 2v) - \beta(u, 2v) = \beta(u, 0),$$

where we have used that $\beta \in (V \times V)'$ for the last equality. Similarly, we can show that $\beta(0, v) = 0$. The linearity of β as a map $V \times V \to \mathbf{F}$ then implies that

$$0 = \beta(u, 0) + \beta(0, v) = \beta(u, v).$$

Thus $\beta = 0$.

Exercise 9.A.4. Suppose V is a real inner product space and β is a bilinear form on V. Show that there exists a unique operator $T \in \mathcal{L}(V)$ such that

$$\beta(u, v) = \langle u, Tv \rangle$$

for all $u, v \in V$.

This exercise states that if V is a real inner product space, then every bilinear form on V is of the form given by the third bullet point in 9.2.

Solution. Let $e_1, \ldots, e_n$ be an orthonormal basis of V and define $T \in \mathcal{L}(V)$ by

$$Te_k = \beta(e_1, e_k)e_1 + \cdots + \beta(e_n, e_k)e_n$$

for $k \in \{1, \ldots, n\}$. Observe that $\langle e_j, Te_k \rangle = \beta(e_j, e_k)$ for any $j, k \in \{1, \ldots, n\}$. It follows that, for any $u = \sum_{k=1}^{n} a_k e_k, v = \sum_{k=1}^{n} b_k e_k \in V$,

$$\langle u, Tv \rangle = \sum_{j=1}^{n} \sum_{k=1}^{n} a_j b_k \langle e_j, Te_k \rangle = \sum_{j=1}^{n} \sum_{k=1}^{n} a_j b_k \beta(e_j, e_k) = \beta(u, v).$$

Thus T is the desired operator.

Exercise 9.A.5. Suppose β is a bilinear form on a real inner product space V and T is the unique operator on V such that $\beta(u, v) = \langle u, Tv \rangle$ for all $u, v \in V$ (see Exercise 4). Show that β is an inner product on V if and only if T is an invertible positive operator on V.

Solution. First suppose that T is an invertible positive operator. To show that β is an inner product on V, we must verify each property in definition 6.2.

Positivity/definiteness. 7.61 shows that $\beta(v, v) = \langle v, Tv \rangle = \langle Tv, v \rangle > 0$ for all non-zero $v \in V$ and certainly $\beta(0, 0) = 0$.

Additivity/homogeneity in first slot/symmetry. These properties are immediate from the corresponding properties of $\langle \cdot, \cdot \rangle$.

Now suppose that β is an inner product on V. The symmetry of β and $\langle \cdot, \cdot \rangle$ gives us

$$\langle Tu, v\rangle = \langle v, Tu\rangle = \beta(v, u) = \beta(u, v) = \langle u, Tv\rangle.$$

Thus T is self-adjoint. Now let $v \in V$ be non-zero and observe that, by the positive-definiteness of β,

$$\langle Tv, v\rangle = \langle v, Tv\rangle = \beta(v, v) > 0.$$

7.61 allows us to conclude that T is a positive invertible operator.

Exercise 9.A.6. Prove or give a counterexample: If ρ is a symmetric bilinear form on V, then

$$\{v \in V : \rho(v, v) = 0\}$$

is a subspace of V.

Solution. This is false. Let ρ be the symmetric bilinear form on $\mathbf{F}^2$ given by

$$\rho((x_1, x_2), (y_1, y_2)) = x_1 y_1 - x_2 y_2.$$

Observe that

$$\rho((1,1),(1,1)) = \rho((1,-1),(1,-1)) = 0 \neq \rho((2,0),(2,0)).$$

Thus $\{v \in V : \rho(v, v) = 0\}$ is not a subspace of $\mathbf{F}^2$.

Exercise 9.A.7. Explain why the proof of 9.13 (diagonalization of symmetric bilinear form by an orthonormal basis on a real inner product space) fails if the hypothesis that $\mathbf{F} = \mathbf{R}$ is dropped.

Solution. Define B and T as in the proof of 9.13. Note that a symmetric matrix with complex entries need not equal its own conjugate transpose nor commute with its own conjugate transpose, e.g.

$$A = \begin{pmatrix} 1 & i \\ i & -1 \end{pmatrix}$$

satisfies $A^{\mathrm{t}} = A, A \neq A^*$, and $AA^* \neq A^*A$. Thus we cannot necessarily claim that T is self-adjoint or normal and hence cannot necessarily apply the complex spectral theorem.

Exercise 9.A.8. Find formulas for $\dim V^{(2)}_{\text{sym}}$ and $\dim V^{(2)}_{\text{alt}}$ in terms of $\dim V$.

Solution. Let $e_1, \ldots, e_n$ be a basis of V. As 9.5 shows, the map $\beta \mapsto \mathcal{M}(\beta)$ is an isomorphism of $V^{(2)}$ onto $\mathbf{F}^{n,n}$. Furthermore, by 9.12, under this isomorphism the symmetric bilinear forms in $V^{(2)}$ correspond exactly to the symmetric matrices in $\mathbf{F}^{n,n}$. Thus to calculate $\dim V^{(2)}_{\text{sym}}$ it will suffice to find the dimension of the subspace of $\mathbf{F}^{n,n}$ consisting of the symmetric matrices; as we showed in Exercise 7.A.16 (b), this subspace has dimension $n(n+1)/2$. It now follows from 9.17 that

$$\dim V_{\text{alt}}^{(2)} = \dim V^{(2)} - \dim V_{\text{sym}}^{(2)} = n^2 - \frac{n(n+1)}{2} = \frac{n(n-1)}{2}.$$

Exercise 9.A.9. Suppose that n is a positive integer and

$$V = \{p \in \mathcal{P}_n(\mathbf{R}) : p(0) = p(1)\}.$$

Define $\alpha : V \times V \to \mathbf{R}$ by

$$\alpha(p, q) = \int_0^1 pq'.$$

Show that α is an alternating bilinear form on V.

Solution. The bilinearity of α follows from the linearity of differentiation and of integration. For any $p \in V$, observe that

$$\int_0^1 pp' = \frac{1}{2} \int_0^1 \left(p^2\right)' = \frac{1}{2} \left[(p(1))^2 - (p(0))^2\right] = 0.$$

Thus α is alternating.

Exercise 9.A.10. Suppose that n is a positive integer and

$$V = \{p \in \mathcal{P}_n(\mathbf{R}) : p(0) = p(1) \text{ and } p'(0) = p'(1)\}.$$

Define $\rho : V \times V \to \mathbf{R}$ by

$$\rho(p, q) = \int_0^1 pq''.$$

Show that ρ is a symmetric bilinear form on V.

Solution. The bilinearity of ρ follows from the linearity of differentiation and of integration. For any $p, q \in V$, observe that

$$\int_0^1 pq'' - qp'' = \int_0^1 [pq']' - [qp']' = [p(1)q'(1) - p(0)q'(0)] - [q(1)p'(1) - q(0)p'(0)] = 0.$$

Thus ρ is symmetric.

9.B. Alternating Multilinear Forms

Exercise 9.B.1. Suppose m is a positive integer. Show that $\dim V^{(m)} = (\dim V)^m$.

Solution. Let $e_1, \ldots, e_n$ be a basis of V and let $\varphi_1, \ldots, \varphi_n$ be the corresponding dual basis of V', so that for any $v \in V$ we have, by 3.114,

$$v = \varphi_1(v)e_1 + \cdots + \varphi_n(v)e_n.$$

For each $(i_1, \ldots, i_m) \in \{1, \ldots, n\}^m$ define an m-linear form $\alpha_{i_1,\ldots,i_m}$ by

$$\alpha_{i_1,\ldots,i_m}(v_1, \ldots, v_m) = \varphi_{i_1}(v_1) \cdots \varphi_{i_m}(v_m).$$

Let $\mathcal{B}$ be the list of all such $\alpha_{i_1,\ldots,i_m}$; we claim that $\mathcal{B}$ is a basis of $V^{(m)}$. For any $\alpha \in V^{(m)}$, observe that

$$\begin{aligned}
\alpha(v_1, \ldots, v_m) &= \alpha\left(\sum_{i_1=1}^{n} \varphi_{i_1}(v_1)e_{i_1}, \ldots, \sum_{i_m=1}^{n} \varphi_{i_m}(v_m)e_{i_m}\right) \\
&= \sum_{i_1=1}^{n} \cdots \sum_{i_m=1}^{n} \alpha\left(e_{i_1}, \ldots, e_{i_m}\right) \varphi_{i_1}(v_1) \cdots \varphi_{i_m}(v_m) \\
&= \sum_{i_1=1}^{n} \cdots \sum_{i_m=1}^{n} \alpha\left(e_{i_1}, \ldots, e_{i_m}\right) \alpha_{i_1,\ldots,i_m}(v_1, \ldots, v_m).
\end{aligned}$$

Thus $\alpha \in \operatorname{span} \mathcal{B}$ and it follows that $V^{(m)} = \operatorname{span} \mathcal{B}$. Now suppose that a linear combination of $\mathcal{B}$ is zero:

$$\sum_{i_1=1}^{n} \cdots \sum_{i_m=1}^{n} c_{i_1,\ldots,i_m} \, \alpha_{i_1,\ldots,i_m}(v_1, \ldots, v_m) = 0 \quad \text{for all } v_1, \ldots, v_m \in V.$$

For any $(k_1, \ldots, k_m) \in \{1, \ldots, n\}^m$, note that

$$\alpha_{i_1,\ldots,i_m}\left(e_{k_1}, \ldots, e_{k_m}\right) = \begin{cases} 1 & \text{if } i_1 = k_1, \ldots, i_m = k_m, \\ 0 & \text{otherwise.} \end{cases}$$

Thus

$$0 = \sum_{i_1=1}^{n} \cdots \sum_{i_m=1}^{n} c_{i_1,\ldots,i_m} \, \alpha_{i_1,\ldots,i_m}\left(e_{k_1}, \ldots, e_{k_m}\right) = c_{k_1,\ldots,k_m}.$$

Hence $\mathcal{B}$ is linearly independent and it follows that $\mathcal{B}$ is a basis of $V^{(m)}$, as claimed.

Now observe that $\mathcal{B}$ is a list of length n^m, since the set $\{1, \ldots, n\}^m$ contains n^m elements. Thus $\dim V^{(m)} = n^m = (\dim V)^m$.

Exercise 9.B.2. Suppose $n \geq 3$ and $\alpha : \mathbf{F}^n \times \mathbf{F}^n \times \mathbf{F}^n \to \mathbf{F}$ is defined by

$$\begin{aligned}\alpha((x_1, ..., x_n), (y_1, ..., y_n), (z_1, ..., z_n)) \\ = x_1 y_2 z_3 - x_2 y_1 z_3 - x_3 y_2 z_1 - x_1 y_3 z_2 + x_3 y_1 z_2 + x_2 y_3 z_1.\end{aligned}$$

Show that α is an alternating 3-linear form on $\mathbf{F}^n$.

Solution. A straightforward calculation shows that α is 3-linear. To see that α is alternating, let $x = (x_1, ..., x_n), y = (y_1, ..., y_n)$, and $z = (z_1, ..., z_n)$. Observe that

$$\alpha(x, x, z) = x_1 x_2 z_3 - x_2 x_1 z_3 - x_3 x_2 z_1 - x_1 x_3 z_2 + x_3 x_1 z_2 + x_2 x_3 z_1 = 0.$$

We can similarly show that $\alpha(x, y, x) = \alpha(x, y, y) = 0$. Thus α is alternating.

Exercise 9.B.3. Suppose m is a positive integer and α is an m-linear form on V such that $\alpha(v_1, ..., v_m) = 0$ whenever $v_1, ..., v_m$ is a list of vectors in V with $v_j = v_{j+1}$ for some $j \in \{1, ..., m-1\}$. Prove that α is an alternating m-linear form on V.

Solution. For any $k \in \{1, ..., m-1\}$, observe that

$$\begin{aligned}0 &= \alpha(v_1, ..., v_k + v_{k+1}, v_k + v_{k+1}, ..., v_m) \\ &= \alpha(v_1, ..., v_k, v_k, ..., v_m) + \alpha(v_1, ..., v_{k+1}, v_{k+1}, ..., v_m) \\ &\qquad + \alpha(v_1, ..., v_k, v_{k+1}, ..., v_m) + \alpha(v_1, ..., v_{k+1}, v_k, ..., v_m) \\ &= \alpha(v_1, ..., v_k, v_{k+1}, ..., v_m) + \alpha(v_1, ..., v_{k+1}, v_k, ..., v_m).\end{aligned}$$

Thus swapping the vectors in any two consecutive slots of $\alpha(v_1, ..., v_m)$ changes the value of α by a factor of -1. Now suppose that $v_1, ..., v_m$ is a list of vectors in V with $v_j = v_k$ for some $1 \leq j < k \leq m$. By performing consecutive swaps in the slots of $\alpha(v_1, ..., v_m)$ if necessary, which only changes the sign of α, we can ensure that v_k appears directly after v_j:

$$\alpha(v_1, ..., v_j, ..., v_k, ..., v_m) = \pm\alpha(v_1, ..., v_j, v_k, ..., v_m) = 0.$$

Thus α is alternating.

Exercise 9.B.4. Prove or give a counterexample: If $\alpha \in V^{(4)}_{\text{alt}}$, then

$$\{(v_1, v_2, v_3, v_4) \in V^4 : \alpha(v_1, v_2, v_3, v_4) = 0\}$$

is a subspace of V^4.

Solution. This is false. For a counterexample, consider $V = \mathbf{R}^4$. By 9.37 there exists a nonzero alternating 4-linear form α on $\mathbf{R}^4$. Let e_1, e_2, e_3, e_4 be the standard basis of $\mathbf{R}^4$ and observe that $\alpha(e_1, e_2, 0, 0) = \alpha(0, 0, e_3, e_4) = 0$. However, $\alpha(e_1, e_2, e_3, e_4) \neq 0$ by 9.39.

Exercise 9.B.5. Suppose m is a positive integer and β is an m-linear form on V. Define an m-linear form α on V by

$$\alpha(v_1, ..., v_m) = \sum_{(j_1,...,j_m) \in \operatorname{perm} m} (\operatorname{sign}(j_1, ..., j_m))\beta\left(v_{j_1}, ..., v_{j_m}\right)$$

for $v_1, ..., v_m \in V$. Explain why $\alpha \in V_{\text{alt}}^{(m)}$.

Solution. Suppose that $v_k = v_\ell$ for some $1 \leq k < \ell \leq m$. We will show that

$$\alpha(v_1, ..., v_k, ..., v_\ell, ..., v_m) = 0.$$

Let $G \subseteq \operatorname{perm} m$ consist of those permutations with k and ℓ in the correct order and let $H = (\operatorname{perm} m) \setminus G$ consist of those permutations with k and ℓ in reverse order. Notice that the map $\Phi : G \to H$ which swaps the position of k and ℓ is a bijection: it is its own inverse. Thus each permutation in H corresponds via Φ to exactly one permutation in G. Notice further that $\operatorname{sign}(\Phi(j_1, ..., j_m)) = -\operatorname{sign}(j_1, ..., j_m)$ by 9.34 and, because $v_k = v_\ell$,

$$\beta\left(v_{j_1}, ..., v_{j_m}\right) = \beta\left(v_{i_1}, ..., v_{i_m}\right) \quad \text{where } (i_1, ..., i_m) = \Phi(j_1, ..., j_m).$$

To summarize, for each $(i_1, ..., i_m) \in H$ we have $(i_1, ..., i_m) = \Phi(j_1, ..., j_m)$ for a unique $(j_1, ..., j_m) \in G$ such that

$$\operatorname{sign}(i_1, ..., i_m) = -\operatorname{sign}(j_1, ..., j_m) \quad \text{and} \quad \beta\left(v_{i_1}, ..., v_{i_m}\right) = \beta\left(v_{j_1}, ..., v_{j_m}\right).$$

It follows that

$$\sum_{(i_1,...,i_m) \in H} (\operatorname{sign}(i_1, ..., i_m))\beta\left(v_{i_1}, ..., v_{i_m}\right) = -\sum_{(j_1,...,j_m) \in G} \operatorname{sign}(j_1, ..., j_m)\beta\left(v_{j_1}, ..., v_{j_m}\right).$$

Thus

$$\begin{aligned}
\alpha(v_1, ..., v_m) &= \sum_{(j_1,...,j_m) \in \operatorname{perm} m} (\operatorname{sign}(j_1, ..., j_m))\beta\left(v_{j_1}, ..., v_{j_m}\right) \\
&= \sum_{(j_1,...,j_m) \in G} \operatorname{sign}(j_1, ..., j_m)\beta\left(v_{j_1}, ..., v_{j_m}\right) \\
&\qquad + \sum_{(i_1,...,i_m) \in H} (\operatorname{sign}(i_1, ..., i_m))\beta\left(v_{i_1}, ..., v_{i_m}\right) \\
&= \sum_{(j_1,...,j_m) \in G} \operatorname{sign}(j_1, ..., j_m)\beta\left(v_{j_1}, ..., v_{j_m}\right) \\
&\qquad - \sum_{(j_1,...,j_m) \in G} \operatorname{sign}(j_1, ..., j_m)\beta\left(v_{j_1}, ..., v_{j_m}\right) \\
&= 0.
\end{aligned}$$

Exercise 9.B.6. Suppose m is a positive integer and β is an m-linear form on V. Define an m-linear form α on V by

$$\alpha(v_1, ..., v_m) = \sum_{(j_1,...,j_m) \in \operatorname{perm} m} \beta\left(v_{j_1}, ..., v_{j_m}\right)$$

for $v_1, ..., v_m \in V$. Explain why

$$\alpha\left(v_{k_1}, ..., v_{k_m}\right) = \alpha(v_1, ..., v_m)$$

for all $v_1, ..., v_m \in V$ and all $(k_1, ..., k_m) \in \operatorname{perm} m$.

Solution. The set of all permutations of $(1, ..., m)$ is exactly the same as the set of all permutations of $(k_1, ..., k_m)$.

Exercise 9.B.7. Give an example of a nonzero alternating 2-linear form α on $\mathbf{R}^3$ and a linearly independent list $v_1, v_2 \in \mathbf{R}^3$ such that $\alpha(v_1, v_2) = 0$.

This exercise shows that 9.39 can fail if the hypothesis that $n = \dim V$ is deleted.

Solution. We can take

$$\alpha((x_1, x_2, x_3), (y_1, y_2, y_3)) = x_1 y_2 - x_2 y_1, \quad v_1 = (1, 0, 0), \quad \text{and} \quad v_2 = (0, 0, 1).$$

9.C. Determinants

Exercise 9.C.1. Prove or give a counterexample:

$$S, T \in \mathcal{L}(V) \quad \Rightarrow \quad \det(S+T) = \det S + \det T.$$

Solution. This is false. Let $S, T \in \mathcal{L}(\mathbf{R}^2)$ be the operators whose matrices with respect to the standard basis are

$$\mathcal{M}(S) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad \mathcal{M}(T) = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \quad \Rightarrow \quad \mathcal{M}(S+T) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

It follows from 9.48 and 9.53 that

$$\det S + \det T = 0 \neq 1 = \det(S+T).$$

Exercise 9.C.2. Suppose the first column of a square matrix A consists of all zeros except possibly the first entry $A_{1,1}$. Let B be the matrix obtained from A by deleting the first row and the first column of A. Show that $\det A = A_{1,1} \det B$.

Solution 1. Suppose that A is an $n \times n$ matrix; we may as well suppose that $n \geq 2$. Let T be the operator on $\mathbf{F}^n$ whose matrix with respect to the standard basis $e_1, \ldots, e_n$ is A^{t}. Let $U = \operatorname{span}(e_2, \ldots, e_n)$ and notice that U is invariant under T since all entries of the first row of A^{t} are zero, except possibly the first entry. Notice further that the matrix of $T|_U$ with respect to $e_2, \ldots, e_n$ is B^{t}. Let α be an alternating n-linear form on $\mathbf{F}^n$ such that $\alpha(e_1, \ldots, e_n) = 1$; see 9.37 for the existence of such an α. If we define β by $\beta(v_2, \ldots, v_n) = \alpha(Te_1, v_2, \ldots, v_n)$, then β is an alternating $(n-1)$-linear form on U. Using 9.56(a), observe that

$$\begin{aligned}
\det A = \det A^{\mathrm{t}} &= \det T \\
&= (\det T)\, \alpha(e_1, e_2, \ldots, e_n) \\
&= \alpha(Te_1, Te_2, \ldots, Te_n) \\
&= \beta(Te_2, \ldots, Te_n) \\
&= (\det T|_U)\, \beta(e_2, \ldots, e_n) \\
&= (\det B^{\mathrm{t}})\, \alpha(Te_1, e_2, \ldots, e_n) \\
&= (\det B)\, \alpha\big(A_{1,1} e_1 + \cdots + A_{1,n} e_n, e_2, \ldots, e_n\big) \\
&= A_{1,1} (\det B)\, \alpha(e_1, \ldots, e_n) \\
&= A_{1,1} \det B.
\end{aligned}$$

Solution 2. Suppose that A is an $n \times n$ matrix; we may as well suppose that $n \geq 2$. Consider the formula for $\det A$ given by 9.46:

$$\det A = \sum_{(j_1,\ldots,j_n)\in \operatorname{perm} n} (\operatorname{sign}(j_1,\ldots,j_n)) A_{j_1,1} \cdots A_{j_n,n}.$$

By assumption we have $A_{j_1,1} = 0$ if $j_1 \neq 1$ and thus

$$\det A = A_{1,1} \sum_{(1,j_2,\ldots,j_n)\in \operatorname{perm} n} (\operatorname{sign}(1,j_2,\ldots,j_n)) A_{j_2,2} \cdots A_{j_n,n}.$$

For $(1, j_2, \ldots, j_n) \in \operatorname{perm} n$ notice that $j_k \geq 2$ for each $k \in \{2, \ldots, n\}$. Thus if we define $i_k = j_{k+1} - 1$ for $k \in \{1, \ldots, n-1\}$, then $(i_1, \ldots, i_{n-1}) \in \operatorname{perm}(n-1)$. Observe that:

- the map sending $(1, j_2, \ldots, j_n) \mapsto (i_1, \ldots, i_{n-1})$ is a bijection between the set of all permutations $(j_1, j_2, \ldots, j_n) \in \operatorname{perm} n$ satisfying $j_1 = 1$ and $\operatorname{perm}(n-1)$;
- $B_{i_k,k} = A_{j_{k+1}-1,k+1}$ for each $k \in \{1, \ldots, n-1\}$;
- $\operatorname{sign}(i_1, \ldots, i_{n-1}) = \operatorname{sign}(1, j_2, \ldots, j_n)$ since 1 is in its natural position in the permutation $(1, j_2, \ldots, j_n)$ and $i_k < i_\ell$ if and only if $j_{k+1} < j_{\ell+1}$.

It follows that

$$\begin{aligned}
\det A &= A_{1,1} \sum_{(1,j_2,\ldots,j_n)\in \operatorname{perm} n} (\operatorname{sign}(1,j_2,\ldots,j_n)) A_{j_2,2} \cdots A_{j_n,n} \\
&= A_{1,1} \sum_{(i_1,\ldots,i_{n-1})\in \operatorname{perm}(n-1)} (\operatorname{sign}(i_1,\ldots,i_{n-1})) B_{i_1,1} \cdots B_{i_{n-1},n-1} = A_{1,1} \det B.
\end{aligned}$$

Exercise 9.C.3. Suppose $T \in \mathcal{L}(V)$ is nilpotent. Prove that $\det(I + T) = 1$.

Solution. By 8.18 there is a basis $e_1, \ldots, e_n$ of V such that the matrix of T with respect to $e_1, \ldots, e_n$ is upper-triangular with each diagonal entry equal to zero. It follows that the matrix of $I + T$ with respect to $e_1, \ldots, e_n$ is upper-triangular with each diagonal entry equal to one. 9.48 and 9.53 allow us to conclude that $\det(I + T) = 1$.

Exercise 9.C.4. Suppose $S \in \mathcal{L}(V)$. Prove that S is unitary if and only if

$$|\det S| = \|S\| = 1.$$

Solution. If S is unitary then $|\det S| = 1$ by 9.58 and $\|S\| = 1$ since $\|Sv\| = \|v\|$ for every $v \in V$.

Conversely, suppose that $|\det S| = \|S\| = 1$. Let $s_1 \geq \cdots \geq s_n \geq 0$ be the singular values of S, so that $s_1 = \|S\| = 1$. Observe that, by 9.60,

$$1 = |\det S| = s_1 \cdots s_n.$$

It follows that $s_1 = \cdots = s_n = 1$, otherwise the right-hand side of the equation above would be strictly less than 1. Thus S is unitary by 7.69.

Exercise 9.C.5. Suppose A is a block upper-triangular matrix

$$A = \begin{pmatrix} A_1 & & * \\ & \ddots & \\ 0 & & A_m \end{pmatrix},$$

where each A_k along the diagonal is a square matrix. Prove that

$$\det A = (\det A_1) \cdots (\det A_m).$$

Solution. It will suffice to prove the case where $m = 2$. A straightforward induction argument will then prove the general case. Suppose therefore that A is a block upper-triangular matrix of the form

$$A = \begin{pmatrix} B & D \\ 0 & C \end{pmatrix},$$

where B is a $k \times k$ matrix and C is an $\ell \times \ell$ matrix, so that A is a $(k+\ell) \times (k+\ell)$ matrix and D is a $k \times \ell$ matrix. Let $e_1, ..., e_{k+\ell}$ be the standard basis of $\mathbf{F}^{k+\ell}$, let

$$U = \operatorname{span}(e_1, ..., e_k), \quad W = \operatorname{span}(e_{k+1}, ..., e_\ell),$$

and let $T \in \mathcal{L}(V)$ and $S \in \mathcal{L}(W)$ be such that

$$\mathcal{M}(T, (e_1, ..., e_{k+\ell})) = A \quad \text{and} \quad \mathcal{M}(S, (e_{k+1}, ..., e_{k+\ell})) = C.$$

Notice that U is invariant under T and $\mathcal{M}(T|_U, (e_1, ..., e_k)) = B$. Let α be a $(k+\ell)$-linear form on $\mathbf{F}^{k+\ell}$ satisfying $\alpha(e_1, ..., e_{k+\ell}) = 1$; see 9.37 for the existence of such an α. If we define $\beta : U \to \mathbf{F}$ and $\gamma : W \to \mathbf{F}$ by

$$\beta(v_1, ..., v_k) = \alpha(v_1, ..., v_k, Te_{k+1}, ..., Te_{k+\ell}),$$
$$\gamma(v_{k+1}, ..., v_{k+\ell}) = \alpha(e_1, ..., e_k, v_{k+1}, ..., v_{k+\ell}),$$

then β is an alternating k-form on U and γ is an alternating ℓ-form on W. Now observe that

$$\begin{aligned} \alpha(e_1, ..., e_k, Te_{k+1}, ..., Te_{k+\ell}) &= \alpha\Bigg(e_1, ..., e_k, \sum_{i_1=1}^{k} D_{i_1,1} e_{i_1} + \sum_{i_1=1}^{\ell} C_{i_1,1} e_{k+i_1}, \\ &\qquad\qquad ..., \sum_{i_\ell=1}^{k} D_{i_\ell,\ell} e_{i_\ell} + \sum_{i_\ell=1}^{\ell} C_{i_\ell,\ell} e_{k+i_\ell}\Bigg) \\ &= \sum_{i_1=1}^{k} \cdots \sum_{i_\ell=1}^{k} D_{i_1,1} \cdots D_{i_\ell,\ell}\, \alpha\big(e_1, ..., e_k, e_{i_1}, ..., e_{i_\ell}\big) \\ &\quad + \sum_{i_1=1}^{k} \cdots \sum_{i_\ell=1}^{k} C_{i_1,1} \cdots C_{i_\ell,\ell}\, \alpha\big(e_1, ..., e_k, e_{k+i_1}, ..., e_{k+i_\ell}\big) \\ &= \sum_{i_1=1}^{k} \cdots \sum_{i_\ell=1}^{k} C_{i_1,1} \cdots C_{i_\ell,\ell}\, \alpha\big(e_1, ..., e_k, e_{k+i_1}, ..., e_{k+i_\ell}\big); \end{aligned}$$

this last line holds since $\alpha\Big(e_1, ..., e_k, e_{i_1}, ..., e_{i_\ell}\Big) = 0$ for any $(i_1, ..., i_\ell) \in \{1, ..., k\}^\ell$ (because α is alternating). It follows that

$$\begin{aligned}
\alpha(e_1, ..., e_k, Te_{k+1}, ..., Te_{k+\ell}) &= \sum_{i_1=1}^{k} \cdots \sum_{i_\ell=1}^{k} C_{i_1,1} \cdots C_{i_\ell,\ell}\, \alpha\Big(e_1, ..., e_k, e_{k+i_1}, ..., e_{k+i_\ell}\Big) \\
&= \alpha\left(e_1, ..., e_k, \sum_{i_1=1}^{\ell} C_{i_1,1} e_{k+i_1}, ..., \sum_{i_\ell=1}^{\ell} C_{i_\ell,\ell} e_{k+i_\ell}\right) \\
&= \alpha(e_1, ..., e_k, Se_{k+1}, ..., Se_{k+\ell}).
\end{aligned}$$

Now observe that

$$\begin{aligned}
\det A &= \det T \\
&= (\det T)\, \alpha(e_1, ..., e_k, e_{k+1}, ..., e_{k+\ell}) \\
&= \alpha(Te_1, ..., Te_k, Te_{k+1}, ..., Te_{k+\ell}) \\
&= \beta((T|_U)e_1, ..., (T|_U)e_k) \\
&= (\det(T|_U))\, \beta(e_1, ..., e_n) \\
&= (\det B)\, \alpha(e_1, ..., e_n, Te_{k+1}, ..., Te_{k+\ell}) \\
&= (\det B)\, \alpha(e_1, ..., e_k, Se_{k+1}, ..., Se_{k+\ell}) \\
&= (\det B)\, \gamma(Se_{k+1}, ..., Se_{k+\ell}) \\
&= (\det B)(\det S)\gamma(e_{k+1}, ..., e_{k+\ell}) \\
&= (\det B)(\det C)\, \alpha(e_1, ..., e_k, e_{k+1}, ..., e_{k+\ell}) \\
&= (\det B)(\det C).
\end{aligned}$$

Exercise 9.C.6. Suppose $A = \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}$ is an n-by-n matrix, with v_k denoting the k^{th} column of A. Show that if $(m_1, ..., m_n) \in \operatorname{perm} n$, then

$$\det\begin{pmatrix} v_{m_1} & \cdots & v_{m_n} \end{pmatrix} = (\operatorname{sign}(m_1, ..., m_n)) \det A.$$

Solution. Swapping pairs of columns of $\begin{pmatrix} v_{m_1} & \cdots & v_{m_n} \end{pmatrix}$ multiplies the determinant by -1 (see 9.57(b)). If N is the number of swaps required until the columns are in the correct order, i.e. $\begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}$, then $\operatorname{sign}(m_1, ..., m_n) = (-1)^N$.

Exercise 9.C.7. Suppose $T \in \mathcal{L}(V)$ is invertible. Let p denote the characteristic polynomial of T and let q denote the characteristic polynomial of T^{-1}. Prove that

$$q(z) = \frac{1}{p(0)} z^{\dim V} p\left(\frac{1}{z}\right)$$

for all nonzero $z \in \mathbf{F}$.

Solution. By definition we have $p(z) = \det(zI - T)$ and $q(z) = \det(zI - T^{-1})$. Now observe that

$$\begin{aligned}
\frac{1}{p(0)} z^{\dim V} p\left(\frac{1}{z}\right) &= \frac{1}{\det(-T)} z^{\dim V} \det\left(\frac{1}{z} I - T\right) \\
&= \det(-T^{-1}) \det(I - zT) \\
&= \det(zI - T^{-1}) \\
&= q(z),
\end{aligned}$$

where the second line holds by 9.50 and the third bullet point of 9.42 and the second line holds by 9.49(a).

Exercise 9.C.8. Suppose $T \in \mathcal{L}(V)$ is an operator with no eigenvalues (which implies that $\mathbf{F} = \mathbf{R}$). Prove that $\det T > 0$.

Solution. Since T has no eigenvalues, it must be the case that $\mathbf{F} = \mathbf{R}$ and $n := \dim V$ is even (by 5.34). Furthermore, the characteristic polynomial p of T must have no real roots. Thus the factorization of p over $\mathbf{C}$ is of the form

$$p(z) = (z - \lambda_1)(z - \overline{\lambda_1}) \cdots (z - \lambda_m)(z - \overline{\lambda_m})$$

for some non-zero complex numbers $\lambda_1, \ldots, \lambda_m$ (see Chapter 4). It follows that the constant term of p is

$$\lambda_1 \overline{\lambda_1} \cdots \lambda_m \overline{\lambda_m} = |\lambda_1|^2 \cdots |\lambda_m|^2 > 0.$$

On the other hand, by 9.65, the constant term of p equals $(-1)^n (\det T) = \det T$. Thus $\det T > 0$.

Exercise 9.C.9. Suppose that V is a real vector space of even dimension, $T \in \mathcal{L}(V)$, and $\det T < 0$. Prove that T has at least two distinct eigenvalues.

Solution. We will prove the contrapositive, i.e. assuming that V is a real vector space of even dimension and $T \in \mathcal{L}(V)$, we will prove that if T has at most one distinct eigenvalue then $\det T \geq 0$. If T has no eigenvalues then Exercise 9.C.8 shows that $\det T > 0$, so suppose that T has exactly one eigenvalue λ and let $\dim V = 2n$ for some positive integer n. It follows that the characteristic polynomial p of T is given by $p(z) = (z - \lambda)^{2n}$. Now observe that

$$\det T = (-1)^{2n} \det T = \det(-T) = p(0) = (-\lambda)^{2n} \geq 0.$$

Exercise 9.C.10. Suppose V is a real vector space of odd dimension and $T \in \mathcal{L}(V)$. Without using the minimal polynomial, prove that T has an eigenvalue.

This result was previously proved without using determinants or the characteristic polynomial—see 5.34.

Solution. The characteristic polynomial p of T is a polynomial of odd degree with real coefficients. It follows from Exercise 4.9 that p has a real zero, i.e. T has an eigenvalue.

Exercise 9.C.11. Prove or give a counterexample: If $\mathbf{F} = \mathbf{R}, T \in \mathcal{L}(V)$, and $\det T > 0$, then T has a square root.

If $\mathbf{F} = \mathbf{C}, T \in \mathcal{L}(V)$ *and* $\det T \neq 0$, *then* T *has a square root (see 8.41).*

Solution. This is false. For a counterexample, let T be the operator on $\mathbf{R}^2$ whose matrix with respect to the standard basis is

$$A = \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix}.$$

Thus $\det T = \det A = 1$ by 9.48. We claim that T has no square root. It will suffice to show that for any matrix

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbf{R}^{2,2}$$

we have $M^2 \neq A$. Indeed, observe that

$$M^2 = \begin{pmatrix} a^2 + bc & b(a+d) \\ c(a+d) & bc + d^2. \end{pmatrix}$$

Equating this to A, we must have $c(a+d) = 0$. If $c = 0$ then $a^2 = -1$, which cannot hold for $a \in \mathbf{R}$, so it must be the case that $a = -d$. But now $b(a+d) = 0 \neq 1$. Thus there does not exist $M \in \mathbf{R}^{2,2}$ satisfying $M^2 = A$.

Exercise 9.C.12. Suppose $S, T \in \mathcal{L}(V)$ and S is invertible. Define $p : \mathbf{F} \to \mathbf{F}$ by

$$p(z) = \det(zS - T).$$

Prove that p is a polynomial of degree $\dim V$ and that the coefficient of $z^{\dim V}$ in this polynomial is $\det S$.

Solution. Let $n = \dim V$ and observe that

$$p(z) = \det(zS - T) = \det(S(zI - S^{-1}T)) = (\det S)(\det(zI - S^{-1}T)) = (\det S)q(z),$$

where q is the characteristic polynomial of $S^{-1}T$. Since $\det S \neq 0$ (because S is invertible) we see that p is a polynomial of degree n. Furthermore, by 9.65, the coefficient of z^n in the polynomial p is $\det S$.

Exercise 9.C.13. Suppose $\mathbf{F} = \mathbf{C}, T \in \mathcal{L}(V)$, and $n = \dim V > 2$. Let $\lambda_1, \ldots, \lambda_n$ denote the eigenvalues of T, with each eigenvalue included as many times as its multiplicity.

(a) Find a formula for the coefficient of z^{n-2} in the characteristic polynomial of T in terms of $\lambda_1, \ldots, \lambda_n$.

(b) Find a formula for the coefficient of z in the characteristic polynomial of T in terms of $\lambda_1, \ldots, \lambda_n$.

Solution.

(a) By 9.62, the characteristic polynomial p of T is given by

$$p(z) = (z - \lambda_1) \cdots (z - \lambda_n).$$

Multiplying p out involves making a binary choice for each factor (either choose z or $-\lambda_k$), summing over all 2^n choices, and collecting like powers. To find the coefficient of z^{n-2}, we should choose z from $n-2$ of the factors; equivalently, we should choose $-\lambda_k$ from 2 of the factors. That is, for each choice of $1 \leq i < j \leq n$ we obtain a contribution of $(-\lambda_i)(-\lambda_j) = \lambda_i \lambda_j$ to the coefficient of z^{n-2}. Summing over all such choices, we see that the coefficient of z^{n-2} in p is

$$\sum_{1 \leq i < j \leq n} \lambda_i \lambda_j.$$

(b) As in part (a), to find the coefficient of z we should choose z from exactly one of the factors, which is equivalent to choosing $-\lambda_k$ from $n-1$ of the factors. In other words, we should omit exactly one of the λ_k's, i.e. for each choice of $1 \leq k \leq n$ we obtain a contribution of

$$(-\lambda_1) \cdots (-\lambda_{k-1})(-\lambda_{k+1}) \cdots (-\lambda_n) = (-1)^{n-1}(\lambda_1 \cdots \lambda_{k-1}\lambda_{k+1} \cdots \lambda_n)$$

to the coefficient of z. If we write $\lambda_1 \cdots \widehat{\lambda_k} \cdots \lambda_n$ to mean the product of all the eigenvalues $\lambda_1, \ldots, \lambda_n$ except for λ_k, then the coefficient of z in p is

$$(-1)^{n-1} \sum_{k=1}^{n} \lambda_1 \cdots \widehat{\lambda_k} \cdots \lambda_n.$$

Exercise 9.C.14. Suppose V is an inner product space and T is a positive operator on V. Prove that

$$\det \sqrt{T} = \sqrt{\det T}.$$

Solution. By 7.38(c) there is an orthonormal basis $e_1, \ldots, e_n$ of V with respect to which the matrix of T is of the form

$$\begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{pmatrix}$$

for some non-negative real numbers $\lambda_1, \ldots, \lambda_n$. It is then clear that the matrix of $\sqrt{T}$ with respect to $e_1, \ldots, e_n$ is

$$\begin{pmatrix} \sqrt{\lambda_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sqrt{\lambda_n} \end{pmatrix}.$$

It follows from 9.48 that

$$\det \sqrt{T} = \sqrt{\lambda_1} \cdots \sqrt{\lambda_n} = \sqrt{\lambda_1 \cdots \lambda_n} = \sqrt{\det T}.$$

Exercise 9.C.15. Suppose V is an inner product space and $T \in \mathcal{L}(V)$. Use the polar decomposition to give a proof that

$$|\det T| = \sqrt{\det(T^* T)}$$

that is different from the proof given earlier (see 9.60).

Solution. By the polar decomposition (7.93) there exists a unitary operator $S \in \mathcal{L}(V)$ such that $T = S\sqrt{T^* T}$. Now observe that

$$|\det T| = \left|\det\left(S\sqrt{T^* T}\right)\right| = \left|(\det S)\left(\det\left(\sqrt{T^* T}\right)\right)\right| = |\det S|\left|\sqrt{\det(T^* T)}\right| = \sqrt{\det(T^* T)},$$

where the second equality follows from 9.49, the third equality follows from Exercise 9.C.14, and the fourth equality follows from 9.58 and 9.59.

Exercise 9.C.16. Suppose $T \in \mathcal{L}(V)$. Define $g : \mathbf{F} \to \mathbf{F}$ by $g(x) = \det(I + xT)$. Show that $g'(0) = \operatorname{tr} T$.

Look for a clean solution to this exercise, without using the explicit but complicated formula for the determinant of a matrix.

Solution. Let $n = \dim V$ and let p be the characteristic polynomial of $-T$. For $x \neq 0$, observe that

$$g(x) = \det(I + xT) = x^n \det\left(\frac{1}{x}I + T\right) = x^n p\left(\frac{1}{x}\right) = 1 + (\operatorname{tr} T)x + \cdots + (\det T)x^n,$$

where we have used 9.65 for the last equality. Since $g(0) = \det I = 1$, we see that this formula holds for all $x \in \mathbf{F}$. Thus $g'(0) = \operatorname{tr} T$.

Exercise 9.C.17. Suppose a, b, c are positive numbers. Find the volume of the ellipsoid

$$\left\{(x, y, z) \in \mathbf{R}^3 : \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} < 1\right\}$$

by finding a set $\Omega \subseteq \mathbf{R}^3$ whose volume you know and an operator T on $\mathbf{R}^3$ such that $T(\Omega)$ equals the ellipsoid above.

Solution. Let e_1, e_2, e_3 be the standard orthonormal basis of $\mathbf{R}^3$, let $\Omega = \{v \in \mathbf{R}^3 : \|v\| < 1\}$, and define $T \in \mathcal{L}(\mathbf{R}^3)$ by $Te_1 = ae_1, Te_2 = be_2$, and $Te_3 = ce_3$. Then as the proof of 7.99 shows,

$$T(\Omega) = E(ae_1, be_2, ce_3) = \left\{(x, y, z) \in \mathbf{R}^3 : \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} < 1\right\}.$$

Thus the volume of the ellipsoid in question is given by 9.61:

$$\operatorname{volume} T(\Omega) = |\det T|(\operatorname{volume} \Omega) = \frac{4abc\pi}{3}.$$

Exercise 9.C.18. Suppose that A is an invertible square matrix. Prove that Hadamard's inequality (9.66) is an equality if and only if each column of A is orthogonal to the other columns.

Solution. Let $v_1, \ldots, v_n$ be the columns of A, let $e_1, \ldots, e_n$ be the result of applying the Gram-Schmidt procedure to $v_1, \ldots, v_n$, and let $A = QR$ be the QR factorization (see 7.58) of A, i.e. Q is the unitary matrix whose columns are $e_1, \ldots, e_n$ and R is the upper-triangular matrix with positive diagonal entries whose $(j, k)^{\text{th}}$ entry is $R_{j,k} = \langle v_k, e_j \rangle$. By studying the Gram-Schmidt procedure, we see that the columns of A are orthogonal to each other if and only if the list $e_1, \ldots, e_n$ is given by

$$e_1 = \frac{v_1}{\|v_1\|}, \quad \ldots, \quad e_n = \frac{v_n}{\|v_n\|}. \tag{1}$$

Furthermore, the proof of Hadamard's inequality (9.66) shows that the inequality is an equality if and only if

$$R_{1,1} \cdots R_{n,n} = \|R_{\cdot,1}\| \cdots \|R_{\cdot,n}\|. \tag{2}$$

Suppose that the columns of A are orthogonal to each other. It follows from (1) and the definition of R that R is a diagonal matrix whose k^{th} diagonal entry is $\|v_k\|$, from which it follows that both sides of (2) are equal to $\|v_1\| \cdots \|v_n\|$. Thus Hadamard's inequality is an equality.

Now suppose that Hadamard's inequality is an equality, so that (2) holds. It follows that each inequality $R_{k,k} \leq \|R_{\cdot,k}\|$ must be an equality, otherwise the left-hand side of (2) would be strictly less than the right-hand side. Thus each off-diagonal entry of R must be zero, i.e. $\langle v_k, e_j \rangle = 0$ for $j \neq k$. It follows that at the k^{th} stage of the Gram-Schmidt procedure we set $e_k = \|v_k\|^{-1} v_k$, so that (1) holds. Thus the columns of A are orthogonal to each other.

Exercise 9.C.19. Suppose V is an inner product space, $e_1, \dots, e_n$ is an orthonormal basis of V, and $T \in \mathcal{L}(V)$ is a positive operator.

(a) Prove that $\det T \leq \prod_{k=1}^{n} \langle Te_k, e_k \rangle$.

(b) Prove that if T is invertible, then the inequality in (a) is an equality if and only if e_k is an eigenvector of T for each $k = 1, \dots, n$.

Solution.

(a) Let B be the matrix of $\sqrt{T}$ with respect to $e_1, \dots, e_n$ and suppose the columns of B are $v_1, \dots, v_n$, so that $\|v_k\| = \|\sqrt{T} e_k\|$ for each $k \in \{1, \dots, n\}$ (the norm on the left-hand side is the usual norm on $\mathbf{F}^n$ and the norm on the right-hand side is the norm on V). Note that $\det T$ and $\det \sqrt{T}$ are non-negative since T and $\sqrt{T}$ are positive operators (see 9.59). It follows from Hadamard's inequality (9.66) that

$$\det \sqrt{T} = \det B \leq \prod_{k=1}^{n} \|v_k\| = \prod_{k=1}^{n} \left\| \sqrt{T} e_k \right\| = \prod_{k=1}^{n} \sqrt{\langle Te_k, e_k \rangle}.$$

Thus

$$\det T = \left(\sqrt{\det T} \right)^2 = \left(\det \sqrt{T} \right)^2 \leq \prod_{k=1}^{n} \langle Te_k, e_k \rangle,$$

where we have used Exercise 9.C.14 for the second equality.

(b) Let $\lambda_1, \dots, \lambda_n$ be the eigenvalues of T; since T is a positive operator, these are also the singular values of T (once sorted into decreasing order; see Exercise 7.E.7). If each e_k is an eigenvector of T then, by 9.60, both sides of the inequality in (a) equal $\lambda_1 \cdots \lambda_n$.

Now suppose that the inequality in (a) is an equality. Since we only used Hadamard's inequality to derive the inequality in (a) and T being invertible implies $\sqrt{T}$ is invertible, Exercise 9.C.19 shows that the columns of B (as defined in (a)) must be orthogonal to each other. It follows that $B^* B$ is a diagonal matrix, since its $(j, k)^{\text{th}}$ entry is $\langle v_j, v_k \rangle$. Now observe that, by the self-adjointness of $\sqrt{T}$,

$$T = \left(\sqrt{T} \right)^2 = \left(\sqrt{T} \right)^* \left(\sqrt{T} \right) \quad \Rightarrow \quad \mathcal{M}(T, (e_1, \dots, e_n)) = B^* B.$$

Thus the matrix of T with respect to $e_1, \dots, e_n$ is diagonal, i.e. each e_k is an eigenvector of T.

Exercise 9.C.20. Suppose A is an n-by-n matrix, and suppose c is such that $|A_{j,k}| \leq c$ for all $j, k \in \{1, \dots, n\}$. Prove that

$$|\det A| \leq c^n n^{n/2}.$$

The formula for the determinant of a matrix (9.46) shows that $|\det A| \leq c^n n!$. *However, the estimate given by this exercise is much better. For example, if* $c = 1$ *and* $n = 100$, *then* $c^n n! \approx 10^{158}$, *but the estimate given by this exercise is the much smaller number* 10^{100}. *If* n *is an integer power of* 2, *then the inequality above is sharp and cannot be improved.*

Solution. Suppose the columns of A are $v_1, \dots, v_n$. For any $k \in \{1, \dots, n\}$, observe that

$$\|v_k\|^2 = \sum_{j=1}^{n} |A_{j,k}|^2 \leq c^2 n \quad \Rightarrow \quad \|v_k\| \leq c n^{1/2}.$$

It follows from this inequality and Hadamard's inequality (9.66) that

$$|\det A| \leq \prod_{k=1}^{n} \|v_k\| \leq c^n n^{n/2}.$$

Exercise 9.C.21. Suppose n is a positive integer and $\delta : \mathbf{C}^{n,n} \to \mathbf{C}$ is a function such that

$$\delta(AB) = \delta(A) \cdot \delta(B)$$

for all $A, B \in \mathbf{C}^{n,n}$ and $\delta(A)$ equals the product of the diagonal entries of A for each diagonal matrix $A \in \mathbf{C}^{n,n}$. Prove that

$$\delta(A) = \det A$$

for all $A \in \mathbf{C}^{n,n}$.

Recall that $\mathbf{C}^{n,n}$ *denotes the set of* n*-by-*n *matrices with entries in* $\mathbf{C}$. *This exercise shows that the determinant is the unique function defined on square matrices that is multiplicative and has the desired behavior on diagonal matrices. This result is analogous to Exercise 10 in Section 8D, which shows that the trace is uniquely determined by its algebraic properties.*

Solution. First suppose that A is not invertible. It follows from Exercise 3.C.5 that there are matrices B, C, and D such that $A = BDC$, where D is diagonal with at least one diagonal entry equal to zero. Thus $\delta(D) = 0$ and hence

$$\delta(A) = \delta(B)\delta(D)\delta(C) = 0 = \det A,$$

where we have used 9.50 for the last equality.

To prove that $\delta(A) = \det A$ for invertible A, let us first prove the following. For $i, j \in \{1, ..., n\}$ such that $i \neq j$ and $\alpha \in \mathbf{C}$, let $C_{i,j}(\alpha)$ be the identity matrix except for an α in the i^{th} row and j^{th} column. For any $A \in \mathbf{C}^{n,n}$, a calculation shows that $AC_{i,j}(\alpha)$ is the matrix obtained from A by adding α times column i to column j. We claim that $\delta\big(C_{i,j}(\alpha)\big) = 1$. For ease of notation, let $C = C_{i,j}(\alpha)$. Note that C is invertible (its inverse is $C_{i,j}(-\alpha)$). Let B be the identity matrix except for $B_{j,j} = \frac{1}{2}$ and let D be the identity matrix except for $D_{j,j} = 2$. A calculation shows that $C = BCDC^{-1}$ and it follows that

$$\delta(C) = \delta(B)\delta(C)\delta(D)\delta(C^{-1}) = \tfrac{1}{2} \cdot \delta(C) \cdot 2 \cdot \delta(C^{-1}) = \delta(CC^{-1}) = \delta(I) = 1,$$

as claimed.

Now suppose that A is invertible. By 5.47 and 3.84, there exists an invertible matrix B and an invertible upper-triangular matrix U such that $A = B^{-1}UB$, from which it follows that

$$\delta(A) = \delta(B^{-1})\delta(U)\delta(B) = \delta(U)\delta(B^{-1}B) = \delta(U)\delta(I) = \delta(U).$$

Since $\det A = \det U$ by 9.52, it will suffice to show that $\delta(U) = \det U$. Suppose that the diagonal entries of U are $\lambda_1, ..., \lambda_n$; note that each λ_k is non-zero since U is invertible and that $\det U = \lambda_1 \cdots \lambda_n$ by 9.48. For any matrix of the form $C_{i,j}(\alpha)$, observe that

$$\delta\big(UC_{i,j}(\alpha)\big) = \delta(U)\delta\big(C_{i,j}(\alpha)\big) = \delta(U).$$

Given that U is upper-triangular with non-zero diagonal entries, we can multiply U on the right by successive matrices of the form $C_{i,j}(\alpha)$ until we obtain a diagonal matrix D with diagonal entries $\lambda_1, ..., \lambda_n$; as we just showed, this has no effect on the value of δ, i.e. $\delta(D) = \delta(U)$. Thus

$$\delta(U) = \delta(D) = \lambda_1 \cdots \lambda_n = \det U,$$

as desired.

9.D. Tensor Products

Exercise 9.D.1. Suppose $v \in V$ and $w \in W$. Prove that $v \otimes w = 0$ if and only if $v = 0$ or $w = 0$.

Solution. If $v = 0$ then $\varphi(v) = 0$ for any $\varphi \in V'$ and if $w = 0$ then $\tau(w) = 0$ for any $\tau \in W'$. Thus if $v = 0$ or $w = 0$ then

$$(v \otimes w)(\varphi, \tau) = \varphi(v)\tau(w) = 0$$

for any $\varphi \in V'$ and any $\tau \in W'$. Hence $v \otimes w = 0$.

If $v \neq 0$ and $w \neq 0$ then, by Exercise 3.F.3, there exist $\varphi \in V'$ and $\tau \in W'$ such that $\varphi(v) = \tau(w) = 1$. It follows that

$$(v \otimes w)(\varphi, \tau) = \varphi(v)\tau(w) = 1.$$

Thus $v \otimes w \neq 0$.

Exercise 9.D.2. Give an example of six distinct vectors $v_1, v_2, v_3, w_1, w_2, w_3 \in \mathbf{R}^3$ such that

$$v_1 \otimes w_1 + v_2 \otimes w_2 + v_3 \otimes w_3 = 0$$

but none of $v_1 \otimes w_1, v_2 \otimes w_2, v_3 \otimes w_3$ is a scalar multiple of another element of this list.

Solution. Let

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} -1 \\ -1 \\ -1 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix}, \quad w_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad w_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad w_3 = \begin{pmatrix} 0 \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}.$$

By identifying $v_k \otimes w_k$ with the matrix $v_k w_k^{\mathrm{t}}$ (see 9.76), we find that

$$v_1 \otimes w_1 = \begin{pmatrix} 1 & -1 & 0 \\ 1 & -1 & 0 \\ 1 & -1 & 0 \end{pmatrix}, \quad v_2 \otimes w_2 = \begin{pmatrix} -1 & 0 & -1 \\ -1 & 0 & -1 \\ -1 & 0 & -1 \end{pmatrix}, \quad v_3 \otimes w_3 = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}.$$

Since the sum of these three matrices is zero and none of them is a scalar multiple of another, $v_1 \otimes w_1, v_2 \otimes w_2, v_3 \otimes w_3$ is the desired list.

Exercise 9.D.3. Suppose that $v_1, \ldots, v_m$ is a linearly independent list in V. Suppose also that $w_1, \ldots, w_m$ is a list in W such that

$$v_1 \otimes w_1 + \cdots + v_m \otimes w_m = 0.$$

Prove that $w_1 = \cdots = w_m = 0$.

Solution. We will prove the contrapositive. Suppose that $w_k \neq 0$ for some $k \in \{1, ..., m\}$. By Exercise 3.F.3 there exists some $\tau \in W'$ such that $\tau(w_k) \neq 0$. As in the proof of 9.74, let $\varphi_1, ..., \varphi_m \in V'$ be such that $\varphi_i(v_j) = 1$ if $i = j$ and $\varphi_i(v_j) = 0$ otherwise. Now observe that

$$[v_1 \otimes w_1 + \cdots + v_m \otimes w_m](\varphi_k, \tau) = \varphi_k(v_1)\tau(w_1) \\ + \cdots + \varphi_k(v_k)\tau(w_k) + \cdots + \varphi_k(v_m)\tau(w_m) = \tau(w_k) \neq 0.$$

Thus $v_1 \otimes w_1 + \cdots + v_m \otimes w_m \neq 0$.

Exercise 9.D.4. Suppose $\dim V > 1$ and $\dim W > 1$. Prove that

$$\{v \otimes w : (v, w) \in V \times W\}$$

is not a subspace of $V \otimes W$.

This exercise implies that if $\dim V > 1$ *and* $\dim W > 1$, *then*

$$\{v \otimes w : (v, w) \in V \times W\} \neq V \otimes W.$$

Solution. Let $e_1, e_2, ..., e_m$ be a basis of V with dual basis $\varphi_1, \varphi_2, ..., \varphi_m$, and let $f_1, f_2, ..., f_n$ be a basis of W with dual basis $\tau_1, \tau_2, ..., \tau_n$. Let

$$U = \{v \otimes w : (v, w) \in V \times W\}.$$

Observe that

$$[e_1 \otimes f_1 + e_2 \otimes f_2](\varphi_i, \tau_j) = \begin{cases} 1 & \text{if } i = j \in \{1, 2\}, \\ 0 & \text{otherwise.} \end{cases}$$

For any $v = a_1 e_1 + a_2 e_2 + \cdots + a_m e_m \in V$ and any $w = b_1 f_1 + b_2 f_2 + \cdots + b_n f_n \in W$, observe that $(v \otimes w)(\varphi_i, \tau_j) = a_i b_j$. Equating $e_1 \otimes f_1 + e_2 \otimes f_2 = v \otimes w$ then gives us the system of equations

$$a_1 b_1 = a_2 b_2 = 1 \quad \text{and} \quad a_1 b_2 = a_2 b_1 = 0,$$

which has no solution. Thus $e_1 \otimes f_1 + e_2 \otimes f_2 \neq v \otimes w$ for any $(v, w) \in V \times W$, which shows that U is not closed under addition and hence is not a subspace of $V \otimes W$.

Exercise 9.D.5. Suppose m and n are positive integers. For $v \in \mathbf{F}^m$ and $w \in \mathbf{F}^n$, identify $v \otimes w$ with an m-by-n matrix as in Example 9.76. With that identification, show that the set

$$\{v \otimes w : v \in \mathbf{F}^m \text{ and } w \in \mathbf{F}^n\}$$

is the set of m-by-n matrices (with entries in $\mathbf{F}$) that have rank at most one.

Solution. Thinking of v and w as column vectors, Example 9.76 shows that we can identify $v \otimes w$ with the m-by-n matrix vw^{t}. The desired result is now immediate from Exercise 3.C.16.

Exercise 9.D.6. Suppose m and n are positive integers. Give a description, analogous to Exercise 5, of the set of m-by-n matrices (with entries in $\mathbf{F}$) that have rank at most two.

Solution. Let M be the collection of m-by-n matrices with entries in $\mathbf{F}$ that have rank at most two. We claim that M can be identified with the set

$$\{u \otimes w + v \otimes x : u, v \in \mathbf{F}^m \text{ and } w, x \in \mathbf{F}^n\}$$

by identifying $u \otimes w$ with uw^{t} (thinking of u and w as column vectors) as in example 9.76. This amounts to showing that

$$M = \{uw^{\mathrm{t}} + vx^{\mathrm{t}} : u, v \in \mathbf{F}^{m,1} \text{ and } w, x \in \mathbf{F}^{n,1}\}.$$

For $u, v \in \mathbf{F}^{m,1}$ and $w, x \in \mathbf{F}^{n,1}$, observe that $uw^{\mathrm{t}} + vx^{\mathrm{t}}$ is the matrix whose k^{th} column is $w_k u + x_k v$, from which it is clear that the span of the columns of $uw^{\mathrm{t}} + vx^{\mathrm{t}}$ is equal to the span of u and v. Thus $uw^{\mathrm{t}} + vx^{\mathrm{t}}$ has rank at most two.

Exercise 9.D.7. Suppose $\dim V > 2$ and $\dim W > 2$. Prove that

$$\{v_1 \otimes w_1 + v_2 \otimes w_2 : v_1, v_2 \in V \text{ and } w_1, w_2 \in W\} \neq V \otimes W.$$

Solution. Suppose $e_1, ..., e_m$ is a basis of V and $f_1, ..., f_n$ is a basis of W, where $m, n \geq 3$. By identifying an m-by-n matrix A with the bilinear functional

$$\sum_{k=1}^{n} \sum_{j=1}^{m} A_{j,k}(e_j \otimes f_k)$$

as in the proof of 9.74, we can identify $V \otimes W$ with $\mathbf{F}^{m,n}$. Under this identification the set

$$E \coloneqq \{v_1 \otimes w_1 + v_2 \otimes w_2 : v_1, v_2 \in V \text{ and } w_1, w_2 \in W\}$$

corresponds to the set of m-by-n matrices that have rank at most two, as we showed in Exercise 9.D.6. Given that $m, n \geq 3$, there exist matrices in $\mathbf{F}^{m,n}$ with rank strictly greater than two. Thus $E \neq V \otimes W$.

Exercise 9.D.8. Suppose $v_1, ..., v_m \in V$ and $w_1, ..., w_m \in W$ are such that

$$v_1 \otimes w_1 + \cdots + v_m \otimes w_m = 0.$$

Suppose that U is a vector space and $\Gamma : V \times W \to U$ is a bilinear map. Show that

$$\Gamma(v_1, w_1) + \cdots + \Gamma(v_m, w_m) = 0.$$

Solution. 9.79(a) shows that there exists a linear map $\hat{\Gamma} : V \otimes W \to U$ such that

$$\hat{\Gamma}(v \otimes w) = \Gamma(v, w)$$

for all $v \in V$ and all $w \in W$. It follows that

$$\Gamma(v_1, w_1) + \cdots + \Gamma(v_m, w_m) = \hat{\Gamma}(v_1 \otimes w_1) + \cdots + \hat{\Gamma}(v_m \otimes w_m)$$
$$= \hat{\Gamma}(v_1 \otimes w_1 + \cdots + v_m \otimes w_m) = \hat{\Gamma}(0) = 0.$$

Exercise 9.D.9. Suppose $S \in \mathcal{L}(V)$ and $T \in \mathcal{L}(W)$. Prove that there exists a unique operator on $V \otimes W$ that takes $v \otimes w$ to $Sv \otimes Tw$ for all $v \in V$ and $w \in W$.

In an abuse of notation, the operator on $V \otimes W$ given by this exercise is often called $S \otimes T$.

Solution. Define a bilinear map $\Gamma : V \times W \to V \otimes W$ by $\Gamma(v, w) = Sv \otimes Tw$. By 9.79, there is a unique operator $\hat{\Gamma} \in \mathcal{L}(V \otimes W)$ such that

$$\hat{\Gamma}(v \otimes w) = \Gamma(v, w) = Sv \otimes Tw.$$

Exercise 9.D.10. Suppose $S \in \mathcal{L}(V)$ and $T \in \mathcal{L}(W)$. Prove that $S \otimes T$ is an invertible operator on $V \otimes W$ if and only if both S and T are invertible operators. Also, prove that if both S and T are invertible operators, then $(S \otimes W)^{-1} = S^{-1} \otimes T^{-1}$, where we are using the notation from the comment after Exercise 9.

Solution. If S and T are both invertible then

$$(S^{-1} \otimes T^{-1})(S \otimes W)(v \otimes w) = (S^{-1} \otimes T^{-1})(Sv \otimes Tw) = S^{-1}Sv \otimes T^{-1}Tw = v \otimes w$$

for any $(v, w) \in V \times W$. Thus $S \otimes W$ is invertible and $(S \otimes W)^{-1} = S^{-1} \otimes T^{-1}$.

Now suppose that S is not invertible (the case where T is not invertible is handled similarly). There exists a non-zero $v \in V$ such that $Sv = 0$. Let $w \in W$ be non-zero (we may as well assume $W \neq 0$) and note that $v \otimes w \neq 0$ by Exercise 9.D.1. Now observe that

$$(S \otimes T)(v \otimes w) = Sv \otimes Tw = 0 \otimes Tw = 0.$$

Thus $S \otimes T$ is not invertible.

Exercise 9.D.11. Suppose V and W are inner product spaces. Prove that if $S \in \mathcal{L}(V)$ and $T \in \mathcal{L}(W)$, then $(S \otimes T)^* = S^* \otimes T^*$, where we are using the notation from the comment after Exercise 9.

Solution. For any $u, v \in V$ and any $w, x \in W$, observe that

$$\begin{aligned}
\langle (S \otimes T)(v \otimes w), u \otimes x \rangle &= \langle Sv \otimes Tw, u \otimes x \rangle \\
&= \langle Sv, u \rangle \langle Tw, x \rangle \\
&= \langle u, S^* u \rangle \langle w, T^* x \rangle \\
&= \langle v \otimes w, S^* u \otimes T^* x \rangle \\
&= \langle v \otimes w, (S^* \otimes T^*)(u \otimes x) \rangle.
\end{aligned}$$

Thus $(S \otimes T)^* = S^* \otimes T^*$.

Exercise 9.D.12. Suppose that $V_1, \ldots, V_m$ are finite-dimensional inner product spaces. Prove that there is a unique inner product on $V_1 \otimes \cdots \otimes V_m$ such that

$$\langle v_1 \otimes \cdots \otimes v_m, u_1 \otimes \cdots \otimes u_m \rangle = \langle v_1, u_1 \rangle \cdots \langle v_m, u_m \rangle$$

for all $(v_1, \ldots, v_m)$ and $(u_1, \ldots, u_m)$ in $V_1 \times \cdots \times V_m$.

Note that the equation above implies that

$$\|v_1 \otimes \cdots \otimes v_m\| = \|v_1\| \times \cdots \times \|v_m\|$$

for all $(v_1, \ldots, v_m) \in V_1 \times \cdots \times V_m$.

Solution. For each $k \in \{1, \ldots, m\}$ let $e_1^k, \ldots, e_{n_k}^k$ be an orthonormal basis of V_k. A very tedious calculation shows that

$$\left\langle \sum_{j_1=1}^{n_1} \cdots \sum_{j_m=1}^{n_m} b_{j_1,\ldots,j_m} \left(e_{j_1}^1 \otimes \cdots \otimes e_{j_m}^m\right), \sum_{j_1=1}^{n_1} \cdots \sum_{j_m=1}^{n_m} c_{j_1,\ldots,j_m} \left(e_{j_1}^1 \otimes \cdots \otimes e_{j_m}^m\right) \right\rangle$$

$$\sum_{j_1=1}^{n_1} \cdots \sum_{j_m=1}^{n_m} b_{j_1,\ldots,j_m} \overline{c_{j_1,\ldots,j_m}}$$

is the desired inner product.

Exercise 9.D.13. Suppose that $V_1, \ldots, V_m$ are finite-dimensional inner product spaces and $V_1 \otimes \cdots \otimes V_m$ is made into an inner product space using the inner product from Exercise 12. Suppose $e_1^k, \ldots, e_{n_k}^k$ is an orthonormal basis of V_k for each $k = 1, \ldots, m$. Show that the list

$$\left\{e_{j_1}^1 \otimes \cdots \otimes e_{j_m}^m\right\}_{j_1=1,\ldots,n_1;\cdots;j_m=1,\ldots,n_m}$$

is an orthonormal basis of $V_1 \otimes \cdots \otimes V_m$.

Solution. The list in question is a basis of $V_1 \otimes \cdots \otimes V_m$ by 9.90. To verify orthonormality, suppose that $i_k, j_k \in \{1, \ldots, n_k\}$ for each $k \in \{1, \ldots, m\}$ and observe that

$$\left\langle e_{i_1}^1 \otimes \cdots \otimes e_{i_m}^m, e_{j_1}^1 \otimes \cdots \otimes e_{j_m}^m \right\rangle = \left\langle e_{i_1}^1, e_{j_1}^1 \right\rangle \cdots \left\langle e_{i_m}^m, e_{j_m}^m \right\rangle = \begin{cases} 1 & \text{if } i_1 = j_1, \ldots, i_m = j_m, \\ 0 & \text{otherwise.} \end{cases}$$