Solutions to Axler's *Linear Algebra Done Right*, 4th Edition

by: blanketism

began: July 4, 2026
completed: November 2, 2026
last edited: November 17, 2026

Contents

Preface **3**

1 Vector Spaces **4**
1A $\mathbf{R}^n$ and $\mathbf{C}^n$ . . . 4
1B Definition of Vector Space . . . 8
1C Subspaces . . . 13

2 Finite-Dimensional Vector Spaces **22**
2A Span and Linear Independence . . . 22
2B Bases . . . 30
2C Dimension . . . 36

3 Linear Maps **47**
3A Vector Space of Linear Maps . . . 47
3B Null Spaces and Ranges . . . 57
3C Matrices . . . 72
3D Invertibility and Isomorphisms . . . 81
3E Products and Quotients of Vector Spaces . . . 91
3F Duality . . . 99

4 Polynomials **114**

5 Eigenvalues and Eigenvectors **119**
5A Invariant Subspaces . . . 119
5B The Minimal Polynomial . . . 134
5C Upper-Triangular Matrices . . . 148
5D Diagonalizable Operators . . . 154
5E Commuting Operators . . . 166

6 Inner Product Spaces **171**
6A Inner Products and Norms . . . 171
6B Orthonormal Bases . . . 189
6C Orthogonal Complements and Minimization Problems . . . 202

7 Operators on Inner Product Spaces **212**
7A Self-Adjoint and Normal Operators . . . 212
7B Spectral Theorem . . . 226
7C Positive Operators . . . 237
7D Isometries, Unitary Operators, and Matrix Factorization . . . 246
7E Singular Value Decomposition . . . 255
7F Consequences of Singular Value Decomposition . . . 263

8 Operators on Complex Vector Spaces **277**
8A Generalized Eigenvectors and Nilpotent Operators . . . 277
8B Generalized Eigenspace Decomposition . . . 286
8C Consequences of Generalized Eigenspace Decomposition . . . 295
8D Trace: A Connection Between Matrices and Operators . . . 300

9 Multilinear Algebra and Determinants **305**
9A Bilinear Forms and Quadratic Forms . . . 305
9B Alternating Linear Forms . . . 309
9C Determinants . . . 313
9D Tensor Products . . . 322

A Appendix: Alternate Proofs of Exercises **327**
A1 Vector Spaces . . . 327
A2 Finite-Dimensional Vector Spaces . . . 328
A3 Linear Maps . . . 329
A4 Polynomials . . . 333
A5 Eigenvalues and Eigenvectors . . . 334
A6 Inner Product Spaces . . . 337
A7 Operators on Inner Product Spaces . . . 339
A8 Generalized Eigenvectors and Nilpotent Operators . . . 343
A9 Multilinear Algebra and Determinants . . . 344

Preface

For any corrections, suggestions, or feedback regarding this solutions manual, readers are encouraged to reach out through the following channels:

- Discord Errata Server: https://discord.gg/qecQk37vfm
- Reddit: https://www.reddit.com/user/ansv9a8fdh3/
- GitHub: https://github.com/blanketism/Axler-Solutions

This solutions manual accompanies the fourth edition of Sheldon Axler's *Linear Algebra Done Right*. While several community-maintained solution sets exist for the 3rd edition, the 4th edition introduces enough new and reorganized material to warrant a dedicated solutions manual.

This manual is designed primarily for students encountering proof-based linear algebra for the first time. Axler's text is aimed at students with no linear algebra experience but have developed mathematical maturity through prior coursework in proof-based classes. Additionally, the text places an emphasis on analysis-inspired reasoning rather than purely computational techniques.

This manual is meant to be a companion to, rather than a replacement for, reading the text and attempting the exercises independently. Linear algebra–particularly in Axler's determinant-free and operator-theoretic approach–is best learned through active engagement with the material. Hence, I strongly encourage readers to attempt each exercise on their own before consulting the solutions provided here.

Each exercise is presented verbatim, along with any author hints or notes. Differently colored exercises, as seen below, indicate selected exercises that I found interesting, challenging, or foundational. Exercises marked with ✔ have alternate solutions in the appendix that offer different methods or additional insights.

Exercise 01: Example Title ✔

Here is an example special exercise denoted with an appendix solution.

Exercise 01 ✔

Here is an example special exercise denoted with an appendix solution.
An appendix exercise always corresponds to the original exercise text.

Despite my best efforts to ensure the accuracy of the solutions presented in this manual, I cannot guarantee that they are free from errors. Corrections, suggestions, and feedback are always welcome. If you notice a mistake or have a question, I encourage you to reach out through the channels listed above.

Finally, I would like to express my gratitude to Sheldon Axler for writing *Linear Algebra Done Right*, a book that has had a profound influence on my understanding and appreciation of linear algebra. This manual is a labor of love, and I hope that it serves as a useful companion for students working through the text and developing a deeper understanding of linear algebra.

1 Vector Spaces

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

Exercise 1A1

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

Solution. Let $\alpha = a + bi$ and $\beta = c + di$. Then

$$\begin{aligned}\alpha + \beta &= (a + bi) + (c + di)\\ &= (a + c) + (b + d)i\\ &= (c + a) + (d + b)i\\ &= (c + di) + (a + bi) = \beta + \alpha.\end{aligned}$$

■

Exercise 1A2

Show that $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$ for all $\alpha, \beta, \gamma \in \mathbf{C}$.

Solution. Let $\alpha = a + bi$, $\beta = c + di$, and $\gamma = e + fi$. Then

$$\begin{aligned}(\alpha + \beta) + \gamma &= ((a + bi) + (c + di)) + (e + fi)\\ &= ((a + c) + (b + d)i) + (e + fi)\\ &= (a + c + e) + (b + d + f)i\\ &= (a + (c + e)) + (b + (d + f))i\\ &= (a + bi) + ((c + di) + (e + fi)) = \alpha + (\beta + \gamma).\end{aligned}$$

■

Exercise 1A3

Show that $(\alpha\beta)\gamma = \alpha(\beta\gamma)$ for all $\alpha, \beta, \gamma \in \mathbf{C}$.

Solution. Let $\alpha = a + bi$, $\beta = c + di$, and $\gamma = e + fi$. Then

$$\begin{aligned}(\alpha\beta)\gamma &= ((a + bi)(c + di))(e + fi)\\ &= ((ac - bd) + (ad + bc)i)(e + fi)\\ &= (ace - bde - adf - bcf) + (acf - bdf + ade + bce)i\\ &= (ace - adf - bde - bcf) + (acf + ade + bce - bdf)i\\ &= (a + bi)((ce - df) + (cf + de)i)\\ &= (a + bi)((c + di)(e + fi)) = \alpha(\beta\gamma).\end{aligned}$$

■

Exercise 1A4

Show that $\gamma(\alpha + \beta) = \gamma\alpha + \gamma\beta$ for all $\alpha, \beta, \gamma \in \mathbf{C}$.

Solution. Let $\alpha = a + bi$, $\beta = c + di$, and $\gamma = e + fi$. Then

$$\begin{aligned}\gamma(\alpha+\beta) &= (e+fi)((a+bi)+(c+di))\\ &= (e+fi)((a+c)+(b+d)i)\\ &= (e(a+c)-f(b+d))+(e(b+d)+f(a+c))i\\ &= (ea-fb+ec-fd)+(eb+fa+ed+fc)i\\ &= (ea-fb)+(eb+fa)i+(ec-fd)+(ed+fc)i\\ &= (e+fi)(a+bi)+(e+fi)(c+di) = \gamma\alpha+\gamma\beta.\end{aligned}$$

■

Exercise 1A5

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

Solution. Let $\alpha = a + bi$. Then $\beta = -a - bi$ satisfies

$$\alpha + \beta = (a+bi)+(-a-bi) = (a-a)+(b-b)i = 0 + 0i = 0.$$

If $\gamma \in \mathbf{C}$ is such that $\alpha + \gamma = 0$, then Exercise 1A1 and Exercise 1A2 give

$$\gamma = \gamma + 0 = \gamma + (\alpha + \beta) = (\gamma + \alpha) + \beta = 0 + \beta = \beta.$$

Therefore, β is unique. ■

Exercise 1A6

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. Let $\alpha = a + bi$ be such that $\alpha \neq 0$, so at least one of a or b is nonzero. Then $a^2 + b^2 \neq 0$. Let $\beta \in \mathbf{C}$ be given by

$$\beta = \frac{a}{a^2+b^2} + \frac{-b}{a^2+b^2}i.$$

Since $(a + bi)(a - bi) = a^2 + b^2$, we have

$$\alpha\beta = (a+bi)\left(\frac{a}{a^2+b^2} + \frac{-b}{a^2+b^2}i\right) = \frac{(a+bi)(a-bi)}{a^2+b^2} = \frac{a^2+b^2}{a^2+b^2} = 1.$$

Moreover, if $\gamma \in \mathbf{C}$ is such that $\alpha\gamma = 1$, then using Exercise 1A3 gives

$$\gamma = \gamma \cdot 1 = \gamma(\alpha\beta) = (\gamma\alpha)\beta = 1 \cdot \beta = \beta.$$

Therefore, β is unique. ■

Exercise 1A7

Show that

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

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

Solution. By direct computation,

$$\left(\frac{-1+\sqrt{3}i}{2}\right)^3 = \frac{(-1)^3+3(-1)^2(\sqrt{3}i)+3(-1)(\sqrt{3}i)^2+(\sqrt{3}i)^3}{8} = \frac{-1+3\sqrt{3}i+9-3\sqrt{3}i}{8} = \frac{8}{8} = 1. \qquad \blacksquare$$

Exercise 1A8

Find two distinct square roots of i.

Solution. Let $z = \frac{1+i}{\sqrt{2}}$. Then

$$z^2 = \frac{1+2i-1}{2} = \frac{2i}{2} = i.$$

Since $z \neq 0$, it follows that $-z \neq z$, and $(-z)^2 = (-1)^2 z^2 = z^2 = i$, so $-z$ is also a square root of i. $\blacksquare$

Exercise 1A9

Find $x \in \mathbf{R}^4$ such that

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

Solution. By direct computation,

$$x = \frac{1}{2}((5,9,-6,8) - (4,-3,1,7)) = \left(\frac{1}{2}, 6, -\frac{7}{2}, \frac{1}{2}\right). \qquad \blacksquare$$

Exercise 1A10

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 such a λ exists, then $\lambda(2-3i) = 12-5i$ gives

$$\lambda = \frac{12-5i}{2-3i} = 3+2i.$$

However, $\lambda(-6+7i) = (3+2i)(-6+7i) = -32+9i \neq -32-9i$, so no such λ exists. $\blacksquare$

Exercise 1A11

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

Solution. Let $x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n), z = (z_1, \ldots, z_n) \in \mathbf{F}^n$. Then

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

$\blacksquare$

Exercise 1A12

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

Solution. Let $a, b \in \mathbf{F}$ and let $x = (x_1, \ldots, x_n) \in \mathbf{F}^n$. Then

$$\begin{aligned}(ab)x &= (ab)(x_1, \ldots, x_n)\\ &= ((ab)x_1, \ldots, (ab)x_n)\\ &= (a(bx_1), \ldots, a(bx_n))\\ &= a(bx_1, \ldots, bx_n)\\ &= a(b(x_1, \ldots, x_n)) = a(bx).\end{aligned}$$

■

Exercise 1A13

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

Solution. Let $x = (x_1, \ldots, x_n) \in \mathbf{F}^n$. Then

$$\begin{aligned}1x &= 1(x_1, \ldots, x_n)\\ &= (1x_1, \ldots, 1x_n)\\ &= (x_1, \ldots, x_n) = x.\end{aligned}$$

■

Exercise 1A14

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

Solution. Let $\lambda \in \mathbf{F}$ and let $x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n) \in \mathbf{F}^n$. Then

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

■

Exercise 1A15

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

Solution. Let $a, b \in \mathbf{F}$ and let $x = (x_1, \ldots, x_n) \in \mathbf{F}^n$. Then

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

■

1B Definition of Vector Space

Exercise 1B1

Prove that $-(-v) = v$ for every $v \in V$.

Solution. Let $v \in V$. By the definition of additive inverse, we have

$$v + (-v) = 0 = -(-v) + (-v),$$

so both v and $-(-v)$ are additive inverses of $-v$. By the uniqueness of additive inverses, we conclude that $-(-v) = v$. ■

Exercise 1B2

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

Solution. If $a = 0$, then we are done. Suppose $a \neq 0$. Then

$$v = 1v = \frac{1}{a}(av) = \frac{1}{a} \cdot 0 = 0.$$ ■

Exercise 1B3

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

Solution. Let $x = \frac{1}{3}(w - v)$. Then

$$v + 3\left(\frac{1}{3}(w - v)\right) = v + (w - v) = w.$$

If $y \in V$ is such that $v + 3y = w$, then

$$v + 3y = w = v + 3x \implies 3y = 3x \implies y = x.$$

Therefore, x is the unique element of V such that $v + 3x = w$. ■

Exercise 1B4

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. Which one?

Solution. The empty set fails to satisfy the requirement that there exists a zero vector $0 \in V$ such that $v + 0 = v$ for all $v \in V$ since it contains no elements. ■

Exercise 1B5

Show that in the definition of a vector space, 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 in V.

Axler Note. *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. Suppose that V satisfies all the vector space axioms except possibly the additive inverse condition. We must prove that the additive inverse condition holds if and only if $0v = 0$ for all $v \in V$.
$\boxed{\Rightarrow}$ Suppose the additive inverse condition holds. By 1.30, $0v = 0$ for all $v \in V$.
$\boxed{\Leftarrow}$ Suppose $0v = 0$ for all $v \in V$. Let $v \in V$. Then

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

Hence, $(-1)v$ is the additive inverse of v. ■

Exercise 1B6

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

$$\begin{aligned} t + \infty = \infty + t = \infty + \infty = \infty, \\ t + (-\infty) = (-\infty) + t = (-\infty) + (-\infty) = -\infty, \\ \infty + (-\infty) = (-\infty) + \infty = 0. \end{aligned}$$

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

Solution. No, $\mathbf{R} \cup \{\infty, -\infty\}$ is not a vector space over $\mathbf{R}$ as it does not satisfy the associative property of addition. Let $t \in \mathbf{R} \setminus \{0\}$. Then

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

■

Exercise 1B7

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. Let $f, g \in V^S$ and let $a \in \mathbf{F}$. Define addition and scalar multiplication on V^S by

$$(f + g)(x) = f(x) + g(x), \quad (af)(x) = a(f(x)),$$

for all $x \in S$. We now verify the vector space axioms for V^S, where the following calculations repeatedly use the commutativity, associativity, and distributive properties of addition and scalar multiplication in V and $\mathbf{F}$.
Commutativity. Let $f, g \in V^S$. Then for all $x \in S$,

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

so $f + g = g + f$.
Associativity. Let $f, g, h \in V^S$ and let $a, b \in \mathbf{F}$. Then for all $x \in S$,

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

so $(f+g)+h = f+(g+h)$. Moreover,

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

so $(ab)f = a(bf)$.

Additive identity. Let $0 \in V$ be the additive identity in V, and define the zero function $0 \in V^S$ by $0(x) = 0$ for all $x \in S$. Then for all $f \in V^S$ and all $x \in S$,

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

so $f + 0 = f$.

Additive inverse. Let $f \in V^S$. Define $-f \in V^S$ by $(-f)(x) = -f(x)$ for all $x \in S$. Then for all $x \in S$,

$$(f + (-f))(x) = f(x) + (-f(x)) = 0 = 0(x),$$

so $f + (-f) = 0$.

Multiplicative identity. Let $1 \in \mathbf{F}$ be the multiplicative identity in $\mathbf{F}$. Then for all $f \in V^S$ and all $x \in S$,

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

so $1f = f$.

Distributive properties. Let $a, b \in \mathbf{F}$ and $f, g \in V^S$. Then for all $x \in S$,

$$(a(f+g))(x) = a((f+g)(x)) = a(f(x)+g(x)) = af(x) + ag(x) = (af + ag)(x),$$

so $a(f+g) = af + ag$.

Also,

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

so $(a+b)f = af + bf$. Therefore, V^S is a vector space. ∎

Exercise 1B8

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 $u_1, u_2, v_1, v_2 \in V$.
- Complex scalar multiplication on $V_{\mathbf{C}}$ is defined by

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

  for all $a, b \in \mathbf{R}$ and $u, v \in V$.

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

Axler Note. *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 prove that $V_{\mathbf{C}}$ is a complex vector space by verifying the vector space axioms.

Commutativity. Let $u_1 + iv_1, u_2 + iv_2 \in V_{\mathbf{C}}$. Then

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

Associativity. Let $u_1 + iv_1, u_2 + iv_2, u_3 + iv_3 \in V_{\mathbf{C}}$ and let $a + ib, c + id \in \mathbf{C}$. Then

$$\begin{aligned}((u_1 + iv_1) + (u_2 + iv_2)) + (u_3 + iv_3) &= ((u_1 + u_2) + i(v_1 + v_2)) + (u_3 + iv_3)\\ &= ((u_1 + u_2) + u_3) + i((v_1 + v_2) + v_3)\\ &= (u_1 + (u_2 + u_3)) + i(v_1 + (v_2 + v_3))\\ &= (u_1 + iv_1) + ((u_2 + u_3) + i(v_2 + v_3))\\ &= (u_1 + iv_1) + ((u_2 + iv_2) + (u_3 + iv_3)).\end{aligned}$$

Moreover,

$$\begin{aligned}((a + ib)(c + id))(u_1 + iv_1) &= ((ac - bd) + i(ad + bc))(u_1 + iv_1)\\ &= ((ac - bd)u_1 - (ad + bc)v_1) + i((ad + bc)u_1 + (ac - bd)v_1)\\ &= (a(cu_1 - dv_1) - b(cv_1 + du_1)) + i(a(cv_1 + du_1) + b(cu_1 - dv_1))\\ &= (a + ib)((c + id)(u_1 + iv_1)).\end{aligned}$$

Additive identity. Let $0 \in V$ be the additive identity in V. Then for any $u + iv \in V_{\mathbf{C}}$,

$$(0 + i0) + (u + iv) = (0 + u) + i(0 + v) = u + iv,$$

so $0 + i0$ is the additive identity in $V_{\mathbf{C}}$.

Additive inverse. Let $u + iv \in V_{\mathbf{C}}$. Since V is a vector space, both u and v have additive inverses $-u, -v \in V$, respectively. Then

$$(u + iv) + ((-u) + i(-v)) = (u + (-u)) + i(v + (-v)) = 0 + i0,$$

so $(-u) + i(-v)$ is the additive inverse of $u + iv$ in $V_{\mathbf{C}}$.

Multiplicative identity. Write $1 = 1 + i0 \in \mathbf{C}$. Then for any $u + iv \in V_{\mathbf{C}}$,

$$(1 + i0)(u + iv) = (1 \cdot u - 0 \cdot v) + i(1 \cdot v + 0 \cdot u) = u + iv.$$

Distributive properties. Let $u_1 + iv_1, u_2 + iv_2 \in V_{\mathbf{C}}$ and let $a + ib, c + id \in \mathbf{C}$. Then

$$\begin{aligned}(a + ib)((u_1 + iv_1) + (u_2 + iv_2)) &= (a + ib)((u_1 + u_2) + i(v_1 + v_2))\\ &= (a(u_1 + u_2) - b(v_1 + v_2)) + i(a(v_1 + v_2) + b(u_1 + u_2))\\ &= (au_1 - bv_1 + au_2 - bv_2) + i(av_1 + bu_1 + av_2 + bu_2)\\ &= (au_1 - bv_1 + i(av_1 + bu_1)) + (au_2 - bv_2 + i(av_2 + bu_2))\\ &= (a + ib)(u_1 + iv_1) + (a + ib)(u_2 + iv_2).\end{aligned}$$

Moreover,

$$\begin{aligned}((a + ib) + (c + id))(u_1 + iv_1) &= ((a + c) + i(b + d))(u_1 + iv_1)\\ &= ((a + c)u_1 - (b + d)v_1) + i((a + c)v_1 + (b + d)u_1)\\ &= (au_1 - bv_1 + cu_1 - dv_1) + i(av_1 + bu_1 + cv_1 + du_1)\\ &= (au_1 - bv_1 + i(av_1 + bu_1)) + (cu_1 - dv_1 + i(cv_1 + du_1))\\ &= (a + ib)(u_1 + iv_1) + (c + id)(u_1 + iv_1).\end{aligned}$$

Therefore, $V_{\mathbf{C}}$ is a complex vector space. ■

1C Subspaces

Lemma 1

Let U be a subspace of V. Suppose $\varphi \in \mathbf{F}^U$ satisfies

$$\varphi(u+v) = \varphi(u) + \varphi(v) \quad \text{and} \quad \varphi(au) = a\varphi(u)$$

for all $u, v \in U$ and $a \in \mathbf{F}$. Then for any $b \in \mathbf{F}$, the set $\{u \in U : \varphi(u) = b\}$ is a subspace of V if and only if $b = 0$.

Proof. Let $U_b = \{u \in U : \varphi(u) = b\}$. Then $\varphi(0) = \varphi(0+0) = \varphi(0) + \varphi(0)$, so $\varphi(0) = 0$.
$\boxed{\Rightarrow}$ Suppose U_b is a subspace of V. By 1.34, $0 \in U_b$, so $b = \varphi(0) = 0$.
$\boxed{\Leftarrow}$ Suppose $b = 0$. Then $0 \in U_0$. Let $u, v \in U_0$ and let $a \in \mathbf{F}$. Then

$$\varphi(u+v) = \varphi(u) + \varphi(v) = 0 + 0 = 0,$$

and

$$\varphi(au) = a\varphi(u) = a0 = 0.$$

Hence, $u + v \in U_0$ and $au \in U_0$, so U_0 is closed under addition and scalar multiplication. Therefore, U_0 is a subspace of V by 1.34. ∎

Exercise 1C1

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 be the subset of $\mathbf{F}^3$ in question, and note that addition and scalar multiplication in $\mathbf{F}^3$ are defined componentwise. Therefore, every φ below satisfies $\varphi(u+v) = \varphi(u) + \varphi(v)$ and $\varphi(au) = a\varphi(u)$ for all $u, v \in \mathbf{F}^3$ and $a \in \mathbf{F}$.

(a) Let $\varphi : \mathbf{F}^3 \to \mathbf{F}$ be defined by $\varphi(x_1, x_2, x_3) = x_1 + 2x_2 + 3x_3$, so $U = \{(x_1, x_2, x_3) \in \mathbf{F}^3 : \varphi(x_1, x_2, x_3) = 0\}$. By Lemma 1, U is a subspace of $\mathbf{F}^3$.
(b) Take the same φ as in part (a), so $U = \{(x_1, x_2, x_3) \in \mathbf{F}^3 : \varphi(x_1, x_2, x_3) = 4\}$. Since $4 \neq 0$, Lemma 1 gives that U is not a subspace of $\mathbf{F}^3$.
(c) Because $1 \cdot 0 \cdot 0 = 0$ and $0 \cdot 1 \cdot 1 = 0$, $(1,0,0), (0,1,1) \in U$. However, $(1,0,0) + (0,1,1) = (1,1,1) \notin U$ since $1 \cdot 1 \cdot 1 \neq 0$. Therefore, U is not closed under addition, so U is not a subspace of $\mathbf{F}^3$ by 1.34.
(d) Let $\varphi : \mathbf{F}^3 \to \mathbf{F}$ be defined by $\varphi(x_1, x_2, x_3) = x_1 - 5x_3$, so $U = \{(x_1, x_2, x_3) \in \mathbf{F}^3 : \varphi(x_1, x_2, x_3) = 0\}$. By Lemma 1, U is a subspace of $\mathbf{F}^3$. ∎

Exercise 1C2

Verify all assertions about subspaces in Example 1.35.

Solution. Let U be the subset in question.

(a) Let $\varphi : \mathbf{F}^4 \to \mathbf{F}$ be defined by $\varphi(x_1, x_2, x_3, x_4) = x_3 - 5x_4$, so $U = \{(x_1, x_2, x_3, x_4) \in \mathbf{F}^4 : \varphi(x_1, x_2, x_3, x_4) = b\}$. Since addition and scalar multiplication in $\mathbf{F}^4$ are defined componentwise, φ satisfies the hypotheses of Lemma 1, so U is a subspace of $\mathbf{F}^4$ if and only if $b = 0$.

(b) The zero function is continuous, so $0 \in U$. Let $f, g \in U$ and let $a \in \mathbf{R}$. Then $f + g$ and af are continuous on $[0,1]$, so $f + g, af \in U$. Therefore, U is a subspace of $\mathbf{R}^{[0,1]}$ by 1.34.

(c) The zero function is differentiable on $\mathbf{R}$, so $0 \in U$. Let $f, g \in U$ and let $a \in \mathbf{R}$. Then $f + g$ and af are differentiable on $\mathbf{R}$, so $f + g, af \in U$. Therefore, U is a subspace of $\mathbf{R}^{\mathbf{R}}$ by 1.34.

(d) Let W be the set of differentiable real-valued functions on $(0,3)$. By the argument in part (c) with $(0,3)$ in place of $\mathbf{R}$, W is a subspace of $\mathbf{R}^{(0,3)}$. Define $\varphi \colon W \to \mathbf{R}$ by $\varphi(f) = f'(2)$, so $U = \{f \in W : \varphi(f) = b\}$. Since differentiation is additive and homogeneous, φ satisfies the hypotheses of Lemma 1, so U is a subspace of $\mathbf{R}^{(0,3)}$ if and only if $b = 0$.

(e) The zero sequence $0 = (0, 0, \ldots)$ has limit 0, so $0 \in U$. Let $z, w \in U$ and let $a \in \mathbf{C}$. Then $z + w \to 0 + 0 = 0$ and $az \to a \cdot 0 = 0$, so $z + w, az \in U$. Therefore, U is a subspace of $\mathbf{C}^{\infty}$ by 1.34. ■

Exercise 1C3

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 of differentiable real-valued functions on $(-4,4)$ and let $W = \{f \in U : f'(-1) = 3f(2)\}$. By the argument in Exercise 1C2(c) with $(-4,4)$ in place of $\mathbf{R}$, U is a subspace of $\mathbf{R}^{(-4,4)}$. Define $\varphi \colon U \to \mathbf{R}$ by $\varphi(f) = f'(-1) - 3f(2)$, so $W = \{f \in U : \varphi(f) = 0\}$. Since differentiation and evaluation are additive and homogeneous, φ satisfies the hypotheses of Lemma 1, so W is a subspace of $\mathbf{R}^{(-4,4)}$. ■

Exercise 1C4

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 of continuous real-valued functions on $[0,1]$ and let $W = \{f \in U : \int_0^1 f = b\}$. By Exercise 1C2(b), U is a subspace of $\mathbf{R}^{[0,1]}$. Define $\varphi \colon U \to \mathbf{R}$ by $\varphi(f) = \int_0^1 f$, so $W = \{f \in U : \varphi(f) = b\}$. Since integration is additive and homogeneous, φ satisfies the hypotheses of Lemma 1, so W is a subspace of $\mathbf{R}^{[0,1]}$ if and only if $b = 0$. ■

Exercise 1C5

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

Solution. Since $i(1,0) = (i,0) \notin \mathbf{R}^2$, $\mathbf{R}^2$ is not closed under scalar multiplication, and therefore it is not a subspace of $\mathbf{C}^2$ by 1.34. ■

Exercise 1C6

(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. Let U be the subset in question.

(a) Yes, U is a subspace of $\mathbf{R}^3$. Note that $a^3 = b^3$ implies $a = b$ for $a, b \in \mathbf{R}$, so $U = \{(a,a,c) : a, c \in \mathbf{R}\}$, which contains $(0,0,0)$. Let $x = (x_1, x_1, x_3), y = (y_1, y_1, y_3) \in U$ and let $t \in \mathbf{R}$. Then $x + y = (x_1 + y_1, x_1 + y_1, x_3 + y_3) \in U$ and $tx = (tx_1, tx_1, tx_3) \in U$. Hence, U is closed under addition and scalar multiplication, so U is a subspace of $\mathbf{R}^3$ by 1.34.

(b) No, U is not a subspace of $\mathbf{C}^3$. Let $\omega = \frac{-1+\sqrt{3}i}{2}$ so that $\omega^3 = 1$ by Exercise 1A7. Then $(1,1,0), (\omega,1,0) \in U$. However,

$$(1+\omega)^3 = \left(\frac{1+\sqrt{3}i}{2}\right)^3 = -1 \neq 8 = 2^3.$$

Therefore, $(1,1,0) + (\omega,1,0) = (1+\omega,2,0) \notin U$, so U is not closed under addition and hence not a subspace of $\mathbf{C}^3$ by 1.34. ■

Exercise 1C7

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. Consider the subset $\mathbf{Q}^2$ of $\mathbf{R}^2$. Then $\mathbf{Q}^2$ contains $(0,0)$, is closed under addition since the sum of any two rational numbers is rational, and is closed under taking additive inverses since the additive inverse of a rational number is rational. However, $\sqrt{2}(1,0) = (\sqrt{2},0) \notin \mathbf{Q}^2$. Hence, $\mathbf{Q}^2$ is not closed under scalar multiplication and is not a subspace of $\mathbf{R}^2$ by 1.34. ■

Exercise 1C8

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. Consider the subset of $\mathbf{R}^2$ given by

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

Then U is nonempty since $(1,0) \in U$, and for all $a, x, y \in \mathbf{R}$, we have $a(x,0) = (ax,0) \in U$ and $a(0,y) = (0,ay) \in U$, so U is closed under scalar multiplication. However, U is not closed under addition since $(1,0) + (0,1) = (1,1) \notin U$. Hence, U is not a subspace of $\mathbf{R}^2$ by 1.34. ■

Exercise 1C9

A function $f : \mathbf{R} \to \mathbf{R}$ is called **periodic** if there exists a positive number p such that $f(x+p) = f(x)$ 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. The set of periodic functions is not a subspace of $\mathbf{R}^{\mathbf{R}}$. Consider the periodic functions $f(x) = \cos x$ and $g(x) = \cos\sqrt{2}x$. Then the sum $f+g$ attains a maximum value of 2 at 0. If $f+g$ were periodic with period p, then $(f+g)(p) = (f+g)(0) = 2$, so $\cos p + \cos(p\sqrt{2}) = 2$. Since $\cos(x) \leq 1$ for all $x \in \mathbf{R}$, it follows that $\cos p = \cos(p\sqrt{2}) = 1$, and therefore $p = 2\pi m$ and $p\sqrt{2} = 2\pi n$ for positive integers m and n. Then

$$(2\pi m)\sqrt{2} = 2\pi n \implies \sqrt{2} = \frac{n}{m},$$

contradicting that $\sqrt{2}$ is irrational. Hence, $f+g$ is not periodic, so the set of periodic functions is not closed under addition and therefore not a subspace of $\mathbf{R}^{\mathbf{R}}$ by 1.34. ■

Exercise 1C10

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. Let $U = V_1 \cap V_2$. By 1.34, $0 \in V_1$ and $0 \in V_2$, so $0 \in U$. Let $u, v \in U$ and let $a \in \mathbf{F}$. By 1.34, $u + v, au \in V_1$ and $u + v, au \in V_2$, so $u + v, au \in U$, and therefore U is a subspace of V by 1.34. ■

Exercise 1C11

Prove that the intersection of every collection of subspaces of V is a subspace of V.

Solution. Let $\mathcal{U}$ be a collection of subspaces of V and let

$$U = \bigcap_{W \in \mathcal{U}} W,$$

where an empty intersection is understood to be V. By 1.34, $0 \in W$ for all $W \in \mathcal{U}$, and thus $0 \in U$. Let $u, v \in U$ and let $a \in \mathbf{F}$. Then $u, v \in W$ for all $W \in \mathcal{U}$. By 1.34, $u + v, au \in W$ for all $W \in \mathcal{U}$, so $u + v, au \in U$. Therefore, U is a subspace of V by 1.34. ■

Exercise 1C12

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. Let U and W be subspaces of V.
$\boxed{\Rightarrow}$ Suppose $U \cup W$ is a subspace of V, and suppose neither $U \subseteq W$ nor $W \subseteq U$. Then there exist $u \in U \setminus W$ and $w \in W \setminus U$. If $u + w \in U$, then $w = (u + w) - u \in U$, a contradiction. Similarly, $u + w \in W$ gives $u = (u + w) - w \in W$, also a contradiction. Therefore, $u + w \notin U \cup W$, so $U \cup W$ is not closed under addition, contradicting 1.34. Thus, one of the subspaces must be contained in the other.
$\boxed{\Leftarrow}$ Suppose $U \subseteq W$. Then $U \cup W = W$, which is a subspace of V. Similarly, if $W \subseteq U$, then $U \cup W = U$, which is a subspace of V. ■

Exercise 1C13

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.

Axler Note. *This exercise is surprisingly harder than* Exercise 1C12*, 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, U_3 be subspaces of V.
$\boxed{\Rightarrow}$ Suppose $U_1 \cup U_2 \cup U_3$ is a subspace of V, but none of the subspaces contains the other two. By assumption, $U_2 \cup U_3 \nsubseteq U_1$ and $U_1 \cup U_3 \nsubseteq U_2$, so there exist $u \in U_1 \setminus (U_2 \cup U_3)$ and $v \in U_2 \setminus (U_1 \cup U_3)$. Since $U_1 \cup U_2 \cup U_3$ is a subspace, $u + av \in U_1 \cup U_2 \cup U_3$ for any nonzero $a \in \mathbf{F}$. If $u + av \in U_1$, then $(u + av) - u = av \in U_1$, which implies $v \in U_1$, a contradiction. Similarly, if $u + av \in U_2$, then $(u + av) - av = u \in U_2$, a contradiction. Therefore, $u + av \in U_3$ for all nonzero $a \in \mathbf{F}$. It follows that $u + v \in U_3$ and $u - v \in U_3$, so $(u + v) + (u - v) = 2u \in U_3$ and thus $u \in U_3$, a contradiction. Hence, one of the subspaces must contain the other two.
$\boxed{\Leftarrow}$ Suppose one of the subspaces contains the other two. Without loss of generality, suppose $U_2 \subseteq U_1$ and $U_3 \subseteq U_1$. Then $U_1 \cup U_2 \cup U_3 = U_1$, which is a subspace of V. ■

Exercise 1C14

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 $U + W = \{(a, b, 2a) \in \mathbf{F}^3 : a, b \in \mathbf{F}\}$.
$\boxed{\subseteq}$ Let $u = (x, -x, 2x) \in U$ and $w = (y, y, 2y) \in W$, and define $a = x + y$ and $b = y - x$. Then

$$u + w = (x + y, y - x, 2(x + y)) = (a, b, 2a),$$

so $u + w \in \{(a, b, 2a) \in \mathbf{F}^3 : a, b \in \mathbf{F}\}$.
$\boxed{\supseteq}$ Let $a, b \in \mathbf{F}$ and let $x = \frac{a-b}{2}$ and $y = \frac{a+b}{2}$. Then

$$(a, b, 2a) = (x + y, y - x, 2(x + y)) = (x, -x, 2x) + (y, y, 2y) \in U + W.$$

In words, $U + W$ is the set of all vectors in three-space whose third component is twice the first. ■

Exercise 1C15

Suppose U is a subspace of V. What is $U + U$?

Solution. We claim that $U + U = U$.
$\boxed{\subseteq}$ Let $u_1 + u_2 \in U + U$, where $u_1, u_2 \in U$. Since U is closed under addition, $u_1 + u_2 \in U$.
$\boxed{\supseteq}$ Let $u \in U$. Then $u = u + 0 \in U + U$ since $0 \in U$. ■

Exercise 1C16

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

Solution. Yes, the operation of addition on subspaces of V is commutative. For every $v \in V$, we have

$$\begin{aligned} v \in U + W &\iff v = u + w \text{ for some } u \in U, w \in W \\ &\iff v = w + u \text{ for some } w \in W, u \in U \\ &\iff v \in W + U. \end{aligned}$$

■

Exercise 1C17

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. Yes, the operation of addition on subspaces of V is associative. For every $v \in V$, we have

$$\begin{aligned} v \in (V_1 + V_2) + V_3 &\iff v = (v_1 + v_2) + v_3 \text{ for some } v_1 \in V_1, v_2 \in V_2, v_3 \in V_3 \\ &\iff v = v_1 + (v_2 + v_3) \text{ for some } v_1 \in V_1, v_2 \in V_2, v_3 \in V_3 \\ &\iff v \in V_1 + (V_2 + V_3). \end{aligned}$$

■

Exercise 1C18

Does the operation of addition on the subspaces of V have an additive identity? Which subspaces have additive inverses?

Solution. Yes, the operation of addition on the subspaces of V has additive identity $\{0\}$. For any subspace U of V and for any $v \in V$,

$$v \in U \iff v = u + 0 \text{ for some } u \in U \iff v \in U + \{0\}.$$

Therefore, $U + \{0\} = U$.

By the first part, $\{0\} + \{0\} = \{0\}$, so $\{0\}$ is its own additive inverse. Suppose U is a subspace of V with an additive inverse W, so $U + W = \{0\}$. By 1.40, $U \subseteq U + W = \{0\}$, so $U = \{0\}$. Therefore, the only subspace that has an additive inverse is the zero subspace $\{0\}$. ∎

Exercise 1C19

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. The statement is false. Consider the vector space $\mathbf{R}^2$ and the sets

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

Observe that $(0, 0) \in V_1$. For any $(x_1, 0), (x_2, 0) \in V_1$ and any $a \in \mathbf{R}$, we have

$$(x_1, 0) + (x_2, 0) = (x_1 + x_2, 0) \in V_1, \quad \text{and} \quad a(x_1, 0) = (ax_1, 0) \in V_1.$$

Therefore, V_1 is a subspace of $\mathbf{R}^2$ by 1.34. The proof for V_2 is similar.

Observe that $(0, 0) \in U$. For any $(z_1, z_1), (z_2, z_2) \in U$ and any $a \in \mathbf{R}$, we have

$$(z_1, z_1) + (z_2, z_2) = (z_1 + z_2, z_1 + z_2) \in U, \quad \text{and} \quad a(z_1, z_1) = (az_1, az_1) \in U.$$

Therefore, U is a subspace of $\mathbf{R}^2$ by 1.34.

It follows from the definitions of V_1, V_2, and U that $V_1 + U \subseteq \mathbf{R}^2$ and $V_2 + U \subseteq \mathbf{R}^2$. Moreover, for any $(x_0, y_0) \in \mathbf{R}^2$, we can write

$$(x_0, y_0) = (x_0 - y_0, 0) + (y_0, y_0) \in V_1 + U, \quad \text{and} \quad (x_0, y_0) = (0, y_0 - x_0) + (x_0, x_0) \in V_2 + U,$$

so $V_1 + U = V_2 + U = \mathbf{R}^2$. However, $V_1 \neq V_2$ since, for example, $(1, 0) \in V_1$ but $(1, 0) \notin V_2$. ∎

Exercise 1C20

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

$$W = \{(0, w_2, 0, w_4) \in \mathbf{F}^4 : w_2, w_4 \in \mathbf{F}\}.$$

Observe that $(0,0,0,0) \in W$. For any $(0,w_2,0,w_4), (0,w_2',0,w_4') \in W$ and any $a \in \mathbf{F}$, we have

$$(0,w_2,0,w_4) + (0,w_2',0,w_4') = (0,w_2+w_2',0,w_4+w_4') \in W, \quad \text{and} \quad a(0,w_2,0,w_4) = (0,aw_2,0,aw_4) \in W.$$

Therefore, W is a subspace of $\mathbf{F}^4$ by 1.34.

Suppose $(x_1,x_2,x_3,x_4) \in U \cap W$. Since $(x_1,x_2,x_3,x_4) \in U$, we have that $x_1 = x_2$ and $x_3 = x_4$. Moreover, since $(x_1,x_2,x_3,x_4) \in W$, we have that $x_1 = 0$ and $x_3 = 0$. Hence, $x_1 = x_2 = x_3 = x_4 = 0$, so $U \cap W = \{0\}$. Let $(y_1,y_2,y_3,y_4) \in \mathbf{F}^4$. Then we can write

$$(y_1,y_2,y_3,y_4) = (y_1,y_1,y_3,y_3) + (0,y_2-y_1,0,y_4-y_3) \in U + W,$$

so $\mathbf{F}^4 = U \oplus W$ by 1.46. ■

Exercise 1C21

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

$$W = \{(0,0,w_3,w_4,w_5) \in \mathbf{F}^5 : w_3,w_4,w_5 \in \mathbf{F}\}.$$

Observe that $(0,0,0,0,0) \in W$. For any $(0,0,w_3,w_4,w_5), (0,0,w_3',w_4',w_5') \in W$ and any $a \in \mathbf{F}$, we have

$$(0,0,w_3,w_4,w_5) + (0,0,w_3',w_4',w_5') = (0,0,w_3+w_3',w_4+w_4',w_5+w_5') \in W,$$

and

$$a(0,0,w_3,w_4,w_5) = (0,0,aw_3,aw_4,aw_5) \in W.$$

Therefore, W is a subspace of $\mathbf{F}^5$ by 1.34.

Suppose $(x_1,x_2,x_3,x_4,x_5) \in U \cap W$. Since $(x_1,x_2,x_3,x_4,x_5) \in U$, we have that $x_3 = x_1 + x_2$, $x_4 = x_1 - x_2$, and $x_5 = 2x_1$. Moreover, since $(x_1,x_2,x_3,x_4,x_5) \in W$, we have that $x_1 = 0$ and $x_2 = 0$. Hence, $x_1 = x_2 = x_3 = x_4 = x_5 = 0$, so $U \cap W = \{0\}$. Let $(y_1,y_2,y_3,y_4,y_5) \in \mathbf{F}^5$. Then we can write

$$(y_1,y_2,y_3,y_4,y_5) = (y_1,y_2,y_1+y_2,y_1-y_2,2y_1) + (0,0,y_3-(y_1+y_2),y_4-(y_1-y_2),y_5-2y_1) \in U + W.$$

Therefore, $\mathbf{F}^5 = U \oplus W$ by 1.46. ■

Exercise 1C22

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

$$W_1 = \{(0,0,w_3,0,0) \in \mathbf{F}^5 : w_3 \in \mathbf{F}\},$$
$$W_2 = \{(0,0,0,w_4,0) \in \mathbf{F}^5 : w_4 \in \mathbf{F}\},$$
$$W_3 = \{(0,0,0,0,w_5) \in \mathbf{F}^5 : w_5 \in \mathbf{F}\}.$$

We show that W_1 is a subspace of $\mathbf{F}^5$. Observe that $(0,0,0,0,0)$ and $(0,0,1,0,0) \in W_1$, so $W_1 \neq \{0\}$. For any

$(0,0,w_3,0,0), (0,0,w_3',0,0) \in W_1$ and any $a \in \mathbf{F}$, we have

$$(0,0,w_3,0,0) + (0,0,w_3',0,0) = (0,0,w_3+w_3',0,0) \in W_1, \quad \text{and} \quad a(0,0,w_3,0,0) = (0,0,aw_3,0,0) \in W_1.$$

Therefore, W_1 is a subspace of $\mathbf{F}^5$ by 1.34. The proofs for W_2 and W_3 are similar. Now suppose $0 = (x, y, x+y, x-y, 2x) + z_1 + z_2 + z_3$ for some $z_1 \in W_1$, $z_2 \in W_2$, and $z_3 \in W_3$. Then $x = y = 0$ by the first two coordinates, so the vector in U is 0. Consequently, the third, fourth, and fifth coordinates must also be 0, so $z_1 = z_2 = z_3 = 0$. Moreover, for any $(y_1, y_2, y_3, y_4, y_5) \in \mathbf{F}^5$, we can write

$$(y_1, y_2, y_3, y_4, y_5) = (y_1, y_2, y_1+y_2, y_1-y_2, 2y_1) + (0, 0, y_3-(y_1+y_2), y_4-(y_1-y_2), y_5-2y_1) \in U + W_1 + W_2 + W_3.$$

Therefore, $\mathbf{F}^5 = U \oplus W_1 \oplus W_2 \oplus W_3$ by 1.45. ■

Exercise 1C23

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 $V_1 = V_2$.

Axler 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 statement is false. Let $V = \mathbf{R}^2$ and V_1, V_2, and U be as in Exercise 1C19, so $V_1 + U = V_2 + U = V$, but $V_1 \neq V_2$. If $(x_1, y_1) \in V_1 \cap U$, then $y_1 = 0$ and $x_1 = y_1$, so $x_1 = y_1 = 0$. Similarly, $(x_2, y_2) \in V_2 \cap U$ implies $x_2 = 0$ and $x_2 = y_2$, so $x_2 = y_2 = 0$. Thus, $V_1 \cap U = V_2 \cap U = \{0\}$, so $V_1 \oplus U = V_2 \oplus U = V$ by 1.46. ■

Exercise 1C24

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 $\mathbf{R}^{\mathbf{R}} = V_e \oplus V_o$.

Manual Remark. *The trick of decomposing a function into its even and odd parts is useful here.*

Solution. Observe that $0 \in V_e$ and $0 \in V_o$. Let $f_1, f_2 \in V_e$ and $a \in \mathbf{R}$. Then for all $x \in \mathbf{R}$,

$$(f_1 + f_2)(-x) = f_1(-x) + f_2(-x) = f_1(x) + f_2(x) = (f_1 + f_2)(x),$$

so $f_1 + f_2 \in V_e$. Also,

$$(af_1)(-x) = af_1(-x) = af_1(x) = (af_1)(x),$$

so $af_1 \in V_e$. Hence, V_e is a subspace of $\mathbf{R}^{\mathbf{R}}$ by 1.34.

Let $f_1, f_2 \in V_o$ and $a \in \mathbf{R}$. Then for all $x \in \mathbf{R}$,

$$(f_1 + f_2)(-x) = f_1(-x) + f_2(-x) = -f_1(x) - f_2(x) = -(f_1 + f_2)(x),$$

so $f_1 + f_2 \in V_o$. Also,

$$(af_1)(-x) = af_1(-x) = a(-f_1(x)) = -(af_1)(x),$$

so $af_1 \in V_o$. Hence, V_o is a subspace of $\mathbf{R}^{\mathbf{R}}$ by 1.34.

Let $f \in \mathbf{R}^{\mathbf{R}}$, and consider the functions f_e and f_o defined by

$$f_e(x) = \frac{f(x) + f(-x)}{2} \quad \text{and} \quad f_o(x) = \frac{f(x) - f(-x)}{2}$$

for all $x \in \mathbf{R}$. Then

$$f_e(-x) = \frac{f(-x) + f(x)}{2} = \frac{f(x) + f(-x)}{2} = f_e(x),$$

and

$$f_o(-x) = \frac{f(-x) - f(x)}{2} = -\frac{f(x) - f(-x)}{2} = -f_o(x),$$

so $f_e \in V_e$ and $f_o \in V_o$. Moreover, for all $x \in \mathbf{R}$,

$$f_e(x) + f_o(x) = \frac{f(x) + f(-x)}{2} + \frac{f(x) - f(-x)}{2} = f(x),$$

so $f = f_e + f_o$. Finally, if $g \in V_e \cap V_o$, then $g(x) = g(-x) = -g(x)$ for all $x \in \mathbf{R}$, so $g = 0$. Hence, $\mathbf{R}^{\mathbf{R}} = V_e \oplus V_o$ by 1.46. ■

2 Finite-Dimensional Vector Spaces

2A Span and Linear Independence

Exercise 2A1

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

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

Consider the list of vectors

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

For any $(z, w, -z - w) \in U$, we may write it as $zv_1 + wv_2$. Therefore, v_1, v_2, v_3, v_4 span U. ∎

Exercise 2A2

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. Let $U = \{v_1, v_2, v_3, v_4\}$ and $W = \{v_1 - v_2, v_2 - v_3, v_3 - v_4, v_4\}$. Clearly, $v_4 \in \operatorname{span} U$. Then we may write

$$\begin{aligned} v_3 &= (v_3 - v_4) + v_4 \in \operatorname{span} W, \\ v_2 &= (v_2 - v_3) + v_3 \in \operatorname{span} W, \\ v_1 &= (v_1 - v_2) + v_2 \in \operatorname{span} W. \end{aligned}$$

Thus, $v_1, v_2, v_3, v_4 \in \operatorname{span} W$, and so $V = \operatorname{span} U \subseteq \operatorname{span} W$. Since W is a list of vectors in V, we have $\operatorname{span} W \subseteq V$. Therefore, $\operatorname{span} W = V$, and so W spans V. ∎

Exercise 2A3

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

Solution. This is a generalization of Exercise 2A2 and a direct consequence of the fact that $v_k = w_k - w_{k-1}$ for $k \in \{2, \ldots, m\}$ and $v_1 = w_1$. Since each v_k is a linear combination of $w_1, \ldots, w_k$, we have $\operatorname{span}(v_1, \ldots, v_m) \subseteq \operatorname{span}(w_1, \ldots, w_m)$. On the other hand, each w_k is a linear combination of $v_1, \ldots, v_k$, so $\operatorname{span}(w_1, \ldots, w_m) \subseteq \operatorname{span}(v_1, \ldots, v_m)$. Therefore, $\operatorname{span}(v_1, \ldots, v_m) = \operatorname{span}(w_1, \ldots, w_m)$. ∎

Exercise 2A4

(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) Let v be the vector in the list. If $v = 0$, then the list is linearly dependent, since $0 \cdot v = 0$. If $v \neq 0$, then the only way to write 0 as a linear combination of v is to take the scalar multiple to be 0. Thus, the list is linearly independent.

(b) Let v_1 and v_2 be the vectors in the list. If v_1 is a scalar multiple of v_2, then there exists a scalar λ such that $v_1 = \lambda v_2$. Then we may write

$$0 = v_1 - \lambda v_2,$$

which shows that the list is linearly dependent. Conversely, if the list is linearly dependent, then there exist scalars α and β, not both zero, such that

$$\alpha v_1 + \beta v_2 = 0.$$

If $\alpha \neq 0$, then

$$v_1 = -\frac{\beta}{\alpha} v_2.$$

Similarly, if $\beta \neq 0$, then

$$v_2 = -\frac{\alpha}{\beta} v_1.$$

Thus, the list is linearly independent if and only if neither vector is a scalar multiple of the other. ■

Exercise 2A5

Find a number t such that

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

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

Solution. To say that the list is not linearly independent is to say that there exist scalars α, β, γ, not all zero, such that

$$\alpha(3, 1, 4) + \beta(2, -3, 5) + \gamma(5, 9, t) = (0, 0, 0).$$

This is equivalent to the following system of equations:

$$\begin{aligned} 0 &= 3\alpha + 2\beta + 5\gamma, \\ 0 &= \alpha - 3\beta + 9\gamma, \\ 0 &= 4\alpha + 5\beta + t\gamma. \end{aligned}$$

From the second equation, we obtain $\alpha = 3\beta - 9\gamma$. Substituting this into the first equation, we have

$$\begin{aligned} 0 &= 3(3\beta - 9\gamma) + 2\beta + 5\gamma \\ &= 11\beta - 22\gamma. \end{aligned}$$

Thus, $\beta = 2\gamma$. Substituting this into the expression for α, we have $\alpha = 3(2\gamma) - 9\gamma = -3\gamma$. Substituting these values into

the third equation, we have

$$\begin{aligned} 0 &= 4(-3\gamma) + 5(2\gamma) + t\gamma \\ &= -12\gamma + 10\gamma + t\gamma \\ &= (t-2)\gamma. \end{aligned}$$

Since $\gamma \neq 0$, we must have $t - 2 = 0$, or $t = 2$. Then the list is not linearly independent when $t = 2$. ■

Exercise 2A6

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. We may proceed similarly to Exercise 2A5 to obtain the following system of equations:

$$\begin{aligned} 0 &= 2\alpha + \beta + 7\gamma, \\ 0 &= 3\alpha - \beta + 3\gamma, \\ 0 &= \alpha + 2\beta + c\gamma. \end{aligned}$$

The second equation gives us $\beta = 3\alpha + 3\gamma$. Substituting this into the first equation, we have

$$\begin{aligned} 0 &= 2\alpha + (3\alpha + 3\gamma) + 7\gamma \\ &= 5\alpha + 10\gamma. \end{aligned}$$

Thus, $\alpha = -2\gamma$. Substituting this into the expression for β, we have $\beta = 3(-2\gamma) + 3\gamma = -3\gamma$. Substituting these values into the third equation, we have

$$\begin{aligned} 0 &= (-2\gamma) + 2(-3\gamma) + c\gamma \\ &= -2\gamma - 6\gamma + c\gamma \\ &= (c-8)\gamma. \end{aligned}$$

Since $\gamma \neq 0$, we must have $c - 8 = 0$, or $c = 8$. Then the list is linearly dependent when $c = 8$.

Conversely, if $c \neq 8$, then $(c - 8)\gamma = 0$ implies $\gamma = 0$. Then $\alpha = -2\gamma = 0$ and $\beta = -3\gamma = 0$, so the only solution to the system of equations is $\alpha = \beta = \gamma = 0$. Thus, the list is linearly independent when $c \neq 8$. ■

Exercise 2A7

(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 $\alpha, \beta \in \mathbf{R}$ such that

$$\alpha(1 + i) + \beta(1 - i) = 0.$$

Then we have

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

Since $0 = 0 + 0i$, we must have $\alpha + \beta = 0$ and $\alpha - \beta = 0$. Solving this system of equations, we find that $\alpha = \beta = 0$. Thus, the list is linearly independent.

(b) Suppose $a + bi, c + di \in \mathbf{C}$ such that

$$(a + bi)(1 + i) + (c + di)(1 - i) = 0.$$

Then we have

$$(a - b + c + d) + (a + b - c + d)i = 0.$$

Then we have the system of equations

$$\begin{aligned} 0 &= a - b + c + d, \\ 0 &= a + b - c + d. \end{aligned}$$

Adding the two equations results in $0 = 2a + 2d$, or $a = -d$. Substituting this into the first equation, we obtain $b = c$. Then we have infinitely many nontrivial solutions to the system of equations, such as $a = 1, b = 0, c = 0, d = -1$. Thus, the list is linearly dependent. ■

Exercise 2A8

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. Let $\alpha, \beta, \gamma, \delta \in \mathbf{F}$ such that

$$\alpha(v_1 - v_2) + \beta(v_2 - v_3) + \gamma(v_3 - v_4) + \delta v_4 = 0.$$

Then we have

$$\alpha v_1 + (-\alpha + \beta)v_2 + (-\beta + \gamma)v_3 + (-\gamma + \delta)v_4 = 0.$$

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

$$\begin{aligned} 0 &= \alpha, \\ 0 &= -\alpha + \beta, \\ 0 &= -\beta + \gamma, \\ 0 &= -\gamma + \delta. \end{aligned}$$

From the first equation, we have $\alpha = 0$. Substituting this into the second equation gives us $\beta = 0$. Substituting this into the third equation gives us $\gamma = 0$. Finally, substituting this into the fourth equation gives us $\delta = 0$. Thus, the only solution to the original equation is $\alpha = \beta = \gamma = \delta = 0$, and so the list is linearly independent. ■

Exercise 2A9

Prove or give a counterexample: If $v_1, v_2, \ldots, v_m$ is a linearly independent list of vectors in V, then $5v_1 - 4v_2, v_2, v_3, \ldots, v_m$ is linearly independent.

Solution. Let $\alpha_1, \ldots, \alpha_m \in \mathbf{F}$ such that

$$\alpha_1(5v_1 - 4v_2) + \alpha_2 v_2 + \cdots + \alpha_m v_m = 0.$$

Then we have

$$5\alpha_1 v_1 + (-4\alpha_1 + \alpha_2)v_2 + \alpha_3 v_3 + \cdots + \alpha_m v_m = 0.$$

Since $v_1, v_2, \ldots, v_m$ is linearly independent, we must have

$$\begin{aligned} 0 &= 5\alpha_1, \\ 0 &= -4\alpha_1 + \alpha_2, \\ 0 &= \alpha_3, \\ &\vdots \\ 0 &= \alpha_m. \end{aligned}$$

From the first equation, we have $\alpha_1 = 0$. Substituting this into the second equation gives us $\alpha_2 = 0$. The remaining equations give us $\alpha_k = 0$ for $k = 3, \ldots, m$. Thus, the only solution to the original equation is $\alpha_k = 0$ for $k = 1, \ldots, m$, and so the list is linearly independent. ■

Exercise 2A10

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

Solution. Let $\alpha_1, \ldots, \alpha_m \in \mathbf{F}$ such that

$$\alpha_1(\lambda v_1) + \cdots + \alpha_m(\lambda v_m) = 0.$$

Then we have

$$\lambda(\alpha_1 v_1 + \cdots + \alpha_m v_m) = 0.$$

Since $\lambda \neq 0$, we must have

$$\alpha_1 v_1 + \cdots + \alpha_m v_m = 0.$$

Since $v_1, v_2, \ldots, v_m$ is linearly independent, we must have $\alpha_k = 0$ for $k = 1, \ldots, m$. Thus, the only solution to the original equation is $\alpha_k = 0$ for $k = 1, \ldots, m$, and so the list is linearly independent. ■

Exercise 2A11

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

Solution. Consider the following counterexample in $\mathbf{R}^2$:

$$v_1 = (1, 0), \quad v_2 = (0, 1), \quad w_1 = (0, 1), \quad w_2 = (1, 0).$$

Then v_1, v_2 and w_1, w_2 are linearly independent lists of vectors in $\mathbf{R}^2$. However,

$$v_1 + w_1 = (1, 1), \quad v_2 + w_2 = (1, 1),$$

and so the list $v_1 + w_1, v_2 + w_2$ is linearly dependent. Thus, the statement is false. ■

Exercise 2A12

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

Solution. If $w = 0$, then $w \in \operatorname{span}(v_1, \ldots, v_m)$ and we are done. Otherwise, assume $w \neq 0$, and let $\alpha_1, \ldots, \alpha_m \in \mathbf{F}$, not

all zero, such that

$$\alpha_1(v_1 + w) + \cdots + \alpha_m(v_m + w) = 0.$$

Then we have

$$(\alpha_1 v_1 + \cdots + \alpha_m v_m) + (\alpha_1 + \cdots + \alpha_m)w = 0.$$

If $\alpha_1 + \cdots + \alpha_m = 0$, then we have

$$\alpha_1 v_1 + \cdots + \alpha_m v_m = 0,$$

which contradicts the linear independence of $v_1, \ldots, v_m$. Thus, we must have $\alpha_1 + \cdots + \alpha_m \neq 0$. Let $\beta = \alpha_1 + \cdots + \alpha_m$. Then we may write

$$w = -\frac{\alpha_1}{\beta}v_1 - \cdots - \frac{\alpha_m}{\beta}v_m.$$

Thus, w is a linear combination of $v_1, \ldots, v_m$, and so $w \in \text{span}(v_1, \ldots, v_m)$. ■

Exercise 2A13

Suppose $v_1, \ldots, v_m$ is linearly independent in V and $w \in V$. Show that

$$v_1, \ldots, v_m, w \text{ is linearly independent} \iff w \notin \text{span}(v_1, \ldots, v_m).$$

Solution.

$\boxed{\Rightarrow}$ We proceed by contrapositive. Suppose $w \in \text{span}(v_1, \ldots, v_m)$, but $v_1, \ldots, v_m, w$ is linearly independent. Then w is a linear combination of $v_1, \ldots, v_m$, contradicting the linear independence of $v_1, \ldots, v_m, w$. Hence, $w \notin \text{span}(v_1, \ldots, v_m)$.

$\boxed{\Leftarrow}$ Conversely, suppose $w \notin \text{span}(v_1, \ldots, v_m)$, but $v_1, \ldots, v_m, w$ is linearly dependent. By the linear dependence lemma, there exists $k \in \{1, \ldots, m\}$ such that $v_k \in \text{span}(v_1, \ldots, v_{k-1})$. This contradicts the linear independence of $v_1, \ldots, v_m$. Hence, $v_1, \ldots, v_m, w$ must be linearly independent. ■

Exercise 2A14

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 the list $v_1, \ldots, v_m$ is linearly independent if and only if the list $w_1, \ldots, w_m$ is linearly independent.

Solution. Suppose $v_1, \ldots, v_m$ is linearly independent. Let $\alpha_1, \ldots, \alpha_m \in \mathbf{F}$ such that

$$\alpha_1 w_1 + \cdots + \alpha_m w_m = 0.$$

Substituting $w_k = v_1 + \cdots + v_k$ for each k, we get

$$\alpha_1 v_1 + \alpha_2(v_1 + v_2) + \alpha_3(v_1 + v_2 + v_3) + \cdots + \alpha_m(v_1 + \cdots + v_m) = 0.$$

Then we may rewrite this as

$$(\alpha_1 + \cdots + \alpha_m)v_1 + (\alpha_2 + \cdots + \alpha_m)v_2 + \cdots + (\alpha_{m-1} + \alpha_m)v_{m-1} + \alpha_m v_m = 0.$$

Since $v_1, \ldots, v_m$ is linearly independent, we must have $\alpha_1 = \cdots = \alpha_m = 0$. Hence, $w_1, \ldots, w_m$ is linearly independent.

Conversely, suppose $w_1, \ldots, w_m$ is linearly independent. Let $\beta_1, \ldots, \beta_m \in \mathbf{F}$ such that

$$\beta_1 v_1 + \cdots + \beta_m v_m = 0.$$

Noting that $v_1 = w_1$ and $v_k = w_k - w_{k-1}$ for $k \in \{2, \dots, m\}$, we may rewrite the equation as

$$\begin{aligned} \beta_1 w_1 + \beta_2(w_2 - w_1) + \cdots + \beta_m(w_m - w_{m-1}) &= 0, \\ (\beta_1 - \beta_2)w_1 + (\beta_2 - \beta_3)w_2 + \cdots + (\beta_{m-1} - \beta_m)w_{m-1} + \beta_m w_m &= 0. \end{aligned}$$

Since $w_1, \dots, w_m$ is linearly independent, we find that the system of equations

$$\begin{aligned} 0 &= \beta_1 - \beta_2, \\ 0 &= \beta_2 - \beta_3, \\ &\vdots \\ 0 &= \beta_{m-1} - \beta_m \\ 0 &= \beta_m, \end{aligned}$$

has the unique solution $\beta_1 = \cdots = \beta_m = 0$. Hence, $v_1, \dots, v_m$ is linearly independent. ■

Exercise 2A15

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

Solution. Since $\mathcal{P}_4(\mathbf{F})$ can be spanned by the list $1, x, x^2, x^3, x^4$, and the length of a linearly independent list cannot exceed the length of a spanning list, there does not exist a list of six polynomials that is linearly independent in $\mathcal{P}_4(\mathbf{F})$. ■

Exercise 2A16

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

Solution. Since the list $1, x, x^2, x^3, x^4$ is a linearly independent list in $\mathcal{P}_4(\mathbf{F})$ and has length 5, and the length of a spanning list cannot be less than the length of a linearly independent list, no list of four polynomials can span $\mathcal{P}_4(\mathbf{F})$. ■

Exercise 2A17

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

Solution. Suppose V is infinite-dimensional. We proceed by induction to construct a sequence of vectors $v_1, v_2, \dots$ in V such that $v_1, \dots, v_m$ is linearly independent for every positive integer m. Since V is infinite-dimensional, there exists a nonzero vector $v_1 \in V$. Suppose we have constructed a linearly independent list of vectors $v_1, \dots, v_m$ in V. If every $v \in V$ is a linear combination of $v_1, \dots, v_m$, then $V = \operatorname{span}(v_1, \dots, v_m)$, which contradicts the assumption that V is infinite-dimensional. Thus, there exists a vector $v_{m+1} \in V$ that is not in $\operatorname{span}(v_1, \dots, v_m)$. Then by Exercise 2A13, the list $v_1, \dots, v_m, v_{m+1}$ is linearly independent. By induction, we can construct a sequence of vectors $v_1, v_2, \dots$ in V such that $v_1, \dots, v_m$ is linearly independent for every positive integer m.

Conversely, suppose there exists a sequence of vectors $v_1, v_2, \dots$ in V such that $v_1, \dots, v_m$ is linearly independent for every positive integer m. Then for every positive integer m, there exists a linearly independent list of vectors in V of length m. Since there is no maximum length for a linearly independent list of vectors in V, no list can span V. Therefore, V is infinite-dimensional. ■

Exercise 2A18

Prove that $\mathbf{F}^\infty$ is infinite-dimensional.

Solution. By Exercise 2A17, it suffices to construct a sequence of vectors $v_1, v_2, \ldots$ in $\mathbf{F}^\infty$ such that $v_1, \ldots, v_m$ is linearly independent for every positive integer m. Let v_k be the vector in $\mathbf{F}^\infty$ whose k-th coordinate is 1 and all other coordinates are 0. Explicitly, we have

$$v_1 = (1, 0, 0, \ldots), \quad v_2 = (0, 1, 0, \ldots), \quad v_3 = (0, 0, 1, \ldots), \quad \ldots$$

Then for any positive integer m, the list $v_1, \ldots, v_m$ is linearly independent, since the only way to write the zero vector as a linear combination of $v_1, \ldots, v_m$ is to take all coefficients to be zero. Thus, by Exercise 2A17, $\mathbf{F}^\infty$ is infinite-dimensional. ■

Exercise 2A19

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

Solution. By Exercise 2A17, it suffices to construct a sequence of continuous real-valued functions $f_1, f_2, \ldots$ on the interval $[0, 1]$ such that $f_1, \ldots, f_m$ is linearly independent for every positive integer m. Let $f_k(x) = x^{k-1}$ for each positive integer k. Then for any positive integer m, let $\alpha_1, \ldots, \alpha_m \in \mathbf{R}$ such that

$$\alpha_1 f_1(x) + \cdots + \alpha_m f_m(x) = 0$$

for all $x \in [0, 1]$. This is equivalent to the polynomial equation

$$\alpha_1 + \alpha_2 x + \cdots + \alpha_m x^{m-1} = 0$$

for all $x \in [0, 1]$. Since a polynomial of degree at most $m - 1$ can have at most $m - 1$ roots unless it is the zero polynomial, we must have $\alpha_1 = \cdots = \alpha_m = 0$. Thus, the list $f_1, \ldots, f_m$ is linearly independent. Therefore, by Exercise 2A17, the real vector space of all continuous real-valued functions on the interval $[0, 1]$ is infinite-dimensional. ■

Exercise 2A20 ✔

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. Note that $\mathcal{P}_m(\mathbf{F})$ is spanned by the list $1, z, \ldots, z^m$ of $m + 1$ vectors, so there cannot be a list of $m + 2$ linearly independent vectors in $\mathcal{P}_m(\mathbf{F})$. Suppose $p_0, \ldots, p_m$ is linearly independent in $\mathcal{P}_m(\mathbf{F})$. Moreover, notice that every linear combination of $p_0, \ldots, p_m$ is a polynomial such that

$$\sum_{i=0}^{m} \alpha_i p_i(2) = 0.$$

Since $q = 1 \in \mathcal{P}_m(\mathbf{F})$ is a polynomial where $q(2) \neq 0$, $q \notin \operatorname{span}(p_0, \ldots, p_n)$. By Exercise 2A13, the list $p_0, \ldots, p_m, q$ is of length $m + 2$ that is linearly independent, which is a contradiction. Hence, $p_0, \ldots, p_m$ cannot be linearly independent in $\mathcal{P}_m(\mathbf{F})$. ■

2B Bases

Exercise 2B1

Find all vector spaces that have exactly one basis.

Manual Remark. *The proof of this exercise fails if* $\mathbf{F}$ *contains only two elements. In this case, an additional vector space with exactly one basis is the vector space of dimension one over* $\mathbf{F}$*, since the only nonzero scalar is* 1*, and any nonzero vector* v *in the vector space is a basis.*

Solution. Let V be a vector space with exactly one basis, call it $v_1, \dots, v_n$. Consider that for $v \in V$, the list containing only v is linearly independent, so we may extend the list v to a basis for V. This means that $v = v_k$ for some $k \in \{1, \dots, n\}$, so every vector in V is one of $v_1, \dots, v_n$. For nonzero $\lambda \in \mathbf{F}$ and $k \in \{1, \dots, n\}$, consider $\lambda v_k \in V$. If $\lambda v_k = v_j$ for some $j \neq k$, then $v_1, \dots, v_n$ is not linearly independent, a contradiction. Therefore, $\lambda v_k = v_k$, and thus $(\lambda - 1)v_k = 0$. Because λ was arbitrary, it must be that $v_k = 0$. Because k was arbitrary, we have $v_k = 0$ for all $k \in \{1, \dots, n\}$, so the basis for V is empty. $\operatorname{span}(\,) = \{0\}$, so the trivial vector space $V = \{0\}$ is the only vector space with exactly one basis. ∎

Exercise 2B2

Verify all assertions in Example 2.27.

Solution.

(a) Suppose $\alpha_1, \dots, \alpha_n \in \mathbf{F}$ such that

$$\alpha_1(1, 0, \dots, 0) + \alpha_2(0, 1, 0, \dots, 0) + \cdots + \alpha_n(0, \dots, 0, 1) = 0.$$

Then each $\alpha_i = 0$, so the list is linearly independent. Moreover, any vector $(x_1, \dots, x_n) \in \mathbf{F}^n$ can be written as

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

so the list spans $\mathbf{F}^n$. Thus, the list is a basis of $\mathbf{F}^n$.

(b) If $(3, 5) = a(1, 2)$ for some $a \in \mathbf{F}$, then we must have $3 = a$ and $5 = 2a$, which is impossible. Thus, the list is linearly independent. Moreover, for any element $(x, y) \in \mathbf{F}^2$, we can solve the system of equations

$$\begin{aligned} x &= 3\alpha + \beta, \\ y &= 5\alpha + 2\beta \end{aligned}$$

for $\alpha, \beta \in \mathbf{F}$. Solving this system of equations, we find that

$$\alpha = 2x - y, \quad \beta = -5x + 3y.$$

Since we can find $\alpha, \beta \in \mathbf{F}$ for any $(x, y) \in \mathbf{F}^2$, the list spans $\mathbf{F}^2$. Thus, the list is a basis of $\mathbf{F}^2$.

(c) Note that $\mathbf{F}^3$ is spanned by the linearly independent list $(1, 0, 0), (0, 1, 0), (0, 0, 1)$. Since the length of a spanning list must be at least the length of a linearly independent list, the list $(1, 2, -4), (7, -5, 6)$ cannot span $\mathbf{F}^3$.

(d) From part (b), we saw that $(1, 2), (3, 5)$ is a basis of $\mathbf{F}^2$. Since $(4, 13) = -5(3, 5) + 19(1, 2)$, the list $(1, 2), (3, 5), (4, 13)$ is linearly dependent. Moreover, since $(1, 2), (3, 5)$ is a basis of $\mathbf{F}^2$, the list $(1, 2), (3, 5), (4, 13)$ spans $\mathbf{F}^2$. Thus, the list is not a basis of $\mathbf{F}^2$.

(e) Let U be the subspace in question. If $\alpha, \beta \in \mathbf{F}$ such that

$$\alpha(1, 1, 0) + \beta(0, 0, 1) = (0, 0, 0),$$

then we must have $\alpha = \beta = 0$, so the list is linearly independent. Moreover, for any $(x, x, y) \in U$, we can write

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

so the list spans U. Thus, the list is a basis of U.

(f) Let U be the subspace in question. If $\alpha, \beta \in \mathbf{F}$ such that

$$\alpha(1, -1, 0) + \beta(1, 0, -1) = (0, 0, 0),$$

then we must have $\alpha = \beta = 0$, so the list is linearly independent. Moreover, for any $(x, y, z) \in U$, note that $x + y + z = 0$, implying that $x = -y - z$. Then we can write

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

so the list spans U. Thus, the list is a basis of U.

(g) If $\alpha_0, \ldots, \alpha_m \in \mathbf{F}$ such that

$$\alpha_0 + \alpha_1 z + \cdots + \alpha_m z^m = 0,$$

then we must have $\alpha_0 = \cdots = \alpha_m = 0$, so the list is linearly independent. Moreover, any polynomial $p(z) \in \mathcal{P}_m(\mathbf{F})$ can be written as

$$p(z) = \alpha_0 + \alpha_1 z + \cdots + \alpha_m z^m,$$

so the list spans $\mathcal{P}_m(\mathbf{F})$. Thus, the list is a basis of $\mathcal{P}_m(\mathbf{F})$. ■

Exercise 2B3

(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 we can rewrite the subspace U as

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

Then we can write any vector in U as a linear combination of the following three vectors:

$$u_1 = (3, 1, 0, 0, 0), \quad u_2 = (0, 0, 7, 1, 0), \quad u_3 = (0, 0, 0, 0, 1).$$

Moreover, these three vectors are linearly independent, since the only way to write the zero vector as a linear combination of u_1, u_2, u_3 is to take all coefficients to be zero. Thus, $\{u_1, u_2, u_3\}$ is a basis of U.

(b) We can extend the basis $\{u_1, u_2, u_3\}$ of U to a basis of $\mathbf{R}^5$ by adding two more vectors that are linearly independent of u_1, u_2, u_3. For example, we can add the vectors

$$u_4 = (1, 0, 0, 0, 0), \quad u_5 = (0, 0, 1, 0, 0),$$

which are linearly independent of u_1, u_2, u_3: since u_1 is the only vector in the list with a nonzero first component,

u_4 cannot be a linear combination of u_1, u_2, u_3. Similarly, since u_2 is the only vector in the list with a nonzero third component, u_5 cannot be a linear combination of u_1, u_2, u_3, u_4. Thus, $\{u_1, u_2, u_3, u_4, u_5\}$ is a basis of $\mathbf{R}^5$.

(c) Let W be the subspace of $\mathbf{R}^5$ spanned by the vectors u_4 and u_5. Then we have

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

since $U \cap W = \{0\}$ and every vector in $\mathbf{R}^5$ can be written as a sum of a vector in U and a vector in W. ■

Exercise 2B4

(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 we can rewrite the subspace U as

$$U = \{(z_1, 6z_1, -2z_4 - 3z_5, z_4, z_5) \in \mathbf{C}^5 : z_1, z_4, z_5 \in \mathbf{C}\}.$$

Then we can write any vector in U as a linear combination of the following three vectors:

$$u_1 = (1, 6, 0, 0, 0), \quad u_2 = (0, 0, -2, 1, 0), \quad u_3 = (0, 0, -3, 0, 1).$$

Moreover, these three vectors are linearly independent, since the only way to write the zero vector as a linear combination of u_1, u_2, u_3 is to take all coefficients to be zero. Thus, $\{u_1, u_2, u_3\}$ is a basis of U.

(b) We can extend the basis $\{u_1, u_2, u_3\}$ of U to a basis of $\mathbf{C}^5$ by adding two more vectors that are linearly independent of u_1, u_2, u_3. For example, we can add the vectors

$$u_4 = (1, 0, 0, 0, 0), \quad u_5 = (0, 0, 1, 0, 0),$$

which are linearly independent of u_1, u_2, u_3: since u_1 is the only vector in the list with a nonzero first component, u_4 cannot be a linear combination of u_1, u_2, u_3. Similarly, u_2 and u_3 have nonzero third components, but any linear combination of u_2 and u_3 will have nonzero fourth and fifth components, so u_5 cannot be a linear combination of u_1, u_2, u_3, u_4. Thus, $\{u_1, u_2, u_3, u_4, u_5\}$ is a basis of $\mathbf{C}^5$.

(c) Let W be the subspace of $\mathbf{C}^5$ spanned by the vectors u_4 and u_5. Then we have

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

since $U \cap W = \{0\}$ and every vector in $\mathbf{C}^5$ can be written as a sum of a vector in U and a vector in W. ■

Exercise 2B5

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 $u_1, \ldots, u_m$ be a basis of U and $w_1, \ldots, w_n$ a basis of W. Then the list $u_1, \ldots, u_m, w_1, \ldots, w_n$ spans V, since any vector in V can be written as a sum of a vector in U and a vector in W. Because every spanning list of V can be

reduced to a basis of V, $u_1, \ldots, u_m, w_1, \ldots, w_n$ can be reduced to a basis of V consisting of vectors in $U \cup W$. ■

Exercise 2B6

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 the list is not a basis of $\mathcal{P}_3(\mathbf{F})$.

Solution. This is not true. Let $p_0 = 1, p_1 = x, p_2 = x^3, p_3 = x^2 + x^3$. Then none of the polynomials p_0, p_1, p_2, p_3 has degree 2. Moreover, the list p_0, p_1, p_2, p_3 is linearly independent, since the only way to write the zero polynomial as a linear combination of p_0, p_1, p_2, p_3 is to take all coefficients to be zero. Finally, noting that $x^2 = p_3 - p_2$, then for any polynomial $p(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 \in \mathcal{P}_3(\mathbf{F})$, we can write

$$p(x) = a_0 p_0 + a_1 p_1 + a_2(p_3 - p_2) + a_3 p_2 = a_0 p_0 + a_1 p_1 + (a_3 - a_2) p_2 + a_2 p_3,$$

which shows that the list spans $\mathcal{P}_3(\mathbf{F})$. Thus, the list p_0, p_1, p_2, p_3 is a basis of $\mathcal{P}_3(\mathbf{F})$. ■

Exercise 2B7

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.

Manual Remark. *The technique of transforming the scalars $\beta_1, \beta_2, \beta_3, \beta_4$ into the scalars $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ is good to remember. You will later learn that this is called a* change of basis.

Solution. Let $\alpha_1, \ldots, \alpha_4 \in \mathbf{F}$ such that

$$\alpha_1(v_1 + v_2) + \alpha_2(v_2 + v_3) + \alpha_3(v_3 + v_4) + \alpha_4 v_4 = 0.$$

Then we may write

$$\alpha_1 v_1 + (\alpha_1 + \alpha_2) v_2 + (\alpha_2 + \alpha_3) v_3 + (\alpha_3 + \alpha_4) v_4 = 0.$$

Since v_1, v_2, v_3, v_4 is a basis of V, we must have

$$\alpha_1 = 0, \quad \alpha_1 + \alpha_2 = 0, \quad \alpha_2 + \alpha_3 = 0, \quad \alpha_3 + \alpha_4 = 0.$$

From the first equation, we have $\alpha_1 = 0$. Substituting this into the second equation gives us $\alpha_2 = 0$. Substituting this into the third equation gives us $\alpha_3 = 0$. Finally, substituting this into the fourth equation gives us $\alpha_4 = 0$. Thus, the list is linearly independent. Moreover, since v_1, v_2, v_3, v_4 is a basis of V, then any vector $v \in V$ can be written as

$$v = \beta_1 v_1 + \beta_2 v_2 + \beta_3 v_3 + \beta_4 v_4$$

for some $\beta_1, \ldots, \beta_4 \in \mathbf{F}$. Using the expressions from proving the linear independence of the list, we may show that there exist $\alpha_1, \ldots, \alpha_4 \in \mathbf{F}$ such that

$$\begin{cases} \beta_1 = \alpha_1, \\ \beta_2 = \alpha_1 + \alpha_2, \\ \beta_3 = \alpha_2 + \alpha_3, \\ \beta_4 = \alpha_3 + \alpha_4. \end{cases}$$

Then $v \in V$ may be written as

$$v = \alpha_1(v_1 + v_2) + \alpha_2(v_2 + v_3) + \alpha_3(v_3 + v_4) + \alpha_4 v_4,$$

where

$$\alpha_1 = \beta_1, \quad \alpha_2 = \beta_2 - \beta_1, \quad \alpha_3 = \beta_3 - \beta_2 + \beta_1, \quad \alpha_4 = \beta_4 - \beta_3 + \beta_2 - \beta_1.$$

Thus, the list spans V. Therefore, the list is a basis of V. ■

Exercise 2B8

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. This is not true. Let $V = \mathbf{R}^4$ and v_1, v_2, v_3, v_4 be the standard basis of $\mathbf{R}^4$. Let U be the subspace of $\mathbf{R}^4$ defined by

$$U = \operatorname{span}(v_1, v_2, v_3 + v_4) = \{(x, y, z, z) \in \mathbf{R}^4 : x, y, z \in \mathbf{R}\}.$$

Then $v_1, v_2 \in U$ and $v_3, v_4 \notin U$. However, v_1, v_2 does not span U, since $(0, 0, 1, 1) \in U$ but cannot be written as a linear combination of v_1 and v_2. Thus, v_1, v_2 is not a basis of U. ■

Exercise 2B9

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. By Exercise 2A14, we know that $v_1, \ldots, v_m$ is linearly independent if and only if $w_1, \ldots, w_m$ is linearly independent. By Exercise 2A3, we know that $v_1, \ldots, v_m$ spans V if and only if $w_1, \ldots, w_m$ spans V. Thus, $v_1, \ldots, v_m$ is a basis of V if and only if $w_1, \ldots, w_m$ is a basis of V. ■

Exercise 2B10

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

$$u_1, \ldots, u_m, w_1, \ldots, w_n$$

is a basis of V.

Solution. Let $v \in V$. Since $V = U \oplus W$, there are unique $u \in U$ and $w \in W$ such that $v = u + w$. Since $u_1, \ldots, u_m$ spans U and $w_1, \ldots, w_n$ spans W, there exists scalars $\alpha_1, \ldots, \alpha_m$ and $\beta_1, \ldots, \beta_n$ such that

$$u = \sum_{i=1}^{m} \alpha_i u_i \quad \text{and} \quad w = \sum_{j=1}^{n} \beta_j w_j.$$

Therefore,

$$v = u + w = \sum_{i=1}^{m} \alpha_i u_i + \sum_{j=1}^{n} \beta_j w_j,$$

which shows that the list spans V. To show linear independence, suppose $v = 0$ so that $v = u + w$ imply $u = w = 0$. Since $u_1, \ldots, u_m$ and $w_1, \ldots, w_n$ are linearly independent, it follows that $\alpha_1 = \cdots = \alpha_m = 0$ and $\beta_1 = \cdots = \beta_n = 0$. Therefore,

the list $u_1, \dots, u_m, w_1, \dots, w_n$ is linearly independent. Combining this with the fact that it spans V, we conclude that it is a basis of V. ■

Exercise 2B11

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).
Axler Note. *Please see Exercise 1B8 for the definition of the complexification* $V_{\mathbf{C}}$.

Solution. Let $v_1, \dots, v_n$ be a basis of V. We claim that $v_1, \dots, v_n$ is also a basis of $V_{\mathbf{C}}$. Let $\alpha_1, \dots, \alpha_n \in \mathbf{C}$ such that

$$\alpha_1 v_1 + \cdots + \alpha_n v_n = 0.$$

Then we may write each $\alpha_k = a_k + b_k i$ for some $a_k, b_k \in \mathbf{R}$. Then we have

$$(a_1 + b_1 i)v_1 + \cdots + (a_n + b_n i)v_n = 0.$$

This implies that

$$(a_1 v_1 + \cdots + a_n v_n) + i(b_1 v_1 + \cdots + b_n v_n) = 0.$$

Since the left-hand side is a sum of two vectors in V, and V is a real vector space, we must have

$$a_1 v_1 + \cdots + a_n v_n = 0 \quad \text{and} \quad b_1 v_1 + \cdots + b_n v_n = 0.$$

Since $v_1, \dots, v_n$ is a basis of V, we have $a_1 = \cdots = a_n = 0$ and $b_1 = \cdots = b_n = 0$. Thus, $\alpha_1 = \cdots = \alpha_n = 0$, which shows that the list is linearly independent. Moreover, let $v \in V_{\mathbf{C}}$. Then we may write $v = u + wi$ for some $u, w \in V$. Since $v_1, \dots, v_n$ is a basis of V, we may write

$$u = \gamma_1 v_1 + \cdots + \gamma_n v_n \quad \text{and} \quad w = \delta_1 v_1 + \cdots + \delta_n v_n$$

for some $\gamma_1, \dots, \gamma_n, \delta_1, \dots, \delta_n \in \mathbf{R}$. Hence,

$$v = u + wi = (\gamma_1 + \delta_1 i)v_1 + \cdots + (\gamma_n + \delta_n i)v_n,$$

which shows that the list spans $V_{\mathbf{C}}$. Therefore, the list is a basis of $V_{\mathbf{C}}$. ■

2C Dimension

Manual Remark. *In Exercise 2C3 through Exercise 2C7, each problem gives a set U that satisfies some linear condition on the polynomials in $\mathcal{P}_4(\mathbf{F})$. In each case, the polynomials we exhibit span U, since expressing a general element of U in terms of the free parameters allowed by the defining condition writes it directly as a linear combination of them. They are also linearly independent: since they have different degrees, a linear combination of them can equal the zero polynomial only if every coefficient is zero, as otherwise the highest-degree term present could not cancel. We will therefore omit this reasoning going forward and simply exhibit a basis for each such U.*

Exercise 2C1

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. Since $\dim(\mathbf{R}^2) = 2$, any subspace of $\mathbf{R}^2$ must have dimension 0, 1, or 2. The only subspace of dimension 0 is $\{0\}$, and the only subspace of dimension 2 is $\mathbf{R}^2$. Any subspace with dimension 1 consists of all scalar multiples of some nonzero $(x, y) \in \mathbf{R}^2$, which is precisely the line with direction (x, y) containing the origin. Thus, the subspaces of $\mathbf{R}^2$ are precisely $\{0\}$, all lines in $\mathbf{R}^2$ containing the origin, and $\mathbf{R}^2$. ■

Exercise 2C2

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. Since $\dim(\mathbf{R}^3) = 3$, any subspace of $\mathbf{R}^3$ must have dimension 0, 1, 2, or 3. The only subspace of dimension 0 is $\{0\}$, and the only subspace of dimension 3 is $\mathbf{R}^3$. Any subspace with dimension 1 consists of all scalar multiples of some nonzero $(x, y, z) \in \mathbf{R}^3$, which is precisely the line containing the origin in the direction of (x, y, z). Any subspace with dimension 2 consists of all linear combinations of two linearly independent vectors in $\mathbf{R}^3$, which is precisely a plane containing the origin. Thus, 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 $\mathbf{R}^3$. ■

Exercise 2C3

(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 $\mathcal{P}_4(\mathbf{F}) = U \oplus W$.

Manual Remark. *Please see remark at the beginning of this section regarding the basis of U.*

Solution.

(a) Let $p = a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 \in U$. Then we have $p(6) = 0$, which gives us the equation

$$a_0 + 6a_1 + 36a_2 + 216a_3 + 1296a_4 = 0.$$

We can solve this equation for a_0 in terms of a_1, a_2, a_3, a_4:

$$a_0 = -6a_1 - 36a_2 - 216a_3 - 1296a_4.$$

Then we can write any polynomial in U as a linear combination of the following four polynomials:

$$u_1 = -6 + x, \quad u_2 = -36 + x^2, \quad u_3 = -216 + x^3, \quad u_4 = -1296 + x^4.$$

(b) We can extend the basis $\{u_1, u_2, u_3, u_4\}$ of U to a basis of $\mathcal{P}_4(\mathbf{F})$ by adding one more polynomial that is linearly independent of u_1, u_2, u_3, u_4. For example, we can add the polynomial

$$u_5 = 1,$$

which is linearly independent of u_1, u_2, u_3, u_4: since u_5 has no higher-degree terms, it cannot be expressed as a linear combination of u_1, u_2, u_3, u_4. Thus, $\{u_1, u_2, u_3, u_4, u_5\}$ is a basis of $\mathcal{P}_4(\mathbf{F})$.

(c) Let $W = \operatorname{span} u_5 = \operatorname{span} 1$. Then we have

$$\mathcal{P}_4(\mathbf{F}) = U + W,$$

since every polynomial in $\mathcal{P}_4(\mathbf{F})$ can be written as a sum of a polynomial in U and a polynomial in W. Moreover, if $p \in U \cap W$, then $p = \alpha$ for some $\alpha \in \mathbf{F}$. Since $p(6) = 0$, we have $\alpha = 0$. Thus, $U \cap W = \{0\}$, and so we have

$$\mathcal{P}_4(\mathbf{F}) = U \oplus W.$$ ■

Exercise 2C4

(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 $\mathcal{P}_4(\mathbf{R}) = U \oplus W$.

Manual Remark. *Please see remark at the beginning of this section regarding the basis of U.*

Solution.

(a) Let $p = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + a_4 x^4 \in U$. Then we have

$$p''(x) = 2a_2 + 6a_3 x + 12a_4 x^2,$$

and so $p''(6) = 0$ gives us the equation

$$2a_2 + 36a_3 + 432a_4 = 0.$$

We can solve this equation for a_2 in terms of a_3, a_4:

$$a_2 = -18a_3 - 216a_4.$$

Then we can write any polynomial in U as a linear combination of the following four polynomials:

$$u_1 = 1, \quad u_2 = x, \quad u_3 = -18x^2 + x^3, \quad u_4 = -216x^2 + x^4.$$

(b) We can extend the basis $\{u_1, u_2, u_3, u_4\}$ of U to a basis of $\mathcal{P}_4(\mathbf{R})$ by adding one more polynomial that is linearly independent of u_1, u_2, u_3, u_4. For example, we can add the polynomial

$$u_5 = x^2,$$

which is linearly independent of u_1, u_2, u_3, u_4: since only u_3 and u_4 have x^2 terms but contain other higher-degree terms, u_5 cannot be expressed as a linear combination of u_1, u_2, u_3, u_4. Thus, $\{u_1, u_2, u_3, u_4, u_5\}$ is a basis of $\mathcal{P}_4(\mathbf{R})$.

(c) Let $W = \operatorname{span} u_5 = \operatorname{span} x^2$. Then we have

$$\mathcal{P}_4(\mathbf{R}) = U + W,$$

since every polynomial in $\mathcal{P}_4(\mathbf{R})$ can be written as a sum of a polynomial in U and a polynomial in W. Moreover, if $p \in U \cap W$, then $p = \alpha x^2$ for some $\alpha \in \mathbf{R}$. Since $p''(6) = 0$, we have $2\alpha = 0$, which implies that $\alpha = 0$. Thus,

$U \cap W = \{0\}$, and so we have

$$\mathcal{P}_4(\mathbf{R}) = U \oplus W.$$

■

Exercise 2C5

(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 $\mathcal{P}_4(\mathbf{F}) = U \oplus W$.

Manual Remark. *Please see remark at the beginning of this section regarding the basis of U.*

Solution.

(a) Let $p = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + a_4 x^4 \in U$. Then we have $p(2) = p(5)$, which gives us the equation

$$a_0 + 2a_1 + 4a_2 + 8a_3 + 16a_4 = a_0 + 5a_1 + 25a_2 + 125a_3 + 625a_4.$$

We can simplify this equation to get

$$-3a_1 - 21a_2 - 117a_3 - 609a_4 = 0.$$

We can solve this equation for a_1 in terms of a_2, a_3, a_4:

$$a_1 = -7a_2 - 39a_3 - 203a_4.$$

Then we can write any polynomial in U as a linear combination of the following four polynomials:

$$u_1 = 1, \quad u_2 = -7x + x^2, \quad u_3 = -39x + x^3, \quad u_4 = -203x + x^4.$$

(b) We can extend the basis $\{u_1, u_2, u_3, u_4\}$ of U to a basis of $\mathcal{P}_4(\mathbf{F})$ by adding one more polynomial that is linearly independent of u_1, u_2, u_3, u_4. For example, we can add the polynomial

$$u_5 = x,$$

which is linearly independent of u_1, u_2, u_3, u_4: since only u_2, u_3, u_4 have x terms but contain other higher-degree terms, u_5 cannot be expressed as a linear combination of u_1, u_2, u_3, u_4. Thus, $\{u_1, u_2, u_3, u_4, u_5\}$ is a basis of $\mathcal{P}_4(\mathbf{F})$.

(c) Let $W = \operatorname{span} u_5 = \operatorname{span} x$. Then we have

$$\mathcal{P}_4(\mathbf{F}) = U + W,$$

since every polynomial in $\mathcal{P}_4(\mathbf{F})$ can be written as a sum of a polynomial in U and a polynomial in W. Moreover, if $p \in U \cap W$, then $p = \alpha x$ for some $\alpha \in \mathbf{F}$. Since $p(2) = p(5)$, we have $2\alpha = 5\alpha$, which implies that $\alpha = 0$. Thus, $U \cap W = \{0\}$, and so we have

$$\mathcal{P}_4(\mathbf{F}) = U \oplus W.$$

■

Exercise 2C6

(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 $\mathcal{P}_4(\mathbf{F}) = U \oplus W$.

Manual Remark. *Please see remark at the beginning of this section regarding the basis of U.*

Solution.

(a) Let $p = a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 \in U$. Then we have $p(2) = p(5) = p(6)$. We know from Exercise 2C5 that the condition $p(2) = p(5)$ gives us the equation

$$a_1 = -7a_2 - 39a_3 - 203a_4.$$

The condition $p(5) = p(6)$ gives us the equation

$$5a_1 + 25a_2 + 125a_3 + 625a_4 = 6a_1 + 36a_2 + 216a_3 + 1296a_4,$$

which simplifies to

$$a_1 + 11a_2 + 91a_3 + 671a_4 = 0.$$

Since we know that $a_1 = -7a_2 - 39a_3 - 203a_4$, we can substitute this into the second equation to get

$$4a_2 + 52a_3 + 468a_4 = 0.$$

We can solve this equation for a_2 in terms of a_3, a_4:

$$a_2 = -13a_3 - 117a_4.$$

Then $a_1 = 52a_3 + 616a_4$. Thus, we can write any polynomial in U as a linear combination of the following three polynomials:

$$u_1 = 1, \quad u_2 = 52x - 13x^2 + x^3, \quad u_3 = 616x - 117x^2 + x^4.$$

(b) We can extend the basis $\{u_1, u_2, u_3\}$ of U to a basis of $\mathcal{P}_4(\mathbf{F})$ by adding two more polynomials that are linearly independent of u_1, u_2, u_3. For example, we can add the polynomials

$$u_4 = x^2, \quad u_5 = x^3,$$

which are linearly independent of u_1, u_2, u_3: since only u_2 and u_3 have x^2 terms but contain other higher-degree terms, u_4 cannot be expressed as a linear combination of u_1, u_2, u_3. Similarly, u_5 has an x^3 term but no higher-degree terms, so it cannot be expressed as a linear combination of u_1, u_2, u_3, u_4. Thus, $\{u_1, u_2, u_3, u_4, u_5\}$ is a basis of $\mathcal{P}_4(\mathbf{F})$.

(c) Let $W = \operatorname{span}(u_4, u_5) = \operatorname{span}(x^2, x^3)$. Then we have

$$\mathcal{P}_4(\mathbf{F}) = U + W,$$

since every polynomial in $\mathcal{P}_4(\mathbf{F})$ can be written as a sum of a polynomial in U and a polynomial in W. Moreover, if $p \in U \cap W$, then $p = \alpha x^2 + \beta x^3$ for some $\alpha, \beta \in \mathbf{F}$. Since $p(2) = p(5) = p(6)$, we have

$$4\alpha + 8\beta = 25\alpha + 125\beta = 36\alpha + 216\beta.$$

Solving this system of equations, we find that $\alpha = \beta = 0$. Thus, $U \cap W = \{0\}$, and so we have

$$\mathcal{P}_4(\mathbf{F}) = U \oplus W.$$

■

Exercise 2C7

(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 $\mathcal{P}_4(\mathbf{R}) = U \oplus W$.

Manual Remark. *Please see remark at the beginning of this section regarding the basis of U.*

Solution.

(a) Let $p = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + a_4 x^4 \in U$. Then we have

$$\begin{aligned} \int_{-1}^{1} p &= \int_{-1}^{1} (a_0 + a_1 x + a_2 x^2 + a_3 x^3 + a_4 x^4)\,\mathrm{d}x \\ &= \int_{-1}^{1} (a_0 + a_2 x^2 + a_4 x^4)\,\mathrm{d}x \\ &= 2\int_{0}^{1} (a_0 + a_2 x^2 + a_4 x^4)\,\mathrm{d}x \\ &= 2\left(a_0 + \frac{a_2}{3} + \frac{a_4}{5}\right) = 0. \end{aligned}$$

We can solve this equation for a_0 in terms of a_2, a_4:

$$a_0 = -\frac{a_2}{3} - \frac{a_4}{5}.$$

Then we can write any polynomial in U as a linear combination of the following four polynomials:

$$u_1 = x, \quad u_2 = x^2 - \frac{1}{3}, \quad u_3 = x^3, \quad u_4 = x^4 - \frac{1}{5}.$$

(b) We can extend the basis $\{u_1, u_2, u_3, u_4\}$ of U to a basis of $\mathcal{P}_4(\mathbf{R})$ by adding one more polynomial that is linearly independent of u_1, u_2, u_3, u_4. For example, we can add the polynomial

$$u_5 = 1,$$

which is linearly independent of u_1, u_2, u_3, u_4: since u_5 has no higher-degree terms, it cannot be expressed as a linear combination of u_1, u_2, u_3, u_4. Thus, $\{u_1, u_2, u_3, u_4, u_5\}$ is a basis of $\mathcal{P}_4(\mathbf{R})$.

(c) Let $W = \operatorname{span} u_5 = \operatorname{span} 1$. Then we have

$$\mathcal{P}_4(\mathbf{R}) = U + W,$$

since every polynomial in $\mathcal{P}_4(\mathbf{R})$ can be written as a sum of a polynomial in U and a polynomial in W. Moreover, if $p \in U \cap W$, then $p = \alpha$ for some $\alpha \in \mathbf{R}$. Since the integral of a constant over a symmetric interval is nonzero unless the function is identically zero, we have $\alpha = 0$. Thus, $U \cap W = \{0\}$, and so we have

$$\mathcal{P}_4(\mathbf{R}) = U \oplus W.$$

■

Exercise 2C8

Suppose $v_1, \ldots, v_m$ is linearly independent in V and $w \in V$. Prove that

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

Solution. Consider the subspaces of V:

$$U = \operatorname{span}(v_1, \dots, v_m), \quad X = \operatorname{span}(v_1 + w, \dots, v_m + w), \quad W = \operatorname{span}(w).$$

Since $v_1, \dots, v_m$ is linearly independent, we have $\dim U = m$. Moreover, $\dim W = 1$. Lastly, note that $v_k = (v_k + w) - w \in X + W$ for each $k = 1, \dots, m$. Therefore, $U \subseteq X + W$, and since U is a subspace of dimension m, we have

$$\dim U \leq \dim(X + W) = \dim X + \dim W - \dim(X \cap W) \leq \dim X + \dim W,$$

which gives

$$m = \dim U \leq \dim X + \dim W = \dim X + 1 \implies \dim X \geq m - 1. \qquad \blacksquare$$

Exercise 2C9

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

Solution. Suppose that

$$\alpha_0 p_0 + \alpha_1 p_1 + \cdots + \alpha_m p_m = 0$$

for some $\alpha_0, \dots, \alpha_m \in \mathbf{F}$. Since p_m is the only polynomial with nonzero coefficient of x^m, we must have $\alpha_m = 0$. Next, since p_{m-1} is the only polynomial with nonzero coefficient of x^{m-1}, we must have $\alpha_{m-1} = 0$. Continuing in this way, we find that $\alpha_0 = \cdots = \alpha_m = 0$. Thus, the list is linearly independent. Moreover, since $\dim \mathcal{P}_m(\mathbf{F}) = m + 1$, any linearly independent list of $m + 1$ vectors in $\mathcal{P}_m(\mathbf{F})$ is a basis. Therefore, $p_0, p_1, \dots, p_m$ is a basis of $\mathcal{P}_m(\mathbf{F})$. $\blacksquare$

Exercise 2C10

Suppose m is a positive integer. For $0 \leq k \leq m$, let

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

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

Solution. Suppose that

$$\alpha_0 p_0 + \alpha_1 p_1 + \cdots + \alpha_m p_m = 0$$

for some $\alpha_0, \dots, \alpha_m \in \mathbf{F}$. Since p_0 is the only polynomial with nonzero constant term, we must have $\alpha_0 = 0$, hence the sum above becomes

$$\alpha_1 p_1 + \cdots + \alpha_m p_m = 0.$$

Next, since p_1 is the only polynomial with nonzero coefficient of x in the sum above, we must have $\alpha_1 = 0$. Continuing in this way, we find that $\alpha_0 = \cdots = \alpha_m = 0$. Thus, the list is linearly independent. Moreover, since $\dim \mathcal{P}_m(\mathbf{F}) = m + 1$, any linearly independent list of $m + 1$ vectors in $\mathcal{P}_m(\mathbf{F})$ is a basis. Therefore, $p_0, \dots, p_m$ is a basis of $\mathcal{P}_m(\mathbf{F})$. $\blacksquare$

Exercise 2C11

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. Let U and W be four-dimensional subspaces of $\mathbf{C}^6$. By the dimension formula for the sum of two subspaces, we have

$$\dim(U + W) = \dim U + \dim W - \dim(U \cap W) = 4 + 4 - \dim(U \cap W) = 8 - \dim(U \cap W).$$

Since $U + W$ is a subspace of $\mathbf{C}^6$, we have $\dim(U + W) \leq 6$. Thus, we have

$$8 - \dim(U \cap W) \leq 6,$$

which implies that $\dim(U \cap W) \geq 2$. Therefore, there exist at least two linearly independent vectors in $U \cap W$, and neither of these vectors is a scalar multiple of the other. ■

Exercise 2C12

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 the dimension formula for the sum of two subspaces, we have

$$\dim(U + W) = \dim U + \dim W - \dim(U \cap W) = 3 + 5 - \dim(U \cap W) = 8 - \dim(U \cap W).$$

Since $U + W = \mathbf{R}^8$, we have $\dim(U + W) = 8$. Thus, we have

$$8 - \dim(U \cap W) = 8,$$

which implies that $\dim(U \cap W) = 0$. Therefore, $U \cap W = \{0\}$, and so we have

$$\mathbf{R}^8 = U \oplus W.$$ ■

Exercise 2C13

Suppose U and W are both five-dimensional subspaces of $\mathbf{R}^8$. Prove that $U \cap W \neq \{0\}$.

Solution. By the dimension formula for the sum of two subspaces, we have

$$\dim(U + W) = \dim U + \dim W - \dim(U \cap W) = 5 + 5 - \dim(U \cap W) = 10 - \dim(U \cap W).$$

Since $U + W$ is a subspace of $\mathbf{R}^8$, we have $\dim(U + W) \leq 8$. Thus, we have

$$10 - \dim(U \cap W) \leq 8,$$

which implies that $\dim(U \cap W) \geq 2$. Therefore, $U \cap W$ is nontrivial, and so we have $U \cap W \neq \{0\}$. ■

Exercise 2C14

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. Let $W = V_1 \cap V_2$. By the dimension formula for the sum of two subspaces, we have

$$\dim(V_1 + V_2) = \dim V_1 + \dim V_2 - \dim W = 7 + 7 - \dim W = 14 - \dim W.$$

Since $V_1 + V_2$ is a subspace of V, we have $\dim(V_1 + V_2) \leq 10$. Thus, we have $14 - \dim W \leq 10$, which implies that $\dim W \geq 4$. Therefore, $V_1 \cap V_2$ is at least four-dimensional. Now, consider the subspace $W \cap V_3 = V_1 \cap V_2 \cap V_3$. By the dimension formula for the sum of two subspaces, we have

$$\dim(W + V_3) = \dim W + \dim V_3 - \dim(W \cap V_3) = \dim W + 7 - \dim(W \cap V_3).$$

Since $W + V_3$ is a subspace of V, we have $\dim(W + V_3) \leq 10$. Thus, we have

$$\dim W + 7 - \dim(W \cap V_3) \leq 10,$$

which implies that $\dim(W \cap V_3) \geq \dim W + 7 - 10 = \dim W - 3$. Since $\dim W \geq 4$, we have $\dim(W \cap V_3) \geq 1$. Therefore, $V_1 \cap V_2 \cap V_3$ is nontrivial, and so we have $V_1 \cap V_2 \cap V_3 \neq \{0\}$. ■

Exercise 2C15

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. Let $W = V_1 \cap V_2$. By the dimension formula for the sum of two subspaces, we have

$$\dim(V_1 + V_2) = \dim V_1 + \dim V_2 - \dim W.$$

Since $V_1 + V_2$ is a subspace of V, we have $\dim(V_1 + V_2) \leq \dim V$. Thus, we have

$$\dim V_1 + \dim V_2 - \dim W \leq \dim V,$$

which implies that $\dim W \geq \dim V_1 + \dim V_2 - \dim V$. Now, consider the subspace $W \cap V_3 = V_1 \cap V_2 \cap V_3$. By the dimension formula for the sum of two subspaces, we have

$$\dim(W + V_3) = \dim W + \dim V_3 - \dim(W \cap V_3).$$

Since $W + V_3$ is a subspace of V, we have $\dim(W + V_3) \leq \dim V$. Thus, we have

$$\dim W + \dim V_3 - \dim(W \cap V_3) \leq \dim V,$$

which implies that $\dim(W \cap V_3) \geq \dim W + \dim V_3 - \dim V$. Since $\dim W \geq \dim V_1 + \dim V_2 - \dim V$, we have

$$\dim(W \cap V_3) \geq (\dim V_1 + \dim V_2 - \dim V) + \dim V_3 - \dim V = \dim V_1 + \dim V_2 + \dim V_3 - 2 \dim V.$$

Since $\dim V_1 + \dim V_2 + \dim V_3 > 2 \dim V$, we have $\dim(W \cap V_3) > 0$. Therefore, $V_1 \cap V_2 \cap V_3$ is nontrivial, and so we have $V_1 \cap V_2 \cap V_3 \neq \{0\}$. ■

Exercise 2C16

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 be a subspace of V with $\dim U = m < n = \dim V$. We can extend a basis of U to a basis of V. Let $\{u_1, \ldots, u_m\}$ be a basis of U, and extend it to a basis of V by adding vectors $v_1, \ldots, v_{n-m}$. Then $\{u_1, \ldots, u_m, v_1, \ldots, v_{n-m}\}$ is a basis of V. For each $i = 1, \ldots, n - m$, let

$$W_i = \operatorname{span}(u_1, \ldots, u_m, v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_{n-m}).$$

By construction, W_i does not contain the extended basis vector v_i. Moreover, each W_i is a subspace of V with dimension $n - 1$, since it is spanned by $n - 1$ linearly independent vectors. Additionally, we have

$$U = W_1 \cap W_2 \cap \cdots \cap W_{n-m},$$

since any vector in the intersection containing nonzero v_i would not be in W_i, leaving behind only the basis vectors in U. Therefore, there exist $n - m$ subspaces of V, each of dimension $n - 1$, whose intersection equals U. ■

Exercise 2C17

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

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

Solution. We will prove the statement by induction on m. For the base case, let $m = 2$. Then the sum $V_1 + V_2$ is finite-dimensional. By the dimension formula for the sum of two subspaces, we have

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

Hence, the statement holds for $m = 2$. Now, suppose that the statement holds for some positive integer m. That is, suppose that if $V_1, \ldots, V_m$ are finite-dimensional subspaces of V, then $V_1 + \cdots + V_m$ is finite-dimensional and

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

We want to show that the statement holds for $m + 1$. Let $V_1, \ldots, V_{m+1}$ be finite-dimensional subspaces of V. By the induction hypothesis, we know that $W = V_1 + \cdots + V_m$ is finite-dimensional and

$$\dim W \leq \dim V_1 + \cdots + \dim V_m.$$

Then we have

$$\dim(W + V_{m+1}) = \dim W + \dim V_{m+1} - \dim(W \cap V_{m+1}).$$

Since $W \cap V_{m+1}$ is a subspace of V, we have $\dim(W \cap V_{m+1}) \geq 0$. Thus, we have

$$\dim(W + V_{m+1}) \leq \dim W + \dim V_{m+1} \leq \dim V_1 + \cdots + \dim V_m + \dim V_{m+1}.$$

Therefore, $V_1 + \cdots + V_{m+1}$ is finite-dimensional and

$$\dim(V_1 + \cdots + V_{m+1}) \leq \dim V_1 + \cdots + \dim V_m + \dim V_{m+1}.$$ ■

Exercise 2C18

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. Let $v_1, \ldots, v_n$ be a basis of V. For each $i = 1, \ldots, n$, let $V_i = \operatorname{span} v_i$. Then each V_i is a one-dimensional subspace of V, since it is spanned by a single nonzero vector. Moreover, we have

$$V = V_1 + \cdots + V_n,$$

and the sum is direct because the basis vectors are linearly independent. Therefore, we have

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

Exercise 2C19

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

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

Then either prove the formula above or give a counterexample.

Solution. One may guess that the given formula above is true by analogy with the formula for the number of elements in the union of three finite sets. However, this erroneously conflates the idea of subspace sums with the idea of set unions. In fact, the formula is not true in general. For example, let $V = \mathbf{R}^2$, and let

$$V_1 = \operatorname{span}((1,0)), \quad V_2 = \operatorname{span}((0,1)), \quad V_3 = \operatorname{span}((1,1)).$$

It is easy to see that $\dim V_1 = \dim V_2 = \dim V_3 = 1$, but $\dim(V_1 \cap V_2) = \dim(V_1 \cap V_3) = \dim(V_2 \cap V_3) = 0$ and $\dim(V_1 \cap V_2 \cap V_3) = 0$. Thus, the right-hand side of the formula is

$$1 + 1 + 1 - 0 - 0 - 0 + 0 = 3,$$

but the left-hand side is $\dim(V_1 + V_2 + V_3) = \dim \mathbf{R}^2 = 2$. Therefore, the formula is not true in general. ■

Exercise 2C20

Prove that if V_1, V_2, and V_3 are subspaces of a finite-dimensional vector space, then

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

Solution. Note that we need to account for all pairs of subspaces and all pairs of sums of subspaces, since the sum of subspaces is not equivalent to the union of sets. We can use the dimension formula repeatedly to obtain the answer. To that end, define the following subspaces:

$$W_1 = V_1 + V_2, \quad W_2 = V_1 + V_3, \quad W_3 = V_2 + V_3.$$

Applying the dimension formula to each of W_1, W_2, W_3, we have

$$\begin{aligned}\dim W_1 &= \dim V_1 + \dim V_2 - \dim(V_1 \cap V_2), \\ \dim W_2 &= \dim V_1 + \dim V_3 - \dim(V_1 \cap V_3), \\ \dim W_3 &= \dim V_2 + \dim V_3 - \dim(V_2 \cap V_3).\end{aligned}$$

Next, we can apply the dimension formula to each of the sums $W_1 + V_3, W_2 + V_2, W_3 + V_1$:

$$\begin{aligned}
\dim(W_1 + V_3) &= \dim W_1 + \dim V_3 - \dim(W_1 \cap V_3), \\
&= \dim V_1 + \dim V_2 - \dim(V_1 \cap V_2) + \dim V_3 - \dim(W_1 \cap V_3) \\
\dim(W_2 + V_2) &= \dim W_2 + \dim V_2 - \dim(W_2 \cap V_2), \\
&= \dim V_1 + \dim V_3 - \dim(V_1 \cap V_3) + \dim V_2 - \dim(W_2 \cap V_2) \\
\dim(W_3 + V_1) &= \dim W_3 + \dim V_1 - \dim(W_3 \cap V_1), \\
&= \dim V_2 + \dim V_3 - \dim(V_2 \cap V_3) + \dim V_1 - \dim(W_3 \cap V_1).
\end{aligned}$$

Note that $W_1 + V_3 = W_2 + V_2 = W_3 + V_1 = V_1 + V_2 + V_3$. Thus, we can add the three equations above and divide by 3 to obtain the desired formula:

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

■

3 Linear Maps

3A Vector Space of Linear Maps

Exercise 3A1

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. Suppose T is linear. Then

$$T(0, 0, 0) = (b, 0),$$

which implies that $b = 0$. Now consider the vector $(1, 1, 1) \in \mathbf{R}^3$. Then

$$T(2, 2, 2) = 2T(1, 1, 1) \implies (2, 12 + 8c) = (2, 12 + 2c).$$

Then $12 + 8c = 12 + 2c$, or $c = 0$.

Conversely, if $b = c = 0$, then $T(x, y, z) = (2x - 4y + 3z, 6x)$. Let $(x_1, y_1, z_1), (x_2, y_2, z_2) \in \mathbf{R}^3$ and $\alpha \in \mathbf{R}$. Then

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

and

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

Therefore, T is linear if and only if $b = c = 0$. ∎

Exercise 3A2

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. Suppose T is linear. Then for any $p \in \mathcal{P}(\mathbf{R})$, we have $T(2p) = 2Tp$. Looking at the first component of $T(2p)$, we have

$$3(2p)(4) + 5(2p)'(6) + b(2p)(1)(2p)(2) = 6p(4) + 10p'(6) + 4bp(1)p(2).$$

On the other hand, the first component of $2Tp$ is

$$2(3p(4) + 5p'(6) + bp(1)p(2)) = 6p(4) + 10p'(6) + 2bp(1)p(2).$$

Comparing the two expressions, we get $4bp(1)p(2) = 2bp(1)p(2)$ for all $p \in \mathcal{P}(\mathbf{R})$. Since there exist polynomials p such that $p(1)p(2) \neq 0$, we must have $b = 0$.

The second component of $T(2p)$ must be equal to the second component of $2Tp$. We have

$$\int_{-1}^{2} x^3(2p)(x)\,\mathrm{d}x + c\sin(2p(0)) = 2\int_{-1}^{2} x^3 p(x)\,\mathrm{d}x + 2c\sin(p(0)).$$

The integral terms may be subtracted, so we get $c\sin(2p(0)) = 2c\sin(p(0))$ for all $p \in \mathcal{P}(\mathbf{R})$. Consider $p(x) = \pi/2$. Then $\sin(2p(0)) = \sin(\pi) = 0$ and $2\sin(p(0)) = 2\sin(\pi/2) = 2$. Therefore, $c = 0$.

Conversely, if $b = c = 0$, let $p, q \in \mathcal{P}(\mathbf{R})$ and $\alpha \in \mathbf{R}$. Then

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

and

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

Therefore, T is linear if and only if $b = c = 0$. ■

Exercise 3A3

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

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

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

Solution. Let $e_1, \ldots, e_n$ be the standard basis of $\mathbf{F}^n$. Then for each $k = 1, \ldots, n$, we have

$$T(e_k) = (A_{1,k}, A_{2,k}, \ldots, A_{m,k})$$

for some scalars $A_{j,k} \in \mathbf{F}$. Now let $(x_1, \ldots, x_n) \in \mathbf{F}^n$. Then

$$\begin{aligned}
T(x_1, \ldots, x_n) &= T(x_1e_1 + \cdots + x_ne_n) \\
&= x_1T(e_1) + \cdots + x_nT(e_n) \\
&= x_1(A_{1,1}, A_{2,1}, \ldots, A_{m,1}) + \cdots + x_n(A_{1,n}, A_{2,n}, \ldots, A_{m,n}) \\
&= (A_{1,1}x_1 + \cdots + A_{1,n}x_n, A_{2,1}x_1 + \cdots + A_{2,n}x_n, \ldots, A_{m,1}x_1 + \cdots + A_{m,n}x_n).
\end{aligned}$$

Therefore,

$$T(x_1, \dots, x_n) = (A_{1,1}x_1 + \cdots + A_{1,n}x_n, A_{2,1}x_1 + \cdots + A_{2,n}x_n, \dots, A_{m,1}x_1 + \cdots + A_{m,n}x_n). \qquad \blacksquare$$

Exercise 3A4

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

Solution. Suppose that $a_1v_1 + \cdots + a_mv_m = 0$ for some scalars $a_1, \dots, a_m \in \mathbf{F}$. Applying T to both sides, we get

$$a_1Tv_1 + \cdots + a_mTv_m = T(0) = 0.$$

Since $Tv_1, \dots, Tv_m$ is linearly independent, it follows that $a_1 = \cdots = a_m = 0$. Therefore, $v_1, \dots, v_m$ is linearly independent. $\blacksquare$

Exercise 3A5

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

Solution. We must first ensure that the operations of addition and scalar multiplication are well-defined, i.e., if $S, T \in \mathcal{L}(V, W)$ and $\alpha \in \mathbf{F}$, then $S + T \in \mathcal{L}(V, W)$ and $\alpha T \in \mathcal{L}(V, W)$. We first show that $S + T$ is linear. Let $v, u \in V$. Then

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

Moreover, for any $v \in V$ and $\alpha \in \mathbf{F}$,

$$\begin{aligned}(S+T)(\alpha v) &= S(\alpha v) + T(\alpha v)\\ &= \alpha Sv + \alpha Tv\\ &= \alpha(Sv + Tv)\\ &= \alpha(S+T)v.\end{aligned}$$

This shows that $S + T$ is linear. Now we show that αT is linear. Let $v, u \in V$. Then

$$\begin{aligned}(\alpha T)(v+u) &= \alpha T(v+u)\\ &= \alpha(Tv + Tu)\\ &= \alpha Tv + \alpha Tu\\ &= (\alpha T)v + (\alpha T)u.\end{aligned}$$

Also, for any $v \in V$ and $\beta \in \mathbf{F}$,

$$\begin{aligned}(\alpha T)(\beta v) &= \alpha T(\beta v)\\ &= \alpha(\beta Tv)\\ &= (\alpha\beta)Tv\\ &= (\beta\alpha)Tv\\ &= \beta(\alpha Tv)\\ &= \beta(\alpha T)v.\end{aligned}$$

Therefore, $S + T \in \mathcal{L}(V, W)$ and $\alpha T \in \mathcal{L}(V, W)$. We now check the vector space axioms:

(1) **Commutativity**: Let $S, T \in \mathcal{L}(V, W)$. Then for any $v \in V$,

$$\begin{aligned}(S + T)v &= Sv + Tv \\ &= Tv + Sv \\ &= (T + S)v.\end{aligned}$$

(2) **Associativity**: Let $R, S, T \in \mathcal{L}(V, W)$. Then for any $v \in V$,

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

Similarly, for any $\alpha, \beta \in \mathbf{F}$ and $T \in \mathcal{L}(V, W)$,

$$\begin{aligned}((\alpha\beta)T)v &= (\alpha\beta)(Tv) \\ &= \alpha(\beta Tv) \\ &= \alpha((\beta T)v) \\ &= (\alpha(\beta T))v.\end{aligned}$$

(3) **Additive identity**: Let $0 \in \mathcal{L}(V, W)$ be the zero map defined by $0v = 0$ for all $v \in V$. Then for any $T \in \mathcal{L}(V, W)$ and $v \in V$,

$$(T + 0)v = Tv + 0v = Tv + 0 = Tv.$$

(4) **Inverse**: Let $T \in \mathcal{L}(V, W)$. Define $-T \in \mathcal{L}(V, W)$ by $(-T)v = -Tv$ for all $v \in V$. Then for any $v \in V$,

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

(5) **Multiplicative identity**: Let $1 \in \mathbf{F}$ be the multiplicative identity. Then for any $T \in \mathcal{L}(V, W)$ and $v \in V$,

$$(1T)v = 1(Tv) = Tv.$$

(6) **Distributive properties**: Let $T, S \in \mathcal{L}(V, W)$ and $\alpha, \beta \in \mathbf{F}$. Then for any $v \in V$,

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

and

$$\begin{aligned}((\alpha + \beta)T)v &= (\alpha + \beta)(Tv) \\ &= \alpha Tv + \beta Tv \\ &= (\alpha T)v + (\beta T)v \\ &= ((\alpha T) + (\beta T))v.\end{aligned}$$

■

Exercise 3A6

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

(a) **Associativity**: $(T_1T_2)T_3 = T_1(T_2T_3)$ whenever T_1, T_2, and T_3 are linear maps such that the products make sense (meaning T_3 maps into the domain of T_2, and T_2 maps into the domain of T_1).

(b) **Identity**: $TI = IT = T$ whenever $T \in \mathcal{L}(V, W)$; here the first I is the identity operator on V and the second I is the identity operator on W.

(c) **Distributive properties**: $(S_1 + S_2)T = S_1T + S_2T$ and $S(T_1 + T_2) = ST_1 + ST_2$ whenever $T, T_1, T_2 \in \mathcal{L}(U, V)$ and $S, S_1, S_2 \in \mathcal{L}(V, W)$.

Solution.

(a) Let $T_1 \in \mathcal{L}(V, W)$, $T_2 \in \mathcal{L}(U, V)$, and $T_3 \in \mathcal{L}(X, U)$. Then for any $x \in X$,

$$\begin{aligned}((T_1T_2)T_3)x &= (T_1T_2)(T_3x)\\ &= T_1(T_2(T_3x))\\ &= T_1((T_2T_3)x)\\ &= (T_1(T_2T_3))x.\end{aligned}$$

(b) Let $T \in \mathcal{L}(V, W)$. Then for any $v \in V$,

$$(TI)v = T(Iv) = Tv \quad \text{and} \quad (IT)v = I(Tv) = Tv.$$

(c) Let $T, T_1, T_2 \in \mathcal{L}(U, V)$ and $S, S_1, S_2 \in \mathcal{L}(V, W)$. Then for any $u \in U$,

$$\begin{aligned}((S_1 + S_2)T)u &= (S_1 + S_2)(Tu)\\ &= S_1(Tu) + S_2(Tu)\\ &= (S_1T)u + (S_2T)u\\ &= (S_1T + S_2T)u,\end{aligned}$$

and

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

■

Exercise 3A7

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. Suppose $\dim V = 1$ and let $v \in V$ be a nonzero vector. Then $\{v\}$ is a basis of V. Let $T \in \mathcal{L}(V)$. Since $Tv \in V$, there exists $\lambda \in \mathbf{F}$ such that $Tv = \lambda v$. Now let $u \in V$. Then there exists $\alpha \in \mathbf{F}$ such that $u = \alpha v$. Therefore,

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

■

Exercise 3A8

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.

Manual Remark. *This was a particularly difficult exercise because all obvious examples of functions that are homogeneous are also additive. The key idea is to utilize composition of functions.*

Solution. Let $\varphi : \mathbf{R}^2 \to \mathbf{R}$ be defined by

$$\varphi(x, y) = \sqrt[3]{x^3 + y^3}.$$

Then for any $a \in \mathbf{R}$ and $(x, y) \in \mathbf{R}^2$,

$$\varphi(a(x, y)) = \varphi(ax, ay) = \sqrt[3]{(ax)^3 + (ay)^3} = \sqrt[3]{a^3(x^3 + y^3)} = a\sqrt[3]{x^3 + y^3} = a\varphi(x, y).$$

However, φ is not additive. For example, let $v = (1, 0)$ and $w = (0, 1)$. Then

$$\varphi(v + w) = \varphi(1, 1) = \sqrt[3]{1^3 + 1^3} = \sqrt[3]{2},$$

but

$$\varphi(v) + \varphi(w) = \varphi(1, 0) + \varphi(0, 1) = \sqrt[3]{1^3 + 0^3} + \sqrt[3]{0^3 + 1^3} = 1 + 1 = 2.$$

Therefore, φ is not linear. ■

Exercise 3A9

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.)

Solution. Let $\varphi : \mathbf{C} \to \mathbf{C}$ be defined by

$$\varphi(x + iy) = x.$$

Then for any $w = x_1 + iy_1$ and $z = x_2 + iy_2$ in $\mathbf{C}$,

$$\varphi(w + z) = \varphi((x_1 + x_2) + i(y_1 + y_2)) = x_1 + x_2 = \varphi(w) + \varphi(z).$$

However, φ is not homogeneous. For example, let $a = i$ and $z = 1$. Then

$$\varphi(iz) = \varphi(i) = 0,$$

but

$$i\varphi(z) = i\varphi(1) = i(1) = i.$$

Therefore, φ is not linear. ■

Exercise 3A10

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 a linear map.

Solution. This is false. Let $q(x) = 1$ and define $T : \mathcal{P}(\mathbf{R}) \to \mathcal{P}(\mathbf{R})$ by $Tp = q \circ p$. Then for any $p_1, p_2 \in \mathcal{P}(\mathbf{R})$,

$$T(p_1 + p_2) = q \circ (p_1 + p_2) = 1 \neq 1 + 1 = Tp_1 + Tp_2.$$

Therefore, T is not linear. ∎

Exercise 3A11

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 T is a scalar multiple of the identity, i.e., $T = \lambda I$ for some $\lambda \in \mathbf{F}$. Then for any $S \in \mathcal{L}(V)$,

$$ST = S(\lambda I) = \lambda S = (\lambda I)S = TS.$$

Conversely, suppose $ST = TS$ for every $S \in \mathcal{L}(V)$. We claim that v and Tv are linearly dependent for every $v \in V$. Suppose not. Then there exists $v \in V$ such that v and Tv are linearly independent. Extend $\{v, Tv\}$ to a basis of V. Define $S \in \mathcal{L}(V)$ by sending v to itself and sending $Tv, v_3, \ldots, v_n$ to 0. Then

$$STv = S(Tv) = 0, \quad TSv = Tv \neq 0,$$

which contradicts the assumption that $ST = TS$. Therefore, v and Tv are linearly dependent for every $v \in V$.

Now let $v \in V$ be nonzero. Then there exists $\lambda_v \in \mathbf{F}$ such that $Tv = \lambda_v v$. We claim that λ_v is independent of the choice of v. Let $v_1, \ldots, v_n$ be a basis of V.

- If $n = 1$, then the claim is immediate, since $Tv_1 = \lambda_{v_1} v_1$.
- If $n \geq 2$, let v_i, v_j be two distinct basis vectors. Then their linear independence implies

$$\lambda_{v_i+v_j} v_i + \lambda_{v_i+v_j} v_j = T(v_i + v_j) = Tv_i + Tv_j = \lambda_{v_i} v_i + \lambda_{v_j} v_j.$$

  Since v_i and v_j are linearly independent, it follows that $\lambda_{v_i+v_j} = \lambda_{v_i} = \lambda_{v_j}$. Therefore, λ_v is independent of the choice of v, and we may denote it by λ.

Let $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$ be any vector in V. Then

$$Tv = T(\alpha_1 v_1 + \cdots + \alpha_n v_n) = \alpha_1 Tv_1 + \cdots + \alpha_n Tv_n = \alpha_1 \lambda v_1 + \cdots + \alpha_n \lambda v_n = \lambda(\alpha_1 v_1 + \cdots + \alpha_n v_n) = \lambda v.$$

Therefore, T is a scalar multiple of the identity. ∎

Exercise 3A12

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. Let $u \in U$ be such that $Su \neq 0$. Since $U \neq V$, there exists $v \in V$ such that $v \notin U$. Because $v = (v+u) - u \notin U$, then $v + u \notin U$ as well. Therefore,

$$T(u+v) = 0, \quad \text{but} \quad Tu + Tv = Su + 0 = Su \neq 0.$$

Hence, T is not linear. ∎

Exercise 3A13

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 $u \in U$.

Solution. Let $u_1, \ldots, u_m$ be a basis of U. Since U is a subspace of V, we can extend this basis to a basis of V, say $u_1, \ldots, u_m, v_1, \ldots, v_n$. Define $T : V \to W$ by

$$Tu_i = Su_i \text{ for all } i = 1, \ldots, m, \quad \text{and} \quad Tv_k = 0 \text{ for all } k = 1, \ldots, n.$$

By the linear map lemma, $T \in \mathcal{L}(V, W)$. Moreover, for any $u \in U$, there exist scalars $\alpha_1, \ldots, \alpha_m \in \mathbf{F}$ such that $u = \sum_{i=1}^{m} \alpha_i u_i$. Then

$$Tu = \sum_{i=1}^{m} \alpha_i Tu_i = \sum_{i=1}^{m} \alpha_i Su_i = Su$$

for all $u \in U$. ∎

Exercise 3A14

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. Let $v_1, \ldots, v_n$ be a basis of V. By Exercise 2A17, there exist a sequence $w_1, w_2, \ldots$ of vectors in W such that $w_1, \ldots, w_m$ is linearly independent for each $m \in \mathbb{N}$. For each $i \in \mathbb{N}$, define $T_i \in \mathcal{L}(V, W)$ by

$$T_i v_j = \begin{cases} w_i & \text{if } j = 1, \\ 0 & \text{if } j \neq 1. \end{cases}$$

In other words, T_i sends the first basis vector v_1 to w_i and all other basis vectors to 0. We now have a sequence of linear maps $T_1, T_2, \ldots$ in $\mathcal{L}(V, W)$. Suppose that

$$\sum_{\ell=1}^{m} \alpha_\ell T_\ell = 0$$

for some scalars $\alpha_1, \ldots, \alpha_m \in \mathbf{F}$. Applying both sides to v_1, we get

$$\sum_{\ell=1}^{m} \alpha_\ell T_\ell v_1 = \sum_{\ell=1}^{m} \alpha_\ell w_\ell = 0.$$

Since $w_1, \ldots, w_m$ is linearly independent, it follows that $\alpha_1 = \cdots = \alpha_m = 0$. It follows that $T_1, \ldots, T_m$ is linearly independent in $\mathcal{L}(V, W)$ for each $m \in \mathbb{N}$. Because there exists a sequence $T_1, T_2, \ldots$ of linearly independent linear maps in $\mathcal{L}(V, W)$ where $T_1, \ldots, T_m$ is linearly independent for each $m \in \mathbb{N}$, it follows by Exercise 2A17 that $\mathcal{L}(V, W)$ is infinite-dimensional. ∎

Exercise 3A15

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

Solution. Since $v_1, \dots, v_m$ is linearly dependent, there exist scalars $a_1, \dots, a_m \in \mathbf{F}$, not all zero, such that

$$a_1 v_1 + \cdots + a_m v_m = 0.$$

Fix some nonzero a_i among $a_1, \dots, a_m$. Let $w_1, \dots, w_m \in W$ be defined by

$$w_k = \begin{cases} 0 & \text{if } k \neq i, \\ w & \text{if } k = i, \end{cases}$$

where w is any nonzero vector in W. Suppose for contradiction that there exists $T \in \mathcal{L}(V, W)$ such that $Tv_k = w_k$ for each $k = 1, \dots, m$. Then

$$T(a_1 v_1 + \cdots + a_m v_m) = a_1 T v_1 + \cdots + a_m T v_m = a_1 w_1 + \cdots + a_m w_m.$$

Since $a_1 v_1 + \cdots + a_m v_m = 0$, we have $T(0) = 0$, so

$$0 = a_1 w_1 + \cdots + a_m w_m.$$

However, by definition of w_k, we have $a_k w_k = a_i w$ if $k = i$ and $a_k w_k = 0$ if $k \neq i$. Therefore, the sum $a_1 w_1 + \cdots + a_m w_m$ is a nonzero scalar multiple of w, which is nonzero since $w \neq 0$. This is a contradiction. Therefore, no such $T \in \mathcal{L}(V, W)$ exists. ∎

Exercise 3A16

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. Let v_1, v_2 be linearly independent vectors in V. We may extend this list to a basis of V, say $v_1, v_2, \dots, v_n$. Define $S, T \in \mathcal{L}(V)$ by

$$Sv_1 = v_2, \quad Sv_2 = v_1, \quad Sv_k = 0 \text{ for } k > 2,$$

and

$$Tv_1 = v_1, \quad Tv_2 = 0, \quad Tv_k = 0 \text{ for } k > 2.$$

Then

$$STv_1 = S(Tv_1) = Sv_1 = v_2,$$

but

$$TSv_1 = T(Sv_1) = Tv_2 = 0.$$

Therefore, $ST \neq TS$. ∎

Exercise 3A17

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

Axler Note. *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. Let $\mathcal{E}$ be a two-sided ideal of $\mathcal{L}(V)$. If $\mathcal{E} = \{0\}$, then we are done. Otherwise, there exists $E \in \mathcal{E}$ such that $E \neq 0$. Then there exists $v_1 \in V$ such that $Ev_1 \neq 0$. Since V is finite-dimensional, we may extend the list $\{v\}_1$ to a basis of V, say $v_1, v_2, \ldots, v_n$. Let $w_1 = Ev_1 \neq 0$. We may also extend the list $\{w\}_1$ to a basis of V, say $w_1, w_2, \ldots, w_n$. Define $T_1 \in \mathcal{L}(V)$ by

$$T_1 w_1 = v_1, \quad T_1 w_k = 0 \text{ for } k > 1.$$

Then $T_1 E \in \mathcal{E}$ since $\mathcal{E}$ is a two-sided ideal. Observe that $T_1 E v_1 = T_1(Ev_1) = T_1 w_1 = v_1 \neq 0$. Therefore, $T_1 E \neq 0$. Consider $S_1 \in \mathcal{L}(V)$ defined by

$$S_1 v_1 = v_1, \quad S_1 v_k = 0 \text{ for } k > 1.$$

Then $T_1 E S_1 \in \mathcal{E}$ since $\mathcal{E}$ is a two-sided ideal. Observe that $T_1 E S_1 v_1 = T_1 E v_1 = T_1 w_1 = v_1 \neq 0$. Therefore, $T_1 E S_1 \neq 0$. We now discuss the construction of the identity map $I \in \mathcal{L}(V)$.

Construct $T_i, S_i \in \mathcal{L}(V)$ for each $i = 1, \ldots, n$ as follows:

$$T_i w_1 = v_i, \quad T_i w_k = 0 \text{ for } k \neq 1,$$

and

$$S_i v_i = v_1, \quad S_i v_k = 0 \text{ for } k \neq i.$$

Then $T_i E S_i \in \mathcal{E}$ and $T_i E S_i v_i = v_i \neq 0$. Since $\mathcal{E}$ is a subspace, it contains all linear combinations of these maps. In particular, for any $v \in V$, we can write $v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n$ for some scalars $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbf{F}$. Observe that

$$T_i E S_i v = T_i E S_i(\alpha_1 v_1) + \cdots + T_i E S_i(\alpha_i v_i) + \cdots + T_i E S_i(\alpha_n v_n) = \alpha_i v_i.$$

Consider the map

$$E^* = T_1 E S_1 + T_2 E S_2 + \cdots + T_n E S_n \in \mathcal{E},$$

Then for any $v = \alpha_1 v_1 + \cdots + \alpha_n v_n \in V$,

$$E^* v = T_1 E S_1 v + T_2 E S_2 v + \cdots + T_n E S_n v = \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = v.$$

Hence, $E^* = I \in \mathcal{E}$. Since $\mathcal{E}$ is a two-sided ideal, $I = E^* \in \mathcal{E}$ implies that for any $T \in \mathcal{L}(V)$, we have $TI = T \in \mathcal{E}$. Therefore, $\mathcal{E} = \mathcal{L}(V)$. ∎

3B Null Spaces and Ranges

Exercise 3B1

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

Solution. Let $T \in \mathcal{L}(\mathbf{R}^5, \mathbf{R}^2)$ be defined by

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

Then $\operatorname{null} T = \{(0, 0, x_3, x_4, x_5) : x_3, x_4, x_5 \in \mathbf{R}\}$, which has dimension 3, and $\operatorname{range} T = \mathbf{R}^2$, which has dimension 2. ■

Exercise 3B2

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

Solution. Let $v \in V$. Then

$$(ST)^2 v = ST(STv) = S(T(STv)).$$

Since $STv \in \operatorname{range} S$ and $\operatorname{range} S \subseteq \operatorname{null} T$, we have $T(STv) = 0$. Therefore,

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

Since v was arbitrary, it follows that $(ST)^2 = 0$. ■

Exercise 3B3

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

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

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

Solution.

(a) The list $v_1, \ldots, v_m$ spans V if and only if $\operatorname{range} T = V$, i.e., T is surjective.
(b) The list $v_1, \ldots, v_m$ is linearly independent if and only if $\operatorname{null} T = \{0\}$, i.e., T is injective. ■

Exercise 3B4

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 $T_1, T_2 \in \mathcal{L}(\mathbf{R}^5, \mathbf{R}^4)$ be defined by

$$T_1(x_1, x_2, x_3, x_4, x_5) = (x_1, x_2, 0, 0)$$

and

$$T_2(x_1, x_2, x_3, x_4, x_5) = (0, 0, x_3, x_4).$$

Then $\operatorname{null} T_1 = \{(0, 0, x_3, x_4, x_5) : x_3, x_4, x_5 \in \mathbf{R}\}$ and $\operatorname{null} T_2 = \{(x_1, x_2, 0, 0, x_5) : x_1, x_2, x_5 \in \mathbf{R}\}$ both have dimension 3, so T_1 and T_2 are in the set $\{T \in \mathcal{L}(\mathbf{R}^5, \mathbf{R}^4) : \dim \operatorname{null} T > 2\}$. However,

$$(T_1 + T_2)(x_1, x_2, x_3, x_4, x_5) = (x_1, x_2, x_3, x_4),$$

which has null space $\operatorname{null}(T_1 + T_2) = \{(0, 0, 0, 0, x_5) : x_5 \in \mathbf{R}\}$ of dimension 1. Therefore,

$$\dim \operatorname{null}(T_1 + T_2) = 1 < 2,$$

so $T_1 + T_2$ is not in the set. Hence, the set is not closed under addition and is not a subspace of $\mathcal{L}(\mathbf{R}^5, \mathbf{R}^4)$. ■

Exercise 3B5

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

Solution. Let v_1, v_2, v_3, v_4 be the standard basis of $\mathbf{R}^4$. Define $T \in \mathcal{L}(\mathbf{R}^4)$ by

$$Tv_1 = v_3, \quad Tv_2 = v_4, \quad Tv_3 = 0, \quad Tv_4 = 0.$$

Then for any $x = (x_1, x_2, x_3, x_4) \in \mathbf{R}^4$,

$$Tx = x_1Tv_1 + x_2Tv_2 + x_3Tv_3 + x_4Tv_4 = x_1v_3 + x_2v_4.$$

Then $\operatorname{range} T = \operatorname{span}(v_3, v_4)$. If $Tx = 0$, then $x_1v_3 + x_2v_4 = 0$, which implies $x_1 = x_2 = 0$. Therefore, $\operatorname{null} T = \{(0, 0, x_3, x_4) : x_3, x_4 \in \mathbf{R}\} = \operatorname{span}(v_3, v_4)$. Hence, $\operatorname{range} T = \operatorname{null} T$. ■

Exercise 3B6

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

Solution. Suppose for contradiction that there exists $T \in \mathcal{L}(\mathbf{R}^5)$ such that $\operatorname{range} T = \operatorname{null} T$. Then by the fundamental theorem of linear maps,

$$\dim \mathbf{R}^5 = \dim \operatorname{null} T + \dim \operatorname{range} T = 2 \dim \operatorname{range} T.$$

Therefore, $5 = 2 \dim \operatorname{range} T$, which is impossible since 5 is odd and $2 \dim \operatorname{range} T$ is even. Hence, no such T exists. ■

Exercise 3B7

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

Solution. Let $\dim V = n$ and $\dim W = m$ with $2 \le n \le m$. Let $v_1, v_2, \ldots, v_n$ be a basis of V, and let $w_1, w_2, \ldots, w_m$ be a basis of W. Define $T_1, T_2 \in \mathcal{L}(V, W)$ by

$$T_1v_1 = w_1, \quad T_1v_k = 0 \text{ for } k \ge 2,$$

and

$$T_2v_1 = 0, \quad T_2v_k = w_k \text{ for } k \ge 2.$$

Then $v_2 \in \operatorname{null} T_1$ and $v_1 \in \operatorname{null} T_2$, so both T_1 and T_2 are not injective. However,

$$(T_1 + T_2)v_1 = T_1v_1 + T_2v_1 = w_1 + 0 = w_1 \neq 0,$$

and for $k \ge 2$,

$$(T_1 + T_2)v_k = T_1v_k + T_2v_k = 0 + w_k = w_k \neq 0.$$

Suppose $v \in V$ is such that $(T_1 + T_2)v = 0$. Then v can be expressed as a linear combination of the basis vectors:

$$v = \alpha_1v_1 + \alpha_2v_2 + \cdots + \alpha_nv_n.$$

Then

$$(T_1 + T_2)v = \alpha_1 w_1 + \alpha_2 w_2 + \cdots + \alpha_n w_n = 0.$$

Since $w_1, w_2, \ldots, w_m$ are linearly independent, it follows that $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$, hence $v = 0$. Therefore, $T_1 + T_2$ is injective. Since T_1 and T_2 are not injective, but their sum is injective, the set $\{T \in \mathcal{L}(V, W) : T \text{ is not injective}\}$ is not closed under addition and thus is not a subspace of $\mathcal{L}(V, W)$. ■

Exercise 3B8

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. We proceed similarly to the previous exercise. Let $\dim V = n$ and $\dim W = m$ with $n \geq m \geq 2$. Let $v_1, v_2, \ldots, v_n$ be a basis of V, and let $w_1, w_2, \ldots, w_m$ be a basis of W. Define $T_1, T_2 \in \mathcal{L}(V, W)$ by

$$T_1 v_1 = w_1, \quad T_1 v_k = 0 \text{ for } k \geq 2,$$

and

$$T_2 v_1 = 0, \quad T_2 v_k = w_k \text{ for } k = 2, \ldots, m, \quad T_2 v_k = 0 \text{ for } k > m.$$

Then $\operatorname{range} T_1 = \operatorname{span} w_1$ and $\operatorname{range} T_2 = \operatorname{span} w_2, \ldots, w_m$, both of which are proper subspaces of W, so T_1 and T_2 are not surjective. However,

$$(T_1 + T_2)v_1 = T_1 v_1 + T_2 v_1 = w_1 + 0 = w_1,$$

and for $k = 2, \ldots, m$,

$$(T_1 + T_2)v_k = T_1 v_k + T_2 v_k = 0 + w_k = w_k.$$

Therefore, $\operatorname{range}(T_1 + T_2) = \operatorname{span} w_1, w_2, \ldots, w_m = W$, so $T_1 + T_2$ is surjective. Since T_1 and T_2 are not surjective, but their sum is surjective, the set $\{T \in \mathcal{L}(V, W) : T \text{ is not surjective}\}$ is not closed under addition and thus is not a subspace of $\mathcal{L}(V, W)$. ■

Exercise 3B9

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

Solution. Suppose $a_1, \ldots, a_n \in \mathbf{F}$ are such that

$$a_1 T v_1 + \cdots + a_n T v_n = 0.$$

By linearity of T, we have

$$T(a_1 v_1 + \cdots + a_n v_n) = 0.$$

Injectivity of T implies that $\operatorname{null} T = \{0\}$, so

$$a_1 v_1 + \cdots + a_n v_n = 0.$$

Since $v_1, \ldots, v_n$ is linearly independent, it follows that $a_1 = \cdots = a_n = 0$. Therefore, $Tv_1, \ldots, Tv_n$ is linearly independent in W. ■

Exercise 3B10

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

Solution. Let $w \in \operatorname{range} T$. Then there exists $v \in V$ such that $Tv = w$. Since $v_1, \ldots, v_n$ spans V, we can write

$$v = a_1 v_1 + \cdots + a_n v_n$$

for some scalars $a_1, \ldots, a_n \in \mathbf{F}$. By linearity of T, we have

$$w = Tv = T(a_1 v_1 + \cdots + a_n v_n) = a_1 T v_1 + \cdots + a_n T v_n.$$

Therefore, w is a linear combination of $Tv_1, \ldots, Tv_n$. Since w was arbitrary in $\operatorname{range} T$, it follows that $Tv_1, \ldots, Tv_n$ spans $\operatorname{range} T$. ∎

Exercise 3B11

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

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

Manual Remark. *Note that this exercise proves the existence of a subspace U of V such that $V = U \oplus \operatorname{null} T$.*

Solution. By 2.33, because $\operatorname{null} T$ is a subspace of V with V finite dimensional, there exists a subspace U of V such that $V = \operatorname{null} T \oplus U$. Then $U \cap \operatorname{null} T = \{0\}$. Additionally, every $v \in V$ can be written uniquely as $v = n + u$ with $n \in \operatorname{null} T$ and $u \in U$. Hence,

$$\begin{aligned} \operatorname{range} T &= \{Tv : v \in V\} \\ &= \{T(n+u) : n \in \operatorname{null} T, u \in U\} \\ &= \{Tn + Tu : n \in \operatorname{null} T, u \in U\} \\ &= \{Tu : u \in U\}. \end{aligned}$$

∎

Exercise 3B12

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. We may write the null space of T as

$$\operatorname{null} T = \{(5x_2, x_2, 7x_4, x_4) : x_2, x_4 \in \mathbf{F}\}.$$

From this, we can see that $\operatorname{null} T$ is spanned by the vectors $(5, 1, 0, 0)$ and $(0, 0, 7, 1)$. Therefore, $\dim \operatorname{null} T = 2$. By the fundamental theorem of linear maps, we have

$$\dim \mathbf{F}^4 = \dim \operatorname{null} T + \dim \operatorname{range} T \implies 4 = 2 + \dim \operatorname{range} T.$$

Solving for $\dim \operatorname{range} T$, we find that $\dim \operatorname{range} T = 2$. Since $\dim \mathbf{F}^2 = 2$, it follows that $\operatorname{range} T = \mathbf{F}^2$. Therefore, T is surjective. ∎

Exercise 3B13

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. By the fundamental theorem of linear maps, we have

$$\dim \mathbf{R}^8 = \dim \operatorname{null} T + \dim \operatorname{range} T \implies 8 = 3 + \dim \operatorname{range} T.$$

Solving for $\dim \operatorname{range} T$, we find that $\dim \operatorname{range} T = 5$. Since $\dim \mathbf{R}^5 = 5$, it follows that $\operatorname{range} T = \mathbf{R}^5$. Therefore, T is surjective. ∎

Exercise 3B14

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. We can express the null space as

$$\operatorname{null} T = \{(3x_2, x_2, x_4, x_4, x_4) : x_2, x_4 \in \mathbf{F}\}.$$

From this, we can see that $\operatorname{null} T$ is spanned by the vectors $(3, 1, 0, 0, 0)$ and $(0, 0, 1, 1, 1)$. Therefore, $\dim \operatorname{null} T = 2$. By the fundamental theorem of linear maps, we have

$$\dim \mathbf{F}^5 = \dim \operatorname{null} T + \dim \operatorname{range} T \implies 5 = 2 + \dim \operatorname{range} T.$$

Solving for $\dim \operatorname{range} T$, we find that $\dim \operatorname{range} T = 3$. However, $\dim \mathbf{F}^2 = 2$, which is less than 3. This is a contradiction, as the dimension of the range cannot exceed the dimension of the target space. Therefore, no such linear map exists. ∎

Exercise 3B15

Suppose there exists a linear map on V whose null space and range are both finite-dimensional. Prove that V is finite-dimensional.

Manual Remark. *The spanning list in the solution is actually a basis of V.*

Solution. Let $T \in \mathcal{L}(V)$ be such that $\operatorname{null} T$ and $\operatorname{range} T$ are both finite-dimensional. Let $u_1, \ldots, u_m$ be a basis of $\operatorname{null} T$, and let $w_1, \ldots, w_n$ be a basis of $\operatorname{range} T$. Select $v_j \in V$ be such that $Tv_j = w_j$ for $j = 1, \ldots, n$. We claim that the list $u_1, \ldots, u_m, v_1, \ldots, v_n$ spans V.

Let $v \in V$. Then $Tv \in \operatorname{range} T$, so we may write

$$Tv = \alpha_1 w_1 + \cdots + \alpha_n w_n$$

for some scalars $\alpha_1, \ldots, \alpha_n \in \mathbf{F}$. Consider the vector

$$u = v - (\alpha_1 v_1 + \cdots + \alpha_n v_n).$$

Then

$$Tu = Tv - (\alpha_1 Tv_1 + \cdots + \alpha_n Tv_n) = Tv - (\alpha_1 w_1 + \cdots + \alpha_n w_n) = 0.$$

Therefore, $u \in \operatorname{null} T$, and we can write

$$u = \beta_1 u_1 + \cdots + \beta_m u_m$$

for some scalars $\beta_1, \ldots, \beta_m \in \mathbf{F}$. Thus,

$$v = u + (\alpha_1 v_1 + \cdots + \alpha_n v_n) = \beta_1 u_1 + \cdots + \beta_m u_m + \alpha_1 v_1 + \cdots + \alpha_n v_n,$$

which shows that the list spans V. Hence, V is finite-dimensional. ■

Exercise 3B16

Suppose V and W are both finite-dimensional. Prove that there exists an injective linear map from V to W if and only if $\dim V \leq \dim W$.

Solution. Suppose there exists an injective linear map $T \in \mathcal{L}(V, W)$. By contraposition of 3.22, an injective linear map $T \in \mathcal{L}(V, W)$ implies that $\dim V \leq \dim W$.

Conversely, suppose $\dim V \leq \dim W$. Let $n = \dim V$ and $m = \dim W$. Choose a basis $\{v_1, \ldots, v_n\}$ of V. Let $\{w_1, \ldots, w_m\}$ be a basis of W. Define a linear map $T \in \mathcal{L}(V, W)$ by

$$Tv_i = w_i \quad \text{for } i = 1, \ldots, n.$$

Since $n \leq m$, we can assign each basis vector of V to a distinct basis vector of W. To show that T is injective, let $v = \alpha_1 v_1 + \cdots + \alpha_n v_n \in V$ such that $Tv = 0$. Then

$$Tv = \alpha_1 Tv_1 + \cdots + \alpha_n Tv_n = \alpha_1 w_1 + \cdots + \alpha_n w_n = 0.$$

Since $w_1, \ldots, w_n$ are linearly independent, it follows that $\alpha_1 = \cdots = \alpha_n = 0$. Therefore, $v = 0$, and T is injective. Hence, there exists an injective linear map from V to W if and only if $\dim V \leq \dim W$. ■

Exercise 3B17

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. Suppose there exists a surjective linear map $T \in \mathcal{L}(V, W)$. By contraposition of 3.24, a surjective linear map $T \in \mathcal{L}(V, W)$ implies that $\dim V \geq \dim W$.

Conversely, suppose $\dim V \geq \dim W$. Let $n = \dim V$ and $m = \dim W$. Choose a basis $\{v_1, \ldots, v_n\}$ of V. Let $\{w_1, \ldots, w_m\}$ be a basis of W. Define a linear map $T \in \mathcal{L}(V, W)$ by

$$Tv_i = w_i \quad \text{for } i = 1, \ldots, m, \quad Tv_i = 0 \quad \text{for } i = m+1, \ldots, n.$$

To show that T is surjective, let $w \in W$. Then w can be expressed as a linear combination of the basis vectors of W:

$$w = \beta_1 w_1 + \cdots + \beta_m w_m.$$

Consider the vector $v = \beta_1 v_1 + \cdots + \beta_m v_m \in V$. Then

$$Tv = \beta_1 Tv_1 + \cdots + \beta_m Tv_m = \beta_1 w_1 + \cdots + \beta_m w_m = w.$$

Since w was arbitrary in W, it follows that $\operatorname{range} T = W$. Therefore, T is surjective. Hence, there exists a surjective linear map from V onto W if and only if $\dim V \geq \dim W$. ■

Exercise 3B18

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. Suppose there exists $T \in \mathcal{L}(V, W)$ such that $\operatorname{null} T = U$. By the fundamental theorem of linear maps, we have

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

Since $\operatorname{range} T$ is a subspace of W, we have $\dim \operatorname{range} T \leq \dim W$. Therefore,

$$\dim V = \dim U + \dim \operatorname{range} T \leq \dim U + \dim W.$$

Rearranging this inequality gives

$$\dim U \geq \dim V - \dim W.$$

Conversely, suppose $\dim U \geq \dim V - \dim W$. Let $n = \dim V$, $m = \dim W$, and $k = \dim U$. Let $u_1, \ldots, u_k$ be a basis of U. Extend this basis to a basis of V by adding vectors $v_1, \ldots, v_{n-k}$. Define a linear map $T \in \mathcal{L}(V, W)$ by

$$Tu_i = 0 \quad \text{for } i = 1, \ldots, k, \quad Tv_j = w_j \quad \text{for } j = 1, \ldots, n-k,$$

where $w_1, \ldots, w_{n-k}$ are linearly independent vectors in W. Since $\dim U = k$ and $\dim V - \dim W = n - m$, we have $k \geq n - m$, which implies that $n - k \leq m$. Therefore, we can choose $w_1, \ldots, w_{n-k}$ to be linearly independent in W. To show that $\operatorname{null} T = U$, let $v \in \operatorname{null} T$. Then $Tv = 0$. Express v as a linear combination of the basis vectors of V:

$$v = \alpha_1 u_1 + \cdots + \alpha_k u_k + \beta_1 v_1 + \cdots + \beta_{n-k} v_{n-k}.$$

Applying T to both sides, we have

$$Tv = \alpha_1 Tu_1 + \cdots + \alpha_k Tu_k + \beta_1 Tv_1 + \cdots + \beta_{n-k} Tv_{n-k} = 0 + \beta_1 w_1 + \cdots + \beta_{n-k} w_{n-k} = 0.$$

Since $w_1, \ldots, w_{n-k}$ are linearly independent, it follows that $\beta_1 = \cdots = \beta_{n-k} = 0$. Therefore,

$$v = \alpha_1 u_1 + \cdots + \alpha_k u_k \in U.$$

Hence, $\operatorname{null} T \subseteq U$. Conversely, if $v \in U$, then $Tv = 0$ by the definition of T. Therefore, $U \subseteq \operatorname{null} T$. Combining these two inclusions, we have $\operatorname{null} T = U$. Thus, there exists $T \in \mathcal{L}(V, W)$ such that $\operatorname{null} T = U$ if and only if $\dim U \geq \dim V - \dim W$. ∎

Exercise 3B19

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 T is injective. Since $\operatorname{range} T$ is a subspace of W, it is finite-dimensional. Let $\dim \operatorname{range} T = n$. If there is no linearly independent list of n vectors in V, then V is finite-dimensional, and $\dim V < n$, which contradicts the fundamental theorem of linear maps. Therefore, there exists a linearly independent list $v_1, \ldots, v_n$ of vectors in V.

Since T is injective, the list $Tv_1, \ldots, Tv_n$ is linearly independent in $\operatorname{range} T$ by Exercise 3B9. Additionally, the list has length n and so is a basis of $\operatorname{range} T$. By the linear map lemma, there exists $S_0 \in \mathcal{L}(\operatorname{range} T, V)$ such that $S_0(Tv_i) = v_i$ for

all $i = 1, \dots, n$. By Exercise 3A13, there exists an extension $S \in \mathcal{L}(W, V)$ of S_0. Since

$$(ST)v_i = S(Tv_i) = S_0(Tv_i) = v_i$$

for all $i = 1, \dots, n$, ST acts as the identity on a basis of V. By the linear map lemma, this map is unique, hence ST is the identity operator on V.

Conversely, suppose there exists $S \in \mathcal{L}(W, V)$ such that ST is the identity operator on V. Let $v \in \operatorname{null} T$. Then $Tv = 0$, and applying S gives

$$STv = S(0) = 0.$$

Since ST is the identity operator on V, we have $v = STv = 0$. Therefore, $\operatorname{null} T = \{0\}$, which means T is injective. ∎

Exercise 3B20

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 T is surjective, and let $w_1, \dots, w_n$ be a basis for W. By surjectivity of T, there are vectors $v_1, \dots, v_n \in V$ where $Tv_i = w_i$ for all $i = 1, \dots, n$. By the linear map lemma, there exists $S \in \mathcal{L}(W, V)$ such that

$$Sw_i = v_i \quad \text{for } i = 1, \dots, n.$$

Then

$$TSw_i = T(Sw_i) = Tv_i = w_i \quad \text{for } i = 1, \dots, n.$$

By the linear map lemma, because TS agrees with the identity on a basis of W, it must be the identity operator on W.

Conversely, suppose there exists $S \in \mathcal{L}(W, V)$ such that TS is the identity operator on W. Let $w \in W$. Then

$$TSw = w.$$

This means that for every $w \in W$, there exists a vector $v = Sw \in V$ such that $Tv = T(Sw) = w$. Therefore, $\operatorname{range} T = W$, which means that T is surjective. ∎

Exercise 3B21

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 $X = \{v \in V : Tv \in U\}$. We first show that X is a subspace of V. Let $v_1, v_2 \in X$ and $\alpha, \beta \in \mathbf{F}$. Then

$$T(\alpha v_1 + \beta v_2) = \alpha Tv_1 + \beta Tv_2.$$

Since $Tv_1, Tv_2 \in U$ and U is a subspace of W, it follows that $\alpha Tv_1 + \beta Tv_2 \in U$. Therefore, $\alpha v_1 + \beta v_2 \in X$, and thus X is a subspace of V.

Because V is finite-dimensional, so are $\operatorname{null} T$ and X. Let $v_1, \dots, v_m$ be a basis for $\operatorname{null} T$. Since U is a subspace of W, $0 \in U$, so $\operatorname{null} T \subseteq X$. We may then extend $v_1, \dots, v_m$ to a basis $v_1, \dots, v_m, v_{m+1}, \dots, v_n$ of X. If $w \in U \cap \operatorname{range} T$, then there exists $v \in V$ such that $Tv = w$, and since $v_1, \dots, v_n$ is a basis of X, we may write

$$w = Tv = T\left(\sum_{i=1}^{n} \alpha_i v_i\right) = \sum_{i=1}^{n} \alpha_i Tv_i = \sum_{i=m+1}^{n} \alpha_i Tv_i$$

for scalars $\alpha_1, \dots, \alpha_n \in \mathbf{F}$. Then $Tv_{m+1}, \dots, Tv_n$ span $U \cap \operatorname{range} T$. Additionally, if $\sum_{i=m+1}^{n} \alpha_i T v_i = 0$, then

$$T\left(\sum_{i=m+1}^{n} \alpha_i v_i\right) = \sum_{i=m+1}^{n} \alpha_i T v_i = 0,$$

which means that $\sum_{i=m+1}^{n} \alpha_i v_i \in \operatorname{null} T$, contradicting the fact that $v_1, \dots, v_n$ was linearly independent unless all $\alpha_{m+1}, \dots, \alpha_n$ are zero. Therefore, $Tv_{m+1}, \dots, Tv_n$ is linearly independent in $U \cap \operatorname{range} T$, so it is a basis for $U \cap \operatorname{range} T$. Hence,

$$\dim(U \cap \operatorname{range} T) = n - m = \dim X - \dim \operatorname{null} T,$$

where rearranging gives the desired formula. ■

Exercise 3B22

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. Let $u \in U$ such that $u \in \operatorname{null} ST$. Then $STu = 0$, which implies that $Tu \in \operatorname{null} S$. Therefore, we can express $\operatorname{null} ST$ as

$$\operatorname{null} ST = \{u \in U : Tu \in \operatorname{null} S\}.$$

Observe that this looks similar to the set in Exercise 3b21, where T is a linear map from U to V and $\operatorname{null} S$ is a subspace of V. By applying the result from Exercise 3b21, we have

$$\dim \operatorname{null} ST = \dim \operatorname{null} T + \dim(\operatorname{null} S \cap \operatorname{range} T).$$

Since $\operatorname{null} S \cap \operatorname{range} T$ is a subspace of $\operatorname{null} S$, we have

$$\dim(\operatorname{null} S \cap \operatorname{range} T) \leq \dim \operatorname{null} S.$$

Therefore,

$$\dim \operatorname{null} ST = \dim \operatorname{null} T + \dim(\operatorname{null} S \cap \operatorname{range} T) \leq \dim \operatorname{null} T + \dim \operatorname{null} S.$$ ■

Exercise 3B23

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. Let $u \in U$ such that $STu \in \operatorname{range} ST$. Then $STu = S(Tu)$, which implies that $Tu \in \operatorname{range} T$. Therefore, we can express $\operatorname{range} ST$ as

$$\operatorname{range} ST = \{S(Tu) : u \in U\}.$$

Since $Tu \in \operatorname{range} T$, we can rewrite this as

$$\operatorname{range} ST = \{Sv : v \in \operatorname{range} T\}.$$

This shows that $\operatorname{range} ST$ is a subset of $\operatorname{range} S$, and hence

$$\dim \operatorname{range} ST \leq \dim \operatorname{range} S.$$

Now let $v_1, \dots, v_m$ be a basis of range T. Then for any $u \in U$, we can express Tu as a linear combination of the basis vectors:

$$Tu = \alpha_1 v_1 + \cdots + \alpha_m v_m$$

for some scalars $\alpha_1, \dots, \alpha_m \in \mathbf{F}$. Applying S to both sides, we have

$$STu = S(Tu) = S(\alpha_1 v_1 + \cdots + \alpha_m v_m) = \alpha_1 Sv_1 + \cdots + \alpha_m Sv_m.$$

This shows that range ST is spanned by the vectors $Sv_1, \dots, Sv_m$. Therefore,

$$\dim \operatorname{range} ST \leq m = \dim \operatorname{range} T.$$

Combining the two inequalities, we have

$$\dim \operatorname{range} ST \leq \min(\dim \operatorname{range} S, \dim \operatorname{range} T).$$ ■

Exercise 3B24

(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) Since $ST = 0$, then $S(Tv) = 0$ for all $v \in V$. This implies that $\operatorname{range} T \subseteq \operatorname{null} S$. By the fundamental theorem of linear maps, we have

$$5 = \dim \operatorname{null} S + \dim \operatorname{range} S \geq \dim \operatorname{range} T + \dim \operatorname{range} S.$$

Now, by Exercise 3B23, we have

$$\dim \operatorname{range} TS \leq \min(\dim \operatorname{range} T, \dim \operatorname{range} S).$$

Suppose that $\dim \operatorname{range} TS > 2$. Then both $\dim \operatorname{range} T$ and $\dim \operatorname{range} S$ must be greater than 2, or equivalently, $\dim \operatorname{range} T \geq 3$ and $\dim \operatorname{range} S \geq 3$. This implies that

$$\dim \operatorname{range} T + \dim \operatorname{range} S \geq 3 + 3 = 6 > 5,$$

which is a contradiction. Therefore, $\dim \operatorname{range} TS \leq 2$.

(b) Let v_1, v_2, v_3, v_4, v_5 be a basis of $\mathbf{F}^5$. Define $S, T \in \mathcal{L}(\mathbf{F}^5)$ by

$$Sv_1 = v_1, \quad Sv_2 = v_2, \quad Sv_3 = Sv_4 = Sv_5 = 0,$$

and

$$Tv_1 = v_3, \quad Tv_2 = v_4, \quad Tv_3 = Tv_4 = Tv_5 = 0.$$

Then $ST = 0$ since $S(Tv_i) = S(0) = 0$ for all $i = 1, \dots, 5$. Now, we compute $\operatorname{range} TS$. We have

$$TSv_1 = T(Sv_1) = Tv_1 = v_3, \quad TSv_2 = T(Sv_2) = Tv_2 = v_4, \quad TSv_3 = TSv_4 = TSv_5 = 0.$$

Therefore, $\operatorname{range} TS = \operatorname{span}(v_3, v_4)$, which has dimension 2. Thus, we have constructed an example of $S, T \in \mathcal{L}(\mathbf{F}^5)$ with $ST = 0$ and $\dim \operatorname{range} TS = 2$. ■

Exercise 3B25

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 $T = ES$.

Manual Remark. *This exercise requires showing that some $R \in \mathcal{L}(\operatorname{range} S, W)$ is well-defined. An issue arises when T is not injective: for some $w \in \operatorname{range} S$, there may be distinct $v_1, v_2 \in V$ such that $Sv_1 = Sv_2 = w$. In this case, we must show that $Tv_1 = Tv_2$.*

Solution. Suppose $\operatorname{null} S \subseteq \operatorname{null} T$. Define a linear map $R : \operatorname{range} S \to W$ by $R(Sv) = Tv$ for all $v \in V$. We need to show that R is well-defined. Suppose $Sv_1 = Sv_2$ for some $v_1, v_2 \in V$. Then $S(v_1 - v_2) = 0$, which implies that $v_1 - v_2 \in \operatorname{null} S$. Since $\operatorname{null} S \subseteq \operatorname{null} T$, we have $v_1 - v_2 \in \operatorname{null} T$, which means that $T(v_1 - v_2) = 0$. Therefore, $Tv_1 = Tv_2$, and hence $R(Sv_1) = R(Sv_2)$. Thus, R is well-defined.

By Exercise 3A13, we can extend R to some $E \in \mathcal{L}(W)$. Then for any $v \in V$, we have

$$Tv = R(Sv) = E(Sv),$$

and hence $T = ES$.

Conversely, suppose there exists $E \in \mathcal{L}(W)$ such that $T = ES$. Let $v \in \operatorname{null} S$. Then $Sv = 0$, and applying E gives

$$Tv = E(Sv) = E(0) = 0.$$

Therefore, $v \in \operatorname{null} T$, and hence $\operatorname{null} S \subseteq \operatorname{null} T$. ∎

Exercise 3B26

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 $S = TE$.

Manual Remark. *From the previous exercise, the reader may be tempted to show that a map is well-defined again. However, this is not necessary here, since we are defining a map on the entire space V, not just on* range S.

Solution. Suppose $\operatorname{range} S \subseteq \operatorname{range} T$. Let $v_1, \ldots, v_n$ be a basis of V. For each $i = 1, \ldots, n$, since $Sv_i \in \operatorname{range} S \subseteq \operatorname{range} T$, there exists $u_i \in V$ such that $Tu_i = Sv_i$. Define a linear map $E \in \mathcal{L}(V)$ by

$$Ev_i = u_i \quad \text{for } i = 1, \ldots, n.$$

Then for any $v \in V$, we can express v as a linear combination of the basis vectors:

$$v = \alpha_1 v_1 + \cdots + \alpha_n v_n$$

for some scalars $\alpha_1, \ldots, \alpha_n \in \mathbf{F}$. Applying S and TE to both sides, we have

$$Sv = \alpha_1 Sv_1 + \cdots + \alpha_n Sv_n = \alpha_1 Tu_1 + \cdots + \alpha_n Tu_n = T(\alpha_1 u_1 + \cdots + \alpha_n u_n) = TEv.$$

Therefore, $S = TE$.

Conversely, suppose there exists $E \in \mathcal{L}(V)$ such that $S = TE$. Let $w \in \operatorname{range} S$. Then there exists $v \in V$ such that $Sv = w$. Applying the definition of S, we have

$$w = Sv = TE(v) = T(E(v)).$$

Therefore, $w \in \operatorname{range} T$, and hence $\operatorname{range} S \subseteq \operatorname{range} T$. ∎

Exercise 3B27

Suppose $P \in \mathcal{L}(V)$ and $P^2 = P$. Prove that $V = \operatorname{null} P \oplus \operatorname{range} P$.

Solution. Let $v \in V$. We can express v as

$$v = (v - Pv) + Pv.$$

Observe that $Pv \in \operatorname{range} P$ by definition. Now, we need to show that $v - Pv \in \operatorname{null} P$. Applying P to $v - Pv$, we have

$$P(v - Pv) = Pv - P^2v = Pv - Pv = 0.$$

Therefore, $v - Pv \in \operatorname{null} P$. This shows that every vector $v \in V$ can be expressed as a sum of a vector in $\operatorname{null} P$ and a vector in $\operatorname{range} P$, which implies that

$$V = \operatorname{null} P + \operatorname{range} P.$$

Next, we need to show that the sum is direct. Suppose $u \in \operatorname{null} P \cap \operatorname{range} P$. Then there exists some $w \in V$ such that $u = Pw$. Since $u \in \operatorname{null} P$, we have

$$Pu = P(Pw) = P^2w = Pw = u.$$

But since $u \in \operatorname{null} P$, we also have $Pu = 0$. Therefore, we have

$$u = Pu = 0.$$

This shows that the intersection of $\operatorname{null} P$ and $\operatorname{range} P$ is trivial, i.e., $\operatorname{null} P \cap \operatorname{range} P = \{0\}$. Hence, the sum is direct, and we conclude that

$$V = \operatorname{null} P \oplus \operatorname{range} P.$$ ■

Exercise 3B28

Suppose $D \in \mathcal{L}(\mathcal{P}(\mathbf{R}))$ is such that $\deg Dp = (\deg p) - 1$ for every nonconstant polynomial $p \in \mathcal{P}(\mathbf{R})$. Prove that D is surjective.

Manual Remark. *Note that $\mathcal{P}(\mathbf{R})$ is infinite-dimensional, so we cannot use the fundamental theorem of linear maps directly. Instead, one should consider finite-dimensional subspaces of $\mathcal{P}(\mathbf{R})$.*

Solution. If $p = 0$, then $p = D(0) \in \operatorname{range} D$ by linearity. Suppose $p \neq 0$ and let $n = \deg p$. Consider the list $D(x), D(x^2), \ldots, D(x^{n+1})$. Since x^k is nonconstant for each $k \in \{1, \ldots, n+1\}$, we have $\deg D(x^k) = k - 1$. Thus, $D(x), D(x^2), \ldots, D(x^{n+1})$ is a list of $n+1$ polynomials with pairwise distinct degrees, which implies linear independence. Since $\dim \mathcal{P}_n(\mathbf{R}) = n + 1$, this list is a basis for $\mathcal{P}_n(\mathbf{R})$. In particular, $p \in \mathcal{P}_n(\mathbf{R})$, so there exist scalars $c_1, \ldots, c_{n+1} \in \mathbf{R}$ such that

$$p = \sum_{k=1}^{n+1} c_k D(x^k) = D\left(\sum_{k=1}^{n+1} c_k x^k\right) \in \operatorname{range} D.$$

Since p was arbitrary, D is surjective. ■

Exercise 3B29 ✔

Suppose $p \in \mathcal{P}(\mathbf{R})$. Prove that there exists a polynomial $q \in \mathcal{P}(\mathbf{R})$ such that $5q'' + 3q' = p$.

Solution. Define $D \in \mathcal{L}(\mathcal{P}(\mathbf{R}))$ by $Dq = 5q'' + 3q'$ for all $q \in \mathcal{P}(\mathbf{R})$. For nonconstant q, we have $\deg Dq = \deg q - 1$. By Exercise 3B28, D is surjective. Therefore, for any polynomial $p \in \mathcal{P}(\mathbf{R})$, there exists a polynomial $q \in \mathcal{P}(\mathbf{R})$ such that $Dq = p$, which means that $5q'' + 3q' = p$. ■

Exercise 3B30

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. Let $v \in V$. We want to show that v can be expressed as a sum of a vector in $\operatorname{null} \varphi$ and a scalar multiple of u. Since $\varphi(u) \neq 0$, we can define a scalar

$$a = \frac{\varphi(v)}{\varphi(u)}.$$

Then consider the vector

$$w = v - au.$$

We have

$$\varphi(w) = \varphi(v - au) = \varphi(v) - a\varphi(u) = \varphi(v) - \frac{\varphi(v)}{\varphi(u)}\varphi(u) = 0.$$

Therefore, $w \in \operatorname{null} \varphi$. Thus, we can express v as

$$v = w + au,$$

where $w \in \operatorname{null} \varphi$ and au is a scalar multiple of u. This shows that every vector in V can be expressed as a sum of a vector in $\operatorname{null} \varphi$ and a scalar multiple of u, which implies that

$$V = \operatorname{null} \varphi + \{au : a \in \mathbf{F}\}.$$

Next, we need to show that the sum is direct. Suppose $w \in \operatorname{null} \varphi \cap \{au : a \in \mathbf{F}\}$. Then there exists a scalar $a \in \mathbf{F}$ such that $w = au$. Since $w \in \operatorname{null} \varphi$, we have

$$0 = \varphi(w) = \varphi(au) = a\varphi(u).$$

Since $\varphi(u) \neq 0$, it follows that $a = 0$. Therefore, $w = 0$, which shows that the intersection of $\operatorname{null} \varphi$ and $\{au : a \in \mathbf{F}\}$ is trivial. Hence, the sum is direct, and we conclude that

$$V = \operatorname{null} \varphi \oplus \{au : a \in \mathbf{F}\}.$$ ■

Exercise 3B31

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. Suppose there exists $T \in \mathcal{L}(V, W)$ such that $\operatorname{null} T = X$ and $\operatorname{range} T = Y$. By the fundamental theorem of linear maps, we have

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

Conversely, suppose $\dim X + \dim Y = \dim V$. Let $n = \dim V$, $k = \dim X$, and $m = \dim Y$. Then $n = k + m$. Choose a basis $\{x_1, \ldots, x_k\}$ of X and extend it to a basis $\{x_1, \ldots, x_k, v_1, \ldots, v_m\}$ of V. Let $\{y_1, \ldots, y_m\}$ be a basis of Y. Define a linear map $T \in \mathcal{L}(V, W)$ by

$$Tx_i = 0 \quad \text{for } i = 1, \ldots, k,$$

and

$$Tv_j = y_j \quad \text{for } j = 1, \ldots, m.$$

Then for any $v \in V$, we can express v as a linear combination of the basis vectors:

$$v = \alpha_1 x_1 + \cdots + \alpha_k x_k + \beta_1 v_1 + \cdots + \beta_m v_m$$

for some scalars $\alpha_1, \ldots, \alpha_k, \beta_1, \ldots, \beta_m \in \mathbf{F}$. Applying T to both sides gives

$$Tv = \alpha_1 T x_1 + \cdots + \alpha_k T x_k + \beta_1 T v_1 + \cdots + \beta_m T v_m = 0 + \beta_1 y_1 + \cdots + \beta_m y_m.$$

Therefore, $Tv \in Y$. Now suppose $y = \gamma_1 y_1 + \cdots + \gamma_m y_m \in Y$. Then

$$T(\gamma_1 v_1 + \cdots + \gamma_m v_m) = \gamma_1 y_1 + \cdots + \gamma_m y_m = y.$$

Hence, range $T = Y$. Moreover, if $Tv = 0$, then $\beta_1 y_1 + \cdots + \beta_m y_m = 0$. Since $y_1, \ldots, y_m$ are linearly independent, it follows that $\beta_1 = \cdots = \beta_m = 0$. Thus, $v \in X$. Additionally, if $v \in X$, then $v = \alpha_1 x_1 + \cdots + \alpha_k x_k$ for some scalars $\alpha_1, \ldots, \alpha_k \in \mathbf{F}$, and hence $Tv = 0$. Therefore, null $T = X$. This shows that there exists a linear map $T \in \mathcal{L}(V, W)$ such that null $T = X$ and range $T = Y$. ■

Exercise 3B32

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 $\varphi = 0$.

Axler Note. *The description of the two-sided ideals of $\mathcal{L}(V)$ given by Exercise 3A17 might be useful.*

Solution. Observe that null φ is a two-sided ideal of $\mathcal{L}(V)$ as follows: if $X \in \text{null } \varphi$ and $Y \in \mathcal{L}(V)$, then

$$\varphi(XY) = \varphi(X)\varphi(Y) = 0 \cdot \varphi(Y) = 0, \quad \varphi(YX) = \varphi(Y)\varphi(X) = \varphi(Y) \cdot 0 = 0.$$

Since $\dim V > 1$, then Exercise 3A16 gives the existence of $S, T \in \mathcal{L}(V)$ such that $ST \neq TS$. Then we have

$$\varphi(ST - TS) = \varphi(S)\varphi(T) - \varphi(T)\varphi(S) = 0.$$

Therefore, $ST - TS \in \text{null } \varphi$. Since null φ is a two-sided ideal of $\mathcal{L}(V)$, then by Exercise 3A17, we have null $\varphi = \mathcal{L}(V)$. Hence, $\varphi = 0$. ■

Exercise 3B33

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 range $T_{\mathbf{C}} = W_{\mathbf{C}}$ if and only if range $T = W$.

Axler Note. *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 $u + vi$ and $w + xi \in V_{\mathbf{C}}$. Then for any $a + bi \in \mathbf{C}$, we have

$$\begin{aligned} T_{\mathbf{C}}((a+bi)(u+vi)+(w+xi)) &= T_{\mathbf{C}}((au-bv+(av+bu)i)+(w+xi)) \\ &= T(au-bv+w)+iT(av+bu+x) \\ &= aTu-bTv+Tw+i(aTv+bTu+Tx) \\ &= (a+bi)(Tu+iTv)+(Tw+iTx) \\ &= (a+bi)T_{\mathbf{C}}(u+vi)+T_{\mathbf{C}}(w+xi). \end{aligned}$$

Therefore, $T_{\mathbf{C}}$ is a complex linear map from $V_{\mathbf{C}}$ to $W_{\mathbf{C}}$.

(b) Suppose $T_{\mathbf{C}}$ is injective. Let $v \in V$ be such that $Tv = 0$. Then

$$T_{\mathbf{C}}(v + i0) = Tv + iT0 = 0 + i0 = 0.$$

Since $T_{\mathbf{C}}$ is injective, we have $v + i0 = 0$, which implies that $v = 0$. Therefore, T is injective.

Conversely, suppose T is injective. Let $u + iv \in V_{\mathbf{C}}$ be such that $T_{\mathbf{C}}(u + iv) = 0$. Then

$$Tu + iTv = 0.$$

This implies that $Tu = 0$ and $Tv = 0$. Since T is injective, we have $u = 0$ and $v = 0$. Therefore, $u + iv = 0$, and hence $T_{\mathbf{C}}$ is injective.

(c) Suppose $\operatorname{range} T_{\mathbf{C}} = W_{\mathbf{C}}$. Let $w \in W$. Then $w + i0 \in W_{\mathbf{C}}$, and since $\operatorname{range} T_{\mathbf{C}} = W_{\mathbf{C}}$, there exists $u + iv \in V_{\mathbf{C}}$ such that

$$T_{\mathbf{C}}(u + iv) = Tu + iTv = w + i0.$$

This implies that $Tu = w$ and $Tv = 0$. Therefore, $w \in \operatorname{range} T$, and hence $\operatorname{range} T = W$.

Conversely, suppose $\operatorname{range} T = W$. Let $w_1 + iw_2 \in W_{\mathbf{C}}$. Since $\operatorname{range} T = W$, there exist $u, v \in V$ such that $Tu = w_1$ and $Tv = w_2$. Then

$$T_{\mathbf{C}}(u + iv) = Tu + iTv = w_1 + iw_2.$$

Therefore, $w_1 + iw_2 \in \operatorname{range} T_{\mathbf{C}}$, and hence $\operatorname{range} T_{\mathbf{C}} = W_{\mathbf{C}}$. ∎

3C Matrices

Exercise 3C1

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. Let $A = \mathcal{M}(T)$ be the matrix with respect to these bases. Then the j-th column of A is the coordinate vector of Tv_j with respect to the basis for W.

Suppose A had fewer than $\dim \operatorname{range} T$ nonzero columns. Since zero columns do not contribute to the span, the columns of A would then span a space of dimension less than $\dim \operatorname{range} T$. This is a contradiction, since fewer than $\dim \operatorname{range} T$ vectors cannot span $\operatorname{range} T$. Therefore, A must have at least $\dim \operatorname{range} T$ nonzero columns, and hence at least $\dim \operatorname{range} T$ nonzero entries. ∎

Exercise 3C2

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 $\dim \operatorname{range} T = 1$. Then there exists a nonzero vector $y_1 \in W$ such that $\operatorname{range} T = \operatorname{span} y_1$. Let $\dim V = n$ and $\dim W = m$. Extend y_1 to a basis $y_1, \ldots, y_m$ of W, and define the following vectors in W:

$$w_1 = y_1 - (y_2 + \cdots + y_m), \quad w_k = y_k \text{ for } k = 2, \ldots, m.$$

It is clear that $w_1 + w_2 + \cdots + w_m = y_1$, and hence $w_1, \ldots, w_m$ spans W. Now suppose for scalars $\alpha_1, \ldots, \alpha_m \in \mathbf{F}$, we have

$$\alpha_1 w_1 + \alpha_2 w_2 + \cdots + \alpha_m w_m = 0.$$

Then we can express this as

$$\alpha_1(y_1 - (y_2 + \cdots + y_m)) + \alpha_2 y_2 + \cdots + \alpha_m y_m = 0.$$

Simplifying, we get

$$(\alpha_1)y_1 + (-\alpha_1 + \alpha_2)y_2 + \cdots + (-\alpha_1 + \alpha_m)y_m = 0.$$

Since $y_1, \ldots, y_m$ are linearly independent, we must have $\alpha_1 = \alpha_2 = \cdots = \alpha_m = 0$. Therefore, $w_1, \ldots, w_m$ are linearly independent and hence form a basis of W.

We now construct a basis of V. By the fundamental theorem of linear maps, we have

$$\dim V = \dim T + \dim \operatorname{range} T = \dim T + 1.$$

Let $u_2, \ldots, u_n$ be a basis of T. Then we can extend this to a basis $u_1, u_2, \ldots, u_n$ of V, where $u_1 \notin T$. Since $\operatorname{range} T = \operatorname{span} y_1$, we have

$$Tu_1 = \beta y_1$$

for some nonzero scalar $\beta \in \mathbf{F}$. If $\beta \neq 1$, we may replace u_1 with $\beta^{-1}u_1$ to ensure that $Tu_1 = y_1$. We construct the following vectors in V:

$$v_1 = u_1, \quad v_k = u_1 + u_k \text{ for } k = 2, \ldots, n.$$

Note that for every $j = 1, \ldots, n$, we have $Tv_j = Tu_1 + Tu_j = y_1 + 0 = y_1$. We claim that $v_1, \ldots, v_n$ form a basis of V. Suppose for scalars $\alpha_1, \ldots, \alpha_n \in \mathbf{F}$, we have

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n = 0.$$

Then we can express this as

$$(\alpha_1 + \alpha_2 + \cdots + \alpha_n)u_1 + \alpha_2 u_2 + \cdots + \alpha_n u_n = 0.$$

Since $u_1, \ldots, u_n$ are linearly independent, we must have $\alpha_1 + \alpha_2 + \cdots + \alpha_n = 0$ and $\alpha_2 = \cdots = \alpha_n = 0$. This implies that $\alpha_1 = 0$, and hence $v_1, \ldots, v_n$ are linearly independent and form a basis of V. Then with respect to the bases $v_1, \ldots, v_n$ of V and $w_1, \ldots, w_m$ of W, we have that $Tv_j = y_1 = w_1 + w_2 + \cdots + w_m$ for each $j = 1, \ldots, n$. Therefore, the matrix of T with respect to these bases has all entries equal to 1.

Conversely, suppose 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. Then for each basis vector v_j of V, we have

$$Tv_j = w$$

for some fixed nonzero vector $w \in W$, where $w = w_1 + \cdots + w_m$ is the sum of the basis vectors of W. Since range T is spanned by the images of the basis vectors of V, then any sum of Tv_j is a scalar multiple of w. Therefore, range T = span w, and hence dim range $T = 1$. ∎

Exercise 3C3

Suppose $v_1, \ldots, v_n$ is a basis of V and $w_1, \ldots, w_m$ 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 $\mathcal{M}(\lambda T) = \lambda \mathcal{M}(T)$.

Solution.

(a) Let $v_1, \ldots, v_n$ be a basis of V and $w_1, \ldots, w_m$ be a basis of W. Then the matrix of $S + T$ with respect to these bases is given by

$$\begin{aligned}
\mathcal{M}(S+T) &= \begin{pmatrix} (S+T)v_1 & (S+T)v_2 & \cdots & (S+T)v_n \end{pmatrix} \\
&= \begin{pmatrix} Sv_1 + Tv_1 & Sv_2 + Tv_2 & \cdots & Sv_n + Tv_n \end{pmatrix} \\
&= \begin{pmatrix} Sv_1 & Sv_2 & \cdots & Sv_n \end{pmatrix} + \begin{pmatrix} Tv_1 & Tv_2 & \cdots & Tv_n \end{pmatrix} \\
&= \mathcal{M}(S) + \mathcal{M}(T).
\end{aligned}$$

(b) Let $v_1, \ldots, v_n$ be a basis of V and $w_1, \ldots, w_m$ be a basis of W. Then the matrix of λT with respect to these bases is given by

$$\begin{aligned}
\mathcal{M}(\lambda T) &= \begin{pmatrix} (\lambda T)v_1 & (\lambda T)v_2 & \cdots & (\lambda T)v_n \end{pmatrix} \\
&= \begin{pmatrix} \lambda T v_1 & \lambda T v_2 & \cdots & \lambda T v_n \end{pmatrix} \\
&= \lambda \begin{pmatrix} T v_1 & T v_2 & \cdots & T v_n \end{pmatrix} \\
&= \lambda \mathcal{M}(T).
\end{aligned}$$

∎

Exercise 3C4

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

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

Solution. Consider the standard basis of $\mathcal{P}_3(\mathbf{R})$. Reorder it to obtain the basis $x^3, x^2, x, 1$. Consider the standard basis of $\mathcal{P}_2(\mathbf{R})$. Reorder it to obtain the basis $x^2, x, 1$. Observe that $D(x^3) = 3x^2$, $D(x^2) = 2x$, $D(x) = 1$, and $D(1) = 0$. However,

we need the matrix of D to have the desired form. To achieve this, we can scale the basis vectors of $\mathcal{P}_3(\mathbf{R})$ and $\mathcal{P}_2(\mathbf{R})$ appropriately. Then we may define the basis for $\mathcal{P}_3(\mathbf{R})$ as

$$\frac{1}{3}x^3, \frac{1}{2}x^2, x, 1$$

while maintaining the standard basis for $\mathcal{P}_2(\mathbf{R})$ as $x^2, x, 1$. Then we have

$$D\left(\frac{1}{3}x^3\right) = x^2, \quad D\left(\frac{1}{2}x^2\right) = x, \quad D(x) = 1, \quad D(1) = 0.$$

Therefore, with respect to the bases, we get the desired matrix representation of D. ■

Exercise 3C5

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 $n = \dim V$ and $m = \dim W$. By the fundamental theorem of linear maps, we have

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

Let $r = \dim \operatorname{range} T$. Let $u_1, \dots, u_{n-r}$ be a basis of $\operatorname{null} T$. Extend this to a basis $v_1, \dots, v_r, u_1, \dots, u_{n-r}$ of V. We claim that $Tv_1, \dots, Tv_r$ are linearly independent. Suppose for scalars $\alpha_1, \dots, \alpha_r \in \mathbf{F}$, we have

$$\alpha_1 Tv_1 + \cdots + \alpha_r Tv_r = 0.$$

Then

$$T(\alpha_1 v_1 + \cdots + \alpha_r v_r) = 0,$$

which implies that $\alpha_1 v_1 + \cdots + \alpha_r v_r \in \operatorname{null} T$. Since $v_1, \dots, v_r, u_1, \dots, u_{n-r}$ is a basis of V, it follows that $\alpha_1 = \cdots = \alpha_r = 0$. Therefore, $Tv_1, \dots, Tv_r$ are linearly independent and form a basis of $\operatorname{range} T$. Extend this to a basis $Tv_1, \dots, Tv_r, w_1, \dots, w_{m-r}$ of W. Then with respect to these bases, Tv_k is the k-th basis vector of W for each $k = 1, \dots, r$, and

$$Tu_j = 0, \quad \text{for } j = 1, \dots, n-r.$$

Therefore, the matrix of T with respect to these bases has 1's in the diagonal entries corresponding to Tv_k and 0's elsewhere, as desired. ■

Exercise 3C6

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.

Solution. Let $n = \dim W$. If $Tv_1 = 0$, then we can choose any basis of W and the first column of $\mathcal{M}(T)$ will be all zeros. If $Tv_1 \neq 0$, then we can extend Tv_1 to a basis of W. Let $w_1 = Tv_1$ and extend it to a basis $w_1, w_2, \dots, w_n$ of W. Then with respect to the bases $v_1, \dots, v_m$ of V and $w_1, \dots, w_n$ of W, we have

$$Tv_1 = w_1$$

and for each $j = 2, \dots, m$, we can express

$$Tv_j = \alpha_{j1}w_1 + \alpha_{j2}w_2 + \cdots + \alpha_{jn}w_n$$

for some scalars $\alpha_{j1}, \alpha_{j2}, \dots, \alpha_{jn} \in \mathbf{F}$. Therefore, the first column of $\mathcal{M}(T)$ has a 1 in the first row and zeros in all other rows, as desired. ■

Exercise 3C7

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.

Solution. Let $m = \dim V$ and $u_1, \dots, u_m$ be a basis for V. If T has trivial w_1 component, meaning that the coefficient of w_1 in Tv is zero for all $v \in V$, then choose any basis of V and the first row of $\mathcal{M}(T)$ will be all zeros. Otherwise, reorder the basis so that Tu_1 has non-trivial w_1 component. Then we can express

$$Tu_1 = \alpha w_1 + \beta_2 w_2 + \cdots + \beta_n w_n$$

for some scalars $\alpha, \beta_2, \dots, \beta_n \in \mathbf{F}$ such that $\alpha \neq 0$. We can scale u_1 by α^{-1} to obtain a new vector $v_1 = \alpha^{-1}u_1$ such that

$$Tv_1 = w_1 + \frac{\beta_2}{\alpha}w_2 + \cdots + \frac{\beta_n}{\alpha}w_n.$$

With the remaining $u_2, \dots, u_m$ basis vectors of V, note that they can be expressed as

$$Tu_j = \alpha_{j1}w_1 + \alpha_{j2}w_2 + \cdots + \alpha_{jn}w_n$$

for some scalars $\alpha_{j1}, \alpha_{j2}, \dots, \alpha_{jn} \in \mathbf{F}$. Consider the set of vectors of V given by

$$v_1 = \alpha^{-1}u_1, \quad v_j = u_j - \alpha_{j1}v_1 \text{ for } j = 2, \dots, m.$$

Note that $Tv_j = Tu_j - \alpha_{j1}Tv_1$ is just a combination of $w_2, \dots, w_n$. We now claim that $v_1, v_2, \dots, v_m$ are linearly independent. Suppose for scalars $\gamma_1, \dots, \gamma_m \in \mathbf{F}$, we have

$$\gamma_1 v_1 + \gamma_2 v_2 + \cdots + \gamma_m v_m = 0.$$

Then we can express this as

$$\gamma_1 v_1 + \gamma_2(u_2 - \alpha_{21}v_1) + \cdots + \gamma_m(u_m - \alpha_{m1}v_1) = 0.$$

Simplifying, we get

$$(\gamma_1 - \gamma_2\alpha_{21} - \cdots - \gamma_m\alpha_{m1})v_1 + \gamma_2 u_2 + \cdots + \gamma_m u_m = 0.$$

Since $v_1, u_2, \dots, u_m$ are linearly independent, we must have $\gamma_1 - \gamma_2\alpha_{21} - \cdots - \gamma_m\alpha_{m1} = 0$ and $\gamma_2 = \cdots = \gamma_m = 0$. This implies that $\gamma_1 = 0$, and hence $v_1, v_2, \dots, v_m$ are linearly independent. Then $v_1, v_2, \dots, v_m$ form a basis of V. It follows that Tv_1 has a 1 in the first row, while Tv_j for $j = 2, \dots, m$ have zeros in the first row. ■

Exercise 3C8

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.

Solution. Let A be an m-by-n matrix and B be an n-by-p matrix. The entry in row j, column k, of the product AB is given by

$$(AB)_{j,k} = \sum_{r=1}^{n} A_{j,r} B_{r,k}.$$

The row vector $A_{j,\cdot}$ is the j-th row of A, which can be expressed as

$$A_{j,\cdot} = (A_{j,1}, A_{j,2}, \ldots, A_{j,n}).$$

When we multiply this row vector by the matrix B, we get a new row vector whose entries are given by

$$(A_{j,\cdot} B)_k = \sum_{r=1}^{n} A_{j,r} B_{r,k}.$$

Comparing this with the expression for $(AB)_{j,k}$, we see that

$$(AB)_{j,k} = (A_{j,\cdot} B)_k$$

for each column index k. Therefore, the entire row vector $(AB)_{j,\cdot}$ is equal to the row vector $A_{j,\cdot} B$. This proves that row j of AB equals (row j of A) times B. ■

Exercise 3C9

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.

Solution. Let $a = \begin{pmatrix} a_1 & \cdots & a_n \end{pmatrix}$ be a 1-by-n matrix and B be an n-by-p matrix. The product aB is a 1-by-p matrix, and its entries are given by

$$(aB)_{1,k} = \sum_{j=1}^{n} a_j B_{j,k}$$

for each column index k.

Now, consider the linear combination of the rows of B given by

$$a_1 B_{1,\cdot} + a_2 B_{2,\cdot} + \cdots + a_n B_{n,\cdot}.$$

The entry in row 1, column k, of this linear combination is

$$(a_1 B_{1,\cdot} + a_2 B_{2,\cdot} + \cdots + a_n B_{n,\cdot})_k = a_1 B_{1,k} + a_2 B_{2,k} + \cdots + a_n B_{n,k} = \sum_{j=1}^{n} a_j B_{j,k}.$$

Comparing this with the expression for $(aB)_{1,k}$, we see that

$$(aB)_{1,k} = (a_1 B_{1,\cdot} + a_2 B_{2,\cdot} + \cdots + a_n B_{n,\cdot})_k$$

for each column index k. Therefore, we conclude that

$$aB = a_1 B_{1,\cdot} + a_2 B_{2,\cdot} + \cdots + a_n B_{n,\cdot},$$

which shows that aB is indeed a linear combination of the rows of B, with the scalars coming from the entries of a. ■

Exercise 3C10

Give an example of 2-by-2 matrices A and B such that $AB \neq BA$.

Solution. Let

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Then we compute the products AB and BA:

$$AB = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 \cdot 0 + 2 \cdot 1 & 1 \cdot 1 + 2 \cdot 0 \\ 3 \cdot 0 + 4 \cdot 1 & 3 \cdot 1 + 4 \cdot 0 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 4 & 3 \end{pmatrix},$$

and

$$BA = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 0 \cdot 1 + 1 \cdot 3 & 0 \cdot 2 + 1 \cdot 4 \\ 1 \cdot 1 + 0 \cdot 3 & 1 \cdot 2 + 0 \cdot 4 \end{pmatrix} = \begin{pmatrix} 3 & 4 \\ 1 & 2 \end{pmatrix}.$$

By comparing the two products, we see that $AB \neq BA$. ■

Exercise 3C11

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. Let A be an m-by-n matrix, and let B and C be n-by-p matrices. Then $B + C$ is an n-by-p matrix, and hence $A(B + C)$ is an m-by-p matrix. The products AB and AC are also defined, and both are m-by-p matrices. Therefore, the sum $AB + AC$ is defined and is also an m-by-p matrix. By Exercise 3C8 and Exercise 3C9, we have

$$A(B + C)_{j,\cdot} = A_{j,\cdot}(B + C) = A_{j,\cdot}B + A_{j,\cdot}C = (AB)_{j,\cdot} + (AC)_{j,\cdot} = (AB + AC)_{j,\cdot}$$

for all $j = 1, \ldots, m$. Therefore, $A(B + C) = AB + AC$.

Similarly, let D and E be m-by-n matrices, and let F be an n-by-p matrix. Then $D + E$ is an m-by-n matrix, and hence $(D + E)F$ is an m-by-p matrix. The products DF and EF are also defined, and both are m-by-p matrices. Therefore, the sum $DF + EF$ is defined and is also an m-by-p matrix. Then the product $(D + E)F$ is defined. By Exercise 3C8 and Exercise 3C9, we have

$$(D + E)F_{j,\cdot} = (D + E)_{j,\cdot}F = D_{j,\cdot}F + E_{j,\cdot}F = (DF)_{j,\cdot} + (EF)_{j,\cdot} = (DF + EF)_{j,\cdot}$$

for all $j = 1, \ldots, m$. Therefore, $(D + E)F = DF + EF$. ■

Exercise 3C12

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

$$(AB)C = A(BC).$$

Manual Remark. *Unfortunately, we cannot use the same trick as in Exercise 3C11 to prove this result.*

Solution. Let A be an m-by-n matrix, B be an n-by-p matrix, and C be a p-by-q matrix. Then the product AB is defined and is an m-by-p matrix. Therefore, the product $(AB)C$ is also defined and is an m-by-q matrix. Going by the definition of matrix multiplication, we have

$$((AB)C)_{j,k} = \sum_{r=1}^{p} (AB)_{j,r} C_{r,k} = \sum_{r=1}^{p} \left(\sum_{s=1}^{n} A_{j,s} B_{s,r} \right) C_{r,k} = \sum_{r=1}^{p} \sum_{s=1}^{n} A_{j,s} B_{s,r} C_{r,k}.$$

On the other hand, the product BC is defined and is an n-by-q matrix. Therefore, the product $A(BC)$ is also defined and is an m-by-q matrix. Again, by the definition of matrix multiplication, we have

$$(A(BC))_{j,k} = \sum_{s=1}^{n} A_{j,s} (BC)_{s,k} = \sum_{s=1}^{n} A_{j,s} \left(\sum_{r=1}^{p} B_{s,r} C_{r,k} \right) = \sum_{s=1}^{n} \sum_{r=1}^{p} A_{j,s} B_{s,r} C_{r,k}.$$

Since the order of summation does not matter, we can interchange the order of the sums in the last expression to get

$$(A(BC))_{j,k} = \sum_{r=1}^{p} \sum_{s=1}^{n} A_{j,s} B_{s,r} C_{r,k}.$$

Comparing the expressions for $((AB)C)_{j,k}$ and $(A(BC))_{j,k}$, we see that they are equal for all j and k. Therefore, $(AB)C = A(BC)$. ∎

Exercise 3C13

Suppose A is an n-by-n matrix and $1 \le j, k \le n$. 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. Note that $A^3 = (AA)A = A^2 A$. By the definition of matrix multiplication, we have

$$(A^3)_{j,k} = (A^2 A)_{j,k} = \sum_{p=1}^{n} (A^2)_{j,p} A_{p,k}.$$

Now, we can express $(A^2)_{j,p}$ as

$$(A^2)_{j,p} = (AA)_{j,p} = \sum_{r=1}^{n} A_{j,r} A_{r,p}.$$

Substituting this back into the expression for $(A^3)_{j,k}$, we get

$$(A^3)_{j,k} = \sum_{p=1}^{n} \left(\sum_{r=1}^{n} A_{j,r} A_{r,p} \right) A_{p,k} = \sum_{p=1}^{n} \sum_{r=1}^{n} A_{j,r} A_{r,p} A_{p,k}.$$

By renaming the indices r and p, we can rewrite this as

$$(A^3)_{j,k} = \sum_{p=1}^{n} \sum_{r=1}^{n} A_{j,p} A_{p,r} A_{r,k},$$

which is the desired expression. ∎

Exercise 3C14

Suppose m and n are positive integers. Prove that the function $A \mapsto A^t$ is a linear map from $\mathbf{F}^{m,n}$ to $\mathbf{F}^{n,m}$.

Solution. Let $A, B \in \mathbf{F}^{m,n}$ and let $\lambda \in \mathbf{F}$. We need to show that $(A+B)^t = A^t + B^t$ and $(\lambda A)^t = \lambda A^t$.

First, consider the sum $A + B$. The entry in row j, column k, of $A + B$ is given by

$$(A+B)_{j,k} = A_{j,k} + B_{j,k}.$$

Taking the transpose, we have

$$((A+B)^t)_{k,j} = (A+B)_{j,k} = A_{j,k} + B_{j,k} = (A^t)_{k,j} + (B^t)_{k,j} = (A^t + B^t)_{k,j}.$$

Therefore, $(A+B)^t = A^t + B^t$.

Next, consider the scalar multiplication λA. The entry in row j, column k, of λA is given by

$$(\lambda A)_{j,k} = \lambda A_{j,k}.$$

Taking the transpose, we have

$$((\lambda A)^t)_{k,j} = (\lambda A)_{j,k} = \lambda A_{j,k} = \lambda (A^t)_{k,j} = (\lambda A^t)_{k,j}.$$

Therefore, $(\lambda A)^t = \lambda A^t$.

Since both properties hold, we conclude that the function $A \mapsto A^t$ is a linear map from $\mathbf{F}^{m,n}$ to $\mathbf{F}^{n,m}$. ■

Exercise 3C15

Prove that if A is an m-by-n matrix and C is an n-by-p matrix, then

$$(AC)^t = C^t A^t.$$

Solution. Let A be an m-by-n matrix and C be an n-by-p matrix. The entry in row j, column k, of the product AC is given by

$$(AC)_{j,k} = \sum_{r=1}^{n} A_{j,r} C_{r,k}.$$

Taking the transpose, we have

$$((AC)^t)_{k,j} = (AC)_{j,k} = \sum_{r=1}^{n} A_{j,r} C_{r,k}.$$

Now, consider the product $C^t A^t$. The entry in row k, column j, of this product is given by

$$(C^t A^t)_{k,j} = \sum_{r=1}^{n} (C^t)_{k,r} (A^t)_{r,j} = \sum_{r=1}^{n} C_{r,k} A_{j,r}.$$

Since multiplication in the field $\mathbf{F}$ is commutative, we can rewrite this as

$$(C^t A^t)_{k,j} = \sum_{r=1}^{n} A_{j,r} C_{r,k},$$

which is exactly the same as the expression for $((AC)^t)_{k,j}$.

Therefore, we conclude that

$$(AC)^t = C^t A^t.$$

■

Exercise 3C16

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. Suppose the rank of A is 1. Then the column rank of A is 1, which means that all columns of A are scalar multiples of a single non-zero column vector. Let $c = (c_1, \ldots, c_m)^t$ be this non-zero column vector. Then for each column k of A, there exists a scalar $d_k \in \mathbf{F}$ such that the k-th column of A is given by $d_k c$. Therefore, we can express the entries of A as

$$A_{j,k} = c_j d_k$$

for all $j = 1, \ldots, m$ and $k = 1, \ldots, n$.

Conversely, 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 all j and k. Since $A \neq 0$, then $(c_1, \ldots, c_m) \neq 0$ and $(d_1, \ldots, d_n) \neq 0$. This means that all columns of A are scalar multiples of the column vector c, and hence the column rank of A is 1. Therefore, the rank of A is 1. ■

Exercise 3C17

Suppose $T \in \mathcal{L}(V)$, and $u_1, \ldots, u_n$ and $v_1, \ldots, v_n$ 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 $\mathbf{F}^{1,n}$.

Solution.

- (a) $\iff$ (b): (a) $\iff$ $\dim \operatorname{null} T = 0$ $\iff$ $\dim \operatorname{range} T = n$ $\iff$ column rank $\mathcal{M}(T) = n$. For n column vectors, rank n implies linear independence $\iff$ (b).
- (b) $\iff$ (c): n columns of $\mathcal{M}(T)$ are linearly independent in $\mathbf{F}^{n,1}$ $\iff$ n columns of $\mathcal{M}(T)$ span $\mathbf{F}^{n,1}$.
- (c) $\iff$ (d): columns span $\mathbf{F}^{n,1}$ $\iff$ rank $\mathcal{M}(T) = n$ by (b) $\iff$ (c). By Theorem 3.57, rank $\mathcal{M}(T)$ = row rank $\mathcal{M}(T)$ $\iff$ rows span $\mathbf{F}^{1,n}$.
- (d) $\iff$ (e): n rows of $\mathcal{M}(T)$ span $\mathbf{F}^{1,n}$ $\iff$ n rows of $\mathcal{M}(T)$ are linearly independent in $\mathbf{F}^{1,n}$. ■

3D Invertibility and Isomorphisms

Exercise 3D1

Suppose $T \in \mathcal{L}(V, W)$ is invertible. Show that T^{-1} is invertible and

$$(T^{-1})^{-1} = T.$$

Solution. Since T is invertible, then $T^{-1} \in \mathcal{L}(W, V)$. By $TT^{-1} = I$ and $T^{-1}T = I$, we have that T^{-1} is invertible and $(T^{-1})^{-1} = T$. ∎

Exercise 3D2

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. Since T and S are invertible, then $T^{-1} \in \mathcal{L}(V, U)$ and $S^{-1} \in \mathcal{L}(W, V)$. By the definition of the inverse, we have that

$$(ST)(T^{-1}S^{-1}) = S(TT^{-1})S^{-1} = SIS^{-1} = SS^{-1} = I,$$

and

$$(T^{-1}S^{-1})(ST) = T^{-1}(S^{-1}S)T = T^{-1}IT = T^{-1}T = I.$$

Thus, ST is invertible and $(ST)^{-1} = T^{-1}S^{-1}$. ∎

Exercise 3D3

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.

- (a) $\implies$ (b): Suppose T is invertible. Consider scalars $\alpha_1, \ldots, \alpha_n$ such that

  $$\alpha_1 Tv_1 + \cdots + \alpha_n Tv_n = 0.$$

  Applying T^{-1} to both sides, we get

  $$\alpha_1 v_1 + \cdots + \alpha_n v_n = 0.$$

  Since $v_1, \ldots, v_n$ is a basis of V, it follows that $\alpha_1 = \cdots = \alpha_n = 0$. Hence, $Tv_1, \ldots, Tv_n$ is linearly independent. Since $Tv_1, \ldots, Tv_n$ has n elements and V is n-dimensional, it follows that $Tv_1, \ldots, Tv_n$ is a basis of V.
- (b) $\implies$ (c): This is trivial.
- (c) $\implies$ (a): Suppose $Tv_1, \ldots, Tv_n$ is a basis of V for some basis $v_1, \ldots, v_n$ of V. Then $\operatorname{range} T = V$. By the fundamental theorem of linear maps, we have that $\dim \operatorname{null} T = 0$. Thus, T is injective. Since T is injective and $\operatorname{range} T = V$, then T is invertible. ∎

Exercise 3D4

Suppose V is finite-dimensional and $\dim V > 1$. Prove that the set of non-invertible linear maps from V to itself is not a subspace of $\mathcal{L}(V)$.

Solution. Let $v_1, \ldots, v_n$ be a basis of V. Define $T, S \in \mathcal{L}(V)$ such that

$$Tv_1 = 0, \quad Tv_k = v_k \text{ for } k = 2, \ldots, n,$$

and

$$Sv_1 = v_1, \quad Sv_k = 0 \text{ for } k = 2, \ldots, n.$$

Then T and S are non-invertible linear maps, since $\operatorname{null} T = \operatorname{span} v_1$ and $\operatorname{null} S = \operatorname{span}(v_2, \ldots, v_n)$. However, $T + S$ is invertible, since

$$(T + S)v_1 = v_1, \quad (T + S)v_k = v_k \text{ for } k = 2, \ldots, n.$$

Hence, $T + S = I$ is invertible. Therefore, the set of non-invertible linear maps from V to itself is not a subspace of $\mathcal{L}(V)$. ∎

Exercise 3D5

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. Suppose there exists an invertible linear map T from V to itself such that $Tu = Su$ for every $u \in U$. If $Su = 0$, then $Tu = 0$, and since T is invertible, $u = 0$. Hence, S is injective.

Conversely, suppose S is injective. Let $u_1, \ldots, u_m$ be a basis of U. Since S is injective, then $Su_1, \ldots, Su_m$ is linearly independent. Extend $u_1, \ldots, u_m$ to a basis $u_1, \ldots, u_n$ of V. Extend $Su_1, \ldots, Su_m$ to a basis $v_1, \ldots, v_n$ of V. Define $T \in \mathcal{L}(V)$ such that

$$Tu_k = Su_k \text{ for } k = 1, \ldots, m, \quad Tu_k = v_k \text{ for } k = m + 1, \ldots, n.$$

Then T is invertible, since $Tu_1, \ldots, Tu_n$ is a basis of V. Thus, there exists an invertible linear map T from V to itself such that $Tu = Su$ for every $u \in U$. ∎

Exercise 3D6

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 $\operatorname{null} S = \operatorname{null} T$. Define a linear map E_0 from $\operatorname{range} T$ to $\operatorname{range} S$ by setting $E_0(Tv) = Sv$ for all $v \in V$. This is well-defined because if $Tv = Tv'$ for some $v, v' \in V$, then $T(v - v') = 0$, so $v - v' \in \operatorname{null} T = \operatorname{null} S$, and thus $S(v - v') = 0$, which implies $Sv = Sv'$. Moreover, if $E_0(Tv) = 0$ for some $v \in V$, then $Sv = 0$, which implies $v \in \operatorname{null} S = \operatorname{null} T$, and thus $Tv = 0$. Hence, E_0 is injective. By Exercise 3D5, we can extend E_0 to an invertible linear map E on all of W, and then we have $S = ET$.

Conversely, suppose there exists an invertible $E \in \mathcal{L}(W)$ such that $S = ET$. If $v \in \operatorname{null} S$, then $Sv = 0$, and since $S = ET$, we have $ETv = 0$. Since E is invertible, $Tv = 0$ so that $v \in \operatorname{null} T$. Hence, $\operatorname{null} S \subseteq \operatorname{null} T$. Similarly, if $v \in \operatorname{null} T$, then $Tv = 0$, and since $S = ET$, we have $Sv = ETv = E0 = 0$. Thus, $v \in \operatorname{null} S$, and $\operatorname{null} T \subseteq \operatorname{null} S$. Therefore, $\operatorname{null} S = \operatorname{null} T$. ∎

Exercise 3D7

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 $\operatorname{range} S = \operatorname{range} T$. Then $\dim \operatorname{null} S = \dim \operatorname{null} T$. Let $u_1, \ldots, u_k$ be a basis of $\operatorname{null} S$ and $v_1, \ldots, v_k$ be a basis of $\operatorname{null} T$. Extend both bases to a basis of V, say $u_1, \ldots, u_n$ and $v_1, \ldots, v_n$. Observe that for $i > k$, we have that

$Su_i \in \operatorname{range} S = \operatorname{range} T$. Then there must exist $y_i \in V$ such that $Ty_i = Su_i$. Define a linear map $E \in \mathcal{L}(V)$ by setting $Eu_i = v_i$ for $i \leq k$ and $Eu_i = y_i$ for $i > k$. To see that $TE = S$, we have for $i \leq k$ that $TEu_i = Tv_i = 0 = Su_i$. Moreover, for $i > k$, we have $TEu_i = Ty_i = Su_i$. To see that E is invertible, suppose we have scalars $\alpha_1, \dots, \alpha_n$ such that

$$\alpha_1 Eu_1 + \alpha_2 Eu_2 + \cdots + \alpha_n Eu_n = 0.$$

Applying T, we get

$$\alpha_{k+1} Su_{k+1} + \cdots + \alpha_n Su_n = 0,$$

where we used the fact that $TE = S$ and $Su_1 = \cdots = Su_k = 0$ because $u_1, \dots, u_k \in \operatorname{null} S$. Since $u_1, \dots, u_n$ is a basis of V, the vectors $u_{k+1}, \dots, u_n$ map to a spanning list of range S. Since $\dim \operatorname{range} S = n - k$, it follows that $Su_{k+1}, \dots, Su_n$ are linearly independent. Then from the previous equation, $\alpha_{k+1} = \cdots = \alpha_n = 0$. Substituting this back into the original equation, we get $\alpha_1 v_1 + \cdots + \alpha_k v_k = 0$, and since $v_1, \dots, v_k$ are linearly independent, we must have $\alpha_1 = \cdots = \alpha_k = 0$. Hence, all the scalars are zero, proving that E is invertible.

Conversely, suppose there exists an invertible $E \in \mathcal{L}(V)$ such that $S = TE$. We have that $\operatorname{range} S \subseteq \operatorname{range} T$ by Exercise 3B26. Because E is invertible, we also have that $T = SE^{-1}$, so $\operatorname{range} T \subseteq \operatorname{range} S$ by the same reasoning. Hence, $\operatorname{range} S = \operatorname{range} T$. Similarly, if $w \in \operatorname{range} T$, then there exists some $v \in V$ such that $Tv = w$. Since E is invertible, there exists some $u \in V$ such that $Eu = v$. Then we have

$$Su = TEu = Tv = w,$$

and thus $w \in \operatorname{range} S$. Hence, $\operatorname{range} T \subseteq \operatorname{range} S$. Therefore, $\operatorname{range} S = \operatorname{range} T$. ■

Exercise 3D8

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 invertible $E_1 \in \mathcal{L}(V)$ and invertible $E_2 \in \mathcal{L}(W)$ such that $S = E_2 T E_1$. We claim that E_1 is an invertible map from $\operatorname{null} S$ to $\operatorname{null} T$. To see this, suppose $v, w \in \operatorname{null} S$ and $E_1 v = E_1 w$. Then $v = w$ since E_1 is invertible, proving injectivity. To see surjectivity, let $u \in \operatorname{null} T$. Since E_1 is invertible, there exists some $v \in V$ such that $E_1 v = u$. Then

$$Sv = E_2 T E_1 v = E_2 T u = E_2 0 = 0,$$

and thus $v \in \operatorname{null} S$. Hence, E_1 is an invertible map from $\operatorname{null} S$ to $\operatorname{null} T$, and thus $\dim \operatorname{null} S = \dim \operatorname{null} T$.

Conversely, suppose $\dim \operatorname{null} S = \dim \operatorname{null} T$. Then we may construct an invertible linear map $E_1 \in \mathcal{L}(V)$ that maps $\operatorname{null} S$ onto $\operatorname{null} T$. We now claim that $\operatorname{null} TE_1 = \operatorname{null} S$. To see this, let $v \in \operatorname{null} TE_1$. This is only true when $TE_1 v = 0$, or $E_1 v \in \operatorname{null} T$. By invertibility of E_1, we have $v \in \operatorname{null} S$. By Exercise 3D6 with $\operatorname{null} S = \operatorname{null} TE_1$, there exists an invertible $E_2 \in \mathcal{L}(W)$ such that $S = E_2 T E_1$. ■

Exercise 3D9

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.

Axler Note. *Here $T|_U$ means the function T restricted to U. Thus, $T|_U$ is the function whose domain is U, with $T|_U(u) = Tu$ for every $u \in U$.*

Solution. Since T is surjective, then $\operatorname{range} T = W$. By Exercise 3B11, there exists a subspace U of V such that $V = \operatorname{null} T \oplus U$. We show that $T|_U$ is an isomorphism of U onto W.

Suppose $u \in U$ such that $T|_U u = 0$. Then $Tu = 0$, and thus $u \in \operatorname{null} T \cap U$. Since $V = \operatorname{null} T \oplus U$, we have $\operatorname{null} T \cap U = \{0\}$, and thus $u = 0$. This shows that $T|_U$ is injective.

To see surjectivity, let $w \in W$. Since T is surjective, there exists $v \in V$ such that $Tv = w$. Write $v = n + u$ with $n \in \operatorname{null} T$ and $u \in U$. Then $Tu = T(v - n) = Tv - Tn = w - 0 = w$. Hence, $T|_U$ is surjective onto W, and therefore an isomorphism of U onto W. ∎

Exercise 3D10

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 $\dim U$.

Axler 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) Consider the zero map $0 \in \mathcal{L}(V, W)$. For this map, $U \subseteq \operatorname{null} 0$ since $\operatorname{null} 0 = V$. Hence, $0 \in \mathcal{E}$, showing that $\mathcal{E}$ is non-empty. To see that $\mathcal{E}$ is closed under addition and scalar multiplication, let $S, T \in \mathcal{E}$ and $\alpha \in \mathbf{F}$. For any $u \in U$, we have $(S + T)u = Su + Tu = 0 + 0 = 0$, so $S + T \in \mathcal{E}$. Similarly, $(\alpha T)u = \alpha(Tu) = \alpha \cdot 0 = 0$, so $\alpha T \in \mathcal{E}$. Therefore, $\mathcal{E}$ is a subspace of $\mathcal{L}(V, W)$.

(b) Consider the linear map $\Phi : \mathcal{L}(V, W) \to \mathcal{L}(U, W)$ defined by $\Phi(T) = T|_U$. Let $T \in \mathcal{L}(V, W)$ such that $\Phi(T) = T|_U = 0$. Then $T|_U = 0$ implies that for every $u \in U$, we have $Tu = 0$, i.e., $U \subseteq \operatorname{null} T$. Hence, $T \in \mathcal{E}$, showing that $\operatorname{null} \Phi \subseteq \mathcal{E}$. Conversely, if $T \in \mathcal{E}$, then $U \subseteq \operatorname{null} T$, which means that $T|_U = 0$, i.e., $T \in \operatorname{null} \Phi$. Therefore, $\operatorname{null} \Phi = \mathcal{E}$, and $\dim \mathcal{E} = \dim \operatorname{null} \Phi$.

We now show that Φ is surjective. Let $S \in \mathcal{L}(U, W)$, and let $u_1, \ldots, u_m$ be a basis for U. Extend this to a basis $u_1, \ldots, u_m, v_1, \ldots, v_n$ for V. Define a linear map $T \in \mathcal{L}(V, W)$ by setting $Tu_i = Su_i$ for $i = 1, \ldots, m$ and $Tv_j = 0$ for $j = 1, \ldots, n$. Then $T|_U = S$, so $\Phi(T) = S$. Hence, Φ is surjective, and $\operatorname{range} \Phi = \mathcal{L}(U, W)$. Therefore, $\dim \operatorname{range} \Phi = \dim \mathcal{L}(U, W) = (\dim U)(\dim W)$. By the fundamental theorem of linear maps,

$$\dim \mathcal{L}(V, W) = \dim \operatorname{null} \Phi + \dim \operatorname{range} \Phi = \dim \mathcal{E} + (\dim U)(\dim W),$$

Observing that $\dim \mathcal{L}(V, W) = (\dim V)(\dim W)$, we get

$$\dim \mathcal{E} = (\dim V)(\dim W) - (\dim U)(\dim W) = (\dim V - \dim U)(\dim W).$$ ∎

Exercise 3D11

Suppose V is finite-dimensional and $S, T \in \mathcal{L}(V)$. Prove that

$$ST \text{ is invertible} \iff S \text{ and } T \text{ are invertible.}$$

Solution. Suppose ST is invertible. Then $\operatorname{null} ST = \{0\}$. If $v \in \operatorname{null} T$, then $STv = S0 = 0$, and thus $v \in \operatorname{null} ST$. Since $\operatorname{null} ST = \{0\}$, we have $v = 0$, and thus $\operatorname{null} T = \{0\}$. Hence, T is injective, and since V is finite-dimensional, T is invertible. Then $S = (ST)T^{-1}$ is a product of invertible linear maps, and thus S is invertible by Exercise 3D2.

Conversely, suppose S and T are invertible. Then ST is a product of invertible linear maps, and thus ST is invertible by Exercise 3D2. ∎

Exercise 3D12

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. Since $STU = I$, then we have that $S(TU) = I$. By Exercise 3B20, then S is surjective, hence invertible with $S^{-1} = TU$. Similarly, $(ST)U = I$. By Exercise 3B19, then U is injective, hence invertible with $U^{-1} = ST$.

$$T(US) = (TU)S = S^{-1}S = I, \quad (US)T = U(ST) = UU^{-1} = I.$$

Hence, T is invertible with $T^{-1} = US$. ■

Exercise 3D13

Show that the result in Exercise 3D12 can fail without the hypothesis that V is finite-dimensional.

Solution. If V were not finite-dimensional, let $v_1, v_2, \ldots$ be an infinite basis for V. Let $S, T, U \in \mathcal{L}(V)$ be defined as follows: let $S = I$,

$$Tv_1 = 0, \quad Tv_k = v_{k-1} \text{ for } k = 2, 3, \ldots$$

and

$$Uv_k = v_{k+1} \text{ for } k = 1, 2, \ldots$$

It is clear that $TUv_k = v_k$ for all $k = 1, 2, \ldots$, so $STU = I$. However, T is not injective since $\operatorname{null} T = \operatorname{span} v_1$, and thus T is not invertible. ■

Exercise 3D14

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. Since RST is surjective, and V is finite dimensional, it follows that RST is injective, hence invertible. By Exercise 3D11, then $RST = (RS)T$ is invertible if and only if RS and T are invertible. In particular, RS is invertible, and applying Exercise 3D11 again, we have that R and S are invertible. Hence, S is injective. ■

Exercise 3D15

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. Since $Tv_1, \ldots, Tv_m$ spans V, we have T surjective. Moreover, V is finite-dimensional, so a surjective linear map on V is also injective, hence invertible. Therefore, for every $v \in V$, there exist scalars $\alpha_1, \ldots, \alpha_m$ such that

$$v = \alpha_1 Tv_1 + \cdots + \alpha_m Tv_m.$$

Applying T^{-1} to both sides, we get

$$T^{-1}v = \alpha_1 v_1 + \cdots + \alpha_m v_m.$$

Since T^{-1} is surjective, every vector $w \in V$ can be written as $T^{-1}v$ for some $v \in V$, and thus $v_1, \ldots, v_m$ spans V. ■

Exercise 3D16

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 $e_1, \ldots, e_n$ be the standard basis of $\mathbf{F}^{n,1}$. Define the k-th column of A to be $Te_k \in \mathbf{F}^{m,1}$ for $k = 1, \ldots, n$. Then for every $x \in \mathbf{F}^{n,1}$, there exist scalars $\alpha_1, \ldots, \alpha_n$ such that

$$x = \alpha_1 e_1 + \cdots + \alpha_n e_n.$$

Applying T to both sides, we have

$$Tx = \alpha_1 Te_1 + \cdots + \alpha_n Te_n = Ax,$$

where the last equality follows from the definition of matrix multiplication. Hence, $Tx = Ax$ for every $x \in \mathbf{F}^{n,1}$. ■

Exercise 3D17

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) Consider $T \in \operatorname{null} \mathcal{A}$. Then $\mathcal{A}(T) = ST = 0$, so T is a linear map from V to $\operatorname{null} S$. Conversely, if T is a linear map from V to $\operatorname{null} S$, then $ST = 0$, so $T \in \operatorname{null} \mathcal{A}$. Hence, $\operatorname{null} \mathcal{A}$ is the set of linear maps from V to $\operatorname{null} S$, and thus $\dim \operatorname{null} \mathcal{A} = (\dim V)(\dim \operatorname{null} S)$.

(b) By fundamental theorem of linear maps, we have

$$\dim \mathcal{L}(V) = \dim \operatorname{null} \mathcal{A} + \dim \operatorname{range} \mathcal{A}.$$

Since $\dim \mathcal{L}(V) = (\dim V)^2$ and $\dim \operatorname{null} \mathcal{A} = (\dim V)(\dim \operatorname{null} S)$, we get

$$(\dim V)^2 = (\dim V)(\dim \operatorname{null} S) + \dim \operatorname{range} \mathcal{A},$$

which implies

$$\dim \operatorname{range} \mathcal{A} = (\dim V)^2 - (\dim V)(\dim \operatorname{null} S) = (\dim V)(\dim \operatorname{range} S).$$ ■

Exercise 3D18

Show that V and $\mathcal{L}(\mathbf{F}, V)$ are isomorphic vector spaces.

Solution. Let $T : \mathcal{L}(\mathbf{F}, V) \to V$ be defined by $T(\alpha) = \alpha(1)$ for $\alpha \in \mathcal{L}(\mathbf{F}, V)$. We first show that T is a linear map. Let $\alpha, \beta \in \mathcal{L}(\mathbf{F}, V)$ and $\lambda \in \mathbf{F}$. Then

$$T(\alpha + \beta) = (\alpha + \beta)(1) = \alpha(1) + \beta(1) = T\alpha + T\beta,$$

and

$$T(\lambda\alpha) = (\lambda\alpha)(1) = \lambda\alpha(1) = \lambda T\alpha.$$

We now show that T is an isomorphism. To see injectivity, let $\alpha, \beta \in \mathcal{L}(\mathbf{F}, V)$ such that $T\alpha = T\beta$. Then $\alpha(1) = \beta(1)$. For any $x \in \mathbf{F}$, we have $\alpha(x) = x\alpha(1) = x\beta(1) = \beta(x)$, so $\alpha = \beta$. Hence, T is injective.

To see surjectivity, let $v \in V$. Define $\alpha \in \mathcal{L}(\mathbf{F}, V)$ by $\alpha(x) = xv$ for $x \in \mathbf{F}$. Then $T\alpha = \alpha(1) = v$, so T is surjective. Therefore, T is an isomorphism, and thus V and $\mathcal{L}(\mathbf{F}, V)$ are isomorphic vector spaces. ■

Exercise 3D19 ✔

Suppose V is 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. Suppose T has the same matrix with respect to every basis of V. Then for any set of bases for V, its matrix must be the same. If $\dim V \leq 1$, then every operator on V is a scalar multiple of the identity operator. Now suppose $\dim V \geq 2$. For any bases $v_1, \ldots, v_n$ and $u_1, \ldots, u_n$ of V, we have

$$\mathcal{M}(T, (v_1, \ldots, v_n)) = \mathcal{M}(T, (u_1, \ldots, u_n)).$$

Hence, we can declare that their matrix is just some A. We now consider the basis $v_1 + v_2, v_2, \ldots, v_n$, which yields the same matrix A. Let $w_1 = v_1 + v_2$ and $w_i = v_i$ for $i \geq 2$. In the new basis, we have the expression

$$Tw_1 = A_{1,1}w_1 + A_{2,1}w_2 + \cdots + A_{n,1}w_n = A_{1,1}v_1 + (A_{1,1} + A_{2,1})v_2 + A_{3,1}v_3 + \cdots + A_{n,1}v_n.$$

This must equal the expression for $Tw_1 = Tv_1 + Tv_2$ in the original basis:

$$Tv_1 + Tv_2 = (A_{1,1} + A_{1,2})v_1 + (A_{2,1} + A_{2,2})v_2 + \cdots + (A_{n,1} + A_{n,2})v_n.$$

Comparing coefficients of $v_1, \ldots, v_n$ in the two expressions, we get

$$\begin{aligned} A_{1,1} &= A_{1,1} + A_{1,2}, \\ A_{1,1} + A_{2,1} &= A_{2,1} + A_{2,2}, \\ A_{j,1} &= A_{j,1} + A_{j,2} \text{ for } j \geq 3. \end{aligned}$$

From this, we obtain that $A_{1,2} = 0, A_{2,2} = A_{1,1}$, and $A_{j,2} = 0$ for $j \geq 3$, i.e., the diagonal elements $A_{1,1}$ and $A_{2,2}$ are equal, while off-diagonal entries in column 2 are 0. Now consider the basis $v_1, v_1 + v_2, \ldots, v_n$. Let $u_1 = v_1, u_2 = v_1 + v_2$, and $u_i = v_i$ for $i \geq 3$. In the new basis, we have

$$\begin{aligned} Tu_2 &= A_{1,2}u_1 + \cdots + A_{n,2}u_n \\ &= A_{1,2}v_1 + A_{2,2}(v_1 + v_2) + A_{3,2}v_3 + \cdots + A_{n,2}v_n \\ &= (A_{1,2} + A_{2,2})v_1 + A_{2,2}v_2 + A_{3,2}v_3 + \cdots + A_{n,2}v_n. \\ &= A_{1,1}v_1 + A_{2,2}v_2. \end{aligned}$$

This must equal the expression for $Tu_2 = Tv_1 + Tv_2$ in the original basis, whose expression is listed above. Comparing coefficients of v_1 and v_2, we get

$$\begin{aligned} A_{1,1} &= A_{1,1} + A_{1,2}, \\ A_{2,2} &= A_{2,1} + A_{2,2}, \\ A_{j,2} &= A_{j,1} + A_{j,2} \quad \text{for } j \geq 3. \end{aligned}$$

From this, we obtain similar results about diagonal entries, but the off-diagonal entries in column 1 must also be 0. We may repeat the first process for $v_1 + v_k$ for $k \geq 3$ to get that entries off the diagonal in column k must also be 0. Continuing this process, we conclude all off-diagonal entries are 0, and all diagonal entries are equal, i.e., A is a scalar multiple of the identity matrix.

Conversely, suppose T is a scalar multiple of the identity operator. Then there exists some $\lambda \in \mathbf{F}$ such that $T = \lambda I$. For any basis of V, the matrix representation of T with respect to that basis is the diagonal matrix with λ on the diagonal. Hence, T has the same matrix with respect to every basis of V. ■

Exercise 3D20 ✔

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. Let $T : \mathcal{P}_n(\mathbf{R}) \to \mathcal{P}_n(\mathbf{R})$ be defined by

$$Tp = (x^2 + x)p'' + 2xp' + p(3)$$

for $p \in \mathcal{P}_n(\mathbf{R})$. We first show that T is a linear map. Let $p, q \in \mathcal{P}_n(\mathbf{R})$ and $\alpha \in \mathbf{R}$. Then

$$T(p + q) = (x^2 + x)(p + q)'' + 2x(p + q)' + (p + q)(3) = Tp + Tq,$$

and

$$T(\alpha p) = (x^2 + x)(\alpha p)'' + 2x(\alpha p)' + (\alpha p)(3) = \alpha Tp.$$

Then $T \in \mathcal{L}(\mathcal{P}_n(\mathbf{R}))$. We now show that T is invertible. To see injectivity, let $p \in \mathcal{P}_n(\mathbf{R})$ such that $Tp = 0$. Then

$$(x^2 + x)p'' + 2xp' + p(3) = 0.$$

Then

$$(x^2 + x)p'' + 2xp' = -p(3).$$

We analyze the degree of p as follows:

- If $\deg p \geq 2$, consider $(x^2 + x)p''$. In particular, $(a_m x^m)'' = a_m m(m-1)x^{m-2}$, hence the coefficient of x^m is $a_m m(m-1)$. Moreover, $2xp'$ has coefficient $2ma_m$. Since $\deg p \geq 2$, then $m \geq 2$, so the total coefficient of x^m is $a_m m(m-1) + 2ma_m = a_m m(m-1+2) = a_m m(m+1)$. Since this is nonzero for $a_m \neq 0$, we get a contradiction with the right-hand side being a constant.
- If $\deg p = 1$, then $p(x) = ax + b$ for some $a, b \in \mathbf{R}$. Then $p''(x) = 0$ and $p'(x) = a$, so the left-hand side is $2ax$. Since the right-hand side is a constant, this is a contradiction.
- If $\deg p = 0$, then $p(x) = c$ for some $c \in \mathbf{R}$. Then $p''(x) = 0$ and $p'(x) = 0$, so the left-hand side is 0, and the right-hand side is $p(3) = c$. Hence, $c = 0$, and thus $p = 0$. Therefore, T is injective.

Since T is injective and $\mathcal{P}_n(\mathbf{R})$ is finite-dimensional, then T is surjective, hence invertible. Then there does exist some $p \in \mathcal{P}_n(\mathbf{R})$ such that $Tp = q$. ■

Exercise 3D21

Suppose n is a positive integer and $A_{j,k} \in \mathbf{F}$ for all $j, k = 1, \dots, n$. 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

$$\begin{aligned} \sum_{k=1}^{n} A_{1,k} x_k &= 0 \\ &\vdots \\ \sum_{k=1}^{n} A_{n,k} x_k &= 0 \end{aligned}$$

(b) For every $c_1, \dots, c_n \in \mathbf{F}$, there exists a solution to the system of equations

$$\begin{aligned} \sum_{k=1}^{n} A_{1,k} x_k &= c_1 \\ &\vdots \\ \sum_{k=1}^{n} A_{n,k} x_k &= c_n \end{aligned}$$

Solution.

(a) Suppose the trivial solution $0 \in \mathbf{F}^{n,1}$ is the only solution to the homogeneous system of equations. Then $\operatorname{null} T = \{0\}$, where $T : \mathbf{F}^{n,1} \to \mathbf{F}^{n,1}$ is defined by $Tx = Ax$ for $x \in \mathbf{F}^{n,1}$. Since $\operatorname{null} T = \{0\}$, then T is injective. Since $\mathbf{F}^{n,1}$ is finite-dimensional, then T is surjective, hence invertible. Then for every $c_1, \dots, c_n \in \mathbf{F}$, there exists some $x \in \mathbf{F}^{n,1}$ such that $Tx = c$, where c is the column vector with entries $c_1, \dots, c_n$. Hence, there exists a solution to the system of equations.

(b) Suppose for every $c_1, \dots, c_n \in \mathbf{F}$, there exists a solution to the system of equations. Then $T : \mathbf{F}^{n,1} \to \mathbf{F}^{n,1}$ defined by $Tx = Ax$ for $x \in \mathbf{F}^{n,1}$ is surjective. Since $\mathbf{F}^{n,1}$ is finite-dimensional, then T is injective. Hence, the trivial solution is the only solution to the homogeneous system of equations. ∎

Exercise 3D22

Suppose $T \in \mathcal{L}(V)$ and $v_1, \dots, v_n$ is a basis of V. Prove that

$$\mathcal{M}(T, (v_1, \dots, v_n)) \text{ is invertible} \iff T \text{ is invertible.}$$

Solution. Suppose $\mathcal{M}(T, (v_1, \dots, v_n))$ is invertible. Then there exists some $A \in \mathbf{F}^{n,n}$ such that $A\mathcal{M}(T, (v_1, \dots, v_n)) = I$. Let U be the linear map from V to V whose matrix with respect to the basis $v_1, \dots, v_n$ is A. Then

$$\mathcal{M}(UT, (v_1, \dots, v_n)) = \mathcal{M}(U, (v_1, \dots, v_n))\mathcal{M}(T, (v_1, \dots, v_n)) = A\mathcal{M}(T, (v_1, \dots, v_n)) = I.$$

Then UT is the identity operator, and thus T is injective. Since V is finite-dimensional, then T is surjective, hence invertible.

Suppose T is invertible. Then T^{-1} exists, and we have

$$\mathcal{M}(T^{-1}T, (v_1, \dots, v_n)) = \mathcal{M}(T^{-1}, (v_1, \dots, v_n))\mathcal{M}(T, (v_1, \dots, v_n)) = I.$$

Hence, $\mathcal{M}(T, (v_1, \dots, v_n))$ is invertible. ∎

Exercise 3D23

Suppose that $u_1, \dots, u_n$ and $v_1, \dots, v_n$ are bases of V. Let $T \in \mathcal{L}(V)$ be such that $Tv_k = u_k$ for each $k = 1, \dots, n$. Prove that

$$\mathcal{M}(T, (v_1, \dots, v_n)) = \mathcal{M}(I, (u_1, \dots, u_n), (v_1, \dots, v_n)).$$

Solution. By definition, column k of $\mathcal{M}(T, (v_1, \dots, v_n))$ is $Tv_k = u_k$ with respect to the basis $v_1, \dots, v_n$. Similarly, column k of $\mathcal{M}(I, (u_1, \dots, u_n), (v_1, \dots, v_n))$ is also u_k with respect to the basis $v_1, \dots, v_n$. ■

Exercise 3D24

Suppose A and B are square matrices of the same size and $AB = I$. Prove that $BA = I$.

Solution. Let $T : \mathbf{F}^{n,1} \to \mathbf{F}^{n,1}$ be defined by $Tx = Ax$ for $x \in \mathbf{F}^{n,1}$. Since $AB = I$, then T is surjective. Since $\mathbf{F}^{n,1}$ is finite-dimensional, surjectivity implies injectivity. Hence, T is invertible, and $T(Bx) = x = T(T^{-1}x)$. Then $Bx = T^{-1}x$ for every $x \in \mathbf{F}^{n,1}$. Therefore,

$$BAx = B(Tx) = T^{-1}(Tx) = x$$

for every $x \in \mathbf{F}^{n,1}$. Hence, $BA = I$. ■

3E Products and Quotients of Vector Spaces

Exercise 3E1

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 $V \times W$.

Solution. Let $G(T)$ denote the graph of T. Now, suppose T is a linear map. Then $G(T)$ is nonempty since $(0, T0) = (0, 0) \in G(T)$. Now, let $(v_1, Tv_1), (v_2, Tv_2) \in G(T)$. Then

$$(v_1 - v_2, Tv_1 - Tv_2) = (v_1, Tv_1) - (v_2, Tv_2) \in G(T)$$

since T is linear. Now, let $(v, Tv) \in G(T)$ and $\alpha \in \mathbf{F}$. Then

$$(\alpha v, \alpha Tv) = \alpha(v, Tv) \in G(T)$$

since T is linear. Thus, $G(T)$ is a subspace of $V \times W$.

Conversely, suppose $G(T)$ is a subspace of $V \times W$. By definition, we have $(v_1, Tv_1), (v_2, Tv_2) \in G(T)$ for all $v_1, v_2 \in V$. Since $G(T)$ is a subspace of $V \times W$, for all $\alpha, \beta \in \mathbf{F}$, we have

$$(\alpha v_1 + \beta v_2, \alpha Tv_1 + \beta Tv_2) = \alpha(v_1, Tv_1) + \beta(v_2, Tv_2) \in G(T).$$

By definition of $G(T)$, we have $\alpha Tv_1 + \beta Tv_2 = T(\alpha v_1 + \beta v_2)$. Thus, T is a linear map. ∎

Exercise 3E2

Suppose that $V_1, \ldots, V_m$ are vector spaces such that $V_1 \times \cdots \times V_m$ is finite-dimensional. Prove that V_k is finite-dimensional for each $k = 1, \ldots, m$.

Solution. Let $V = V_1 \times \cdots \times V_m$, and consider $T \in \mathcal{L}(V, V_k)$ defined by $T(v_1, \ldots, v_m) = v_k$. Then T is a linear map since for all $\alpha, \beta \in \mathbf{F}$ and all $(v_1, \ldots, v_m), (w_1, \ldots, w_m) \in V$, we have

$$\begin{aligned} T(\alpha(v_1, \ldots, v_m) + \beta(w_1, \ldots, w_m)) &= T(\alpha v_1 + \beta w_1, \ldots, \alpha v_m + \beta w_m) \\ &= \alpha v_k + \beta w_k \\ &= \alpha T(v_1, \ldots, v_m) + \beta T(w_1, \ldots, w_m). \end{aligned}$$

Moreover, T is surjective since for any $v_k \in V_k$, we have $v_k = T(0, \ldots, 0, v_k, 0, \ldots, 0)$ where v_k is in the k-th coordinate. Since V is finite-dimensional and T is a linear map that is surjective, it follows that V_k is finite-dimensional. ∎

Exercise 3E3

Suppose $V_1, \ldots, V_m$ 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. Let $V = V_1 \times \cdots \times V_m$. Let $L = \mathcal{L}(V_1, W) \times \cdots \times \mathcal{L}(V_m, W)$. Define $T \in \mathcal{L}(\mathcal{L}(V, W), L)$ by

$$T(S) = (S \circ \iota_1, \ldots, S \circ \iota_m)$$

where $\iota_k \in \mathcal{L}(V_k, V)$ is the map defined as

$$\iota_k(v_k) = (0, \ldots, 0, v_k, 0, \ldots, 0)$$

with $\iota_k(v_k)$ having v_k in the k-th coordinate and 0 in all other coordinates. Then T is linear since for all $\alpha, \beta \in \mathbf{F}$ and all $S_1, S_2 \in \mathcal{L}(V, W)$, we have

$$\begin{aligned} T(\alpha S_1 + \beta S_2) &= ((\alpha S_1 + \beta S_2) \circ \iota_1, \ldots, (\alpha S_1 + \beta S_2) \circ \iota_m) \\ &= (\alpha(S_1 \circ \iota_1) + \beta(S_2 \circ \iota_1), \ldots, \alpha(S_1 \circ \iota_m) + \beta(S_2 \circ \iota_m)) \\ &= \alpha T(S_1) + \beta T(S_2), \end{aligned}$$

hence T is a linear map. Now suppose $T(S) = 0$. Then $S \circ \iota_k = 0$ for all $k = 1, \ldots, m$. Moreover, for all $(v_1, \ldots, v_m) \in V$, we have

$$S(v_1, \ldots, v_m) = S(\iota_1(v_1) + \cdots + \iota_m(v_m)) = S \circ \iota_1(v_1) + \cdots + S \circ \iota_m(v_m) = 0.$$

It follows that $S(v_1, \ldots, v_m) = 0$ for all $(v_1, \ldots, v_m) \in V$. Thus, $S = 0$, and hence T is injective.

Now, let $(S_1, \ldots, S_m) \in L$. Define $S \in \mathcal{L}(V, W)$ by $S(v_1, \ldots, v_m) = S_1 v_1 + \cdots + S_m v_m$. Then $T(S) = (S_1, \ldots, S_m)$. Thus, T is surjective. Therefore, T is an isomorphism. ∎

Exercise 3E4

Suppose $W_1, \ldots, W_m$ 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.

Manual Remark. *Note the difference between this exercise and Exercise 3E3. Here, the domain is decomposed, while in Exercise 3E3, the codomain is decomposed. Hence, the use of the maps ι_k and π_k in the solutions to these exercises are reversed.*

Solution. Let $W = W_1 \times \cdots \times W_m$. Let $L = \mathcal{L}(V, W_1) \times \cdots \times \mathcal{L}(V, W_m)$. Define $T \in \mathcal{L}(\mathcal{L}(V, W), L)$ by

$$T(S) = (\pi_1 \circ S, \ldots, \pi_m \circ S)$$

where $\pi_k \in \mathcal{L}(W, W_k)$ is the map defined as

$$\pi_k(w_1, \ldots, w_m) = w_k$$

Then T is linear since for all $\alpha, \beta \in \mathbf{F}$ and all $S_1, S_2 \in \mathcal{L}(V, W)$, we have

$$\begin{aligned} T(\alpha S_1 + \beta S_2) &= (\pi_1 \circ (\alpha S_1 + \beta S_2), \ldots, \pi_m \circ (\alpha S_1 + \beta S_2)) \\ &= (\alpha(\pi_1 \circ S_1) + \beta(\pi_1 \circ S_2), \ldots, \alpha(\pi_m \circ S_1) + \beta(\pi_m \circ S_2)) \\ &= \alpha T(S_1) + \beta T(S_2), \end{aligned}$$

hence T is a linear map. Now suppose $T(S) = 0$. Then $\pi_k \circ S = 0$ for all $k = 1, \ldots, m$. Moreover, for all $v \in V$, we have

$$Sv = (\pi_1(Sv), \ldots, \pi_m(Sv)) = (0, \ldots, 0) = 0.$$

It follows that $Sv = 0$ for all v, hence $S = 0$, and thus T is injective.

Now let $(S_1, \ldots, S_m) \in L$. Define $S \in \mathcal{L}(V, W)$ by $Sv = (S_1 v, \ldots, S_m v)$. Then $T(S) = (S_1, \ldots, S_m)$. Thus, T is surjective. Therefore, T is an isomorphism. ∎

Exercise 3E5

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. Let $e_1, \ldots, e_m$ be the standard basis of $\mathbf{F}^m$. Define $T \in \mathcal{L}(V^m, \mathcal{L}(\mathbf{F}^m, V))$ by

$$T(v_1, \ldots, v_m)(\lambda_1, \ldots, \lambda_m) = \lambda_1 v_1 + \cdots + \lambda_m v_m.$$

Then T is linear since for all $\alpha, \beta \in \mathbf{F}$ and all $v = (v_1, \ldots, v_m), w = (w_1, \ldots, w_m) \in V^m$, we have

$$\begin{aligned} T(\alpha v + \beta w)(\lambda_1, \ldots, \lambda_m) &= T(\alpha v_1 + \beta w_1, \ldots, \alpha v_m + \beta w_m)(\lambda_1, \ldots, \lambda_m) \\ &= \lambda_1(\alpha v_1 + \beta w_1) + \cdots + \lambda_m(\alpha v_m + \beta w_m) \\ &= \alpha(\lambda_1 v_1 + \cdots + \lambda_m v_m) + \beta(\lambda_1 w_1 + \cdots + \lambda_m w_m) \\ &= \alpha T(v)(\lambda_1, \ldots, \lambda_m) + \beta T(w)(\lambda_1, \ldots, \lambda_m), \end{aligned}$$

hence T is a linear map. Now suppose $T(v) = 0$. Then $T(v)(e_k) = 0$ for all $k = 1, \ldots, m$. It follows that $v_k = 0$ for all $k = 1, \ldots, m$, hence $v = 0$, and thus T is injective.

Now let $S \in \mathcal{L}(\mathbf{F}^m, V)$. Define $v \in V^m$ by $v_k = Se_k$ for all $k = 1, \ldots, m$. Then $T(v) = S$. Thus, T is surjective. Therefore, T is an isomorphism. ∎

Exercise 3E6

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. Let $u \in U$. Then $v + u \in v + U = x + W$. Thus, there exists $w \in W$ such that $v + u = x + w$. Note that $v \in x + W$ implies the existence of some $w' \in W$ such that $v = x + w'$, or $x - v = -w' \in W$. Similarly, $x \in v + U$ implies the existence of some $u' \in U$ such that $x = v + u'$, or $v - x = -u' \in U$. It follows that $u = (x - v) + w \in W$, hence $U \subseteq W$. Now let $w \in W$. Then $x + w \in x + W = v + U$. Thus, there exists $u \in U$ such that $x + w = v + u$. It follows that $w = (v - x) + u \in U$, hence $W \subseteq U$. Therefore, $U = W$. ∎

Exercise 3E7

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 A is a translate of U. Then there exists $v \in \mathbf{R}^3$ such that $A = v + U$. Let $v = (x_0, y_0, z_0)$. Then

$$\begin{aligned} A &= v + U \\ &= \{(x_0, y_0, z_0) + (x, y, z) \in \mathbf{R}^3 : 2x + 3y + 5z = 0\} \\ &= \{(x', y', z') \in \mathbf{R}^3 : 2(x' - x_0) + 3(y' - y_0) + 5(z' - z_0) = 0\} \\ &= \{(x', y', z') \in \mathbf{R}^3 : 2x' + 3y' + 5z' = 2x_0 + 3y_0 + 5z_0\}. \end{aligned}$$

Conversely, suppose there exists $c \in \mathbf{R}$ such that

$$A = \{(x, y, z) \in \mathbf{R}^3 : 2x + 3y + 5z = c\}.$$

Let $v = (x_0, y_0, z_0)$ where x_0, y_0, z_0 are any real numbers such that $2x_0 + 3y_0 + 5z_0 = c$. Then

$$\begin{aligned} v + U &= (x_0, y_0, z_0) + U \\ &= \{(x_0, y_0, z_0) + (x, y, z) \in \mathbf{R}^3 : 2x + 3y + 5z = 0\} \\ &= A. \end{aligned}$$

■

Exercise 3E8

(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) Suppose $T \in \mathcal{L}(V, W)$ and $c \in W$. Let $A = \{x \in V : Tx = c\}$. If A is nonempty, then there exists $v \in V$ such that $Tv = c$. We claim that $A = v + \operatorname{null} T$. Let $x \in A$. Then $Tx = c = Tv$, hence $T(x - v) = 0$, and thus $x - v \in \operatorname{null} T$. It follows that $x \in v + \operatorname{null} T$, hence $A \subseteq v + \operatorname{null} T$. Now let $y \in v + \operatorname{null} T$. Then there exists $u \in \operatorname{null} T$ such that $y = v + u$. It follows that $Ty = T(v + u) = Tv + Tu = c + 0 = c$, hence $y \in A$, and thus $v + \operatorname{null} T \subseteq A$. Therefore, $A = v + \operatorname{null} T$.

(b) A system of linear equations such as 3.27 can be written in the form $Tx = c$ for some linear map T and some vector c. By part (a), the set of solutions to this system of linear equations is either the empty set or is a translate of $\operatorname{null} T$, which is a subspace of $\mathbf{F}^n$. ■

Exercise 3E9

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 A is a translate of some subspace of V. Then there exists $v \in V$ and a subspace U of V such that $A = v + U$. Let $v + u, v + u' \in A$. Then for any $\lambda \in \mathbf{F}$,

$$\lambda(v + u) + (1 - \lambda)(v + u') = v + \underbrace{(\lambda u + (1 - \lambda)u')}_{\in U} \in A.$$

Conversely, suppose $\lambda v + (1 - \lambda)w \in A$ for all $v, w \in A$ and all $\lambda \in \mathbf{F}$. Fix some $v_0 \in A$, and let $U = \{u \in V : v_0 + u \in A\}$. We first claim that U is a subspace of V. Since $v_0 = v_0 + 0 \in A$, we have $0 \in U$. Let $u_1, u_2 \in U$. Then $v_0 + u_1, v_0 + u_2 \in A$. Now let $\lambda \in \mathbf{F}$. Then

$$\begin{aligned} \lambda(v_0 + u_1) + (1 - \lambda)(v_0 + u_2) &= v_0 + \lambda u_1 + (1 - \lambda)u_2 \\ &= v_0 + (\lambda u_1 + (1 - \lambda)u_2) \in A. \end{aligned}$$

It follows that $\lambda u_1 + (1 - \lambda)u_2 \in U$, hence U is a subspace of V. Now we claim that $A = v_0 + U$. Let $w \in A$. By definition, $v_0 + (w - v_0) = w \in A$, hence $w - v_0 \in U$, and thus $w \in v_0 + U$, so $A \subseteq v_0 + U$. Now let $v_0 + u \in v_0 + U$. Then $v_0 + u \in A$ by definition of U, hence $v_0 + U \subseteq A$. Therefore, $A = v_0 + U$. ■

Exercise 3E10

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 the empty set, then we are done. Suppose that $A_1 \cap A_2$ is nonempty. Let $x, y \in A_1 \cap A_2$ and $\lambda \in \mathbf{F}$. Then $x, y \in A_1$, hence $\lambda x + (1-\lambda)y \in A_1$. Moreover, $x, y \in A_2$, hence $\lambda x + (1-\lambda)y \in A_2$. It follows that $\lambda x + (1-\lambda)y \in A_1 \cap A_2$. Thus, by Exercise 3E9, $A_1 \cap A_2$ is a translate of some subspace of V. ■

Exercise 3E11 ✔

Suppose $U = \{(x_1, x_2, \ldots) \in \mathbf{F}^\infty : x_k \neq 0 \text{ for only finitely many } k\}$.

(a) Show that U is a subspace of $\mathbf{F}^\infty$.

(b) Prove that $\mathbf{F}^\infty / U$ is infinite-dimensional.

Solution.

(a) We claim that U can be reformulated as the set of vectors $(x_1, x_2, \ldots) \in \mathbf{F}^\infty$ such that there exists an index N such that $x_k = 0$ for all $k > N$. To see this, suppose $(x_1, x_2, \ldots) \in U$. Then there exists a finite set of indices K such that $x_k = 0$ for all $k \notin K$. If K is empty, let $N = 0$. Otherwise, let $N = \max K$. Then $x_k = 0$ for all $k > N$.

Conversely, suppose there exists an index N such that $x_k = 0$ for all $k > N$. Then the set of indices $\{1, 2, \ldots, N\}$ is finite and contains all indices where $x_k \neq 0$, hence $(x_1, x_2, \ldots) \in U$. Thus, the two formulations of U are equivalent.

We now show that U is a subspace of $\mathbf{F}^\infty$. First, note that the zero vector $(0, 0, \ldots) \in U$ since it has no nonzero coordinates. Now let $u = (u_1, u_2, \ldots), v = (v_1, v_2, \ldots) \in U$. Then there exist indices N_u and N_v such that $u_k = 0$ for all $k > N_u$ and $v_k = 0$ for all $k > N_v$. Let $N = \max(N_u, N_v)$. Then for all $k > N$, we have $u_k = 0$ and $v_k = 0$, hence $(u+v)_k = u_k + v_k = 0 + 0 = 0$. Thus, $u + v \in U$. Now let $\alpha \in \mathbf{F}$. Then for all $k > N_u$, we have $(\alpha u)_k = \alpha u_k = \alpha \cdot 0 = 0$, hence $\alpha u \in U$. Therefore, U is a subspace of $\mathbf{F}^\infty$.

(b) For any given positive integer m, define the sets B_i for $1 \leq i \leq m$ as follows:

$$B_i = \{i + km : k \in \mathbb{N}\} = \{i, i+m, i+2m, \ldots\}.$$

Then $B_1, \ldots, B_m$ are pairwise disjoint subsets of $\mathbb{N}$. For each $i = 1, \ldots, m$, define

$$v_i = (v_{i,1}, v_{i,2}, \ldots) \in \mathbf{F}^\infty$$

such that $v_{i,j} = 1$ if $j \in B_i$ and $v_{i,j} = 0$ otherwise. We claim that $v_1 + U, \ldots, v_m + U$ are linearly independent in $\mathbf{F}^\infty / U$.

Let $\lambda_1, \ldots, \lambda_m \in \mathbf{F}$ such that $\lambda_1(v_1 + U) + \cdots + \lambda_m(v_m + U) = U$. Then $\lambda_1 v_1 + \cdots + \lambda_m v_m \in U$. Then we can show that $\lambda_i = 0$ for all $i = 1, \ldots, m$ as follows. For each i, consider the j-th coordinate of v_i for $j \in B_i$. Since v_i has 1's in all coordinates indexed by B_i and 0's elsewhere, and B_i is disjoint from B_k for $k \neq i$, the j-th coordinate of $\lambda_1 v_1 + \cdots + \lambda_m v_m$ is λ_i. But this sum is in U, which means it has only finitely many nonzero coordinates. Since B_i is infinite, it follows that $\lambda_i = 0$. Hence, all $\lambda_i = 0$, proving that $v_1 + U, \ldots, v_m + U$ are linearly independent. Since m was arbitrary, $\mathbf{F}^\infty / U$ is infinite-dimensional. ■

Exercise 3E12

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) For any element $\lambda_1 v_1 + \cdots + \lambda_m v_m \in A$, we can write

$$\lambda_1 v_1 + \cdots + \lambda_m v_m = v_1 + \lambda_2(v_2 - v_1) + \cdots + \lambda_m(v_m - v_1) = v_1 + \sum_{i=2}^{m} \lambda_i(v_i - v_1),$$

where $\lambda_1 = 1 - (\lambda_2 + \cdots + \lambda_m)$. Let $U = \operatorname{span}(v_2 - v_1, \ldots, v_m - v_1)$. Then $A = v_1 + U$, hence A is a translate of the subspace U of V.

(b) By Exercise 3E9, since B is a translate of some subspace of V, we have $\lambda v + (1 - \lambda)w \in B$ for all $v, w \in B$ and all $\lambda \in \mathbf{F}$. Since $\{v_1, \ldots, v_m\} \subseteq B$, it follows that any linear combination of $v_1, \ldots, v_m$ with coefficients summing to 1 is in B. Hence, $A \subseteq B$.

(c) The subspace $U = \operatorname{span}(v_2 - v_1, \ldots, v_m - v_1)$ has dimension at most $m - 1$, since it is spanned by $m - 1$ vectors. Therefore, $A = v_1 + U$ is a translate of a subspace of V of dimension less than m. ■

Exercise 3E13

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. Since V/U is finite-dimensional, let $v_1 + U, \ldots, v_m + U$ be a basis of V/U. Let $W = \operatorname{span}(v_1, \ldots, v_m)$. We show that $V = U \oplus W$. First, note that any $v \in V$ can be written as $v = u + w$ for some $u \in U$ and $w \in W$. Indeed, since $v + U \in V/U$, we can write

$$v + U = \lambda_1(v_1 + U) + \cdots + \lambda_m(v_m + U)$$

for some scalars $\lambda_1, \ldots, \lambda_m$. This gives

$$v - (\lambda_1 v_1 + \cdots + \lambda_m v_m) \in U.$$

So setting $u = v - (\lambda_1 v_1 + \cdots + \lambda_m v_m) \in U$ and $w = \lambda_1 v_1 + \cdots + \lambda_m v_m \in W$, we have $v = u + w$. To see that the sum is direct, suppose $u + w = 0$ with $u \in U$ and $w \in W$. Then $w = -u \in W \cap U$. Since $w \in W$, we can write $w = \lambda_1 v_1 + \cdots + \lambda_m v_m$ for some scalars $\lambda_1, \ldots, \lambda_m$. Since we established that $w \in U$, we have $w + U = 0 + U$. But $v_1 + U, \ldots, v_m + U$ are linearly independent in V/U, hence $\lambda_1 = \cdots = \lambda_m = 0$, and thus $w = 0$. It follows that $u = 0$ as well, proving that the sum is direct. Therefore, $V = U \oplus W$.

Consider the map $T : V \to U \times (V/U)$ given by $Tv = (u, v + U)$, where $v = u + w$ with $u \in U$ and $w \in W$. This map is well-defined because the decomposition $v = u + w$ is unique. It is linear, injective, and surjective, hence an isomorphism. Therefore, $V \cong U \times (V/U)$. ■

Exercise 3E14

Suppose U and W are subspaces of V and $V = U \oplus W$. Suppose $w_1, \ldots, w_m$ is a basis of W. Prove that $w_1 + U, \ldots, w_m + U$ is a basis of V/U.

Solution. Since $V = U \oplus W$, every $v \in V$ can be written uniquely as $v = u + w$ with $u \in U$ and $w \in W$. We claim that $w_1 + U, \ldots, w_m + U$ is a basis of V/U. First, we show that they span V/U. Let $v + U \in V/U$. Write $v = u + w$ with $u \in U$ and $w \in W$. Then $v + U = (u + w) + U = w + U$. Since $w \in W$, we can write $w = \lambda_1 w_1 + \cdots + \lambda_m w_m$ for some scalars $\lambda_1, \ldots, \lambda_m$. Hence, $v + U = \lambda_1(w_1 + U) + \cdots + \lambda_m(w_m + U)$, showing that $w_1 + U, \ldots, w_m + U$ span V/U.

Next, we show linear independence. Suppose $\lambda_1(w_1 + U) + \cdots + \lambda_m(w_m + U) = 0 + U$. Then $\lambda_1 w_1 + \cdots + \lambda_m w_m \in U$. But $\lambda_1 w_1 + \cdots + \lambda_m w_m \in W$ and $U \cap W = \{0\}$, so $\lambda_1 w_1 + \cdots + \lambda_m w_m = 0$. Since $w_1, \ldots, w_m$ are linearly independent, we have $\lambda_1 = \cdots = \lambda_m = 0$. Therefore, $w_1 + U, \ldots, w_m + U$ are linearly independent, and hence form a basis of V/U. ■

Exercise 3E15

Suppose U is a subspace of V and $v_1 + U, \dots, v_m + U$ is a basis of V/U and $u_1, \dots, u_n$ is a basis of U. Prove that $v_1, \dots, v_m, u_1, \dots, u_n$ is a basis of V.

Solution. First, we show that $v_1, \dots, v_m, u_1, \dots, u_n$ span V. Let $v \in V$. Then $v + U \in V/U$, so we can write

$$v + U = \lambda_1(v_1 + U) + \cdots + \lambda_m(v_m + U)$$

for some scalars $\lambda_1, \dots, \lambda_m$. This gives $v - (\lambda_1 v_1 + \cdots + \lambda_m v_m) \in U$. Since $u_1, \dots, u_n$ is a basis of U, we can write $v - (\lambda_1 v_1 + \cdots + \lambda_m v_m) = \mu_1 u_1 + \cdots + \mu_n u_n$ for some scalars $\mu_1, \dots, \mu_n$. Hence, $v = \lambda_1 v_1 + \cdots + \lambda_m v_m + \mu_1 u_1 + \cdots + \mu_n u_n$, showing that $v_1, \dots, v_m, u_1, \dots, u_n$ span V.

Next, we show linear independence. Suppose $\lambda_1 v_1 + \cdots + \lambda_m v_m + \mu_1 u_1 + \cdots + \mu_n u_n = 0$. Then $\lambda_1 v_1 + \cdots + \lambda_m v_m \in U$. Then

$$\lambda_1(v_1 + U) + \cdots + \lambda_m(v_m + U) = 0 + U.$$

Since $v_1 + U, \dots, v_m + U$ are linearly independent in V/U, we must have $\lambda_1 = \cdots = \lambda_m = 0$. Then the original equation reduces to $\mu_1 u_1 + \cdots + \mu_n u_n = 0$, and since $u_1, \dots, u_n$ are linearly independent, we have $\mu_1 = \cdots = \mu_n = 0$. Therefore, $v_1, \dots, v_m, u_1, \dots, u_n$ are linearly independent, and hence form a basis of V. ∎

Exercise 3E16

Suppose $\varphi \in \mathcal{L}(V, \mathbf{F})$ and $\varphi \neq 0$. Prove that $\dim V/(\operatorname{null} \varphi) = 1$.

Solution. From 3.107 in the text, we know that $V/\operatorname{null} \varphi \cong \operatorname{range} \varphi$. Since $\varphi \neq 0$, we have $\operatorname{range} \varphi \neq \{0\}$, and since $\operatorname{range} \varphi$ is a subspace of $\mathbf{F}$, it follows that $\dim \operatorname{range} \varphi = 1$. This implies $\operatorname{range} \varphi = \mathbf{F}$, and hence $\dim V/(\operatorname{null} \varphi) = \dim \operatorname{range} \varphi = 1$. ∎

Exercise 3E17

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. Since $\dim V/U = \dim \mathbf{F} = 1$, there exists some isomorphism $\psi \in \mathcal{L}(V/U, \mathbf{F})$. Let $\pi : V \to V/U$ be the quotient map. Define $\varphi = \psi \circ \pi \in \mathcal{L}(V, \mathbf{F})$. Suppose $\varphi(v) = 0$. Then $\psi(\pi(v)) = 0$, which implies $\pi(v) = 0$ since ψ is an isomorphism. Hence, $v \in U$, and we have $\operatorname{null} \varphi \subseteq U$. Conversely, if $v \in U$, then $\pi(v) = 0$, so $\varphi(v) = \psi(\pi(v)) = 0$. Therefore, $\operatorname{null} \varphi = U$. ∎

Exercise 3E18

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) Suppose W is a finite-dimensional subspace of V and $V = U + W$. Let $\pi : V \to V/U$ be the quotient map, and consider the restriction $\pi|_W : W \to V/U$. To see that $\operatorname{range}(\pi|_W) = V/U$, note that for any $v + U \in V/U$, we can write $v = u + w$ for some $u \in U$ and $w \in W$. Then $\pi(v) = \pi(u + w) = \pi(u) + \pi(w) = 0 + \pi(w) = \pi(w)$, so $v + U = \pi(w) \in \operatorname{range}(\pi|_W)$. Hence, $\operatorname{range}(\pi|_W) = V/U$. From the fundamental theorem of linear maps, we have

$$\dim W = \dim \operatorname{null} \pi|_W + \dim \operatorname{range}(\pi|_W) = \dim \operatorname{null} \pi|_W + \dim V/U \geq \dim V/U.$$

(b) Let $v_1 + U, \dots, v_m + U$ be a basis of V/U. We claim that $v_1, \dots, v_m$ are linearly independent. Suppose $\lambda_1 v_1 + \cdots + \lambda_m v_m = 0$ for some scalars $\lambda_1, \dots, \lambda_m$. Then

$$\lambda_1 v_1 + \cdots + \lambda_m v_m \in U.$$

But $\lambda_1(v_1 + U) + \cdots + \lambda_m(v_m + U) = 0$ in V/U, and since $v_1 + U, \dots, v_m + U$ are linearly independent, it follows that $\lambda_1 = \cdots = \lambda_m = 0$. Hence, $v_1, \dots, v_m$ are linearly independent.

Let $W = \operatorname{span}\{v_1, \dots, v_m\}$. Then $\dim W = m = \dim V/U$. For any $v \in V$, we can write

$$v + U = \lambda_1(v_1 + U) + \cdots + \lambda_m(v_m + U).$$

This implies that

$$v - (\lambda_1 v_1 + \cdots + \lambda_m v_m) \in U,$$

so $v \in U + W$. Hence, $V = U + W$. To see that $U \cap W = \{0\}$, let $w \in U \cap W$. Then $w \in W$, so $w = \lambda_1 v_1 + \cdots + \lambda_m v_m$ for some scalars $\lambda_1, \dots, \lambda_m$. But $w \in U$, so $w + U = 0$ in V/U. Hence,

$$\lambda_1(v_1 + U) + \cdots + \lambda_m(v_m + U) = 0$$

in V/U. Since $v_1 + U, \dots, v_m + U$ are linearly independent, it follows that $\lambda_1 = \cdots = \lambda_m = 0$. Hence, $w = 0$, and we have $U \cap W = \{0\}$. Therefore, $V = U \oplus W$. ∎

Exercise 3E19

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. Suppose $S \in \mathcal{L}(V/U, W)$ is such that $T = S \circ \pi$. Then for any $u \in U$, we have $T(u) = S(\pi(u)) = S(0) = 0$, so $U \subseteq \operatorname{null} T$.

Conversely, suppose $U \subseteq \operatorname{null} T$. Define $S : V/U \to W$ by $S(v + U) = T(v)$. This is well-defined because if $v + U = v' + U$, then $v - v' \in U \subseteq \operatorname{null} T$, so $T(v) = T(v')$. It is straightforward to check that S is linear and that $T = S \circ \pi$. ∎

3F Duality

Manual Remark. *For the rest of the solutions manual, unless otherwise stated, we will be using the Kronecker delta notation, i.e., we define $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ if $i \neq j$.*

Exercise 3F1

Explain why each linear functional is surjective or is the zero map.

Solution. Let $\varphi \in V'$. If $\varphi = 0$, then it is the zero map. Otherwise, there exists $v \in V$ such that $\varphi(v) \neq 0$, say $\varphi(v) = c \neq 0$. Then for any $x \in \mathbf{F}$, we have

$$\varphi\left(\frac{x}{c}v\right) = \frac{x}{c}\varphi(v) = \frac{x}{c}c = x.$$

Hence, φ is surjective. ∎

Exercise 3F2

Give three distinct examples of linear functionals on $\mathbf{R}^{[0,1]}$.

Solution. Consider the linear functionals on $\mathbf{R}^{[0,1]}$ defined by

$$\varphi_1(f) = f(0), \quad \varphi_2(f) = f(1), \quad \varphi_3(f) = f\left(\frac{1}{2}\right).$$ ∎

Exercise 3F3

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. Since V is finite-dimensional, we can extend v to a basis of V, say $v, v_2, \ldots, v_n$. Define $\varphi \in V'$ by setting $\varphi(v) = 1$ and $\varphi(v_i) = 0$ for $i = 2, \ldots, n$. This defines a linear functional because it is defined on a basis and extended linearly. Clearly, $\varphi(v) = 1$. ∎

Exercise 3F4

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. Since $U \neq V$, there exists $v_1 \in V$ such that $v_1 \notin U$. Extend a basis of U to a basis of V, say $u_1, \ldots, u_k, v_1, \ldots, v_m$. Define $\varphi \in V'$ by setting $\varphi(u_i) = 0$ for all $i = 1, \ldots, k$ and $\varphi(v_j) = 0$ for all $j = 2, \ldots, m$ but $\varphi(v_1) = 1$. This defines a linear functional because it is defined on a basis and extended linearly. Clearly, $\varphi(u) = 0$ for all $u \in U$ and $\varphi \neq 0$. ∎

Exercise 3F5

Suppose $T \in \mathcal{L}(V, W)$ and $w_1, \ldots, w_m$ is a basis of range T. Hence for each $v \in V$, there exist unique numbers $\varphi_1(v), \ldots, \varphi_m(v)$ such that

$$Tv = \varphi_1(v)w_1 + \cdots + \varphi_m(v)w_m,$$

thus defining functions $\varphi_1, \ldots, \varphi_m$ from V to $\mathbf{F}$. Show that each of the functions $\varphi_1, \ldots, \varphi_m$ is a linear functional on V.

Solution. Let $v \in V$. Since $w_1, \ldots, w_m$ is a basis of range T, there exist unique scalars $\varphi_1(v), \ldots, \varphi_m(v)$ such that

$$Tv = \varphi_1(v)w_1 + \cdots + \varphi_m(v)w_m.$$

To show that each φ_i is a linear functional, let $v, u \in V$ and $a, b \in \mathbf{F}$. Then

$$T(av + bu) = aTv + bTu = a(\varphi_1(v)w_1 + \cdots + \varphi_m(v)w_m) + b(\varphi_1(u)w_1 + \cdots + \varphi_m(u)w_m).$$

By the uniqueness of the representation in terms of the basis $w_1, \ldots, w_m$, we must have

$$\varphi_i(av + bu) = a\varphi_i(v) + b\varphi_i(u) \quad \text{for each } i = 1, \ldots, m.$$

Hence, each φ_i is a linear functional on V. ∎

Exercise 3F6

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. Suppose $\operatorname{null} \varphi \subseteq \operatorname{null} \beta$. If $\varphi = 0$, then $\operatorname{null} \varphi = V$, and hence $\operatorname{null} \beta = V$, which implies $\beta = 0$. In this case, we can take $c = 0$ and have $\beta = c\varphi$. Now assume $\varphi \neq 0$. Pick $v_0 \in V$ such that $\varphi(v_0) \neq 0$. Define

$$c = \frac{\beta(v_0)}{\varphi(v_0)}.$$

For any $v \in V$, consider

$$v - \frac{\varphi(v)}{\varphi(v_0)}v_0.$$

This vector is in $\operatorname{null} \varphi$, so it is also in $\operatorname{null} \beta$. Hence,

$$\beta\left(v - \frac{\varphi(v)}{\varphi(v_0)}v_0\right) = 0 \implies \beta(v) = \frac{\varphi(v)}{\varphi(v_0)}\beta(v_0) = c\varphi(v).$$

Therefore, $\beta = c\varphi$.

Conversely, if $\beta = c\varphi$, then clearly $\operatorname{null} \varphi \subseteq \operatorname{null} \beta$ because if $\varphi(v) = 0$, then $\beta(v) = c\varphi(v) = 0$. ∎

Exercise 3F7

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. Define a map $\Phi : (V_1 \times \cdots \times V_m)' \to V_1' \times \cdots \times V_m'$ by

$$\Phi(\varphi) = (\varphi_1, \ldots, \varphi_m),$$

where $\varphi_i \in V_i'$ is defined by $\varphi_i(v_i) = \varphi(0, \ldots, 0, v_i, 0, \ldots, 0)$, with v_i in the ith position.

First, we show that Φ is linear. Let $\varphi, \psi \in (V_1 \times \cdots \times V_m)'$ and $a, b \in \mathbf{F}$. Then

$$\Phi(a\varphi + b\psi) = ((a\varphi + b\psi)_1, \ldots, (a\varphi + b\psi)_m) = (a\varphi_1 + b\psi_1, \ldots, a\varphi_m + b\psi_m) = a\Phi(\varphi) + b\Phi(\psi).$$

Next, we show that Φ is bijective. For injectivity, suppose $\Phi(\varphi) = 0$. Then $\varphi_i = 0$ for all i, which implies $\varphi(0, \ldots, 0, v_i, 0, \ldots, 0) = 0$ for all $v_i \in V_i$. By linearity, φ must be zero on all of $V_1 \times \cdots \times V_m$, so $\varphi = 0$. For surjectivity, given $(\varphi_1, \ldots, \varphi_m) \in V_1' \times \cdots \times V_m'$, define $\varphi \in (V_1 \times \cdots \times V_m)'$ by

$$\varphi(v_1, \ldots, v_m) = \varphi_1(v_1) + \cdots + \varphi_m(v_m),$$

which is linear by linearity of each φ_i. Then $\Phi(\varphi) = (\varphi_1, \ldots, \varphi_m)$, so Φ is surjective.

Therefore, Φ is a linear isomorphism, and $(V_1 \times \cdots \times V_m)'$ and $V_1' \times \cdots \times V_m'$ are isomorphic vector spaces. ∎

Exercise 3F8

Suppose $v_1, \dots, v_n$ is a basis of V and $\varphi_1, \dots, \varphi_n$ is the dual basis of V'. Define $\Gamma : V \to \mathbf{F}^n$ and $\Lambda : \mathbf{F}^n \to V$ by

$$\Gamma(v) = (\varphi_1(v), \dots, \varphi_n(v)) \quad \text{and} \quad \Lambda(a_1, \dots, a_n) = a_1 v_1 + \cdots + a_n v_n.$$

Explain why Γ and Λ are inverses of each other.

Solution. To show that Γ and Λ are inverses of each other, we need to check that $\Gamma(\Lambda(a_1, \dots, a_n)) = (a_1, \dots, a_n)$ for all $(a_1, \dots, a_n) \in \mathbf{F}^n$ and $\Lambda(\Gamma(v)) = v$ for all $v \in V$.

First, let $(a_1, \dots, a_n) \in \mathbf{F}^n$. Then

$$\Gamma(\Lambda(a_1, \dots, a_n)) = \Gamma(a_1 v_1 + \cdots + a_n v_n) = (\varphi_1(a_1 v_1 + \cdots + a_n v_n), \dots, \varphi_n(a_1 v_1 + \cdots + a_n v_n)).$$

By the definition of the dual basis, we have $\varphi_i(v_j) = \delta_{ij}$ for all i, j. Therefore,

$$\varphi_i(a_1 v_1 + \cdots + a_n v_n) = a_i.$$

Therefore,

$$\Gamma(\Lambda(a_1, \dots, a_n)) = (a_1, \dots, a_n).$$

Next, let $v \in V$. Write $v = b_1 v_1 + \cdots + b_n v_n$ for some scalars $b_1, \dots, b_n$. Then

$$\Lambda(\Gamma(v)) = \Lambda(\varphi_1(v), \dots, \varphi_n(v)) = \Lambda(b_1, \dots, b_n) = b_1 v_1 + \cdots + b_n v_n = v.$$

Hence, Γ and Λ are inverses of each other. ■

Exercise 3F9

Suppose m is a positive integer. Show that the dual basis of the basis $1, x, \dots, x^m$ of $\mathcal{P}_m(\mathbf{R})$ is $\varphi_0, \varphi_1, \dots, \varphi_m$, where

$$\varphi_k(p) = \frac{p^{(k)}(0)}{k!}.$$

Axler Note. *Here $p^{(k)}$ denotes the kth derivative of p, with the understanding that the 0th derivative of p is p.*

Solution. Let $p \in \mathcal{P}_m(\mathbf{R})$. We can write $p(x) = a_0 + a_1 x + \cdots + a_m x^m$ for some scalars $a_0, \dots, a_m$. Then for each $k = 0, 1, \dots, m$, we have

$$\begin{aligned}
p(x) &= a_0 + a_1 x + a_2 x^2 + \cdots + a_m x^m \\
p'(x) &= a_1 + 2a_2 x + \cdots + m a_m x^{m-1} \\
p''(x) &= 2a_2 + \cdots + m(m-1) a_m x^{m-2} \\
&\vdots \\
p^{(k)}(x) &= k! a_k + \cdots + \frac{m!}{(m-k)!} a_m x^{m-k}
\end{aligned}$$

Then

$$\varphi_k(p) = \frac{p^{(k)}(0)}{k!} = a_k.$$

Therefore, $\varphi_k(x^j) = \delta_{kj}$ for all $0 \le k, j \le m$, which shows that $\varphi_0, \varphi_1, \dots, \varphi_m$ is indeed the dual basis of $1, x, \dots, x^m$. ■

Exercise 3F10

Suppose m is a positive integer.

(a) Show that $1, x-5, \dots, (x-5)^m$ is a basis of $\mathcal{P}_m(\mathbf{R})$.

(b) What is the dual basis of the basis in (a)?

Solution.

(a) Note that $1, x-5, \dots, (x-5)^m$ is a list of $m+1$ polynomials in $\mathcal{P}_m(\mathbf{R})$ with differing degrees, and since $\mathcal{P}_m(\mathbf{R})$ has dimension $m+1$, this list forms a basis of $\mathcal{P}_m(\mathbf{R})$.

(b) Substitute $y = x - 5$. Then the basis $1, x-5, \dots, (x-5)^m$ becomes $1, y, \dots, y^m$. By Exercise 3F9, the dual basis of $1, y, \dots, y^m$ is $\varphi_0, \varphi_1, \dots, \varphi_m$, where

$$\varphi_k(p) = \frac{p^{(k)}(0)}{k!}.$$

Since $y = x - 5$, then

$$\varphi_k(p) = \frac{p^{(k)}(5)}{k!}.$$

■

Exercise 3F11

Suppose $v_1, \dots, v_n$ is a basis of V and $\varphi_1, \dots, \varphi_n$ 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 $v \in V$. Since $\varphi_1, \dots, \varphi_n$ is the dual basis of $v_1, \dots, v_n$, we have $\varphi_i(v_j) = \delta_{ij}$ for all $1 \le i, j \le n$. Let $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$ for some scalars $\alpha_1, \dots, \alpha_n$. Then

$$\begin{aligned}(\psi(v_1)\varphi_1 + \cdots + \psi(v_n)\varphi_n)(v) &= \psi(v_1)\varphi_1(v) + \cdots + \psi(v_n)\varphi_n(v)\\ &= \psi(v_1)\alpha_1 + \cdots + \psi(v_n)\alpha_n\\ &= \psi(\alpha_1 v_1 + \cdots + \alpha_n v_n)\\ &= \psi(v)\end{aligned}$$

Since this holds for all $v \in V$, we conclude that

$$\psi = \psi(v_1)\varphi_1 + \cdots + \psi(v_n)\varphi_n.$$

■

Exercise 3F12

Suppose $S, T \in \mathcal{L}(V, W)$.

(a) Prove that $(S+T)' = S' + T'$.

(b) Prove that $(\lambda T)' = \lambda T'$ for all $\lambda \in \mathbf{F}$.

Solution.

(a) Let $\psi \in W'$. Then for any $v \in V$, we have

$$\begin{aligned}((S+T)'(\psi))(v) &= \psi((S+T)(v))\\ &= \psi(S(v) + T(v))\\ &= \psi(S(v)) + \psi(T(v))\\ &= (S'(\psi))(v) + (T'(\psi))(v)\\ &= (S'(\psi) + T'(\psi))(v).\end{aligned}$$

Since this holds for all $v \in V$, we conclude that

$$(S+T)'(\psi) = S'(\psi) + T'(\psi).$$

Since this holds for all $\psi \in W'$, we have

$$(S+T)' = S' + T'.$$

(b) Let $\psi \in W'$. Then for any $v \in V$, we have

$$\begin{aligned}((\lambda T)'(\psi))(v) &= \psi((\lambda T)(v)) \\ &= \psi(\lambda T(v)) \\ &= \lambda\psi(T(v)) \\ &= (\lambda T'(\psi))(v).\end{aligned}$$

Since this holds for all $v \in V$, we conclude that

$$(\lambda T)'(\psi) = \lambda T'(\psi).$$

Since this holds for all $\psi \in W'$, we have

$$(\lambda T)' = \lambda T'.$$

■

Exercise 3F13

Show that the dual map of the identity operator on V is the identity operator on V'.

Solution. Let $\psi \in V'$. Let I' denote the dual map of the identity operator I on V. Then for any $v \in V$, we have

$$\begin{aligned}(I'(\psi))(v) &= \psi(I(v)) \\ &= \psi(v).\end{aligned}$$

Since this holds for all $v \in V$, we conclude that

$$I'(\psi) = \psi.$$

Since this holds for all $\psi \in V'$, then I' must be the identity operator on V'. ■

Exercise 3F14

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) Let $(x, y, z) \in \mathbf{R}^3$. Then we have

$$\begin{aligned}(T'(\varphi_1))((x,y,z)) &= \varphi_1(T(x,y,z)) \\ &= \varphi_1(4x+5y+6z, 7x+8y+9z) \\ &= 4x+5y+6z, \\ (T'(\varphi_2))((x,y,z)) &= \varphi_2(T(x,y,z)) \\ &= \varphi_2(4x+5y+6z, 7x+8y+9z) \\ &= 7x+8y+9z.\end{aligned}$$

(b) We can write $T'(\varphi_1)$ and $T'(\varphi_2)$ as linear combinations of ψ_1, ψ_2, ψ_3 as follows:

$$\begin{aligned}T'(\varphi_1) &= 4\psi_1 + 5\psi_2 + 6\psi_3, \\ T'(\varphi_2) &= 7\psi_1 + 8\psi_2 + 9\psi_3.\end{aligned}$$

■

Exercise 3F15

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) Let $p \in \mathcal{P}(\mathbf{R})$. Then we have

$$\begin{aligned}(T'(\varphi))(p) &= \varphi(T(p)) \\ &= (Tp)'(4) \\ &= (2xp(x) + x^2 p'(x) + p'''(x))\big|_{x=4}. \\ &= 8p(4) + 16p'(4) + p'''(4).\end{aligned}$$

(b) Let $p(x) = x^3$. Then we have

$$\begin{aligned}(T'(\varphi))(x^3) &= \varphi(T(x^3)) \\ &= \int_0^1 T(x^3)\,\mathrm{d}x \\ &= \int_0^1 (x^2 \cdot x^3 + (x^3)'')\,\mathrm{d}x \\ &= \int_0^1 (x^5 + 6x)\,\mathrm{d}x \\ &= \frac{1}{6} + \frac{18}{6} \\ &= \frac{19}{6}.\end{aligned}$$

■

Exercise 3F16

Suppose W is finite-dimensional and $T \in \mathcal{L}(V, W)$. Prove that

$$T' = 0 \iff T = 0.$$

Solution. Suppose $T' = 0$. Then for every $\varphi \in W'$ and every $v \in V$, we have

$$(T'(\varphi))(v) = \varphi(T(v)) = 0.$$

If $T(v) \neq 0$ for some $v \in V$, then $T(v) \in W$. Since W is finite-dimensional, then by Exercise 3F3, there exists $\psi \in W'$ such that $\psi(T(v)) \neq 0$, which contradicts the fact that $(T'(\psi))(v) = \psi(T(v)) = 0$. Hence, $T(v) = 0$ for every $v \in V$, and so $T = 0$.

Conversely, suppose $T = 0$. Then for every $\varphi \in W'$ and every $v \in V$, we have

$$(T'(\varphi))(v) = \varphi(T(v)) = \varphi(0) = 0.$$

Hence, $T' = 0$. ∎

Exercise 3F17

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. Suppose T is invertible. Then there exists $S \in \mathcal{L}(W, V)$ such that $ST = I_V$ and $TS = I_W$. Taking the dual of both sides, we get $T'S' = I_{W'}$ and $S'T' = I_{V'}$, which shows that T' is invertible with inverse S'.

Conversely, suppose T' is invertible. Then T' is an isomorphism. Then

$$\dim V = \dim V' = \dim W' = \dim W.$$

By injectivity of T', we have $\operatorname{null} T' = \{0\}$. Since $\operatorname{null} T' = (\operatorname{range} T)^0$, we have $(\operatorname{range} T)^0 = \{0\}$ so that $\operatorname{range} T = W$. Hence, T is surjective. Since $\dim V = \dim W$, we conclude that T is injective as well. Therefore, T is invertible. ∎

Exercise 3F18

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 defined by $\Phi(T) = T'$. We need to show that Φ is linear, injective, and surjective.

Linearity: For $T_1, T_2 \in \mathcal{L}(V, W)$ and $a \in \mathbf{F}$, we have

$$\Phi(T_1 + T_2) = (T_1 + T_2)' = T_1' + T_2' = \Phi(T_1) + \Phi(T_2),$$

and

$$\Phi(aT_1) = (aT_1)' = aT_1' = a\Phi(T_1).$$

Hence, Φ is linear.

To see that Φ is injective, suppose $\Phi(T) = T' = 0$. By Exercise 3F16, this implies $T = 0$. Hence, Φ is injective.

To see that Φ is surjective, note that since V and W are finite-dimensional, $\dim \mathcal{L}(V, W) = \dim V \cdot \dim W = \dim W' \cdot \dim V' = \dim \mathcal{L}(W', V')$. A linear map between finite-dimensional vector spaces of the same dimension that is injective is also surjective. Hence, Φ is surjective and therefore an isomorphism. ∎

Exercise 3F19

Suppose $U \subseteq V$. Explain why

$$U^0 = \{\varphi \in V' : U \subseteq \operatorname{null} \varphi\}.$$

Solution. By definition, the annihilator U^0 of U is the set of all linear functionals $\varphi \in V'$ such that $\varphi(u) = 0$ for all $u \in U$. Equivalently, this means that $U \subseteq \operatorname{null} \varphi$. ■

Exercise 3F20

Suppose V is finite-dimensional and U is a subspace of V. Show that

$$U = \{v \in V : \varphi(v) = 0 \text{ for every } \varphi \in U^0\}.$$

Manual Remark. *This says that* $U = (U^0)^0$.

Solution. Let $S = \{v \in V : \varphi(v) = 0 \text{ for every } \varphi \in U^0\}$. We need to show that $S = U$.

First, if $v \in U$, then for any $\varphi \in U^0$, we have $\varphi(v) = 0$ by definition of the annihilator. Hence, $v \in S$, so $U \subseteq S$.

Conversely, suppose $v \in S$. If $v \notin U$, then by the construction done in Exercise 3F4 with $v_1 = v$, there exists a linear functional $\varphi \in V'$ such that $\varphi(u) = 0$ for all $u \in U$ and $\varphi(v) \neq 0$. But this φ is in U^0, which contradicts $v \in S$. Hence, $v \in U$, and therefore $S \subseteq U$.

Combining both inclusions, we get $S = U$, as desired. ■

Exercise 3F21

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.

(a) Suppose $W^0 \subseteq U^0$. If $u \in U$, then for any $\varphi \in W^0$, we have $\varphi(u) = 0$ because $\varphi \in U^0$. Hence, by Exercise 3F20, we have $u \in W$, so $U \subseteq W$. Conversely, if $U \subseteq W$ and $\varphi \in W^0$, then $\varphi(w) = 0$ for all $w \in W$, and in particular for all $u \in U$. Hence, $\varphi \in U^0$, so $W^0 \subseteq U^0$.

(b) Suppose $W^0 = U^0$. Then $W^0 \subseteq U^0$ and $U^0 \subseteq W^0$. By part (a), this implies that $U \subseteq W$ and $W \subseteq U$. Hence, $U = W$. Conversely, if $U = W$, then clearly $U^0 = W^0$. ■

Exercise 3F22

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) Let $\varphi \in (U + W)^0$. Then for any $u \in U$ and $w \in W$, we have $\varphi(u + w) = 0$. In particular, taking $w = 0$ gives $\varphi(u) = 0$ for all $u \in U$, so $\varphi \in U^0$. Similarly, taking $u = 0$ gives $\varphi(w) = 0$ for all $w \in W$, so $\varphi \in W^0$. Hence, $\varphi \in U^0 \cap W^0$, and we have shown that $(U + W)^0 \subseteq U^0 \cap W^0$.

Conversely, let $\varphi \in U^0 \cap W^0$. Then for any $u \in U$ and $w \in W$, we have $\varphi(u) = 0$ and $\varphi(w) = 0$, so $\varphi(u + w) = \varphi(u) + \varphi(w) = 0$. Hence, $\varphi \in (U + W)^0$, and we have shown that $U^0 \cap W^0 \subseteq (U + W)^0$. Combining both inclusions, we get $(U + W)^0 = U^0 \cap W^0$.

(b) From part (a), we know that $(U+W)^0 = U^0 \cap W^0$. Then

$$\dim(U+W)^0 = \dim V - \dim(U+W).$$

Applying the dimension formula for a sum of subspaces, we have

$$\dim(U^0+W^0) = \dim U^0 + \dim W^0 - \dim(U^0 \cap W^0).$$

Since $(U+W)^0 = U^0 \cap W^0$, we have $\dim(U^0 \cap W^0) = \dim(U+W)^0 = \dim V - \dim(U+W)$. Therefore,

$$\dim(U^0+W^0) = \dim U^0 + \dim W^0 - (\dim V - \dim(U+W)).$$

Using the fact that $\dim U^0 = \dim V - \dim U$ and $\dim W^0 = \dim V - \dim W$, we get

$$\begin{aligned}
\dim(U^0+W^0) &= \dim U^0 + \dim W^0 - \dim(U^0 \cap W^0) \\
&= (\dim V - \dim U) + (\dim V - \dim W) - (\dim V - \dim(U+W)) \\
&= \dim V - \dim U + \dim V - \dim W - \dim V + \dim(U+W) \\
&= \dim V - (\dim U + \dim W - \dim(U+W)) \\
&= \dim V - \dim(U \cap W).
\end{aligned}$$

Hence, $\dim(U^0+W^0) = \dim(U \cap W)^0$. To see equality, we just need to prove one inclusion. If $\varphi = \varphi_U + \varphi_W \in U^0+W^0$, then for any $v \in U \cap W$, we have $\varphi(v) = \varphi_U(v) + \varphi_W(v) = 0+0 = 0$, so $\varphi \in (U \cap W)^0$. Therefore, $U^0+W^0 \subseteq (U \cap W)^0$. Since the dimensions are equal, we conclude that $(U \cap W)^0 = U^0 + W^0$. ∎

Exercise 3F23

Suppose V is finite-dimensional and $\varphi_1, \dots, \varphi_m \in V'$. Prove that the following three sets are equal to each other.

(a) $\operatorname{span}(\varphi_1, \dots, \varphi_m)$

(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. For ease of notation, let A, B, C be the sets in the exercise, i.e.,

$$\begin{aligned}
A &= \operatorname{span}(\varphi_1, \dots, \varphi_m), \\
B &= ((\operatorname{null}\varphi_1) \cap \cdots \cap (\operatorname{null}\varphi_m))^0, \\
C &= \{\varphi \in V' : (\operatorname{null}\varphi_1) \cap \cdots \cap (\operatorname{null}\varphi_m) \subseteq \operatorname{null}\varphi\}.
\end{aligned}$$

We will show that $A = B$ and $B = C$. Starting off with $B = C$, note that by definition,

$$\begin{aligned}
B &= ((\operatorname{null}\varphi_1) \cap \cdots \cap (\operatorname{null}\varphi_m))^0 \\
&= \{\varphi \in V' : (\operatorname{null}\varphi_1) \cap \cdots \cap (\operatorname{null}\varphi_m) \subseteq \operatorname{null}\varphi\} = C,
\end{aligned}$$

where the reformulation of the annihilator subspace was shown in Exercise 3F19.

To show that $A = B$, we may use Exercise 3F22 (b) repeatedly to rewrite B as

$$B = (\operatorname{null}\varphi_1)^0 + \cdots + (\operatorname{null}\varphi_m)^0.$$

By Exercise 3F6, we have $(\operatorname{null}\varphi_i)^0 = \operatorname{span}(\varphi_i)$ for each $i = 1, \dots, m$. Therefore,

$$B = (\operatorname{null}\varphi_1)^0 + \cdots + (\operatorname{null}\varphi_m)^0 = \operatorname{span}(\varphi_1) + \cdots + \operatorname{span}(\varphi_m) = \operatorname{span}(\varphi_1, \dots, \varphi_m) = A.$$

We conclude that $A = B$. ∎

Exercise 3F24

Suppose V is finite-dimensional and $v_1, \dots, v_m \in V$. Define a linear map $\Gamma : V' \to \mathbf{F}^m$ by $\Gamma(\varphi) = (\varphi(v_1), \dots, \varphi(v_m))$.

(a) Prove that $v_1, \dots, v_m$ spans V if and only if Γ is injective.

(b) Prove that $v_1, \dots, v_m$ is linearly independent if and only if Γ is surjective.

Solution.

(a) Suppose $v_1, \dots, v_m$ spans V. If $\Gamma(\varphi) = 0$, then $\varphi(v_i) = 0$ for all $i = 1, \dots, m$. Since $v_1, \dots, v_m$ spans V, it follows that $\varphi(v) = 0$ for all $v \in V$, i.e., $\varphi = 0$. Hence, Γ is injective.

Conversely, suppose Γ is injective. If $v_1, \dots, v_m$ does not span V, then by Exercise 3F4, there exists a nonzero $\varphi \in V'$ that vanishes on $v_1, \dots, v_m$. But then $\Gamma(\varphi) = 0$, contradicting the injectivity of Γ. Hence, $v_1, \dots, v_m$ must span V.

(b) Suppose $v_1, \dots, v_m$ is linearly independent. Extend this list to a basis $v_1, \dots, v_m, v_{m+1}, \dots, v_n$ of V. Define linear functionals $\varphi_1, \dots, \varphi_m \in V'$ by setting $\varphi_i(v_j) = \delta_{ij}$ for $i, j = 1, \dots, m$ and $\varphi_i(v_j) = 0$ for $j = m+1, \dots, n$. Then for any $(a_1, \dots, a_m) \in \mathbf{F}^m$, the linear functional $\varphi = a_1\varphi_1 + \cdots + a_m\varphi_m$ satisfies $\Gamma(\varphi) = (a_1, \dots, a_m)$. Hence, Γ is surjective.

Conversely, suppose Γ is surjective. If $v_1, \dots, v_m$ is not linearly independent, then there exist scalars $a_1, \dots, a_m$, not all zero, such that $a_1v_1 + \cdots + a_mv_m = 0$. Consider some index j such that $a_j \neq 0$. Since Γ is surjective, there exists $\varphi \in V'$ such that $\varphi(v_j) = 1$ and $\varphi(v_i) = 0$ for all $i \neq j$. Then

$$0 = \varphi(a_1v_1 + \cdots + a_mv_m) = a_1\varphi(v_1) + \cdots + a_m\varphi(v_m) = a_j \neq 0,$$

a contradiction. Hence, $v_1, \dots, v_m$ must be linearly independent. ∎

Exercise 3F25

Suppose V is finite-dimensional and $\varphi_1, \dots, \varphi_m \in V'$. Define a linear map $\Gamma : V \to \mathbf{F}^m$ by $\Gamma(v) = (\varphi_1(v), \dots, \varphi_m(v))$.

(a) Prove that $\varphi_1, \dots, \varphi_m$ spans V' if and only if Γ is injective.

(b) Prove that $\varphi_1, \dots, \varphi_m$ is linearly independent if and only if Γ is surjective.

Solution.

(a) Suppose $\varphi_1, \dots, \varphi_m$ spans V'. If $\Gamma(v) = 0$, then $\varphi_i(v) = 0$ for all $i = 1, \dots, m$. Since $\varphi_1, \dots, \varphi_m$ spans V', then by Exercise 3F3, it follows that $\psi(v) = 0$ for all $\psi \in V'$, i.e., $v = 0$. Hence, Γ is injective.

Conversely, suppose Γ is injective. Then $\operatorname{null}\Gamma = \{0\}$. By definition, we have

$$\operatorname{null}\Gamma = \operatorname{null}\varphi_1 \cap \cdots \cap \operatorname{null}\varphi_m.$$

By Exercise 3F23, we have

$$\operatorname{span}(\varphi_1, \dots, \varphi_m) = ((\operatorname{null}\varphi_1) \cap \cdots \cap (\operatorname{null}\varphi_m))^0 = (\operatorname{null}\Gamma)^0 = \{\psi \in V' : \operatorname{null}\Gamma \subseteq \operatorname{null}\psi\} = V'.$$

Hence, $\varphi_1, \dots, \varphi_m$ spans V'.

(b) Suppose $\varphi_1, \dots, \varphi_m$ is linearly independent. Then $\dim \operatorname{span}(\varphi_1, \dots, \varphi_m) = m$. By the fundamental theorem of

linear maps and Exercise 3F23, we have

$$\dim \operatorname{range} \Gamma = \dim V - \dim \operatorname{null} \Gamma = \dim V - \dim((\operatorname{null} \varphi_1) \cap \cdots \cap (\operatorname{null} \varphi_m)) = m.$$

Since $\operatorname{range} \Gamma$ is a subspace of $\mathbf{F}^m$ and has dimension m, we conclude that $\operatorname{range} \Gamma = \mathbf{F}^m$, so Γ is surjective.

Conversely, suppose Γ is surjective. Then $\operatorname{range} \Gamma = \mathbf{F}^m$, so $\dim \operatorname{range} \Gamma = m$. By the fundamental theorem of linear maps, we have

$$\dim \operatorname{range} \Gamma = \dim V - \dim \operatorname{null} \Gamma = \dim V - \dim((\operatorname{null} \varphi_1) \cap \cdots \cap (\operatorname{null} \varphi_m)).$$

By Exercise 3F23, we have

$$\operatorname{span}(\varphi_1, \ldots, \varphi_m) = ((\operatorname{null} \varphi_1) \cap \cdots \cap (\operatorname{null} \varphi_m))^0,$$

so $\dim \operatorname{span}(\varphi_1, \ldots, \varphi_m) = \dim V - \dim((\operatorname{null} \varphi_1) \cap \cdots \cap (\operatorname{null} \varphi_m)) = m$. Hence, $\varphi_1, \ldots, \varphi_m$ is linearly independent. ■

Exercise 3F26

Suppose V is finite-dimensional and Ω is a subspace of V'. Prove that

$$\Omega = \{v \in V : \varphi(v) = 0 \text{ for every } \varphi \in \Omega\}^0.$$

Manual Remark. *This says* $\Omega = (\Omega^0)^0$.

Solution. Let $S = \{v \in V : \varphi(v) = 0 \text{ for every } \varphi \in \Omega\}$. We need to show that $\Omega = S^0$. Let $\varphi_1, \ldots, \varphi_m$ be a basis of Ω. Then $S = (\operatorname{null} \varphi_1) \cap \cdots \cap (\operatorname{null} \varphi_m)$. By Exercise 3F23, we have

$$\Omega = \operatorname{span}(\varphi_1, \ldots, \varphi_m) = ((\operatorname{null} \varphi_1) \cap \cdots \cap (\operatorname{null} \varphi_m))^0 = S^0.$$ ■

Exercise 3F27

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. Let $W = \{p \in \mathcal{P}_5(\mathbf{R}) : p(8) = 0\}$. We need to show that $\operatorname{range} T = W$.

First, if $p \in \operatorname{range} T$, then there exists $q \in \mathcal{P}_5(\mathbf{R})$ such that $T(q) = p$. Since $\varphi \in \operatorname{null} T'$, we have

$$\varphi(T(q)) = (T'(\varphi))(q) = 0.$$

But $\varphi(T(q)) = T(q)(8) = p(8)$, so $p(8) = 0$. Hence, $p \in W$, and we have shown that $\operatorname{range} T \subseteq W$.

Now we compare the dimensions of $\operatorname{range} T$ and W. Recall that $(\operatorname{range} T)^0 = \operatorname{null} T' = \operatorname{span}(\varphi)$, so $\dim(\operatorname{range} T)^0 = 1$. Since $\mathcal{P}_5(\mathbf{R})$ has dimension 6, we have $\dim \operatorname{range} T = 6 - 1 = 5$. On the other hand, W is the set of all polynomials in $\mathcal{P}_5(\mathbf{R})$ that vanish at 8. Using the fundamental theorem of linear maps and the fact that $\varphi \neq 0$ is surjective by Exercise 3F1, we have $\dim W = \dim \mathcal{P}_5(\mathbf{R}) - 1 = 5$. Hence, $\operatorname{range} T$ and W are subspaces of $\mathcal{P}_5(\mathbf{R})$ with the same dimension, and $\operatorname{range} T \subseteq W$. We conclude that $\operatorname{range} T = W$, as desired. ■

Exercise 3F28

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. Since $\varphi_1, \dots, \varphi_m$ is linearly independent, we have $\dim \operatorname{span}(\varphi_1, \dots, \varphi_m) = m$. By Exercise 3F23, we have

$$\operatorname{span}(\varphi_1, \dots, \varphi_m) = ((\operatorname{null}\varphi_1) \cap \cdots \cap (\operatorname{null}\varphi_m))^0.$$

Hence, $\dim((\operatorname{null}\varphi_1) \cap \cdots \cap (\operatorname{null}\varphi_m))^0 = m$. Since V is finite-dimensional, we have

$$\dim((\operatorname{null}\varphi_1) \cap \cdots \cap (\operatorname{null}\varphi_m)) = \dim V - \dim((\operatorname{null}\varphi_1) \cap \cdots \cap (\operatorname{null}\varphi_m))^0 = \dim V - m,$$

as desired. ∎

Exercise 3F29

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) From $\operatorname{null} T' = \operatorname{span}(\varphi)$ and $\operatorname{null} T' = (\operatorname{range} T)^0$, we have by Exercise 3F20 that

$$\operatorname{range} T = ((\operatorname{range} T)^0)^0 = (\operatorname{null} T')^0 = (\operatorname{span}(\varphi))^0 = \operatorname{null}\varphi,$$

where the last equality follows as such:

$$v \in \operatorname{null}\varphi \iff \varphi(v) = 0 \iff \alpha\varphi(v) = 0 \text{ for all } \alpha \in \mathbf{F} \iff v \in (\operatorname{span}(\varphi))^0.$$

(b) Note that $\operatorname{range} T' = \operatorname{span}(\psi)$ and $\operatorname{range} T' = (\operatorname{null} T)^0$. Then $(\operatorname{null} T)^0 = \operatorname{span}(\psi)$. By Exercise 3F20, we have

$$\operatorname{null} T = ((\operatorname{null} T)^0)^0 = (\operatorname{span}(\psi))^0 = \operatorname{null}\psi,$$

where the last equality follows as such:

$$v \in \operatorname{null}\psi \iff \psi(v) = 0 \iff \alpha\psi(v) = 0 \text{ for all } \alpha \in \mathbf{F} \iff v \in (\operatorname{span}(\psi))^0.$$

∎

Exercise 3F30

Suppose V is finite-dimensional and $\varphi_1, \dots, \varphi_n$ is a basis of V'. Show that there exists a basis of V whose dual basis is $\varphi_1, \dots, \varphi_n$.

Solution. Let $\varphi_1, \dots, \varphi_n$ be a basis of V'. From Exercise 3F25, there exists a map $\Gamma : V \to \mathbf{F}^n$ that is an isomorphism, since $\varphi_1, \dots, \varphi_n$ is a basis for V'. Let $e_1, \dots, e_n$ be the standard basis of $\mathbf{F}^n$. Since Γ is an isomorphism, there exist $v_1, \dots, v_n \in V$ such that $\Gamma(v_i) = e_i$ for each $i = 1, \dots, n$. We claim that $v_1, \dots, v_n$ is a basis of V whose dual basis is $\varphi_1, \dots, \varphi_n$.

To see that $v_1, \ldots, v_n$ is a basis of V, we show linear independence. Suppose $a_1, \ldots, a_n$ are scalars such that $a_1 v_1 + \cdots + a_n v_n = 0$. Applying Γ to both sides gives

$$a_1 e_1 + \cdots + a_n e_n = \Gamma(a_1 v_1 + \cdots + a_n v_n) = \Gamma(0) = 0.$$

Since $e_1, \ldots, e_n$ is a basis of $\mathbf{F}^n$, it follows that $a_1 = \cdots = a_n = 0$. Hence, $v_1, \ldots, v_n$ is linearly independent. Since $\dim V = n$, we conclude that $v_1, \ldots, v_n$ is a basis of V. To see that the dual basis of $v_1, \ldots, v_n$ is $\varphi_1, \ldots, \varphi_n$, we need to show that $\varphi_i(v_j) = \delta_{ij}$ for each $i, j = 1, \ldots, n$. Since $\Gamma(v_j) = e_j$, we have

$$\varphi_i(v_j) = (\Gamma(v_j))_i = (e_j)_i = \delta_{ij},$$

as desired. ∎

Exercise 3F31

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 i' is an isomorphism from V'/U^0 onto U'.

Solution.

(a) Note that if $\varphi \in \operatorname{null} i'$, then $i'(\varphi) = 0$, so $\varphi(i(u)) = 0$ for all $u \in U$. Since $i(u) = u$, we have $\varphi(u) = 0$ for all $u \in U$, so $\varphi \in U^0$. Hence, $\operatorname{null} i' \subseteq U^0$. Conversely, if $\varphi \in U^0$, then $\varphi(u) = 0$ for all $u \in U$, so $i'(\varphi) = 0$. Hence, $U^0 \subseteq \operatorname{null} i'$, and we conclude that $\operatorname{null} i' = U^0$.

(b) To see $\operatorname{range} i' \subseteq U'$, note that if $\varphi \in V'$, then $i'(\varphi) \in U'$ since $i'(\varphi)(u) = \varphi(i(u)) = \varphi(u)$ for all $u \in U$. Hence, $\operatorname{range} i' \subseteq U'$.

Since V is finite-dimensional, we have $\dim V' = \dim V$. By the fundamental theorem of linear maps, we have

$$\dim \operatorname{range} i' = \dim V' - \dim \operatorname{null} i' = \dim V - \dim U^0.$$

By the annihilator dimension formula, we have $\dim U^0 = \dim V - \dim U$. Hence,

$$\dim \operatorname{range} i' = \dim V - (\dim V - \dim U) = \dim U = \dim U'.$$

Since $\operatorname{range} i'$ is a subspace of U' and has the same dimension, we conclude that $\operatorname{range} i' = U'$, as desired.

(c) From part (a), we have $\operatorname{null} i' = U^0$. Then we know that V'/U^0 is isomorphic to $\operatorname{range} i'$. Then the induced map $\tilde{i'} : V'/U^0 \to \operatorname{range} i'$ defined by $\tilde{i'}(\varphi + U^0) = i'(\varphi)$ is an isomorphism. Since $\operatorname{range} i' = U'$ by part (b), we conclude that i' induces an isomorphism from V'/U^0 onto U', as desired. ∎

Exercise 3F32

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''.

Axler Note. *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) To see that Λ is linear, let $v, w \in V$ and $a \in \mathbf{F}$. Then for any $\varphi \in V'$, we have

$$\begin{aligned}(\Lambda(v+w))(\varphi) &= \varphi(v+w) = \varphi(v) + \varphi(w) = (\Lambda v)(\varphi) + (\Lambda w)(\varphi),\\ (\Lambda(av))(\varphi) &= \varphi(av) = a\varphi(v) = a(\Lambda v)(\varphi).\end{aligned}$$

Hence, Λ is linear.

(b) Let $v \in V$ and $\varphi \in V'$. Then

$$\begin{aligned}((T'' \circ \Lambda)(v))(\varphi) &= (T''(\Lambda v))(\varphi)\\ &= (\Lambda v)(T'(\varphi))\\ &= T'(\varphi)(v)\\ &= \varphi(T(v))\\ &= (\Lambda(T(v)))(\varphi) = ((\Lambda \circ T)(v))(\varphi).\end{aligned}$$

Hence, $T'' \circ \Lambda = \Lambda \circ T$.

(c) Since V is finite-dimensional, we have $\dim V = \dim V'$. Then $\dim V'' = \dim(V')' = \dim V' = \dim V$. By the fundamental theorem of linear maps, we have

$$\dim \operatorname{range} \Lambda = \dim V - \dim \operatorname{null} \Lambda.$$

To see that Λ is injective, suppose $\Lambda v = 0$ for some $v \in V$. Then for any $\varphi \in V'$, we have $\varphi(v) = (\Lambda v)(\varphi) = 0$. Since this holds for all $\varphi \in V'$ by Exercise 3F3, it follows that $v = 0$. Hence, Λ is injective, so $\operatorname{null} \Lambda = \{0\}$. Therefore, $\dim \operatorname{range} \Lambda = \dim V - 0 = \dim V$. Since $\operatorname{range} \Lambda$ is a subspace of V'' and has the same dimension as V'', we conclude that $\operatorname{range} \Lambda = V''$, so Λ is surjective. Combining injectivity and surjectivity, we conclude that Λ is an isomorphism from V onto V'', as desired. ∎

Exercise 3F33

Suppose U is a subspace of V. Let $\pi : V \to V/U$ be the usual quotient map. Thus, $\pi' \in \mathcal{L}((V/U)', V')$.

(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.

Solution.

(a) Suppose $\pi'(\varphi) = 0$ for some $\varphi \in (V/U)'$. Then $\varphi(\pi(v)) = 0$ for all $v \in V$. Since $\pi(v) = v + U$, we have $\varphi(v + U) = 0$ for all $v \in V$. Since this holds for all $v \in V$ by surjectivity of π, it follows that $\varphi = 0$. Hence, π' is injective.

(b) To see range $\pi' \subseteq U^0$, note that if $\varphi \in (V/U)'$, then for any $u \in U$, we have $\pi'(\varphi)(u) = \varphi(\pi(u)) = \varphi(0 + U) = 0$. Hence, range $\pi' \subseteq U^0$.

To see the other inclusion, let $\psi \in U^0$. We want to find $\varphi \in (V/U)'$ such that $\pi'(\varphi) = \psi$. Define $\varphi : V/U \to \mathbf{F}$ by $\varphi(v + U) = \psi(v)$ for each $v \in V$. We need to check that φ is well-defined. If $v + U = w + U$, then $v - w \in U$, so $\psi(v - w) = 0$, which implies that $\psi(v) = \psi(w)$. Hence, φ is well-defined. It is also linear since for any $a, b \in \mathbf{F}$ and $v, w \in V$, we have

$$\begin{aligned}\varphi(a(v + U) + b(w + U)) &= \varphi((av + bw) + U) = \psi(av + bw) \\ &= a\psi(v) + b\psi(w) = a\varphi(v + U) + b\varphi(w + U).\end{aligned}$$

Finally, we have $\pi'(\varphi)(v) = \varphi(\pi(v)) = \varphi(v + U) = \psi(v)$ for all $v \in V$, so $\pi'(\varphi) = \psi$. Hence, $U^0 \subseteq$ range π', and we conclude that range $\pi' = U^0$.

(c) From part (a), we have π' is injective. From part (b), we have range $\pi' = U^0$. Then the induced map $\tilde{\pi}' : (V/U)' \to U^0$ defined by $\tilde{\pi}'(\varphi) = \pi'(\varphi)$ is an isomorphism. Hence, π' is an isomorphism from $(V/U)'$ onto U^0, as desired. ■

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) $|\operatorname{Re} z| \leq |z|$ and $|\operatorname{Im} z| \leq |z|$

(g) $|\overline{z}| = |z|$

(h) $|wz| = |w||z|$

Solution. In this problem, let $z = a + bi$ and $w = c + di$ with $a, b, c, d \in \mathbf{R}$. Then

(a) $z + \overline{z} = (a + bi) + (a - bi) = 2a = 2 \operatorname{Re} z$.

(b) $z - \overline{z} = (a + bi) - (a - bi) = 2bi = 2(\operatorname{Im} z)i$.

(c) $z\overline{z} = (a + bi)(a - bi) = a^2 + b^2 = |z|^2$.

(d) $\overline{w+z} = \overline{(c+di)+(a+bi)} = \overline{(a+c)+(b+d)i} = (a+c) - (b+d)i = (a - bi) + (c - di) = \overline{w} + \overline{z}$ and $\overline{wz} = \overline{(c+di)(a+bi)} = \overline{(ac-bd)+(ad+bc)i} = (ac-bd) - (ad+bc)i = (c - di)(a - bi) = \overline{w}\overline{z}$.

(e) $\overline{\overline{z}} = \overline{a - bi} = a + bi = z$.

(f) $|\operatorname{Re} z| = |a| \leq \sqrt{a^2 + b^2} = |z|$ and $|\operatorname{Im} z| = |b| \leq \sqrt{a^2 + b^2} = |z|$.

(g) $|\overline{z}| = |a - bi| = \sqrt{a^2 + b^2} = |z|$.

(h) $|wz| = |(c + di)(a + bi)| = |(ac - bd) + (ad + bc)i| = \sqrt{(ac - bd)^2 + (ad + bc)^2} = \sqrt{(a^2 + b^2)(c^2 + d^2)} = |a + bi||c + di| = |z||w|$. ■

Exercise 4-2

Prove that if $w, z \in \mathbf{C}$, then $||w| - |z|| \leq |w - z|$.

Axler Note. *The inequality above is called the* **reverse triangle inequality**.

Solution. Let $w, z \in \mathbf{C}$. By the triangle inequality, we have

$$|w| = |(w - z) + z| \leq |w - z| + |z|.$$

Rearranging gives $|w| - |z| \leq |w - z|$. Similarly,

$$|z| = |(z - w) + w| \leq |z - w| + |w| = |w - z| + |w|,$$

which gives $|z| - |w| \leq |w - z|$, or equivalently $-|w - z| \leq |w| - |z|$. Combining these inequalities, we get

$$||w| - |z|| \leq |w - z|.$$ ■

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. Let $v \in V$. Then there exist $s, t \in \mathbf{R}$ such that

$$\varphi(v) = s + ti.$$

Then

$$\sigma(v) = \operatorname{Re}\varphi(v) = s \quad \text{and} \quad \sigma(iv) = \operatorname{Re}\varphi(iv) = \operatorname{Re}(i(s+ti)) = \operatorname{Re}(-t+si) = -t.$$

Therefore,

$$\sigma(v) - i\sigma(iv) = s - i(-t) = s + ti = \varphi(v). \qquad \blacksquare$$

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. No, the set is not a subspace of $\mathcal{P}(\mathbf{F})$. To see this, consider two polynomials $p, q \in \mathcal{P}(\mathbf{F})$ with $p(x) = x^m$ and $q(x) = -x^m + x^{m-1}$. Then $p + q = x^{m-1}$, which has degree $m-1$ and is not in the set. Therefore, the set is not closed under addition and hence not a subspace. $\blacksquare$

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. No, the set is not a subspace of $\mathcal{P}(\mathbf{F})$. To see this, consider two polynomials $p, q \in \mathcal{P}(\mathbf{F})$ with $p(x) = x^2$ and $q(x) = x - x^2$. Then $p + q = x$, which has degree 1 and is not even. Therefore, the set is not closed under addition and hence not a subspace. $\blacksquare$

Exercise 4-6

Suppose that m and n are positive integers with $m \le n$, and suppose $\lambda_1, \dots, \lambda_m \in \mathbf{F}$. 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. Consider the polynomial

$$p(x) = (x - \lambda_1)^{n-m+1}(x - \lambda_2)\cdots(x - \lambda_m).$$

This polynomial has degree n, vanishes at $\lambda_1, \dots, \lambda_m$, and has no other zeros. Therefore, it satisfies the required conditions. $\blacksquare$

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 $k = 1, \ldots, m+1$.

Axler Note. *This result can be proved without using linear algebra. However, try to find the clearer, shorter proof that uses some linear algebra.*

Solution. Consider the linear map $T : \mathcal{P}_m(\mathbf{F}) \to \mathbf{F}^{m+1}$ defined by

$$T(p) = (p(z_1), \ldots, p(z_{m+1})).$$

Since $\dim \mathcal{P}_m(\mathbf{F}) = m+1$ and $\dim \mathbf{F}^{m+1} = m+1$, it suffices to show that T is injective to conclude that it is also surjective. Suppose $T(p) = 0$, i.e., $p(z_k) = 0$ for all $k = 1, \ldots, m+1$. Since p has degree at most m and has $m+1$ distinct zeros, it must be the zero polynomial. Therefore, T is injective, and hence bijective. This shows that for any $(w_1, \ldots, w_{m+1}) \in \mathbf{F}^{m+1}$, there exists a unique $p \in \mathcal{P}_m(\mathbf{F})$ such that $T(p) = (w_1, \ldots, w_{m+1})$, which is exactly the desired polynomial. ■

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. Suppose p has m distinct zeros, say $z_1, \ldots, z_m$. If p and p' have a common zero z_k, then z_k would be a multiple zero of p, contradicting the assumption that all zeros are distinct. Hence, p and p' have no zeros in common.

Conversely, suppose p and p' have no zeros in common. If p had a multiple zero z, then $p(z) = 0$ and $p'(z) = 0$, which would contradict the assumption. Since p has degree m, it can have at most m zeros by the fundamental theorem of algebra. Therefore, p must have exactly m distinct zeros. ■

Exercise 4-9 ✔

Prove that every polynomial of odd degree with real coefficients has a real zero.

Solution. Let $p(x)$ be a polynomial of odd degree with real coefficients. Since non-real zeros come in pairs, there are an even number of non-real zeros. Therefore, there must be at least one real zero. ■

Exercise 4-10

For $p \in \mathcal{P}(\mathbf{R})$, define $Tp : \mathbf{R} \to \mathbf{R}$ by

$$(Tp)(x) = \begin{cases} \frac{p(x)-p(3)}{x-3} & \text{if } x \neq 3, \\ p'(3) & \text{if } x = 3 \end{cases}$$

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 $p \in \mathcal{P}(\mathbf{R})$. Since $p(x) - p(3)$ has a zero at $x = 3$, we can factor it as $(x-3)q(x)$ for some polynomial $q \in \mathcal{P}(\mathbf{R})$. Then for $x \neq 3$,

$$(Tp)(x) = \frac{p(x) - p(3)}{x - 3} = q(x),$$

and differentiating p gives $p'(3) = q(3)$. Since $Tp(3) = p'(3) = q(3)$, we have $Tp(x) = q(x)$ for all $x \in \mathbf{R}$. Therefore, $Tp \in \mathcal{P}(\mathbf{R})$.

Let $p_1, p_2 \in \mathcal{P}(\mathbf{R})$ and $\alpha, \beta \in \mathbf{R}$. Let $p_1(x) - p_1(3) = (x-3)q_1(x)$ and $p_2(x) - p_2(3) = (x-3)q_2(x)$ for some $q_1, q_2 \in \mathcal{P}(\mathbf{R})$. Then

$$(\alpha p_1 + \beta p_2)(x) = (\alpha p_1(3) + \beta p_2(3)) + (x-3)(\alpha q_1(x) + \beta q_2(x)).$$

Hence, $T(\alpha p_1 + \beta p_2)(x) = \alpha q_1(x) + \beta q_2(x) = \alpha T p_1(x) + \beta T p_2(x)$ for all $x \in \mathbf{R}$. Therefore, T is a linear map. ■

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. Let $p \in \mathcal{P}(\mathbf{C})$ have the form $p(z) = a_0 + a_1 z + \cdots + a_n z^n$ with $a_k \in \mathbf{C}$. Then

$$\overline{p(\overline{z})} = \overline{\sum_{i=0}^{n} a_i \overline{z}^i} = \sum_{i=0}^{n} \overline{a}_i z^i.$$

Therefore,

$$q(z) = p(z)\overline{p(\overline{z})} = \left(\sum_{i=0}^{n} a_i z^i\right)\left(\sum_{j=0}^{n} \overline{a}_j z^j\right) = \sum_{k=0}^{2n}\left(\sum_{i+j=k} a_i \overline{a}_j\right) z^k.$$

Let b_m denote the coefficient of z^m in $q(z)$. Then

$$b_m = \sum_{i+j=m} a_i \overline{a}_j = \sum_{i+j=m} \overline{a}_j a_i = \overline{b_m}.$$

Hence, $b_m \in \mathbf{R}$ for each $m = 0, 1, \ldots, 2n$. Therefore, q is a polynomial with 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 $x_0, x_1, \ldots, x_m$ with $p(x_k) \in \mathbf{R}$ for each $k = 0, 1, \ldots, m$. Prove that all coefficients of p are real.

Solution. Let $p(x_k) = y_k \in \mathbf{R}$ for each $k = 0, 1, \ldots, m$. By Exercise 4-7 with $\mathbf{F} = \mathbf{R}$, there exists a polynomial $q \in \mathcal{P}_m(\mathbf{R})$ that satisfies $q(x_k) = y_k$ for each $k = 0, 1, \ldots, m$. Note that this polynomial must be the same as the original $p \in \mathcal{P}_m(\mathbf{C})$ because a polynomial of degree at most m is uniquely determined by its values at $m+1$ distinct points. Therefore, all coefficients of p are real. ■

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) We first show that U is a subspace of $\mathcal{P}(\mathbf{F})$. Clearly, $0 \in U$ since $0 = p \cdot 0$. If $u_1, u_2 \in U$, then $u_1 = pq_1$ and $u_2 = pq_2$ for some $q_1, q_2 \in \mathcal{P}(\mathbf{F})$. For any scalars $\alpha, \beta \in \mathbf{F}$, we have $\alpha u_1 + \beta u_2 = \alpha p q_1 + \beta p q_2 = p(\alpha q_1 + \beta q_2) \in U$. Therefore, U is closed under addition and scalar multiplication, and hence U is a subspace of $\mathcal{P}(\mathbf{F})$.

Now let $r \in \mathcal{P}(\mathbf{F})$. By the division algorithm for polynomials, there exist unique polynomials $q_0 \in \mathcal{P}(\mathbf{F})$ and $s \in \mathcal{P}(\mathbf{F})$ with $\deg s < \deg p$ such that $r = pq_0 + s$. This shows that every polynomial in $\mathcal{P}(\mathbf{F})$ can be written as an element of U plus a polynomial of degree less than $\deg p$. Consider the map $T : \mathcal{P}_{\deg p-1}(\mathbf{F}) \to \mathcal{P}(\mathbf{F})/U$ defined by $T(s) = s + U$. This map is linear and surjective, and $\operatorname{null} T$ is trivial since if $s + U = U$, then $s \in U$, which implies $s = pq$ for some $q \in \mathcal{P}(\mathbf{F})$. Then

$$\deg s = \deg p + \deg q < \deg p.$$

Therefore, $q = 0$, hence $s = 0$, and T is an isomorphism. We conclude that $\dim \mathcal{P}(\mathbf{F})/U = \dim \mathcal{P}_{\deg p-1}(\mathbf{F}) = \deg p$.

(b) A basis of $\mathcal{P}(\mathbf{F})/U$ is given by the images of the standard basis of $\mathcal{P}_{\deg p-1}(\mathbf{F})$ under T, namely $\{1 + U, x + U, \ldots, x^{\deg p-1} + U\}$. ∎

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 $rp + sq = 1$.

Manual Remark. *In part (a), one may also observe that a multiple of p and q must have* $\deg p + \deg q$, *but* $\deg rp = \deg r + \deg p \leq n - 1 + m$.

Solution.

(a) To show that T is injective, suppose that $T(r, s) = rp + sq = 0$ for some $(r, s) \in \mathcal{P}_{n-1}(\mathbf{C}) \times \mathcal{P}_{m-1}(\mathbf{C})$. Then $rp = -sq$ implies rp is a polynomial multiple of q, so every zero of q must be a zero of rp. Because p and q have no zeros in common, every zero of q must be a zero of r. But $\deg r < \deg q$, so it must be that $r = 0$. Because $r = 0$, we have $-sq = 0$. Then since q is nonconstant, it must be that $s = 0$. Therefore, if $(r, s) \in \operatorname{null} T$, we have $(r, s) = 0$, so T is injective.

(b) To show that T is surjective, note that $\dim(\mathcal{P}_{n-1}(\mathbf{C}) \times \mathcal{P}_{m-1}(\mathbf{C})) = n + m = \dim \mathcal{P}_{m+n-1}(\mathbf{C})$. Since T is a linear map between vector spaces of the same finite dimension, and we have already shown that T is injective, it follows that T is also surjective.

(c) Since T is surjective, for the polynomial $1 \in \mathcal{P}_{m+n-1}(\mathbf{C})$, there exist $r \in \mathcal{P}_{n-1}(\mathbf{C})$ and $s \in \mathcal{P}_{m-1}(\mathbf{C})$ such that $T(r, s) = rp + sq = 1$. ∎

5 Eigenvalues and Eigenvectors

5A Invariant Subspaces

Exercise 5A1

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) If $U \subseteq \operatorname{null} T$, then for any $u \in U$, we have $T(u) = 0 \in U$, so U is invariant under T.

(b) If $\operatorname{range} T \subseteq U$, then for any $u \in U$, $T(u) \in \operatorname{range} T \subseteq U$, so U is invariant under T. ■

Exercise 5A2

Suppose that $T \in \mathcal{L}(V)$ and $V_1, \dots, V_m$ are subspaces of V invariant under T. Prove that $V_1 + \cdots + V_m$ is invariant under T.

Solution. Let $v \in V_1 + \cdots + V_m$. Then $v = v_1 + \cdots + v_m$ for some $v_i \in V_i$. Since each V_i is invariant under T, we have $T(v_i) \in V_i$ for each i. Therefore,

$$T(v) = T(v_1 + \cdots + v_m) = T(v_1) + \cdots + T(v_m) \in V_1 + \cdots + V_m,$$

so $V_1 + \cdots + V_m$ is invariant under T. ■

Exercise 5A3

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 $\{U_i\}_{i \in I}$ be a collection of subspaces of V invariant under T, and let

$$U = \bigcap_{i \in I} U_i.$$

Let $u \in U$. Then $u \in U_i$ for all $i \in I$. Since each U_i is invariant under T, we have $T(u) \in U_i$ for all $i \in I$. Therefore, $T(u) \in U$, so U is invariant under T. ■

Exercise 5A4

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. If $U = \{0\}$, then we are done. Now suppose $U \neq \{0\}$ so that there exists nonzero $v_1 \in U$. Extend v_1 to a basis $v_1, \dots, v_m$ of V, and let $v \in V$. Define the linear operator $T \in \mathcal{L}(V)$ by $T(v_1) = v$ and $T(v_i) = 0$ for $i = 2, \dots, m$. Since U is invariant under T, we have $T(v_1) = v \in U$. But v was an arbitrary element of V, so $U = V$. Therefore, $U = \{0\}$ or $U = V$. ■

Exercise 5A5

Suppose $T \in \mathcal{L}(\mathbf{R}^2)$ is defined by $T(x, y) = (-3y, x)$. Find the eigenvalues of T.

Solution. If (x, y) is an eigenvector of T with eigenvalue λ, then we have the system of equations

$$\begin{cases} -3y = \lambda x, \\ x = \lambda y \end{cases}$$

Solving the system, we get $\lambda^2 = -3$ which has no real solutions. Therefore, T has no real eigenvalues. ■

Exercise 5A6

Define $T \in \mathcal{L}(\mathbf{F}^2)$ by $T(w, z) = (z, w)$. Find all eigenvalues and eigenvectors of T.

Solution. If (w, z) is an eigenvector of T with eigenvalue λ, then we have

$$(z, w) = \lambda(w, z),$$

which gives the system

$$\begin{cases} z = \lambda w, \\ w = \lambda z. \end{cases}$$

Solving, we get $\lambda^2 = 1$, so $\lambda = 1$ or $\lambda = -1$. If $\lambda = 1$, then $z = w$, so the eigenvectors are of the form (w, w). If $\lambda = -1$, then $z = -w$, so the eigenvectors are of the form $(w, -w)$. ■

Exercise 5A7

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. If (z_1, z_2, z_3) is an eigenvector of T with eigenvalue λ, then we have

$$(2z_2, 0, 5z_3) = \lambda(z_1, z_2, z_3),$$

which gives the system

$$\begin{cases} 2z_2 = \lambda z_1, \\ 0 = \lambda z_2, \\ 5z_3 = \lambda z_3. \end{cases}$$

Solving, we get the eigenvalues $\lambda = 0$ and $\lambda = 5$. For $\lambda = 0$, the eigenvectors are of the form $(z_1, 0, 0)$. For $\lambda = 5$, the eigenvectors are of the form $(0, 0, z_3)$. ■

Exercise 5A8

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. If v is an eigenvector of P with eigenvalue λ, then $Pv = \lambda v$. Applying P again, we get $P^2v = P(\lambda v) = \lambda Pv = \lambda^2 v$. But $P^2 = P$, so $P^2v = Pv = \lambda v$. Therefore, $\lambda^2 v = \lambda v$, which gives $(\lambda^2 - \lambda)v = 0$. Since $v \neq 0$, we must have $\lambda^2 - \lambda = 0$, so $\lambda(\lambda - 1) = 0$. Hence, $\lambda = 0$ or $\lambda = 1$. ■

Exercise 5A9

Define $T : \mathcal{P}(\mathbf{R}) \to \mathcal{P}(\mathbf{R})$ by $Tp = p'$. Find all eigenvalues and eigenvectors of T.

Solution. Let $p \in \mathcal{P}(\mathbf{R})$ such that $p' = \lambda p$ for some $\lambda \in \mathbf{R}$.

- If $\lambda \neq 0$, suppose $\deg p = 0$. Then p is a constant polynomial, and $p' = 0 \neq \lambda p$, so $\deg p \geq 1$. If $\deg p = n \geq 1$, then $\deg p' = n - 1$, so $\deg p' \neq \deg p$, which contradicts $p' = \lambda p$.
- If $\lambda = 0$, then $p' = 0$, which implies that p is a constant polynomial. Therefore, the eigenvectors corresponding to the eigenvalue $\lambda = 0$ are all nonzero constant polynomials. ■

Exercise 5A10

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. Let $p(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4$ be an eigenvector of T with eigenvalue λ. Then

$$(Tp)(x) = xp'(x) = x(a_1 + 2a_2x + 3a_3x^2 + 4a_4x^3) = a_1x + 2a_2x^2 + 3a_3x^3 + 4a_4x^4.$$

Equating coefficients with $\lambda p(x) = \lambda a_0 + \lambda a_1x + \lambda a_2x^2 + \lambda a_3x^3 + \lambda a_4x^4$ yields the system:

$$\lambda a_0 = 0, \quad \lambda a_1 = a_1, \quad \lambda a_2 = 2a_2, \quad \lambda a_3 = 3a_3, \quad \lambda a_4 = 4a_4.$$

For each $n \in \{0, 1, 2, 3, 4\}$, the equation $\lambda a_n = na_n$ is satisfied with $a_n \neq 0$ if and only if $\lambda = n$. When $\lambda = n$, all other coefficients must be zero, so the eigenvector is a nonzero scalar multiple of x^n. Therefore, the eigenvalues of T are $0, 1, 2, 3, 4$, with corresponding eigenvectors being nonzero scalar multiples of $1, x, x^2, x^3, x^4$, respectively. ■

Exercise 5A11

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. Since V is finite-dimensional, the set of eigenvalues of T is finite. If T has no eigenvalues, then $T - \lambda I$ is invertible for all $\lambda \in \mathbf{F}$, and we can choose any $\delta > 0$. Otherwise, let $\lambda_1, \ldots, \lambda_m$ be the eigenvalues of T. Define

$$\delta = \min_{\lambda_j \neq \alpha} |\alpha - \lambda_j|.$$

If α is the only eigenvalue, then we can select any $\delta > 0$ since any λ satisfying $0 < |\alpha - \lambda| < \delta$ will not be equal to α, and hence $T - \lambda I$ will be invertible. Otherwise, we have for any $\lambda \in \mathbf{F}$ such that $0 < |\alpha - \lambda| < \delta$, we have $\lambda \neq \lambda_j$ for all $j = 1, \ldots, m$, so λ is not an eigenvalue of T. Hence, $T - \lambda I$ is invertible for all such λ. ■

Exercise 5A12

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. Let $v = u + w \in V$ be an eigenvector of P with eigenvalue λ. Then $P(v) = P(u + w) = u = \lambda v = \lambda(u + w)$. Comparing the U and W components, we get $u = \lambda u$ and $0 = \lambda w$. If $\lambda \neq 0$, then $w = 0$ and $u = \lambda u$ implies $\lambda = 1$. If $\lambda = 0$, then $u = 0$ and w can be any nonzero vector in W. Therefore, the eigenvalues are $\lambda = 1$ and $\lambda = 0$, with corresponding eigenvectors being the nonzero vectors in U for $\lambda = 1$ and the nonzero vectors in W for $\lambda = 0$. ■

Exercise 5A13

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) Let λ be an eigenvalue of T with eigenvector $v \in V$, so $T(v) = \lambda v$. Then

$$(S^{-1}TS)(S^{-1}v) = S^{-1}T(v) = S^{-1}(\lambda v) = \lambda(S^{-1}v).$$

Thus, $S^{-1}v$ is an eigenvector of $S^{-1}TS$ with eigenvalue λ. This shows that every eigenvalue of T is an eigenvalue of $S^{-1}TS$. Conversely, if λ is an eigenvalue of $S^{-1}TS$ with eigenvector u, then Su is an eigenvector of T with eigenvalue λ. Therefore, T and $S^{-1}TS$ have the same eigenvalues.

(b) The relation between the eigenvectors of T and the eigenvectors of $S^{-1}TS$ is given by $v \mapsto S^{-1}v$. Specifically, if v is an eigenvector of T with eigenvalue λ, then $S^{-1}v$ is an eigenvector of $S^{-1}TS$ with the same eigenvalue λ. Conversely, if u is an eigenvector of $S^{-1}TS$ with eigenvalue λ, then Su is an eigenvector of T with eigenvalue λ. ■

Exercise 5A14

Give an example of an operator on $\mathbf{R}^4$ that has no (real) eigenvalues.

Solution. Consider the linear operator on $\mathbf{R}^4$ given by the matrix

$$T = \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Letting T act on $(x_1, x_2, x_3, x_4) \in \mathbf{R}^4$, we have

$$T(x_1, x_2, x_3, x_4) = (-x_2, x_1, -x_4, x_3).$$

Then any eigenvector $(x_1, x_2, x_3, x_4) \in \mathbf{R}^4$ with eigenvalue λ must satisfy

$$(-x_2, x_1, -x_4, x_3) = \lambda(x_1, x_2, x_3, x_4).$$

This yields $\lambda^2 + 1 = 0$, which has no solution in $\mathbf{R}$. ■

Exercise 5A15

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.

$$\begin{aligned} \lambda \text{ is an eigenvalue of } T &\iff T - \lambda I \text{ is not invertible} \\ &\iff (T - \lambda I)' \text{ is not invertible} \\ &\iff T' - \lambda I \text{ is not invertible} \\ &\iff \lambda \text{ is an eigenvalue of } T' \end{aligned}$$

■

Exercise 5A16

Suppose $v_1, \dots, v_n$ is a basis of V and $T \in \mathcal{L}(V)$. Prove that if λ is an eigenvalue of T, then

$$|\lambda| \le n \max\{|\mathcal{M}(T)_{j,k}| : 1 \le j, k \le n\},$$

where $\mathcal{M}(T)_{j,k}$ denotes the entry in row j, column k of the matrix of T with respect to the basis $v_1, \dots, v_n$.
Axler Note. *See Exercise 6A19 for a different bound on* $|\lambda|$.

Solution. Let λ be an eigenvalue of T with eigenvector $u \in V$, so $Tu = \lambda u$. Let $\mathcal{M}(T)$ be the matrix of T with respect to the basis $v_1, \dots, v_n$. Then

$$\lambda u_j = (\mathcal{M}(T)u)_j = \sum_{k=1}^{n} \mathcal{M}(T)_{j,k} u_k \quad \text{for each } j = 1, \dots, n.$$

Taking absolute values, we get

$$|\lambda||u_j| = \left| \sum_{k=1}^{n} \mathcal{M}(T)_{j,k} u_k \right| \le \sum_{k=1}^{n} |\mathcal{M}(T)_{j,k}||u_k| \le n \max\{|\mathcal{M}(T)_{j,k}| : 1 \le j, k \le n\} \max\{|u_k| : 1 \le k \le n\}.$$

Choosing j such that $|u_j| = \max\{|u_k| : 1 \le k \le n\}$, we obtain

$$|\lambda| \le n \max\{|\mathcal{M}(T)_{j,k}| : 1 \le j, k \le n\}. \qquad \blacksquare$$

Exercise 5A17

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 $T_{\mathbf{C}}$.
Axler Note. *See Exercise 3B33 for the definition of* $T_{\mathbf{C}}$.

Solution. Suppose λ is an eigenvalue of T with eigenvector $v \in V$, so $Tv = \lambda v$. Then in the complexification, we have $T_{\mathbf{C}}(v + i0) = Tv + iT0 = \lambda v + i0 = \lambda(v + i0)$, showing that λ is an eigenvalue of $T_{\mathbf{C}}$ with eigenvector $v + i0$.

Conversely, suppose λ is an eigenvalue of $T_{\mathbf{C}}$ with eigenvector $v + iw \in V_{\mathbf{C}}$, where $v, w \in V$. Then

$$T_{\mathbf{C}}(v + iw) = Tv + iTw = \lambda(v + iw) = \lambda v + i\lambda w.$$

Equating real and imaginary parts, we get

$$Tv = \lambda v \quad \text{and} \quad Tw = \lambda w.$$

Since $\lambda \in \mathbf{R}$, if $v \neq 0$, then v is an eigenvector of T corresponding to λ. If $v = 0$, then $w \neq 0$ and $Tw = \lambda w$, so w is an eigenvector of T corresponding to λ. This shows that λ is an eigenvalue of T. $\blacksquare$

Exercise 5A18

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 λ is an eigenvalue of $T_{\mathbf{C}}$ with eigenvector $v + iw \in V_{\mathbf{C}}$, where $v, w \in V$. Then

$$T_{\mathbf{C}}(v + iw) = Tv + iTw = \lambda(v + iw) = \lambda v + i\lambda w.$$

Moreover, we have

$$T_{\mathbf{C}}(v - iw) = Tv - iTw.$$

Note that $v + iw = 0$ if and only if both $v = w = 0$. Then $v + iw \neq 0$ implies $v - iw \neq 0$. Let $\lambda = a + bi$ with $a, b \in \mathbf{R}$. Then equating real and imaginary parts, we have

$$Tv = av - bw \quad \text{and} \quad Tw = bv + aw.$$

Then

$$\begin{aligned} T_{\mathbf{C}}(v - iw) &= Tv - iTw \\ &= (av - bw) - i(bv + aw) \\ &= av - ibv - bw - iaw \\ &= (a - ib)(v - iw) \\ &= \overline{\lambda}(v - iw). \end{aligned}$$

Since $v - iw \neq 0$, this shows that $\overline{\lambda}$ is an eigenvalue of $T_{\mathbf{C}}$ with eigenvector $v - iw$.

Conversely, if $\overline{\lambda}$ is an eigenvalue of $T_{\mathbf{C}}$, then by the same argument, λ is an eigenvalue of $T_{\mathbf{C}}$. This completes the proof. ■

Exercise 5A19

Show that the forward shift operator $T \in \mathcal{L}(\mathbf{F}^{\infty})$ defined by

$$T(z_1, z_2, \ldots) = (0, z_1, z_2, \ldots)$$

has no eigenvalues.

Solution. Suppose $\lambda \in \mathbf{F}$ is an eigenvalue of T with eigenvector $(z_1, z_2, \ldots) \in \mathbf{F}^{\infty}$. Then

$$T(z_1, z_2, \ldots) = (0, z_1, z_2, \ldots) = \lambda(z_1, z_2, \ldots).$$

Equating the first component, we get $0 = \lambda z_1$, so either $\lambda = 0$ or $z_1 = 0$. If $\lambda = 0$, then the second component gives $z_1 = 0$, and the third component gives $z_2 = 0$, and so on. By induction, all $z_k = 0$, contradicting that the eigenvector is nonzero. If $\lambda \neq 0$, then $z_1 = 0$, and the second component gives $z_2 = 0$, and so on. Again, all $z_k = 0$, a contradiction. Therefore, T has no eigenvalues. ■

Exercise 5A20

Define the backward shift operator $S \in \mathcal{L}(\mathbf{F}^{\infty})$ by

$$S(z_1, z_2, z_3, \ldots) = (z_2, z_3, \ldots).$$

(a) Show that every element of $\mathbf{F}$ is an eigenvalue of S.
(b) Find all eigenvectors of S.

Solution.

(a) Let $\alpha \in \mathbf{F}$. Consider the vector $(1, \alpha, \alpha^2, \ldots) \in \mathbf{F}^{\infty}$. Then

$$S(1, \alpha, \alpha^2, \ldots) = (\alpha, \alpha^2, \alpha^3, \ldots) = \alpha(1, \alpha, \alpha^2, \ldots),$$

showing that α is an eigenvalue of S with eigenvector $(1, \alpha, \alpha^2, \ldots)$. Since $\alpha \in \mathbf{F}$ was arbitrary, every element of $\mathbf{F}$ is an eigenvalue of S.

(b) Let $\alpha \in \mathbf{F}$ and consider the eigenvalue α of S. Any eigenvector corresponding to α must satisfy

$$S(z_1, z_2, z_3, \ldots) = \alpha(z_1, z_2, z_3, \ldots),$$

which gives

$$(z_2, z_3, z_4, \ldots) = (\alpha z_1, \alpha z_2, \alpha z_3, \ldots).$$

Equating components, we get $z_2 = \alpha z_1$, $z_3 = \alpha z_2 = \alpha^2 z_1$, and in general $z_k = \alpha^{k-1} z_1$ for $k \geq 1$. Therefore, all eigenvectors corresponding to α are of the form $(z_1, \alpha z_1, \alpha^2 z_1, \ldots) = z_1(1, \alpha, \alpha^2, \ldots)$ with $z_1 \neq 0$. ■

Exercise 5A21

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) Suppose $\lambda \in \mathbf{F}$ with $\lambda \neq 0$ is an eigenvalue of T with eigenvector $v \in V$. Then

$$Tv = \lambda v.$$

Since T is invertible, we can apply T^{-1} to both sides to get

$$v = T^{-1}(\lambda v) = \lambda T^{-1} v.$$

Dividing both sides by λ (which is nonzero), we obtain

$$T^{-1} v = \frac{1}{\lambda} v.$$

This shows that $\frac{1}{\lambda}$ is an eigenvalue of T^{-1} with the same eigenvector v.

Conversely, suppose $\mu \in \mathbf{F}$ with $\mu \neq 0$ is an eigenvalue of T^{-1} with eigenvector $v \in V$. Then

$$T^{-1} v = \mu v.$$

Applying T to both sides gives

$$v = \mu T v \implies Tv = \frac{1}{\mu} v.$$

This shows that $\frac{1}{\mu}$ is an eigenvalue of T with the same eigenvector v.

Therefore, λ is an eigenvalue of T if and only if $\frac{1}{\lambda}$ is an eigenvalue of T^{-1}.

(b) Part (a) shows that if λ is an eigenvalue of T, then $\frac{1}{\lambda}$ is an eigenvalue of T^{-1} with the same eigenvector. Conversely, if μ is an eigenvalue of T^{-1}, then $\frac{1}{\mu}$ is an eigenvalue of T with the same eigenvector. Therefore, T and T^{-1} have the same eigenvectors. ■

Exercise 5A22

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. Consider the vectors $u + w$ and $u - w$. Since u and w are both nonzero, at least one of $u + w$ or $u - w$ is nonzero. We have

$$T(u + w) = Tu + Tw = 3w + 3u = 3(u + w),$$

and

$$T(u - w) = Tu - Tw = 3w - 3u = -3(u - w).$$

Hence, one of $u + w$ or $u - w$ is an eigenvector with eigenvalue 3 or -3, respectively. ■

Exercise 5A23

Suppose V is finite-dimensional and $S, T \in \mathcal{L}(V)$. Prove that ST and TS have the same eigenvalues.

Solution. Let λ be an eigenvalue of ST with eigenvector $v \in V$, so that

$$STv = \lambda v.$$

If $\lambda = 0$, then ST is not invertible. Since V is finite-dimensional, TS is also not invertible, so 0 is an eigenvalue of TS. Otherwise, if $\lambda \neq 0$, let $w = Tv$. Then

$$Sw = S(Tv) = STv = \lambda v \neq 0 \implies w \neq 0.$$

Now, we have

$$TSw = T(STv) = T(\lambda v) = \lambda Tv = \lambda w.$$

Therefore, w is an eigenvector of TS with eigenvalue λ. Then every eigenvalue of ST is also an eigenvalue of TS. By symmetry, the same argument shows that if λ is an eigenvalue of TS, then it is also an eigenvalue of ST. Hence, ST and TS have the same eigenvalues. ■

Exercise 5A24

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) Suppose the sum of the entries in each row of A equals 1. Let $v = (1, 1, \ldots, 1)^t \in \mathbf{F}^n$. Then

$$Tv = Av = v,$$

showing that v is an eigenvector of T corresponding to the eigenvalue 1.

(b) Suppose the sum of the entries in each column of A equals 1. By part (a), this is equivalent to saying that the sum of the entries in each row of A^t equals 1, hence 1 is an eigenvalue of the linear operator defined by A^t. But the

eigenvalues of A^t are the same as the eigenvalues of A because $A - \lambda I$ is not invertible if and only if $A^t - \lambda I$ is not invertible. Therefore, 1 is an eigenvalue of T. ■

Exercise 5A25

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. Let $Tu = \lambda u$ and $Tw = \mu w$ for some scalars $\lambda, \mu \in \mathbf{F}$. Since $u + w$ is also an eigenvector of T, we have $T(u + w) = \nu(u + w)$ for some scalar $\nu \in \mathbf{F}$. Then

$$\lambda u + \mu w = Tu + Tw = T(u + w) = \nu(u + w) = \nu u + \nu w.$$

Suppose u and w correspond to different eigenvalues. Then the list u, w is linearly independent, and we can equate coefficients of u and w above to obtain $\lambda = \nu = \mu$, which is a contradiction. Therefore, u and w must correspond to the same eigenvalue. ■

Exercise 5A26

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. Let $v \in V$ be any nonzero vector. By assumption, v is an eigenvector of T, so $Tv = \lambda_v v$ for some scalar $\lambda_v \in \mathbf{F}$. Now, let $u \in V$ be another nonzero vector. Consider the vector $u + v$. If u and v are linearly dependent, then without loss of generality, let $u = \alpha v$ for some scalar $\alpha \in \mathbf{F}$. Then $Tu = \lambda_u u = \lambda_u \alpha v$, and $Tu = T(\alpha v) = \alpha Tv = \alpha \lambda_v v$. Equating these, we get $\lambda_u \alpha v = \alpha \lambda_v v$, and since $\alpha \neq 0$ and $v \neq 0$, it follows that $\lambda_u = \lambda_v$.

The remaining case is when $u + v$ are linearly independent. In this case, $u + v$ is an eigenvector of T, so $T(u + v) = \lambda_{u+v}(u + v)$ for some scalar $\lambda_{u+v} \in \mathbf{F}$. On the other hand, $T(u + v) = Tu + Tv = \lambda_u u + \lambda_v v$. Equating these, we get $\lambda_{u+v} u + \lambda_{u+v} v = \lambda_u u + \lambda_v v$. Since u and v are linearly independent, we can equate coefficients to get $\lambda_{u+v} = \lambda_u = \lambda_v$.

Since u and v were arbitrary nonzero vectors, it follows that there exists a scalar $\lambda \in \mathbf{F}$ such that $Tv = \lambda v$ for all nonzero $v \in V$. Hence, $T = \lambda I$. ■

Exercise 5A27 ✔

Suppose that V is finite-dimensional and $k \in \{1, \dots, \dim V - 1\}$. 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. Let $v_1 \in V$ be nonzero. Our goal is to intersect enough k-dimensional subspaces to isolate $\operatorname{span} v_1$. We proceed as follows. Let $\dim V = n$, and extend v_1 to a basis $v_1, \dots, v_n$ of V. For each $i = 2, \dots, n$, let

$$U_i = \operatorname{span}(v_1, w_2, \dots, w_k), \quad w_2, \dots, w_k \in \{v_2, \dots, v_n\} \setminus \{v_i\},$$

i.e., each U_i must contain v_1 as well as $k - 1$ vectors from the basis except v_i. This choice is possible, since $k \leq n - 1$, so there are enough vectors in $\{v_2, \dots, v_n\} \setminus \{v_i\}$ to choose $k - 1$ of them. The intersection of all such U_i will be exactly $\operatorname{span} v_1$. Since each U_i is invariant under T, we have $Tv_1 \in U_i$ for all i. Therefore, $Tv_1 \in \operatorname{span} v_1$, which means Tv_1 is a scalar multiple of v_1. Since v_1 was arbitrary, it follows that T is a scalar multiple of the identity operator. ■

Exercise 5A28

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. Let $\lambda_1, \dots, \lambda_m$ be distinct eigenvalues of T with corresponding eigenvectors $v_1, \dots, v_m$. Since the eigenvalues are distinct, at most one eigenvalue must be 0. Without loss of generality, assume $\lambda_1 = 0$ if 0 is an eigenvalue. Then $\lambda_2, \dots, \lambda_m$ are nonzero eigenvalues of T. For $j = 2, \dots, m$, we have $Tv_j = \lambda_j v_j$, so $T(\lambda_j^{-1} v_j) = v_j$, which implies $v_j \in \operatorname{range} T$. Therefore, $v_2, \dots, v_m$ are linearly independent vectors in $\operatorname{range} T$, so $m - 1 \leq \dim \operatorname{range} T$. Hence, $m \leq 1 + \dim \operatorname{range} T$. ∎

Exercise 5A29

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 $Tx - 9x = (-4, 5, \sqrt{7})$.

Solution. Since $\dim \mathbf{R}^3 = 3$, the eigenvalues of T are exactly $-4, 5$, and $\sqrt{7}$. It follows that 9 is not an eigenvalue of T, and $\operatorname{null}(T - 9I) = \{0\}$. Then $T - 9I$ is injective, and since $\mathbf{R}^3$ is finite-dimensional, $T - 9I$ is also surjective. Therefore, for any vector in $\mathbf{R}^3$, there exists $x \in \mathbf{R}^3$ such that $(T - 9I)x = (-4, 5, \sqrt{7})$. ∎

Exercise 5A30

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. Let v be an eigenvector of T corresponding to the eigenvalue λ. Then

$$(T - 2I)(T - 3I)(T - 4I)v = (T - 2I)(T - 3I)(\lambda - 4)v = (T - 2I)(\lambda - 3)(\lambda - 4)v = (\lambda - 2)(\lambda - 3)(\lambda - 4)v = 0.$$

Since $v \neq 0$, it must be that $(\lambda - 2)(\lambda - 3)(\lambda - 4) = 0$. Therefore, $\lambda = 2$, $\lambda = 3$, or $\lambda = 4$. ∎

Exercise 5A31

Give an example of $T \in \mathcal{L}(\mathbf{R}^2)$ such that $T^4 = -I$.

Solution. Consider the rotation matrix by $\pi/4$ radians:

$$T = \begin{pmatrix} \cos(\pi/4) & -\sin(\pi/4) \\ \sin(\pi/4) & \cos(\pi/4) \end{pmatrix}.$$

Then

$$T^4 = \begin{pmatrix} \cos(\pi) & -\sin(\pi) \\ \sin(\pi) & \cos(\pi) \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -I.$$ ∎

Exercise 5A32

Suppose $T \in \mathcal{L}(V)$ has no eigenvalues and $T^4 = I$. Prove that $T^2 = -I$.

Manual Remark. *The proof of this is a little easier if we may use minimal polynomials.*

Solution. Note that

$$T^4 - I = (T^2 - I)(T^2 + I) = (T - I)(T + I)(T^2 + I) = 0.$$

Since T has no eigenvalues, $T - I$ and $T + I$ are both injective, hence $T^2 - I$ is injective. Then for any $v \in V$, we have

$$(T^2 - I)(T^2 + I)v = 0,$$

which implies $(T^2 + I)v = 0$ since $T^2 - I$ is injective. Therefore, $T^2 + I = 0$, or $T^2 = -I$. ∎

Exercise 5A33

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.

(a) Suppose T is injective. Then if $T^m v = 0$, we have $T(T^{m-1}v) = 0$, which implies $T^{m-1}v = 0$ since T is injective. Repeating this argument, we eventually get $v = 0$. Hence, T^m is injective.

Conversely, suppose T^m is injective. If $Tv = 0$, then $T^m v = 0$, which implies $v = 0$. Hence, T is injective.

(b) Suppose T is surjective. Then for any $v \in V$, there exists $w \in V$ such that $Tw = v$. Similarly, there exists $w_2 \in V$ such that $Tw_2 = w$, and so on. Repeating this m times, we find $u \in V$ such that $T^m u = v$. Hence, T^m is surjective.

Conversely, suppose T^m is surjective. Then for any $v \in V$, there exists $u \in V$ such that $T^m u = v$. Let $w = T^{m-1}u$. Then $Tw = T(T^{m-1}u) = T^m u = v$. Hence, T is surjective. ■

Exercise 5A34

Suppose V is finite-dimensional and $v_1, \dots, v_m \in V$. Prove that the list $v_1, \dots, v_m$ is linearly independent if and only if there exists $T \in \mathcal{L}(V)$ such that $v_1, \dots, v_m$ are eigenvectors of T corresponding to distinct eigenvalues.

Solution. Suppose $v_1, \dots, v_m$ is linearly independent. Since V is finite-dimensional, we can extend this list to a basis $v_1, \dots, v_m, v_{m+1}, \dots, v_n$ of V. Define $T \in \mathcal{L}(V)$ by setting $Tv_i = \lambda_i v_i$ for $i = 1, \dots, m$, where $\lambda_1, \dots, \lambda_m$ are distinct scalars, and choose $\lambda_{m+1}, \dots, \lambda_n$ arbitrarily (distinct from each other and from $\lambda_1, \dots, \lambda_m$). Then $v_1, \dots, v_m$ are eigenvectors of T corresponding to distinct eigenvalues.

Conversely, suppose there exists $T \in \mathcal{L}(V)$ such that $v_1, \dots, v_m$ are eigenvectors corresponding to distinct eigenvalues $\lambda_1, \dots, \lambda_m$. By 5.11, eigenvectors corresponding to distinct eigenvalues are linearly independent. Hence, $v_1, \dots, v_m$ is linearly independent. ■

Exercise 5A35

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 $\mathbf{R}$.

Axler 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. Suppose there exists a linear combination

$$a_1 e^{\lambda_1 x} + \cdots + a_n e^{\lambda_n x} = 0$$

for all $x \in \mathbf{R}$. Let $V = \operatorname{span}(e^{\lambda_1 x}, \dots, e^{\lambda_n x})$ and define $D \in \mathcal{L}(V)$ by $Df = f'$. Then $De^{\lambda_i x} = \lambda_i e^{\lambda_i x}$, so $e^{\lambda_i x}$ is an eigenvector of D corresponding to the eigenvalue λ_i. Since the λ_i are distinct, by 5.11, the eigenvectors corresponding to distinct eigenvalues are linearly independent. Hence, $a_1 = \cdots = a_n = 0$, proving that $e^{\lambda_1 x}, \dots, e^{\lambda_n x}$ are linearly independent. ■

Exercise 5A36

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. Suppose there exists a linear combination

$$a_1 \cos(\lambda_1 x) + \cdots + a_n \cos(\lambda_n x) = 0$$

for all $x \in \mathbf{R}$. Let $V = \operatorname{span}(\cos(\lambda_1 x), \sin(\lambda_1 x), \ldots, \cos(\lambda_n x), \sin(\lambda_n x))$ and define $D \in \mathcal{L}(V)$ by $Df = f'$. Then $D^2 \cos(\lambda_i x) = -\lambda_i^2 \cos(\lambda_i x)$, so $\cos(\lambda_i x)$ is an eigenvector of D^2 corresponding to the eigenvalue $-\lambda_i^2$. Since the λ_i are distinct and positive, the $-\lambda_i^2$ are distinct, and by 5.11, the eigenvectors corresponding to distinct eigenvalues are linearly independent. Hence, $a_1 = \cdots = a_n = 0$, proving that $\cos(\lambda_1 x), \ldots, \cos(\lambda_n x)$ are linearly independent. ■

Exercise 5A37

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 λ is an eigenvalue of T with corresponding eigenvector $v \in V$, so $Tv = \lambda v$. Consider a nonzero $\varphi \in V'$ and consider $S \in \mathcal{L}(V)$ defined by $Sw = \varphi(w)v$ for all $w \in V$. Then $S \neq 0$ and for all $w \in V$,

$$TSw = T(\varphi(w)v) = \varphi(w)Tv = \varphi(w)\lambda v = \lambda\varphi(w)v = \lambda Sw.$$

Hence, λ is an eigenvalue of $\mathcal{A}$.

Conversely, if λ is an eigenvalue of $\mathcal{A}$ with corresponding eigenvector $S \in \mathcal{L}(V)$, then $\mathcal{A}(S) = TS = \lambda S$. Since $S \neq 0$, there exists $v \in V$ such that $Sv \neq 0$. Then $TSv = \lambda Sv$, so λ is an eigenvalue of T. Therefore, the set of eigenvalues of T equals the set of eigenvalues of $\mathcal{A}$. ■

Exercise 5A38

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) To show that the definition of T/U makes sense, suppose $v + U = v' + U$ for some $v, v' \in V$. Then $v - v' \in U$, and since U is invariant under T, we have $T(v - v') \in U$, so $Tv - Tv' \in U$, which implies $Tv + U = Tv' + U$. Hence, the definition of T/U is well-defined.

To show that T/U is an operator on V/U, note that for any $v + U \in V/U$, $(T/U)(v + U) = Tv + U \in V/U$. Linearity follows from the linearity of T: for any $\alpha, \beta \in \mathbf{F}$ and $v + U, w + U \in V/U$,

$$\begin{aligned}(T/U)(\alpha(v + U) + \beta(w + U)) &= (T/U)(\alpha v + \beta w + U) \\ &= T(\alpha v + \beta w) + U \\ &= \alpha Tv + \beta Tw + U \\ &= \alpha(T/U)(v + U) + \beta(T/U)(w + U)\end{aligned}$$

(b) Let λ be an eigenvalue of T/U with eigenvector $v+U \neq 0$. Then $v \notin U$ and $(T-\lambda I)v \in U$. Consider $(T-\lambda I)|_U \in \mathcal{L}(U)$.

- If $(T-\lambda I)|_U$ is not injective, then λ is an eigenvalue of $T|_U$ and hence of T.
- If $(T-\lambda I)|_U$ is injective, then by finite-dimensionality it is surjective onto U. Thus, there exists $u \in U$ with $(T-\lambda I)u = (T-\lambda I)v$. Consequently, $(T-\lambda I)(v-u) = 0$ with $v-u \neq 0$ (since $v \notin U$ and $u \in U$), showing λ is an eigenvalue of T.

In either case, λ is an eigenvalue of T. ■

Exercise 5A39

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 T has an eigenvalue λ with corresponding eigenvector $v \in V$. Consider $\operatorname{range}(T-\lambda I)$. Then $\dim \operatorname{range}(T-\lambda I) \leq \dim V - 1$. Extend $\operatorname{range}(T-\lambda I)$ to a subspace U of V of dimension $\dim V - 1$. Then U is invariant under T because for any $u \in U$, $Tu = (T-\lambda I)u + \lambda u \in U$. Hence, there exists a subspace of V of dimension $\dim V - 1$ that is invariant under T.

Conversely, suppose there exists a subspace U of V of dimension $\dim V - 1$ that is invariant under T. Then V/U is one-dimensional, and the induced operator T/U on V/U must have an eigenvalue λ. By Exercise 5A38 (b), any eigenvalue of T/U is also an eigenvalue of T. Hence, T has an eigenvalue. ■

Exercise 5A40

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. We first prove by induction that $(STS^{-1})^k = ST^kS^{-1}$ for all non-negative integers k. For $k = 0$, we have $(STS^{-1})^0 = I = ST^0S^{-1}$. Suppose the result holds for some $k = n$, so that $(STS^{-1})^n = ST^nS^{-1}$. Then

$$\begin{aligned}(STS^{-1})^{n+1} &= (STS^{-1})^n(STS^{-1}) \\ &= (ST^nS^{-1})(STS^{-1}) \\ &= ST^n(S^{-1}S)TS^{-1} \\ &= ST^nTS^{-1} \\ &= ST^{n+1}S^{-1}.\end{aligned}$$

By induction, the result holds for all non-negative integers k.

Now let $p(z) = a_0 + a_1z + \cdots + a_mz^m$ be any polynomial. Then

$$\begin{aligned}p(STS^{-1}) &= a_0I + a_1(STS^{-1}) + \cdots + a_m(STS^{-1})^m \\ &= a_0I + a_1(STS^{-1}) + \cdots + a_m(ST^mS^{-1}) \\ &= S(a_0I + a_1T + \cdots + a_mT^m)S^{-1} \\ &= Sp(T)S^{-1}.\end{aligned}$$

Hence, the result holds for all polynomials p. ■

Exercise 5A41

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. Let $u \in U$. Since U is invariant under T, we have $T(u) \in U$. By induction, suppose $T^k(u) \in U$ for some non-negative integer k. Then $T^{k+1}(u) = T(T^k(u)) \in U$ since U is invariant under T. Hence, $T^k(u) \in U$ for all non-negative integers k. Now let $p(z) = a_0 + a_1 z + \cdots + a_m z^m$ be any polynomial. Then

$$p(T)(u) = a_0 u + a_1 T(u) + \cdots + a_m T^m(u) \in U,$$

since each term is in U. Therefore, U is invariant under $p(T)$ for every polynomial $p \in \mathcal{P}(\mathbf{F})$. ■

Exercise 5A42 ✔

Define $T \in \mathcal{L}(\mathbf{F}^n)$ by $T(x_1, x_2, x_3, \ldots, x_n) = (x_1, 2x_2, 3x_3, \ldots, nx_n)$.

(a) Find all eigenvalues and eigenvectors of T.

(b) Find all subspaces of $\mathbf{F}^n$ that are invariant under T.

Solution.

(a) Let λ be an eigenvalue of T with corresponding eigenvector $v = (x_1, x_2, \ldots, x_n) \in \mathbf{F}^n$. Then $T(v) = \lambda v$, which means

$$(x_1, 2x_2, 3x_3, \ldots, nx_n) = (\lambda x_1, \lambda x_2, \ldots, \lambda x_n).$$

Comparing components, we get $\lambda x_i = i x_i$ for each $i = 1, 2, \ldots, n$. Hence, if $x_i \neq 0$, then $\lambda = i$. Therefore, the eigenvalues of T are $1, 2, \ldots, n$, and the corresponding eigenvectors are nonzero multiples of the standard basis vectors $e_1, e_2, \ldots, e_n$.

(b) Let U be an invariant subspace of $\mathbf{F}^n$ under T. Let

$$u = \sum_{i=1}^{n} \alpha_i e_i \in U,$$

where e_i are the standard basis vectors of $\mathbf{F}^n$ and $\alpha_i \in \mathbf{F}$. Since U is invariant under T, we have

$$Tu = \sum_{i=1}^{n} \alpha_i T(e_i) = \sum_{i=1}^{n} \alpha_i i e_i \in U.$$

Consider some index i such that $\alpha_i \neq 0$. Then

$$u_1 = Tu - iu = \sum_{j=1}^{n} \alpha_j j e_j - i \sum_{j=1}^{n} \alpha_j e_j = \sum_{j=1}^{n} \alpha_j (j - i) e_j \in U.$$

The i-th term in this sum is $\alpha_i (i - i) e_i = 0$, hence the e_i term is eliminated from u_1. Consider another index j of u_1 such that $\alpha_j \neq 0$. Then we can form

$$u_2 = Tu_1 - ju_1 = \sum_{k=1}^{n} \alpha_k (k - i) k e_k - j \sum_{k=1}^{n} \alpha_k (k - i) e_k = \sum_{k=1}^{n} \alpha_k (k - i)(k - j) e_k \in U.$$

The j-th term in this sum is $\alpha_j (j - i)(j - j) e_j = 0$, and u_2 now has all terms except the i-th and j-th terms. By repeating this process for each index with a nonzero coefficient, we can isolate each $\alpha_i e_i$ in U whenever $\alpha_i \neq 0$. This shows that for any $u \in U$, all the standard basis vectors e_i corresponding to nonzero coefficients in u must also be in U. Therefore, U must be the span of some subset of the standard basis vectors $\{e_1, e_2, \ldots, e_n\}$. ■

Exercise 5A43 ✔

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 T is the zero transformation or a scalar multiple of the identity, the result trivially holds. Otherwise, because $\dim V > 1$, there is at least one nonzero vector $v_1 \in V$ that is not an eigenvector of T. Let $Tv_1 = v_2$, so that v_1 and v_2 are linearly independent, and extend v_1, v_2 to a basis $v_1, \ldots, v_n$ for V.

Now suppose that $\{p(T) : p \in \mathcal{P}(\mathbf{F})\} = \mathcal{L}(V)$ and consider the map $S \in \mathcal{L}(V)$ defined by $Sv_1 = 0$ and $Sv_i = v_i$ for $i = 2, \ldots, n$. Because $S \in \mathcal{L}(V)$, there must be a polynomial $p \in \mathcal{P}(\mathbf{F})$ such that $p(T) = S$. Because the null space of $p(T)$ is invariant under T, it must be that the null space of S is invariant under T. This is impossible, though, as $Sv_1 = 0$ and $STv_1 = Sv_2 = v_2 \neq 0$. Therefore, there is no $T \in \mathcal{L}(V)$ such that $\{p(T) : p \in \mathcal{P}(\mathbf{F})\} = \mathcal{L}(V)$. ■

5B The Minimal Polynomial

Exercise 5B1

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. Suppose 9 is an eigenvalue of T^2. Then there exists a nonzero vector $v \in V$ such that $T^2 v = 9v$. This implies $(T - 3I)(T + 3I)v = 0$. Let $w = (T + 3I)v$, and suppose $w \neq 0$. Then

$$(T - 3I)w = Tw - 3w = 0 \implies Tw = 3w,$$

showing that 3 is an eigenvalue of T. If $w = 0$, then $(T + 3I)v = 0$, which implies $Tv = -3v$, showing that -3 is an eigenvalue of T.

Conversely, suppose 3 is an eigenvalue of T. Then there exists a nonzero vector $v \in V$ such that $Tv = 3v$. Applying T again, we get $T^2 v = T(3v) = 3Tv = 3 \cdot 3v = 9v$, showing that 9 is an eigenvalue of T^2. Similarly, if -3 is an eigenvalue of T, then 9 is an eigenvalue of T^2. ∎

Exercise 5B2

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 subspace of V invariant under T. If $U = \{0\}$, we are done. Suppose U is nonzero and finite-dimensional. By invariance of U under T, $T|_U \in \mathcal{L}(U)$, and $T|_U$ has an eigenvalue $\lambda \in \mathbf{C}$ with corresponding eigenvector $v \in U$, $v \neq 0$. But then v is also an eigenvector of T with eigenvalue λ, contradicting the assumption that T has no eigenvalues. Hence, any nonzero invariant subspace U must be infinite-dimensional. ∎

Exercise 5B3

Suppose n is a positive integer and $T \in \mathcal{L}(\mathbf{F}^n)$ is defined by

$$T(x_1, \ldots, x_n) = (x_1 + \cdots + x_n, \ldots, x_1 + \cdots + x_n).$$

(a) Find all eigenvalues and eigenvectors of T.
(b) Find the minimal polynomial of T.

Solution.

(a) If $n = 1$, then $T(x_1) = (x_1)$, and the only eigenvalue is 1 with corresponding eigenvectors being all nonzero vectors in $\mathbf{F}^1$. Otherwise, suppose $T(x_1, \ldots, x_n) = \lambda(x_1, \ldots, x_n)$ for some $\lambda \in \mathbf{F}$ and $(x_1, \ldots, x_n) \neq 0$. Then

$$\lambda x_j = x_1 + \cdots + x_n \quad \text{for all } j = 1, \ldots, n.$$

This implies that $\lambda x_1 = \lambda x_2 = \cdots = \lambda x_n$. If $\lambda \neq 0$, then $x_1 = x_2 = \cdots = x_n$, and we can write $(x_1, \ldots, x_n) = (x_1, \ldots, x_1) = x_1(1, \ldots, 1)$. Substituting this into the eigenvalue equation gives

$$\lambda(x_1, \ldots, x_1) = T(x_1, \ldots, x_1) = (nx_1, \ldots, nx_1),$$

which implies $\lambda = n$. If $\lambda = 0$, then $x_1 + \cdots + x_n = 0$, and the eigenvectors are all nonzero vectors $(x_1, \ldots, x_n)$ such that $x_1 + \cdots + x_n = 0$. Hence, the eigenvalues of T are 0 and n, with corresponding eigenvectors as described.

(b) If $n = 1$, the minimal polynomial is $z - 1$.

For $n > 1$, we know from (a) that the eigenvalues are 0 and n. Then the minimal polynomial is a polynomial multiple of $z(z-n)$. Letting $x = x_1 + \cdots + x_n$, then

$$T(x_1, \ldots, x_n) = (x, \ldots, x), \quad (T - nI)(x_1, \ldots, x_n) = (x - nx_1, \ldots, x - nx_n).$$

Observe that $(x - nx_1) + \cdots + (x - nx_n) = nx - n(x_1 + \cdots + x_n) = 0$. Therefore, $(T - nI)(x_1, \ldots, x_n)$ lies in null T, and applying T again gives

$$T(T - nI)(x_1, \ldots, x_n) = 0.$$

Hence, $z(z-n)$ is a polynomial multiple of the minimal polynomial, and since it must be monic, the minimal polynomial of T is $z(z-n)$. ■

Exercise 5B4

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 α is an eigenvalue of $p(T)$. Then there exists a nonzero vector $v \in V$ such that $p(T)v = \alpha v$. Then $p(T) - \alpha I$ is not injective. If $p(z) - \alpha = 0$, then $\alpha = p(\lambda)$ for any $\lambda \in \mathbf{C}$, and we are done. If $p(z) - \alpha = \beta$ for some constant $\beta \in \mathbf{F}$, then $p(T) - \alpha I = \beta I$ is injective, which contradicts the fact that $p(T) - \alpha I$ is not injective. We may now assume that $p(z) - \alpha$ is a nonconstant polynomial. By the Fundamental Theorem of Algebra, we can factor $p(z) - \alpha$ as

$$p(z) - \alpha = c(z - \lambda_1)(z - \lambda_2) \cdots (z - \lambda_m)$$

for some $c \in \mathbf{C} \setminus \{0\}$ and $\lambda_1, \ldots, \lambda_m \in \mathbf{C}$. Then

$$p(T) - \alpha I = c(T - \lambda_1 I)(T - \lambda_2 I) \cdots (T - \lambda_m I).$$

Since $p(T) - \alpha I$ is not injective, at least one of the factors on the right-hand side is not injective. Therefore, $T - \lambda_j I$ is not injective for some j, which means λ_j is an eigenvalue of T. Then $p(\lambda_j) - \alpha = 0$, so $\alpha = p(\lambda_j)$ for some eigenvalue λ_j of T.

Conversely, suppose $\alpha = p(\lambda)$ for some eigenvalue λ of T with corresponding eigenvector v. Let $p(z) = a_m z^m + \cdots + a_1 z + a_0$. Since $T^m v = \lambda^m v$, we have

$$p(T)v = a_m T^m v + \cdots + a_1 Tv + a_0 v = a_m \lambda^m v + \cdots + a_1 \lambda v + a_0 v = p(\lambda)v = \alpha v.$$

Hence, α is an eigenvalue of $p(T)$. ■

Exercise 5B5

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. Consider the operator $T \in \mathcal{L}(\mathbf{R}^2)$ defined by $T(x, y) = (-y, x)$, which represents a counterclockwise rotation by 90°. Observe that T has no real eigenvalues as follows: If (x, y) is an eigenvector of T with eigenvalue $\lambda \in \mathbf{R}$, then $T(x, y) = \lambda(x, y)$ implies $(-y, x) = (\lambda x, \lambda y)$. This gives the system of equations:

$$-y = \lambda x \quad \text{and} \quad x = \lambda y.$$

Since x and y cannot both be zero, one can work out that $\lambda^2 = -1$, which has no real solutions. Therefore, T has no real eigenvalues. Now consider the polynomial $p(z) = z^2 + 1$. Then

$$p(T)(x, y) = T^2(x, y) + (x, y) = (-x, -y) + (x, y) = (0, 0).$$

Thus, 0 is an eigenvalue of $p(T)$ with corresponding eigenvector $(1,0)$ (or any nonzero vector in $\mathbf{R}^2$). However, there is no real eigenvalue λ of T such that $p(\lambda) = 0$, since the eigenvalues of T are complex. This example demonstrates that the result in Exercise 5B4 does not hold when $\mathbf{C}$ is replaced with $\mathbf{R}$. ■

Exercise 5B6

Suppose $T \in \mathcal{L}(\mathbf{F}^2)$ is defined by $T(w,z) = (-z,w)$. Find the minimal polynomial of T.

Solution. By the computation done in Exercise 5B5, we have $T^2(w,z) = (-w,-z)$, so $T^2 + I = 0$. Therefore, the polynomial $p(z) = z^2 + 1$ annihilates T. To see that this is the minimal polynomial, suppose $T = \alpha I$. Then $T(1,0) = (0,1) \neq \alpha(1,0)$ for any $\alpha \in \mathbf{F}$, so $T \neq \alpha I$. Hence, the minimal polynomial of T is $p(z) = z^2 + 1$. ■

Exercise 5B7

(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.

Axler Hint. *Show that if S is invertible and $p \in \mathcal{P}(\mathbf{F})$, then $p(TS) = S^{-1}p(ST)S$.*

Solution.

(a) Consider the operators $S, T \in \mathcal{L}(\mathbf{F}^2)$ defined by $S(w,z) = (0,w)$ and $T(w,z) = (w,0)$. Then

$$ST(w,z) = S(w,0) = (0,w) \quad \text{and} \quad TS(w,z) = T(0,w) = (0,0).$$

The minimal polynomial of ST is $p(z) = z^2$, since $ST \neq 0$ but $(ST)^2 = 0$, while the minimal polynomial of TS is $q(z) = z$. Hence, the minimal polynomials of ST and TS are not equal.

(b) Without loss of generality, suppose S is invertible, and note that $ST = STSS^{-1}$. Let p be the minimal polynomial of TS. By Exercise 5A40, we have

$$p(ST) = p(S(TS)S^{-1}) = Sp(TS)S^{-1} = 0,$$

so p annihilates ST. Similarly, let q be the minimal polynomial of ST. Then

$$q(TS) = q(S^{-1}(ST)S) = S^{-1}q(ST)S = 0$$

as $TS = S^{-1}(ST)S$. Therefore, since p annihilates ST, p divides q, and since q annihilates TS, q divides p, so $p = q$. ■

Exercise 5B8

Suppose $T \in \mathcal{L}(\mathbf{R}^2)$ is the operator of counterclockwise rotation by $1°$. Find the minimal polynomial of T.

Axler Note. *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. Since T is not a scalar multiple of the identity operator, the minimal polynomial of T cannot have degree 1. Moreover, $\dim \mathbf{R}^2 = 2$ implies that the degree of the minimal polynomial of T must be 2. Therefore, the minimal polynomial of T is of the form

$$p(z) = z^2 + bz + c.$$

Consider the point $(1,0) \in \mathbf{R}^2$. Note that $T(x,y) = (x\cos 1^\circ - y\sin 1^\circ, x\sin 1^\circ + y\cos 1^\circ)$. Then

$$T(1,0) = (\cos 1^\circ, \sin 1^\circ), \quad T^2(1,0) = (\cos 2^\circ, \sin 2^\circ).$$

Since $p(T) = 0$, we have

$$p(T)(1,0) = T^2(1,0) + bT(1,0) + c(1,0) = (\cos 2^\circ + b\cos 1^\circ + c, \sin 2^\circ + b\sin 1^\circ) = (0,0).$$

This gives the system of equations:

$$\begin{cases} \cos 2^\circ + b\cos 1^\circ + c = 0, \\ \sin 2^\circ + b\sin 1^\circ = 0. \end{cases}$$

From this, we get the values of b and c:

$$b = -\frac{\sin 2^\circ}{\sin 1^\circ} = -2\cos 1^\circ, \quad c = -\cos 2^\circ - b\cos 1^\circ = 1.$$

Therefore, the minimal polynomial of T is $p(z) = z^2 - 2\cos 1^\circ z + 1$. ■

Exercise 5B9

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. Let A be the matrix of T with respect to some basis of V. Since all entries of A are rational numbers, A^k also has rational entries for any positive integer k. Then the minimal polynomial $p(z) = z^m + a_{m-1}z^{m-1} + \cdots + a_1 z + a_0$ of T in terms of the matrix A satisfies

$$p(A) = A^m + a_{m-1}A^{m-1} + \cdots + a_1 A + a_0 I = 0.$$

Then this gives a system of $(\dim V)^2$ linear equations in the coefficients $a_0, a_1, \ldots, a_{m-1}$ with rational coefficients. By uniqueness of the minimal polynomial, this system has a unique solution. Since the system has rational coefficients and a unique solution, the solution must also be rational. Therefore, all coefficients of the minimal polynomial of T are rational numbers. ■

Exercise 5B10

Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and $v \in V$. Prove that

$$\operatorname{span}(v, Tv, \ldots, T^m v) = \operatorname{span}(v, Tv, \ldots, T^{\dim V - 1}v)$$

for all integers $m \geq \dim V - 1$.

Solution. Let $n = \dim V$ and let $p(z) = z^d + a_{d-1}z^{d-1} + \cdots + a_0$ be the minimal polynomial of T. Then $d \leq n$ and $p(T) = 0$. For any $v \in V$,

$$p(T)v = 0 \implies T^d v = -a_{d-1}T^{d-1}v - \cdots - a_1 Tv - a_0 v \implies T^d v \in \operatorname{span}\left(v, Tv, \ldots, T^{d-1}v\right).$$

Let $U = \operatorname{span}\left(v, Tv, \ldots, T^{d-1}v\right)$.

We claim that U is invariant under T. Take any $u = c_0 v + c_1 Tv + \cdots + c_{d-1}T^{d-1}v \in U$. Then,

$$Tu = c_0 Tv + c_1 T^2 v + \cdots + c_{d-2}T^{d-1}v + c_{d-1}T^d v.$$

Every term on the right-hand side except $T^d v$ is automatically in the spanning set of U. We also showed that $T^d v \in U$ earlier, so $Tu \in U$. Hence, $T(U) \subseteq U$.

Because $v \in U$ and U is invariant under T, we see that $T^m v \in U$ for every $m \geq 0$ (this is immediate by induction). So, for any $m \geq n-1$, we have that

$$\operatorname{span}(v, Tv, \ldots, T^m v) \subseteq U \subseteq \operatorname{span}(v, Tv, \ldots, T^{n-1} v)$$

by using $d - 1 \leq n - 1$ to obtain the second inclusion. Since $\{v, Tv, \ldots, T^{n-1}v\} \subseteq \{v, Tv, \ldots, T^m v\}$, the reverse inclusion is immediate, and so

$$\operatorname{span}(v, Tv, \ldots, T^m v) = \operatorname{span}(v, Tv, \ldots, T^{n-1} v). \qquad \blacksquare$$

Exercise 5B11

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) We compute T^2:

$$T^2 = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \begin{pmatrix} a & c \\ b & d \end{pmatrix} = \begin{pmatrix} a^2 + bc & ac + cd \\ ab + bd & bc + d^2 \end{pmatrix}.$$

Then

$$\begin{aligned} T^2 - (a+d)T + (ad - bc)I &= \begin{pmatrix} a^2 + bc & ac + cd \\ ab + bd & bc + d^2 \end{pmatrix} - (a+d)\begin{pmatrix} a & c \\ b & d \end{pmatrix} + (ad - bc)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} a^2 + bc - (a+d)a + ad - bc & ac + cd - (a+d)c \\ ab + bd - (a+d)b & bc + d^2 - (a+d)d + ad - bc \end{pmatrix} \\ &= \begin{pmatrix} a^2 + bc - a^2 - ad + ad - bc & ac + cd - ac - cd \\ ab + bd - ab - bd & bc + d^2 - d^2 - ad + ad - bc \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \end{aligned}$$

(b) If $b = c = 0$ and $a = d$, then the matrix of T is

$$\begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix},$$

which is a scalar multiple of the identity matrix. In this case, the minimal polynomial of T is $z - a$.

Otherwise, the polynomial $z^2 - (a+d)z + (ad - bc)$ annihilates T by part (a). To see that this is the minimal polynomial, we need to check that T is not a scalar multiple of the identity matrix. If T were a scalar multiple of the identity matrix, then we would have $b = c = 0$ and $a = d$, which contradicts our assumption. Therefore, the minimal polynomial of T has degree 2 and is $z^2 - (a+d)z + (ad - bc)$ in this case. $\blacksquare$

Exercise 5B12

Define $T \in \mathcal{L}(\mathbf{F}^n)$ by $T(x_1, x_2, x_3, \ldots, x_n) = (x_1, 2x_2, 3x_3, \ldots, nx_n)$. Find the minimal polynomial of T.

Solution. For each $k \in \{1, \ldots, n\}$, the standard basis vector e_k satisfies $Te_k = ke_k$, so each k is an eigenvalue. Let

$p(z) = (z-1)(z-2)\cdots(z-n)$. So,

$$p(T)e_k = p(k)e_k = 0 \cdot e_k = 0 \text{ for all } k \in \{1, \ldots, n\},$$

and since the e_k form a basis of V, $p(T) = 0$. Therefore, the minimal polynomial $m_T(z)$ divides $p(z)$.

Now suppose $m_T(z)$ is the minimal polynomial of T. For each eigenvalue k,

$$0 = m_T(T)e_k = m_T(k)e_k \implies m_T(k) = 0.$$

Thus, $(z-k)$ divides $m_T(z)$ for every $k \in \{1, \ldots, n\}$. Because the factors are distinct, $p(z)$ divides $m_T(z)$.

Since both $p(z)$ and $m_T(z)$ are monic, $p(z) = m_T(z)$. Hence, the minimal polynomial is $p(z) = (z-1)\cdots(z-n)$. ■

Exercise 5B13

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 m be the minimal polynomial of T with degree d. By the Division Algorithm for polynomials, we can write

$$p(z) = q(z)m(z) + r(z),$$

where $q, r \in \mathcal{P}(\mathbf{F})$ and $\deg r < d$. Applying this to T, we have

$$p(T) = q(T)m(T) + r(T).$$

Since $m(T) = 0$, it follows that

$$p(T) = r(T).$$

Thus, there exists a polynomial r such that $p(T) = r(T)$ and $\deg r < d$.

To show uniqueness, suppose there exists another polynomial $s \in \mathcal{P}(\mathbf{F})$ such that $p(T) = s(T)$ and $\deg s < d$. Then we have

$$r(T) = p(T) = s(T).$$

This implies that $(r-s)(T) = 0$. Since $\deg(r-s) < d$, it follows that $r - s$ must be the zero polynomial (otherwise, it would contradict the minimality of m). Hence, $r = s$, proving uniqueness. ■

Exercise 5B14

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. Let $p(z) = 4 + 5z - 6z^2 - 7z^3 + 2z^4 + z^5$ be the minimal polynomial of T. We want to find the minimal polynomial of T^{-1}. Since $p(z)$ is the minimal polynomial of T, we have

$$p(T) = 4I + 5T - 6T^2 - 7T^3 + 2T^4 + T^5 = 0.$$

To find the minimal polynomial of T^{-1}, apply T^{-1} to both sides of the equation until we express it in terms of T^{-1}:

$$4T^{-5} + 5T^{-4} - 6T^{-3} - 7T^{-2} + 2T^{-1} + I = 0.$$

Then the polynomial $q(z) = 4z^5 + 5z^4 - 6z^3 - 7z^2 + 2z + 1$ annihilates T^{-1}. To make it monic, we can divide by the leading

coefficient:

$$q(z) = z^5 + \frac{5}{4}z^4 - \frac{3}{2}z^3 - \frac{7}{4}z^2 + \frac{1}{2}z + \frac{1}{4}.$$

Suppose T^{-1} has a minimal polynomial $m(z)$ with $\deg m < 5$, hence $m(T^{-1}) = 0$. Applying $T^{\deg m}$ to both sides gives $n(T) = 0$ for some polynomial $n(z)$ of degree less than 5, which contradicts the minimality of $p(z)$. Therefore, the minimal polynomial of T^{-1} is $q(z)$. ∎

Exercise 5B15

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. Since V is a finite-dimensional complex vector space, the operator T has at least one eigenvalue $\lambda_0 \in \mathbf{C}$. By Exercise 5A11, there exists $\delta > 0$ such that $T - \lambda I$ is invertible for all $\lambda \in \mathbf{C}$ with $0 < |\lambda - \lambda_0| < \delta$. This implies that for all such λ, we have $\dim \operatorname{range}(T - \lambda I) = \dim V$. However, at the eigenvalue λ_0, we have $\dim \operatorname{range}(T - \lambda_0 I) < \dim V$ since $\dim \operatorname{null}(T - \lambda_0 I) > 0$. Therefore,

$$\lim_{\lambda \to \lambda_0} f(\lambda) \neq f(\lambda_0),$$

which shows that f is not continuous at λ_0. ∎

Exercise 5B16

Suppose $a_0, \dots, a_{n-1} \in \mathbf{F}$. 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

$$a_0 + a_1 z + \cdots + a_{n-1} z^{n-1} + z^n.$$

Solution. Let $p(z) = a_0 + a_1 z + \cdots + a_{n-1} z^{n-1} + z^n$. Let $e_1, \dots, e_n$ be the standard basis of $\mathbf{F}^n$. We will show that $p(T) = 0$ and that no polynomial of lower degree annihilates T. First, we compute $T^k e_1$ for $k = 0, 1, \dots, n$:

$$T^0 e_1 = e_1, \quad T^1 e_1 = e_2, \quad T^2 e_1 = e_3, \quad \dots, \quad T^{n-1} e_1 = e_n, \quad T^n e_1 = -a_0 e_1 - a_1 e_2 - \cdots - a_{n-1} e_n.$$

Then we have

$$\begin{aligned} p(T) e_1 &= a_0 T^0 e_1 + a_1 T^1 e_1 + \cdots + a_{n-1} T^{n-1} e_1 + T^n e_1 \\ &= a_0 e_1 + a_1 e_2 + \cdots + a_{n-1} e_n - a_0 e_1 - a_1 e_2 - \cdots - a_{n-1} e_n \\ &= 0. \end{aligned}$$

Since $T^{j-1}e_1 = e_j$ for $j = 1, 2, \ldots, n$, and $p(T)$ commutes with T, we have

$$p(T)e_j = p(T)T^{j-1}e_1 = T^{j-1}p(T)e_1 = T^{j-1}0 = 0$$

for all $j = 1, 2, \ldots, n$. Since $\{e_1, e_2, \ldots, e_n\}$ is a basis for $\mathbf{F}^n$, it follows that $p(T) = 0$.

To show that p is the minimal polynomial, suppose there exists a polynomial $q(z) = b_0 + b_1 z + \cdots + b_{m-1}z^{m-1} + z^m$ of degree $m < n$ such that $q(T) = 0$. Then we would have

$$q(T)e_1 = b_0 e_1 + b_1 e_2 + \cdots + b_{m-1}e_m + T^m e_1 = 0.$$

However, $T^m e_1 = e_{m+1}$, which is not in the span of $\{e_1, e_2, \ldots, e_m\}$ since $m < n$. This leads to a contradiction, as e_{m+1} cannot be expressed as a linear combination of $e_1, e_2, \ldots, e_m$. Therefore, no polynomial of degree less than n annihilates T, and the minimal polynomial of T is indeed $p(z) = a_0 + a_1 z + \cdots + a_{n-1}z^{n-1} + z^n$. ■

Exercise 5B17

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 $p(z)$ be the minimal polynomial of T. Then we have

$$p(T) = 0.$$

Consider the polynomial $q(z) = p(z + \lambda)$. Then we have

$$q(T - \lambda I) = p((T - \lambda I) + \lambda I) = p(T) = 0.$$

Thus, q annihilates $T - \lambda I$.

To show that q is the minimal polynomial, suppose there exists a polynomial $r(z)$ of degree less than $\deg q$ such that $r(T - \lambda I) = 0$. Then we can write

$$s(z) = r(z - \lambda)$$

for some polynomial $s(z)$ of degree less than $\deg p$. Then we have

$$s(T) = r(T - \lambda I) = 0.$$

However, this contradicts the minimality of p, since $\deg s < \deg p$ and $s(T) = 0$. Therefore, no polynomial of degree less than $\deg q$ annihilates $T - \lambda I$, and the minimal polynomial of $T - \lambda I$ is indeed $q(z) = p(z + \lambda)$. ■

Exercise 5B18

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 $q(z) = \lambda^{\deg p} p(\frac{z}{\lambda})$.

Solution. Let $p(z)$ be the minimal polynomial of T. Then we have

$$p(T) = 0.$$

Consider the polynomial $q(z) = \lambda^{\deg p} p(\frac{z}{\lambda})$. Then we have

$$q(\lambda T) = \lambda^{\deg p} p\left(\frac{\lambda T}{\lambda}\right) = \lambda^{\deg p} p(T) = 0.$$

Moreover, q has leading term $\lambda^{\deg p}(\frac{z}{\lambda})^{\deg p} = z^{\deg p}$, so q is monic and has the same degree as p. To show that q is the minimal polynomial, suppose there exists a polynomial $r(z)$ of degree less than $\deg q$ such that $r(\lambda T) = 0$. Define $s(z) = r(\lambda z)$. Then we have

$$s(T) = r(\lambda T) = 0.$$

However, $\deg s = \deg r < \deg q = \deg p$, which contradicts the minimality of p. Therefore, no polynomial of degree less than $\deg q$ annihilates λT, and the minimal polynomial of λT is indeed $q(z) = \lambda^{\deg p} p(\frac{z}{\lambda})$. ∎

Exercise 5B19

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 $p(z)$ be the minimal polynomial of T with degree d. Let $\mathcal{I} = \{I, T, T^2, \ldots, T^{d-1}\}$. We will show that $\mathcal{I}$ is a basis for $\mathcal{E}$.

First, we show that $\dim \mathcal{E} \leq d$. By the Division Algorithm for polynomials, for any polynomial $q \in \mathcal{P}(\mathbf{F})$, we can write

$$q(z) = s(z)p(z) + r(z),$$

where $s, r \in \mathcal{P}(\mathbf{F})$ and $\deg r < d$. Then we have

$$q(T) = s(T)p(T) + r(T) = 0 + r(T) = r(T).$$

Since $\deg r < d$, the set $\mathcal{I}$ spans $\mathcal{E}$. Hence, $\dim \mathcal{E} \leq d$.

Next, we show that $\dim \mathcal{E} \geq d$. Suppose

$$a_0 I + a_1 T + a_2 T^2 + \cdots + a_{d-1} T^{d-1} = 0$$

for scalars $a_0, a_1, \ldots, a_{d-1} \in \mathbf{F}$ not all zero. Then the polynomial

$$q(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_{d-1} z^{d-1}$$

annihilates T, contradicting the minimality of $p(z)$ since $\deg q < d$. Therefore, the set $\mathcal{I}$ is linearly independent in $\mathcal{E}$, which implies that $\dim \mathcal{E} \geq d$.

Combining both inequalities, we conclude that $\dim \mathcal{E} = d$, which is the degree of the minimal polynomial of T. ∎

Exercise 5B20

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. The eigenvalues of T are 3, 5, and 8, so each is a root of p. Hence,

$$p(z) = (z - 3)^{m_1}(z - 5)^{m_2}(z - 8)^{m_3}$$

where $m_1, m_2, m_3 \geq 1$. Because T acts on a 4-dimensional space, $\deg p \leq 4$, and so $m_1 + m_2 + m_3 \leq 4$. Since $m_2, m_3 \geq 1$, we have $m_1 \leq 4 - 1 - 1 = 2$. By similar logic, $m_1, m_2, m_3 \leq 2$. This means $p(z)$ divides $(z - 3)^2(z - 5)^2(z - 8)^2$. Applying this polynomial to T yields

$$(T - 3I)^2(T - 5I)^2(T - 8I)^2 = 0.$$ ∎

Exercise 5B21

Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Prove that the minimal polynomial of T has degree at most $1 + \dim \operatorname{range} T$.

Axler Note. *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 $r = \dim \operatorname{range} T$. Observe that $\operatorname{range} T$ is invariant under T, and we can consider the restriction of T to $\operatorname{range} T$, denoted by $T|_{\operatorname{range} T}$. Let $p(z)$ be the minimal polynomial of T and let $q(z)$ be the minimal polynomial of $T|_{\operatorname{range} T}$. Then $\deg q \le r$ since $T|_{\operatorname{range} T}$ is an operator on a space of dimension r.

We now show that the polynomial $zq(z)$ annihilates T. For any $v \in V$, we know that $Tv \in \operatorname{range} T$. By invariance, we know that $w \in \operatorname{range} T$ implies $(T|_{\operatorname{range} T})^k w \in \operatorname{range} T$ for any $k \ge 0$. Then

$$q(T)w = q(T|_{\operatorname{range} T})w = 0.$$

Therefore, for any $v \in V$, we have

$$zq(z)(T)v = Tq(T)v = q(T)Tv = 0,$$

where we used the fact that T commutes with $q(T)$ since $q(T)$ is a polynomial in T. This shows that $zq(z)$ annihilates T, and hence the minimal polynomial $p(z)$ of T divides $zq(z)$. Since $\deg q \le r$, we have $\deg p \le 1 + r = 1 + \dim \operatorname{range} T$. ∎

Exercise 5B22

Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Prove that T is invertible if and only if $I \in \operatorname{span}(T, T^2, \dots, T^{\dim V})$.

Solution. Suppose T is invertible. Let $q(z)$ be the minimal polynomial of T. Since T is invertible, 0 is not an eigenvalue of T, and thus $q(0) \neq 0$. We may write $q(z)$ as

$$q(z) = z^m + a_{m-1}z^{m-1} + \cdots + a_1 z + a_0,$$

where $a_0 \neq 0$ and $m \le \dim V$. Then we have

$$0 = q(T) = T^m + a_{m-1}T^{m-1} + \cdots + a_1 T + a_0 I.$$

Rearranging gives

$$I = -\frac{1}{a_0}(T^m + a_{m-1}T^{m-1} + \cdots + a_1 T) \in \operatorname{span}(T, T^2, \dots, T^m) \subseteq \operatorname{span}(T, T^2, \dots, T^{\dim V}).$$

Hence, if T is invertible, then $I \in \operatorname{span}(T, T^2, \dots, T^{\dim V})$.

Conversely, suppose $I \in \operatorname{span}(T, T^2, \dots, T^{\dim V})$. Then there exist scalars $a_1, a_2, \dots, a_{\dim V}$ such that

$$I = a_1 T + a_2 T^2 + \cdots + a_{\dim V} T^{\dim V}.$$

Rearranging gives

$$0 = -I + a_1 T + a_2 T^2 + \cdots + a_{\dim V} T^{\dim V} = p(T),$$

where $p(z) = -1 + a_1 z + a_2 z^2 + \cdots + a_{\dim V} z^{\dim V}$. Since the minimal polynomial of T divides any polynomial that annihilates T, the minimal polynomial of T must also have a nonzero constant term. Therefore, 0 is not an eigenvalue of T, and hence T is invertible. ∎

Exercise 5B23

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. Let $W = \operatorname{span}(v, Tv, \ldots, T^{n-1}v)$. To show that W is invariant under T, we may show that $Tv, T^2v, \ldots, T^nv \in W$. Observe that $Tv, T(Tv) = T^2v, T(T^2v) = T^3v, \ldots, T(T^{n-2}v) = T^{n-1}v$ are all in W by definition. It remains to show that $T^nv \in W$.

Let $p(z)$ be the minimal polynomial of T. Let $m = \deg p$. Since $m \leq n$, we have

$$p(T)v = T^mv + a_{m-1}T^{m-1}v + \cdots + a_1Tv + a_0v = 0.$$

Rearranging gives

$$T^mv = -a_{m-1}T^{m-1}v - \cdots - a_1Tv - a_0v \in W.$$

If $m = n$, then we are done. If $m < n$, then we can apply T^{n-m} to both sides to get

$$T^nv = -a_{m-1}T^{n-1}v - \cdots - a_1T^{n-m+1}v - a_0T^{n-m}v \in W,$$

where we used the fact that $T^{n-m}v, T^{n-m+1}v, \ldots, T^{n-1}v \in W$ since $n - m < n$. Therefore, $T^nv \in W$, and hence W is invariant under T. ∎

Exercise 5B24

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 p be the minimal polynomial of T. The eigenvalues of T are 5 and 6, so each is a root of p. Hence,

$$p(z) = (z - 5)^{m_1}(z - 6)^{m_2}$$

where $m_1, m_2 \geq 1$. Because $\deg p \leq \dim V$, we have $m_1 + m_2 \leq \dim V$. Since $m_2 \geq 1$, we get $m_1 \leq \dim V - 1$, and similarly $m_2 \leq \dim V - 1$. This means $p(z)$ divides $(z - 5)^{\dim V - 1}(z - 6)^{\dim V - 1}$, so

$$(T - 5I)^{\dim V - 1}(T - 6I)^{\dim V - 1} = 0.$$ ∎

Exercise 5B25

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.

Axler Note. *The quotient operator T/U was defined in* Exercise 5A38.

Solution.

(a) Let $p(z)$ be the minimal polynomial of T and let $q(z)$ be the minimal polynomial of T/U. Recall that T/U is defined by $(T/U)(v+U) = Tv+U$ for all $v \in V$. Inductively, we have $(T/U)^k(v+U) = T^kv+U$ for all $k \geq 0$. Then we have

$$p(T/U)(v+U) = p(T)v+U = 0+U = U,$$

where we used the fact that $p(T) = 0$. Therefore, $p(T/U) = 0$, which implies that $q(z)$ divides $p(z)$. Hence, the minimal polynomial of T is a polynomial multiple of the minimal polynomial of the quotient operator T/U.

(b) Let $p(z)$ be the minimal polynomial of T, let $q(z)$ be the minimal polynomial of $T|_U$, and let $r(z)$ be the minimal polynomial of T/U. By definition,

$$r(T/U)(v+U) = r(T)v+U = U,$$

where we used the fact that $r(T/U) = 0$. Then $r(T)v \in U$ for all $v \in V$. Moreover, invariance of U under T implies that $q(T)u = q(T|_U)u = 0$ for all $u \in U$. Since $r(T)v \in U$, and $q(T) = q(T|_U)$ on U, we have

$$q(T)r(T)v = q(T|_U)r(T)v = 0$$

for all $v \in V$. Therefore, $q(T)r(T) = 0$, which implies $p(z)$ divides $q(z)r(z)$. Hence, the product of the minimal polynomials of $T|_U$ and T/U is a polynomial multiple of the minimal polynomial of T. ■

Exercise 5B26

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 λ be an eigenvalue of T with corresponding eigenvector $v \in V \setminus \{0\}$. Then we have

$$Tv = \lambda v.$$

If $v \in U$, then λ is an eigenvalue of $T|_U$. If $v \notin U$, then consider the coset $v+U \in V/U$. We have

$$(T/U)(v+U) = Tv+U = \lambda v+U = \lambda(v+U),$$

which shows that λ is an eigenvalue of T/U.

Conversely, suppose λ is an eigenvalue of $T|_U$ with corresponding eigenvector $u \in U \setminus \{0\}$. Then we have

$$T|_U u = Tu = \lambda u,$$

which shows that λ is also an eigenvalue of T.

Now suppose λ is an eigenvalue of T/U with corresponding eigenvector $v+U \in V/U \setminus \{0\}$. From Exercise 5B25 (a), we know that the minimal polynomial of T is a multiple of the minimal polynomial of T/U. Since λ is an eigenvalue of T/U, it must be a root of the minimal polynomial of T/U. Therefore, λ is also a root of the minimal polynomial of T, which implies that λ is an eigenvalue of T. ■

Exercise 5B27

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.

Axler Note. *The complexification $T_{\mathbf{C}}$ was defined in* Exercise 3B33.

Solution. Let $p(z)$ be the minimal polynomial of T. Then we have

$$p(T) = 0.$$

Consider the complexification $T_{\mathbf{C}}$ of T. By definition, $T_{\mathbf{C}}$ applied on $v + iw \in V_{\mathbf{C}}$ is given by

$$T_{\mathbf{C}}(v + iw) = T(v) + iT(w)$$

for all $v, w \in V$. By induction, we can extend the polynomial $p(z)$ to act on $V_{\mathbf{C}}$ by defining

$$p(T_{\mathbf{C}})(v + iw) = p(T)v + ip(T)w.$$

Since $p(T) = 0$, it follows that

$$p(T_{\mathbf{C}})(v + iw) = 0 + i0 = 0$$

for all $v, w \in V$. Therefore, $p(T_{\mathbf{C}}) = 0$, which shows that the minimal polynomial of $T_{\mathbf{C}}$ divides $p(z)$.

Conversely, let $q(z)$ be the minimal polynomial of $T_{\mathbf{C}}$. We can split $q(z)$ into its real and imaginary parts, say $q(z) = r(z) + is(z)$, where $r(z), s(z) \in \mathcal{P}(\mathbf{R})$. Then for any $v \in V$ which can be viewed as $v + i0 \in V_{\mathbf{C}}$, we have

$$q(T_{\mathbf{C}})(v + i0) = r(T)(v + i0) + is(T)(v + i0) = r(T)v + is(T)v.$$

Since $q(T_{\mathbf{C}}) = 0$, it follows that $r(T)v = 0$ and $s(T)v = 0$ for all $v \in V$. Therefore, both $r(T)$ and $s(T)$ must be 0, which implies that the minimal polynomial of T divides $r(z)$ and $s(z)$. Since $q(z) = r(z) + is(z)$, it follows that the minimal polynomial of T divides $q(z)$. Combining both directions, we conclude that the minimal polynomial of $T_{\mathbf{C}}$ equals the minimal polynomial of T. ∎

Exercise 5B28

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.

Axler Note. *The dual map T' was defined in Section 3F, 3.118.*

Solution. Let p be the minimal polynomial of T. For any $\varphi \in V'$, we have

$$p(T')(\varphi) = \varphi \circ p(T) = \varphi \circ 0 = 0 \implies p(T') = 0.$$

Hence, the minimal polynomial of T' divides p.

Conversely, let q be the minimal polynomial of T', so $q(T') = 0$. Then for all $\varphi \in V'$ and all $v \in V$,

$$0 = \left(q(T')\varphi\right)(v) = \varphi\left(q(T)v\right).$$

Thus, for each fixed v, $\varphi(q(T)v) = 0$ for all $\varphi \in V'$. Since $\varphi(w) = 0$ for all $\varphi \in V'$ implies $w = 0$, this forces $q(T)v = 0$ for all v; i.e., $q(T) = 0$. Therefore, p divides q.

Both p and q divide each other and are monic, so they are equal. ∎

Exercise 5B29 ✔

Show that every operator on a finite-dimensional vector space of dimension at least two has an invariant subspace of dimension two.

Axler Note. Exercise 5C6 *will give an improvement of this result when* $\mathbf{F} = \mathbf{C}$.

Solution. Let V be a finite-dimensional vector space with $\dim V \geq 2$, and let $T \in \mathcal{L}(V)$. Let $p(z)$ be the minimal polynomial of T. Over $\mathbf{F}$, every nonconstant polynomial factors into linear and irreducible quadratic factors. We split the proof into two cases based on the factorization of $p(z)$.

(1) Suppose $p(z)$ has a linear factor $z - \lambda$ for some $\lambda \in \mathbf{F}$. If $\deg p(z) = 1$, then $p(z) = z - \lambda$, which implies $T - \lambda I = 0$, or simply $T = \lambda I$. Since T is just a scalar multiple of the identity, every single subspace of V is invariant under T. Because $\dim V \geq 2$, we can choose any two linearly independent vectors $v, w \in V$, and $\operatorname{span}(v, w)$ will form a 2-dimensional invariant subspace.

If $\deg p(z) \geq 2$, then $p(z)/(z - \lambda)$ has degree at least 1 and must itself contain either a linear factor or an irreducible quadratic factor. The latter is handled by Case 2 below, so we consider the two linear cases:

- Repeated linear factor: $p(z) = (z - \lambda)^2 q(z)$. Since $p(z)$ is the minimal polynomial, there exists $v \in V$ such that $(T - \lambda I)q(T)v \neq 0$ but $(T - \lambda I)^2 q(T)v = 0$. Let $w = q(T)v$ and $u = (T - \lambda I)w$. Then $u \neq 0$ and $(T - \lambda I)u = (T - \lambda I)^2 w = 0$, which implies $Tu = \lambda u$. Thus, u is an eigenvector corresponding to the eigenvalue λ. Moreover, $Tw = (T - \lambda I)w + \lambda w = u + \lambda w$, which shows that $Tw \in \operatorname{span}(u, w)$.

  To confirm linear independence, suppose $u = \alpha w$ for some $\alpha \in \mathbf{F}$. Then $(T - \lambda I)w = \alpha w$, so w is an eigenvector of T with eigenvalue $\lambda + \alpha$. But then $(T - \lambda I)^2 w = \alpha^2 w$. Since $(T - \lambda I)^2 w = 0$ and $w \neq 0$, this forces $\alpha = 0$, contradicting our deduction that $\alpha \neq 0$ (since $u \neq 0$). Therefore, u and w are linearly independent, and $\operatorname{span}(u, w)$ is a 2-dimensional invariant subspace of V under T.
- Distinct linear factor: $p(z) = (z - \lambda)(z - \mu)g(z)$ with $\lambda \neq \mu$. Then T has two distinct eigenvalues, yielding two linearly independent eigenvectors v and u. The subspace $\operatorname{span}(v, u)$ is 2-dimensional, and since $Tv = \lambda v \in \operatorname{span}(v, u)$ and $Tu = \mu u \in \operatorname{span}(v, u)$, it is explicitly invariant.

(2) Suppose $p(z)$ has an irreducible quadratic factor, say $z^2 + az + b$ for some $a, b \in \mathbf{F}$. This means we can write $p(z) = (z^2 + az + b)q(z)$ for some polynomial $q(z)$. Since $p(z)$ is the minimal polynomial, $q(T)$ is not the zero operator, meaning there exists some vector $x \in V$ such that $q(T)x \neq 0$.

Let $u_1 = q(T)x$. Evaluating the minimal polynomial identity gives:

$$0 = p(T)x = (T^2 + aT + bI)q(T)x = (T^2 + aT + bI)u_1.$$

Rearranging this provides the operator relation $T^2 u_1 = -bu_1 - au_2$, where we define $u_2 = Tu_1$.

The vectors u_1 and u_2 must be linearly independent. If $Tu_1 = cu_1$ for some $c \in \mathbf{F}$, then $(T^2 + aT + bI)u_1 = (c^2 + ac + b)u_1 = 0$. Since $u_1 \neq 0$, this implies $c^2 + ac + b = 0$, meaning c is a root of $z^2 + az + b$ over $\mathbf{F}$, contradicting its irreducibility.

Let $U = \operatorname{span}(u_1, u_2)$, which has dimension two. Checking the images of our basis elements under T yields:

$$T(u_1) = u_2 \in U$$
$$T(u_2) = T^2 u_1 = -bu_1 - au_2 \in U.$$

Since T maps the generating elements of U back into U, the subspace U is a 2-dimensional invariant subspace of V under T.

■

5C Upper-Triangular Matrices

Exercise 5C1

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.

Manual Remark. *The matrix in this exercise represents the* $90°$ *counterclockwise rotation of* $\mathbf{R}^2$.

Solution. We claim that the statement is false. Consider the linear operator $T \in \mathcal{L}(\mathbf{R}^2)$ with matrix representation

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

with respect to the standard basis. Then

$$A^2 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}^2 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix},$$

which is an upper-triangular matrix with respect to the standard basis of $\mathbf{R}^2$. However, T does not have an upper-triangular matrix with respect to any basis of $\mathbf{R}^2$, since it has no real eigenvalues and thus cannot be represented in upper-triangular form over $\mathbf{R}$. ∎

Exercise 5C2

Suppose A and B are upper-triangular matrices of the same size, with $\alpha_1, \dots, \alpha_n$ on the diagonal of A and $\beta_1, \dots, \beta_n$ on the diagonal of B.

(a) Show that $A + B$ is an upper-triangular matrix with $\alpha_1 + \beta_1, \dots, \alpha_n + \beta_n$ on the diagonal.

(b) Show that AB is an upper-triangular matrix with $\alpha_1\beta_1, \dots, \alpha_n\beta_n$ on the diagonal.

Axler Note. *The results in this exercise are used in the proof of 5.81.*

Solution.

(a) Let $C = A + B$. Since A and B are upper-triangular, for any entry c_{ij} of C where $i > j$, we have

$$c_{ij} = a_{ij} + b_{ij} = 0 + 0 = 0,$$

which shows that C is upper-triangular. The diagonal entries of C are given by

$$c_{ii} = a_{ii} + b_{ii} = \alpha_i + \beta_i,$$

for $i = 1, \dots, n$.

(b) Let $D = AB$. Since A and B are upper-triangular, for any entry d_{ij} of D where $i > j$, we have the following sum for the (i, j)-th entry of D:

$$d_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}.$$

Consider a_{ik} when $k < i$. Since A is upper-triangular, $a_{ik} = 0$. Now consider b_{kj} when $k > j$. Since B is upper-triangular, $b_{kj} = 0$. Therefore, for $i > j$, all terms in the sum are zero, and we have $d_{ij} = 0$ for $i > j$, showing that D is upper-triangular.

To see that the diagonal entries of D are $\alpha_1\beta_1, \dots, \alpha_n\beta_n$, we compute the diagonal entries:

$$d_{ii} = \sum_{k=1}^{n} a_{ik} b_{ki},$$

for the i-th diagonal entry. Since A and B are upper-triangular, then $k < i$ implies $a_{ik} = 0$ and $k > i$ implies $b_{ki} = 0$. Therefore, the only nonzero term in the sum occurs when $k = i$, giving

$$d_{ii} = a_{ii}b_{ii} = \alpha_i\beta_i,$$

for $i = 1, \ldots, n$. ■

Exercise 5C3 ✔

Suppose $T \in \mathcal{L}(V)$ is invertible and $v_1, \ldots, v_n$ is a basis of V with respect to which the matrix of T is upper triangular, with $\lambda_1, \ldots, \lambda_n$ on the diagonal. Show that the matrix of T^{-1} is also upper triangular with respect to the basis $v_1, \ldots, v_n$, with

$$\frac{1}{\lambda_1}, \ldots, \frac{1}{\lambda_n}$$

on the diagonal.

Solution. Let $V_k = \operatorname{span}(v_1, \ldots, v_k)$ for each $k \in \{1, \ldots, n\}$. Because $\mathcal{M}(T, (v_1, \ldots, v_n))$ is upper-triangular, $T(V_k) \subseteq V_k$ for every k; thus each V_k is invariant under T. Since T is invertible, its restriction to the (finite-dimensional) V_k is injective, hence surjective onto V_k. Therefore, $T(V_k) = V_k$. Applying T^{-1} yields $V_k = T^{-1}(V_k)$, meaning each V_k is also invariant under T^{-1}. So, $\mathcal{M}(T^{-1}, (v_1, \ldots, v_n))$ is upper-triangular.

Let $\mu_1, \ldots, \mu_n$ be the diagonal entries of $\mathcal{M}(T^{-1}, (v_1, \ldots, v_n))$. Since $TT^{-1} = I$ is upper-triangular, we have from Exercise 5C2 (b) that $\lambda_k\mu_k = 1$ for each k, giving $\mu_k = \frac{1}{\lambda_k}$. ■

Exercise 5C4

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.

Axler Note. *This exercise and the exercise below show that 5.41 fails without the hypothesis that an upper-triangular matrix is under consideration.*

Solution. Consider the linear operator $T \in \mathcal{L}(\mathbf{R}^2)$ with matrix representation

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

with respect to the standard basis. A is invertible with inverse

$$A^{-1} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$ ■

Exercise 5C5

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. Consider the linear operator $T \in \mathcal{L}(\mathbf{R}^2)$ with matrix representation

$$A = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}$$

with respect to the standard basis. Since $T(1, -1) = (0, 0)$, T is not invertible. ■

Exercise 5C6

Suppose $\mathbf{F} = \mathbf{C}$, V is finite-dimensional, and $T \in \mathcal{L}(V)$. Prove that if $k \in \{1, \dots, \dim V\}$, then V has a k-dimensional subspace invariant under T.

Solution. Recall that every operator on a finite-dimensional complex vector space has an upper-triangular matrix with respect to some basis. Let $v_1, \dots, v_n$ be such a basis. Then, for each $k \in \{1, \dots, n\}$,

$$Tv_j \in \operatorname{span}(v_1, \dots, v_j) \subseteq \operatorname{span}(v_1, \dots, v_k)$$

for all $j \leq k$. Hence, $\operatorname{span}(v_1, \dots, v_k)$ is a k-dimensional subspace that is invariant under T. ∎

Exercise 5C7

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 $p(T)v = 0$ for any polynomial p, so we can take $p_v(x) = 1$. Otherwise, suppose $v \neq 0$. Consider the list $v, Tv, T^2v, \dots, T^{\dim V}v$. Since V is finite-dimensional, this list is linearly dependent. Then there exist scalars $a_0, \dots, a_m$ not all zero such that

$$a_0 v + a_1 Tv + \cdots + a_m T^m v = 0.$$

Here, m denotes the smallest integer such that there exist scalars $a_0, \dots, a_m$ not all zero satisfying the above equation. Since m is minimal, we have $a_m \neq 0$. Let $p_v(x)$ be the monic polynomial after dividing by a_m:

$$p_v(x) = x^m + \frac{a_{m-1}}{a_m} x^{m-1} + \cdots + \frac{a_0}{a_m}.$$

Then $p_v(T)v = 0$. Suppose q is another monic polynomial of smallest degree such that $q(T)v = 0$. Then $p_v - q$ is a polynomial of smaller degree than p_v such that $(p_v - q)(T)v = 0$, which contradicts the minimality of p_v.

(b) Let m_T be the minimal polynomial of T. Then $m_T(T) = 0$, so $m_T(T)v = 0$. By the division algorithm for polynomials, there exist unique polynomials q_0 and r such that

$$m_T = p_v q_0 + r,$$

where $r = 0$ or $\deg r < \deg p_v$. Then

$$0 = m_T(T)v = (p_v q_0 + r)(T)v = p_v(T)q_0(T)v + r(T)v = q_0(T)p_v(T)v + r(T)v = r(T)v,$$

since $p_v(T)v = 0$. If $r \neq 0$, then r is a polynomial of smaller degree than p_v such that $r(T)v = 0$, which contradicts the minimality of p_v. Therefore, $r = 0$, and we have $m_T = p_v q_0$, which shows that m_T is a polynomial multiple of p_v. ∎

Exercise 5C8

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) Suppose there exists a basis of V with respect to which T has an upper-triangular matrix. Then the diagonal entries of this matrix are the eigenvalues of T. Let $p(z) = z^2 + 2z + 2$. Then $p(T)$ also has an upper-triangular matrix with respect to the same basis, and the diagonal entries of this matrix are $p(\lambda)$, where λ is an eigenvalue of T. Since $p(T)v = 0$ for some nonzero vector v, $p(T)$ is not injective. Then $p(T)$ is not invertible, so $p(\lambda) = 0$ for some eigenvalue $\lambda \in \mathbf{R}$ of T. However, $p(z)$ has no real roots, contradicting that $p(\lambda) = 0$ for some real eigenvalue λ.

(b) Utilizing the same deductions as in part (a), we conclude that there exists an eigenvalue λ of T such that $p(\lambda) = 0$. The roots of $p(z)$ are $-1 + i$ and $-1 - i$, so either $-1 + i$ or $-1 - i$ must be an eigenvalue of T. Since the diagonal entries of an upper-triangular matrix are the eigenvalues of the corresponding operator, either $-1 + i$ or $-1 - i$ must appear on the diagonal of A. ■

Exercise 5C9

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. Since B represents a linear operator on a finite-dimensional complex vector space, let T be the linear operator corresponding to B, i.e., $B = \mathcal{M}(T, (e_1, \ldots, e_n))$, where $(e_1, \ldots, e_n)$ is the standard basis of $\mathbf{C}^n$. Since T has an upper-triangular matrix with respect to some basis of the vector space, let A be the change of basis matrix from the new basis to the standard basis. Then $A^{-1}BA$ is the matrix representation of T with respect to this new basis, which is upper-triangular. ■

Exercise 5C10

Suppose $T \in \mathcal{L}(V)$ and $v_1, \ldots, v_n$ is a basis of V. Show that the following are equivalent.

(a) The matrix of T with respect to $v_1, \ldots, v_n$ is lower triangular.

(b) $\operatorname{span}(v_k, \ldots, v_n)$ is invariant under T for each $k = 1, \ldots, n$.

(c) $Tv_k \in \operatorname{span}(v_k, \ldots, v_n)$ for each $k = 1, \ldots, n$.

Axler Note. *A square matrix is called* ***lower triangular*** *if all entries above the diagonal are* 0.

Solution.

- Suppose the matrix of T with respect to $v_1, \ldots, v_n$ is lower triangular. Fix some $k \in \{1, \ldots, n\}$. For $j \in \{1, \ldots, n\}$,

$$Tv_j \in \operatorname{span}(v_j, \ldots, v_n)$$

because the matrix is lower triangular. If $j \geq k$, then $\operatorname{span}(v_j, \ldots, v_n) \subseteq \operatorname{span}(v_k, \ldots, v_n)$. Therefore, $Tv_j \in \operatorname{span}(v_k, \ldots, v_n)$ for all $j \geq k$, which shows that $\operatorname{span}(v_k, \ldots, v_n)$ is invariant under T.

- Suppose $\operatorname{span}(v_k, \ldots, v_n)$ is invariant under T for each $k = 1, \ldots, n$. Then for each k, we have $Tv_k \in \operatorname{span}(v_k, \ldots, v_n)$, which shows that Tv_k can be expressed as a linear combination of $v_k, \ldots, v_n$.

- Suppose $Tv_k \in \operatorname{span}(v_k, \dots, v_n)$ for each $k = 1, \dots, n$. Then

$$Tv_k = \sum_{j=k}^{n} A_{j,k} v_j$$

for some scalars $A_{j,k}$. Then $A_{i,k} = 0$ for all $i < k$, which shows that the matrix of T with respect to $v_1, \dots, v_n$ is lower triangular. ■

Exercise 5C11

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. Since T has an upper-triangular matrix with respect to some basis of V, let $v_1, \dots, v_n$ be such a basis. Then $\operatorname{span}(v_1, \dots, v_k)$ is invariant under T for each $k = 1, \dots, n$. Consider the reverse basis $w_1 = v_n, w_2 = v_{n-1}, \dots, w_n = v_1$. Then $\operatorname{span}(w_k, \dots, w_n) = \operatorname{span}(v_1, \dots, v_{n-k+1})$ is invariant under T for each $k = 1, \dots, n$. By Exercise 5C10, the matrix of T with respect to the basis $w_1, \dots, w_n$ is lower triangular. ■

Exercise 5C12

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.

Axler Note. *The quotient operator T/U was defined in* Exercise 5A38.

Solution.

(a) Since T has an upper-triangular matrix with respect to some basis of V, then its minimal polynomial is a product of linear factors. Since U is invariant under T, the minimal polynomial of $T|_U$ divides the minimal polynomial of T, and thus is also a product of linear factors. Therefore, $T|_U$ has an upper-triangular matrix with respect to some basis of U.

(b) Since the minimal polynomial of T is a polynomial multiple of the minimal polynomial of T/U, the same argument as in part (a) shows that T/U has an upper-triangular matrix with respect to some basis of V/U. ■

Exercise 5C13

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 $u_1, \dots, u_m$ be a basis of U with respect to which $T|_U$ has an upper-triangular matrix. Let $v_1 + U, \dots, v_n + U$ be a basis of V/U with respect to which T/U has an upper-triangular matrix. Then the list $u_1, \dots, u_m, v_1, \dots, v_n$ is a basis of V. Since $T|_U$ has an upper-triangular matrix with respect to the basis $u_1, \dots, u_m$, we have

$$Tu_j = \sum_{k=1}^{j} A_{k,j} u_k$$

for some scalars $A_{k,j}$ and for each $j = 1, \dots, m$. Since T/U has an upper-triangular matrix with respect to the basis $v_1 + U, \dots, v_n + U$, we have

$$(T/U)(v_j + U) = \sum_{k=1}^{j} B_{k,j}(v_k + U)$$

for some scalars $B_{k,j}$ and for each $j = 1, \dots, n$. This implies that

$$Tv_j - \sum_{k=1}^{j} B_{k,j} v_k \in U,$$

so there exist scalars $C_{l,j}$ such that

$$Tv_j - \sum_{k=1}^{j} B_{k,j} v_k = \sum_{l=1}^{m} C_{l,j} u_l.$$

Therefore,

$$Tv_j = \sum_{l=1}^{m} C_{l,j} u_l + \sum_{k=1}^{j} B_{k,j} v_k,$$

which shows that the matrix of T with respect to the basis $u_1, \dots, u_m, v_1, \dots, v_n$ is upper-triangular. ■

Exercise 5C14

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'.

Manual Remark. *This can be argued by noting that $\mathcal{M}(T^t) = \mathcal{M}(T)^t$, so the matrix of T' with respect to the dual basis is lower-triangular. Then the reversed dual basis gives an upper-triangular matrix for T'.*

Solution. By Exercise 5B28, the minimal polynomials of T and T' are the same. Therefore, the following are equivalent:

- T has an upper-triangular matrix with respect to some basis of V.
- The minimal polynomial of T is a product of linear factors.
- The minimal polynomial of T' is a product of linear factors.
- T' has an upper-triangular matrix with respect to some basis of V'. ■

5D Diagonalizable Operators

Exercise 5D1

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.

Manual Remark. *The roots ω in part (b) are the roots obtained in* Exercise 1C6.

Solution.

(a) If $T^4 = I$, then the minimal polynomial of T divides $z^4 - 1 = (z-1)(z+1)(z-i)(z+i)$, which has no repeated roots. Hence, the minimal polynomial of T has no repeated roots, so T is diagonalizable.

(b) If $T^4 = T$, then the minimal polynomial of T divides $z^4 - z = z(z-1)(z-\omega)(z-\omega^2)$. This polynomial has no repeated roots, so the minimal polynomial of T has no repeated roots, and hence T is diagonalizable.

(c) Let $T \in \mathcal{L}(\mathbf{C}^2)$ be defined by

$$T(z_1, z_2) = (z_2, 0).$$

Then $T^2(z_1, z_2) = (0, 0)$, so $T^4 = T^2$. However, T is not diagonalizable because the only eigenvalue of T is 0, and $\dim E(0, T) = 1 < 2 = \dim \mathbf{C}^2$. ■

Exercise 5D2

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 A is a diagonal matrix of T with respect to some basis of V. Let $\lambda \in \mathbf{F}$. Let m denote the number of times λ appears on the diagonal of A. Then $A - \lambda I$ is a diagonal matrix of $T - \lambda I$ with respect to the same basis, with exactly m zeros on the diagonal. A diagonal matrix with m zeros on the diagonal has null space of dimension m, so $\dim \operatorname{null}(T - \lambda I) = m$. Since $E(\lambda, T) = \operatorname{null}(T - \lambda I)$, we have $\dim E(\lambda, T) = m$. ■

Exercise 5D3

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. Suppose T is diagonalizable. Let $\lambda_1, \dots, \lambda_m$ denote the distinct eigenvalues of T. Then

$$V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T).$$

If 0 is not an eigenvalue of T, then $\operatorname{null} T = \{0\}$ and $\operatorname{range} T = V$, so $V = \operatorname{null} T \oplus \operatorname{range} T$. If 0 is an eigenvalue of T, then we may assume $\lambda_1 = 0$. Let

$$U = E(\lambda_2, T) \oplus \cdots \oplus E(\lambda_m, T).$$

Since $\operatorname{null} T = E(0, T)$, we have $V = \operatorname{null} T \oplus U$. It remains to show that $U = \operatorname{range} T$.

Let $u \in U$. Then $u = u_2 + \cdots + u_m$ for some $u_k \in E(\lambda_k, T)$. Since $\lambda_k \neq 0$ for $k = 2, \dots, m$, we have $u_k = \lambda_k^{-1} T(u_k)$ for each $k = 2, \dots, m$. Hence,

$$u = u_2 + \cdots + u_m = \lambda_2^{-1} T(u_2) + \cdots + \lambda_m^{-1} T(u_m) = T(\lambda_2^{-1} u_2 + \cdots + \lambda_m^{-1} u_m) \in \operatorname{range} T.$$

Thus, $U \subseteq \operatorname{range} T$.

Now let $v \in \operatorname{range} T$. Note that eigenspaces are invariant as follows: If $u \in E(\lambda_k, T)$, then $T(u) = \lambda_k u \in E(\lambda_k, T)$. Since U is a direct sum of eigenspaces, it is invariant under T. Then $v = T(w)$ for some $w \in V$. Since $V = \operatorname{null} T \oplus U$, we may write $w = w_1 + w_2$ for some $w_1 \in \operatorname{null} T$ and $w_2 \in U$. Then

$$v = T(w) = T(w_1 + w_2) = T(w_1) + T(w_2) = 0 + T(w_2) = T(w_2) \in U.$$

Thus, $\operatorname{range} T \subseteq U$. Since $U \subseteq \operatorname{range} T$ and $\operatorname{range} T \subseteq U$, we have $U = \operatorname{range} T$. Therefore, $V = \operatorname{null} T \oplus U = \operatorname{null} T \oplus \operatorname{range} T$. ∎

Exercise 5D4

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.

(a) $\implies$ (b): If $V = \operatorname{null} T \oplus \operatorname{range} T$, then $V = \operatorname{null} T + \operatorname{range} T$.

(b) $\implies$ (c): If $V = \operatorname{null} T + \operatorname{range} T$, then $\dim V = \dim \operatorname{null} T + \dim \operatorname{range} T - \dim(\operatorname{null} T \cap \operatorname{range} T)$. Since $\dim V = \dim \operatorname{null} T + \dim \operatorname{range} T$ by the fundamental theorem of linear maps, we have $\dim(\operatorname{null} T \cap \operatorname{range} T) = 0$, so $\operatorname{null} T \cap \operatorname{range} T = \{0\}$.

(c) $\implies$ (a): If $\operatorname{null} T \cap \operatorname{range} T = \{0\}$, then

$$\begin{aligned}\dim(\operatorname{null} T + \operatorname{range} T) &= \dim \operatorname{null} T + \dim \operatorname{range} T - \dim(\operatorname{null} T \cap \operatorname{range} T) \\ &= \dim \operatorname{null} T + \dim \operatorname{range} T.\end{aligned}$$

By the fundamental theorem of linear maps, we have $\dim V = \dim \operatorname{null} T + \dim \operatorname{range} T$. Hence,

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

Since $\operatorname{null} T + \operatorname{range} T \subseteq V$ and $\dim(\operatorname{null} T + \operatorname{range} T) = \dim V$, we have $\operatorname{null} T + \operatorname{range} T = V$. Along with the assumption that $\operatorname{null} T \cap \operatorname{range} T = \{0\}$, we have $V = \operatorname{null} T \oplus \operatorname{range} T$. ∎

Exercise 5D5

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 T is diagonalizable. Let $\lambda \in \mathbf{C}$. Since T is diagonalizable, we have

$$V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)$$

for some distinct eigenvalues $\lambda_1, \ldots, \lambda_m$ of T. If λ is not an eigenvalue of T, then $\operatorname{null}(T - \lambda I) = \{0\}$ and $\operatorname{range}(T - \lambda I) = V$, so

$$V = \operatorname{null}(T - \lambda I) \oplus \operatorname{range}(T - \lambda I).$$

If λ is an eigenvalue of T, then we may assume $\lambda_1 = \lambda$. Let

$$U = E(\lambda_2, T) \oplus \cdots \oplus E(\lambda_m, T).$$

Then $V = E(\lambda, T) \oplus U$. It remains to show that $U = \text{range}(T - \lambda I)$.

By dimension counting, we have

$$\dim U = \dim V - \dim E(\lambda, T) = \dim V - \dim \text{null}(T - \lambda I) = \dim \text{range}(T - \lambda I).$$

Since each $u_k \in E(\lambda_k, T)$ satisfies $T(u_k) = \lambda_k u_k$, we have

$$(T - \lambda I)(u_k) = (\lambda_k - \lambda)u_k \in E(\lambda_k, T).$$

To see that $T - \lambda I$ is injective on U, suppose $u \in U$ and $(T - \lambda I)(u) = 0$. Then $u = u_2 + \cdots + u_m$ for some $u_k \in E(\lambda_k, T)$, and

$$\begin{aligned} 0 &= (T - \lambda I)(u) \\ &= (T - \lambda I)(u_2 + \cdots + u_m) \\ &= (T - \lambda I)(u_2) + \cdots + (T - \lambda I)(u_m) \\ &= (\lambda_2 - \lambda)u_2 + \cdots + (\lambda_m - \lambda)u_m. \end{aligned}$$

Since $\lambda_2, \ldots, \lambda_m$ are distinct from λ, we have $\lambda_k - \lambda \neq 0$ for each $k = 2, \ldots, m$. Since the sum above is a direct sum, we have $(\lambda_k - \lambda)u_k = 0$ for each $k = 2, \ldots, m$, so $u_k = 0$ for each $k = 2, \ldots, m$. Hence, $u = u_2 + \cdots + u_m = 0$, so $T - \lambda I$ is injective on U. Since $\dim U = \dim \text{range}(T - \lambda I)$ and U is invariant under $T - \lambda I$, we have $U = \text{range}(T - \lambda I)$.

Conversely, suppose $V = \text{null}(T - \lambda I) \oplus \text{range}(T - \lambda I)$ for every $\lambda \in \mathbf{C}$. Then $\text{null}(T - \lambda I) \cap \text{range}(T - \lambda I) = \{0\}$ for every $\lambda \in \mathbf{C}$. Note that $T - \lambda I$ restricted to $\text{range}(T - \lambda I)$ is injective, so λ is not an eigenvalue of $T|_{\text{range}(T-\lambda I)}$. Then the minimal polynomial of $T|_{\text{range}(T-\lambda I)}$ does not have λ as a root. Suppose $(z - \lambda)^2$ divides the minimal polynomial of T. Then there is some $v \in V$ such that $(T - \lambda I)^2 v = 0$, but $(T - \lambda I)v \neq 0$. Then $(T - \lambda I)v \in \text{range}(T - \lambda I)$, and $v \in \text{null}(T - \lambda I) \cap \text{range}(T - \lambda I) = \{0\}$, which is a contradiction. Hence, $(z - \lambda)^2$ does not divide the minimal polynomial of T for any $\lambda \in \mathbf{C}$, so the minimal polynomial of T has no repeated roots. Since the minimal polynomial of T has no repeated roots, T is diagonalizable. ■

Exercise 5D6

Suppose $T \in \mathcal{L}(\mathbf{F}^5)$ and $\dim E(8, T) = 4$. Prove that $T - 2I$ or $T - 6I$ is invertible.

Solution. Suppose $T - 2I$ and $T - 6I$ are not invertible. Then 2 and 6 are eigenvalues of T, so $\dim E(2, T) \geq 1$ and $\dim E(6, T) \geq 1$. Since 8 is also an eigenvalue of T, we have

$$\dim E(2, T) + \dim E(6, T) + \dim E(8, T) \geq 1 + 1 + 4 = 6 > 5 = \dim \mathbf{F}^5,$$

which is a contradiction. Hence, $T - 2I$ or $T - 6I$ is invertible. ■

Exercise 5D7

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. Suppose $v \in E(\lambda, T)$. Then $T(v) = \lambda v$. Since T is invertible, we have $v = T^{-1}(\lambda v) = \lambda T^{-1}(v)$, so $T^{-1}(v) = \frac{1}{\lambda} v$. Hence,

$$v \in E(\lambda^{-1}, T^{-1}),$$

so $E(\lambda, T) \subseteq E(\lambda^{-1}, T^{-1})$.

Now suppose $v \in E(\lambda^{-1}, T^{-1})$. Then

$$T^{-1}(v) = \lambda^{-1} v \implies v = T(\lambda^{-1} v) \implies v = \lambda^{-1} T(v) \implies T(v) = \lambda v.$$

Hence, $v \in E(\lambda, T)$, so $E(\lambda^{-1}, T^{-1}) \subseteq E(\lambda, T)$. Then $E(\lambda, T) = E(\lambda^{-1}, T^{-1})$. ∎

Exercise 5D8

Suppose V is finite-dimensional and $T \in \mathcal{L}(V)$. Let $\lambda_1, \dots, \lambda_m$ 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. For every nonzero eigenvalue λ of T and every $v \in E(\lambda, T)$, we have $v = T(\lambda^{-1} v) \in \operatorname{range} T$, so $E(\lambda, T) \subseteq \operatorname{range} T$. Since eigenspaces corresponding to distinct eigenvalues are linearly independent, we have

$$E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T) \subseteq \operatorname{range} T.$$

It follows that

$$\dim E(\lambda_1, T) + \cdots + \dim E(\lambda_m, T) = \dim(E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)) \leq \dim \operatorname{range} T.$$ ∎

Exercise 5D9

Suppose $R, T \in \mathcal{L}(\mathbf{F}^3)$ each have $2, 6, 7$ as eigenvalues. Prove that there exists an invertible operator $S \in \mathcal{L}(\mathbf{F}^3)$ such that $R = S^{-1} T S$.

Solution. Since R and T each have 3 distinct eigenvalues, they are both diagonalizable. Let r_2, r_6, r_7 be eigenvectors of R with $R r_k = k r_k$ for $k = 2, 6, 7$, and let t_2, t_6, t_7 be eigenvectors of T with $T t_k = k t_k$ for $k = 2, 6, 7$. Each list of 3 eigenvectors is linearly independent, so they each form a basis of $\mathbf{F}^3$. Define $S \in \mathcal{L}(\mathbf{F}^3)$ by $S(r_k) = t_k$ for $k = 2, 6, 7$. Then S is invertible because it maps a basis of $\mathbf{F}^3$ to a basis of $\mathbf{F}^3$. For each $k = 2, 6, 7$, we have

$$S^{-1} T S(r_k) = S^{-1} T(t_k) = S^{-1}(k t_k) = k S^{-1}(t_k) = k r_k = R(r_k).$$

Since $S^{-1} T S$ and R agree on a basis of $\mathbf{F}^3$, we have $R = S^{-1} T S$. ∎

Exercise 5D10

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 $R = S^{-1} T S$.

Solution. Define R to be the operator whose matrix with respect to the standard basis of $\mathbf{F}^4$ has diagonal entries $2, 6, 7, 7$ and all other entries 0. Then R has eigenvalues $2, 6, 7$ and no other eigenvalues. Define T to be the operator whose matrix with respect to the standard basis of $\mathbf{F}^4$ has diagonal entries $2, 6, 7, 2$ and all other entries 0. Then T has eigenvalues $2, 6, 7$ and no other eigenvalues. If there did exist an invertible operator $S \in \mathcal{L}(\mathbf{F}^4)$ such that $R = S^{-1} T S$, then

$$R v = S^{-1} T S v = \lambda v \implies T(Sv) = \lambda S v$$

for every eigenvalue λ of R and every eigenvector $v \in E(\lambda, R)$. Hence, the map $S|_{E(\lambda,R)}$ maps $E(\lambda, R)$ to $E(\lambda, T)$. Since S is invertible, $S|_{E(\lambda,R)}$ is injective, so $\dim E(\lambda, R) \leq \dim E(\lambda, T)$. However, $\dim E(7, R) = 2 > 1 = \dim E(7, T)$, so there does not exist an invertible operator $S \in \mathcal{L}(\mathbf{F}^4)$ such that $R = S^{-1}TS$. ■

Exercise 5D11

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. Consider the operator $T \in \mathcal{L}(\mathbf{C}^3)$ whose matrix with respect to the standard basis of $\mathbf{C}^3$ is

$$\begin{pmatrix} 6 & 1 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 7 \end{pmatrix}.$$

Observe that 6 and 7 are the only eigenvalues of T since the matrix is upper triangular with diagonal entries 6, 6, 7. If T had a diagonal matrix with respect to some basis of $\mathbf{C}^3$, then $\dim E(6, T) + \dim E(7, T) = 3$. However, $E(6, T)$ is the null space of $T - 6I$, which has matrix

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Observe that the null space of this matrix is spanned by $(1, 0, 0)$, so $\dim E(6, T) = 1$. Similarly, $E(7, T)$ is the null space of $T - 7I$, which has matrix

$$\begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Observe that the null space of this matrix is spanned by $(0, 0, 1)$, so $\dim E(7, T) = 1$. Hence, $\dim E(6, T) + \dim E(7, T) = 1 + 1 = 2 < 3$, so T does not have a diagonal matrix with respect to any basis of $\mathbf{C}^3$. ■

Exercise 5D12

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 T cannot be diagonalized, it cannot have 3 distinct eigenvalues. Then 6 and 7 must be the only eigenvalues of T. In particular, 8 is not an eigenvalue of T, so $T - 8I$ is surjective. Then there exists $(z_1, z_2, z_3) \in \mathbf{C}^3$ such that

$$(T - 8I)(z_1, z_2, z_3) = (6, 7, 13) \implies T(z_1, z_2, z_3) = (6 + 8z_1, 7 + 8z_2, 13 + 8z_3).$$ ■

Exercise 5D13

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. Let A be a diagonal matrix with distinct entries on the diagonal, and let B be a matrix of the same size as A. Suppose $AB = BA$. Let $A_{j,j}$ denote the j-th diagonal entry of A, and let $B_{j,k}$ denote the entry of B in row j and column k. Then

$$(AB)_{j,k} = A_{j,j}B_{j,k}, \quad (BA)_{j,k} = B_{j,k}A_{k,k}.$$

Since $AB = BA$, we have

$$A_{j,j}B_{j,k} = B_{j,k}A_{k,k}.$$

If $j \neq k$, then $A_{j,j} - A_{k,k} \neq 0$ because the diagonal entries of A are distinct. Hence,

$$(A_{j,j} - A_{k,k})B_{j,k} = 0 \implies B_{j,k} = 0.$$

Thus, all off-diagonal entries of B are zero, so B is a diagonal matrix.

Conversely, if B is a diagonal matrix, then it is straightforward to verify that $AB = BA$. ■

Exercise 5D14

(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) Consider the operator $T \in \mathcal{L}(\mathbf{C}^2)$ defined by

$$T(z_1, z_2) = (z_2, 0).$$

Then $T^2(z_1, z_2) = (0, 0)$, so T^2 is diagonalizable. However, T is not diagonalizable because the only eigenvalue of T is 0, and $\dim E(0, T) = 1 < 2 = \dim \mathbf{C}^2$.

(b) Suppose T is diagonalizable. Then a basis of V consists of eigenvectors of T, and each eigenvector of T is also an eigenvector of T^k. Hence, T^k is diagonalizable.

Conversely, suppose T^k is diagonalizable. Let $\lambda_1, \ldots, \lambda_m$ denote the distinct eigenvalues of T^k. Then the minimal polynomial of T^k has no repeated roots, so we may write

$$p(z) = (z - \lambda_1) \cdots (z - \lambda_m).$$

Consider the polynomial

$$q(z) = (z^k - \lambda_1) \cdots (z^k - \lambda_m).$$

Note that $q(T) = p(T^k) = 0$, so the minimal polynomial of T divides q. Since T is invertible, T^k is also invertible. Then $\lambda_j \neq 0$ for each $j = 1, \ldots, m$, so q has no repeated roots. Hence, the minimal polynomial of T has no repeated roots, so T is diagonalizable. ■

Exercise 5D15

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.

Axler Note. *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.

(a) $\implies$ (b) Suppose T is diagonalizable. Then the minimal polynomial of T has no repeated roots, so there does not exist $\lambda \in \mathbf{C}$ such that p is a polynomial multiple of $(z - \lambda)^2$.

(b) $\implies$ (c) Suppose there does not exist $\lambda \in \mathbf{C}$ such that p is a polynomial multiple of $(z - \lambda)^2$. Then p has no repeated roots, so we may let

$$p(z) = (z - \lambda_1) \cdots (z - \lambda_m)$$

for some distinct $\lambda_1, \dots, \lambda_m \in \mathbf{C}$. Then evaluating the derivative p' at λ_k gives

$$p'(\lambda_k) = \prod_{j \neq k} (\lambda_k - \lambda_j) \neq 0$$

by the distinctness of $\lambda_1, \dots, \lambda_m$. Hence, p and p' have no zeros in common.

(c) $\implies$ (d) Suppose p and its derivative p' have no zeros in common. Then the greatest common divisor of p and p' has no zeros, so it is a constant polynomial.

(d) $\implies$ (a) Suppose the greatest common divisor of p and p' is the constant polynomial 1. If p had a repeated root $\lambda \in \mathbf{C}$, then p would be a polynomial multiple of $(z - \lambda)^2$, so p' would be a polynomial multiple of $(z - \lambda)$. Hence, p and p' would have a zero in common, which is a contradiction. Thus, p has no repeated roots, so T is diagonalizable. ∎

Exercise 5D16

Suppose that $T \in \mathcal{L}(V)$ is diagonalizable. Let $\lambda_1, \dots, \lambda_m$ denote the distinct eigenvalues of T. Prove that a subspace U of V is invariant under T if and only if there exist subspaces $U_1, \dots, U_m$ 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 U is invariant under T. Let $U_k = U \cap E(\lambda_k, T)$ for each $k = 1, \dots, m$. Then $U_k \subseteq E(\lambda_k, T)$ for each k. For each k, define the polynomial p_k by

$$p_k(z) = \prod_{j \neq k} \frac{z - \lambda_j}{\lambda_k - \lambda_j}.$$

Then $p_k(\lambda_j) = 0$ for $j \neq k$ and $p_k(\lambda_k) = 1$. Let $u \in U$ with

$$u = u_1 + \cdots + u_m, \quad u_k \in E(\lambda_k, T).$$

Then $p_k(T)(u) = u_k \in U$ for each k, so $u_k \in U_k$ for each k. Hence, $U = U_1 \oplus \cdots \oplus U_m$.

Conversely, suppose there exist subspaces $U_1, \dots, U_m$ of V such that $U_k \subseteq E(\lambda_k, T)$ for each k and $U = U_1 \oplus \cdots \oplus U_m$. Then for each $u \in U$, we have

$$u = u_1 + \cdots + u_m, \quad u_k \in U_k.$$

Then

$$T(u) = T(u_1 + \cdots + u_m) = T(u_1) + \cdots + T(u_m) = \lambda_1 u_1 + \cdots + \lambda_m u_m \in U,$$

so U is invariant under T. ∎

Exercise 5D17

Suppose V is finite-dimensional. Prove that $\mathcal{L}(V)$ has a basis consisting of diagonalizable operators.

Solution. Let $n = \dim V$. Let $\{v_1, \dots, v_n\}$ be a basis of V. For each $j, k \in \{1, \dots, n\}$, define $T_{j,k} \in \mathcal{L}(V)$ by

$$T_{j,k}(v_\ell) = \begin{cases} v_j & \text{if } \ell = k, \\ 0 & \text{if } \ell \neq k. \end{cases}$$

If $j = k$, let $D_j = T_{j,j}$. Then D_j is diagonalizable because $D_j(v_j) = v_j$ and $D_j(v_\ell) = 0$ for $\ell \neq j$. If $j \neq k$, let $N_{j,k} = D_j + T_{j,k}$. Then $N_{j,k}$ is diagonalizable because $N_{j,k}(v_j) = v_j$, $N_{j,k}(v_k) = v_j$, and $N_{j,k}(v_\ell) = 0$ for $\ell \neq j, k$. Then the eigenvalues of $N_{j,k}$ are 0 with corresponding eigenvector $v_j - v_k$, and 1 with $n - 2$ corresponding eigenvectors v_ℓ for $\ell \neq j, k$, so $N_{j,k}$ is diagonalizable.

Note that the set

$$\{D_1, \dots, D_n\} \cup \{N_{j,k} : j, k \in \{1, \dots, n\}, j \neq k\}$$

spans $\mathcal{L}(V)$. For any $T \in \mathcal{L}(V)$, we have

$$T = \sum_{j,k=1}^{n} \alpha_{j,k} T_{j,k} = \sum_{j \neq k} \alpha_{j,k}(N_{j,k} - D_j) + \sum_{j=1}^{n} \alpha_{j,j} D_j,$$

which shows that the set above spans $\mathcal{L}(V)$.

Moreover, note that $\dim \mathcal{L}(V) = n^2$ and the set above has $n + n(n-1) = n^2$ elements, so it is a basis of $\mathcal{L}(V)$ consisting of diagonalizable operators. ∎

Exercise 5D18

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.

Solution. Since T is diagonalizable, V has a basis $v_1, \dots, v_n$ of eigenvectors of T. Let $\lambda_1, \dots, \lambda_n$ be the corresponding eigenvalues of T. Note that $v_1 + U, \dots, v_n + U$ span V/U. Moreover, for each j, we have

$$(T/U)(v_j + U) = T(v_j) + U = \lambda_j v_j + U = \lambda_j (v_j + U),$$

so V/U is spanned by eigenvectors of T/U. We may reduce this spanning set to a basis of V/U consisting of eigenvectors of T/U, so T/U is diagonalizable. ∎

Exercise 5D19

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.

Axler Note. *See Exercise 5C13 for an analogous statement about upper-triangular matrices.*

Solution. This statement is false. Consider the operator $T \in \mathcal{L}(\mathbf{F}^2)$ whose matrix with respect to the standard basis of $\mathbf{F}^2$ is

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Let $U = \operatorname{span}\{(1,0)\}$. Then U is invariant under T since $T(1,0) = (0,0) \in U$. The operator $T|_U$ is diagonalizable because it is the zero operator on a one-dimensional space. The quotient space V/U is one-dimensional, so T/U is diagonalizable. However, T is not diagonalizable because the only eigenvalue of T is 0, and $\dim E(0,T) = 1 < 2 = \dim \mathbf{F}^2$. ■

Exercise 5D20

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. Suppose T is diagonalizable. Then there exists a basis of V consisting of eigenvectors of T. Let $\{v_1, \dots, v_n\}$ be a basis of V consisting of eigenvectors of T, and let $\lambda_1, \dots, \lambda_n$ be the corresponding eigenvalues. Let $\{\varphi_1, \dots, \varphi_n\}$ be the dual basis of V' corresponding to $\{v_1, \dots, v_n\}$. Then for each $k = 1, \dots, n$, we have

$$T'(\varphi_k)(v_j) = \varphi_k(T(v_j)) = \varphi_k(\lambda_j v_j) = \lambda_j \varphi_k(v_j) = \begin{cases} \lambda_k & j = k, \\ 0 & j \neq k, \end{cases}$$

so $T'(\varphi_k) = \lambda_k \varphi_k$. Hence, $\{\varphi_1, \dots, \varphi_n\}$ is a basis of V' consisting of eigenvectors of T', so T' is diagonalizable.

Conversely, suppose T' is diagonalizable. Then there exists a basis of V' consisting of eigenvectors of T'. Using the same reasoning as above, $(T')'$ is diagonalizable with the same eigenvalues as T'. From Exercise 3F32, we know that V is isomorphic to V'' with the map $\Lambda : V \to V''$ defined by

$$\Lambda(v)(\varphi) = \varphi(v)$$

for each $v \in V$ and $\varphi \in V'$. Then $(T')'(\Lambda(v)) = \Lambda(T(v))$ for each $v \in V$, so T has a basis of eigenvectors corresponding to the same eigenvalues as T', so T is diagonalizable. ■

Exercise 5D21

The **Fibonacci sequence** $F_0, F_1, F_2, \ldots$ is defined by

$$F_0 = 0, \quad F_1 = 1, \quad \text{and } F_n = F_{n-2} + F_{n-1} \text{ for } n \geq 2.$$

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

$$\frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n.$$

Axler Note. *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****.*

Manual Remark. *Part (e) uses the triangle inequality and creating "zeros" in the right places to show that the second term in the formula for F_n is less than $1/2$ in absolute value.*

Solution.

(a) We prove the statement by induction on n. For the base case $n = 0$, we have $T^0(0, 1) = (0, 1) = (F_0, F_1)$. Now suppose the statement holds for some nonnegative integer n. Then

$$\begin{aligned} T^{n+1}(0, 1) &= T(T^n(0, 1)) \\ &= T(F_n, F_{n+1}) \\ &= (F_{n+1}, F_n + F_{n+1}) \\ &= (F_{n+1}, F_{n+2}), \end{aligned}$$

so the statement holds for $n + 1$.

(b) Suppose $T(x, y) = \lambda(x, y)$ for some $\lambda \in \mathbf{R}$ and $(x, y) \in \mathbf{R}^2$ with $(x, y) \neq (0, 0)$. Then

$$(y, x + y) = (\lambda x, \lambda y),$$

so $y = \lambda x$ and $x + y = \lambda y$. We then have

$$x + \lambda x = \lambda^2 x \implies (1 + \lambda - \lambda^2)x = 0.$$

If $x = 0$, then $y = \lambda x = 0$, which is a contradiction. Hence, $x \neq 0$, so we have

$$\lambda^2 - \lambda - 1 = 0 \implies \lambda = \frac{1 \pm \sqrt{5}}{2}.$$

(c) Let

$$\lambda_1 = \frac{1+\sqrt{5}}{2}, \quad \lambda_2 = \frac{1-\sqrt{5}}{2}.$$

Using $y = \lambda x$, we have the basis $\{(1, \lambda_1), (1, \lambda_2)\}$ of $\mathbf{R}^2$ consisting of eigenvectors of T.

(d) Let v_1 and v_2 be the eigenvectors of T corresponding to the eigenvalues λ_1 and λ_2, respectively. Since v_1, v_2 form a basis of $\mathbf{R}^2$, we may write $(0,1) = \alpha v_1 + \beta v_2$ for some $\alpha, \beta \in \mathbf{R}$:

$$(0,1) = \alpha\left(1, \frac{1+\sqrt{5}}{2}\right) + \beta\left(1, \frac{1-\sqrt{5}}{2}\right) = (\alpha+\beta, \alpha\lambda_1 + \beta\lambda_2).$$

Then $\alpha + \beta = 0$ and $\alpha\lambda_1 + \beta\lambda_2 = 1$, so $\beta = -\alpha$ and $\alpha(\lambda_1 - \lambda_2) = 1$. Note that $\lambda_1 - \lambda_2 = \sqrt{5}$, so $\alpha = 1/\sqrt{5}$ and $\beta = -1/\sqrt{5}$. Then

$$\begin{aligned} T^n(0,1) &= T^n\left(\frac{1}{\sqrt{5}}v_1 - \frac{1}{\sqrt{5}}v_2\right) \\ &= \frac{1}{\sqrt{5}}T^n(v_1) - \frac{1}{\sqrt{5}}T^n(v_2) \\ &= \frac{1}{\sqrt{5}}\lambda_1^n v_1 - \frac{1}{\sqrt{5}}\lambda_2^n v_2 \\ &= \frac{1}{\sqrt{5}}\left(\lambda_1^n - \lambda_2^n, \lambda_1^{n+1} - \lambda_2^{n+1}\right). \end{aligned}$$

Hence, we have

$$F_n = \frac{1}{\sqrt{5}}\left(\lambda_1^n - \lambda_2^n\right) = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right].$$

(e) Observe that $\frac{1}{\sqrt{5}} < \frac{1}{2}$. Moreover,

$$|\lambda_2| = \frac{\sqrt{5}-1}{2} < 1$$

since $\sqrt{5} < 3$. Then for all $n \geq 0$, we have $|\lambda_2^n| \leq 1$, so

$$\left|\frac{1}{\sqrt{5}}\lambda_2^n\right| < \frac{1}{2} \implies \left|\frac{1}{\sqrt{5}}\lambda_1^n - F_n\right| < \frac{1}{2}.$$

Let m be an integer such that $m \neq F_n$. Then $|m - F_n| \geq 1$, so

$$\left|\frac{1}{\sqrt{5}}\lambda_1^n - m\right| = \left|\frac{1}{\sqrt{5}}\lambda_1^n - F_n + F_n - m\right| \geq |F_n - m| - \left|\frac{1}{\sqrt{5}}\lambda_1^n - F_n\right| > 1 - \frac{1}{2} = \frac{1}{2}.$$

Since the distance from $\frac{1}{\sqrt{5}}\lambda_1^n$ to any integer other than F_n is greater than $1/2$, we conclude that F_n is the integer closest to $\frac{1}{\sqrt{5}}\lambda_1^n$. ■

Exercise 5D22

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.

Axler Note. *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. By Theorem 5.67, each eigenvalue λ of T lies in at least one of the Gershgorin disks

$$D_j = \left\{ z \in \mathbf{C} : |z - A_{j,j}| \leq \sum_{\substack{k=1 \\ k \neq j}}^{n} |A_{j,k}| \right\}.$$

If $\lambda = 0$ is an eigenvalue of T, then 0 lies in at least one of the Gershgorin disks D_j. However, by the given inequality, we have

$$|0 - A_{j,j}| = |A_{j,j}| > \sum_{\substack{k=1 \\ k \neq j}}^{n} |A_{j,k}|,$$

so $0 \notin D_j$ for each $j \in \{1, \ldots, n\}$. Hence, 0 is not an eigenvalue of T, so T is invertible. ■

Exercise 5D23

Suppose the definition of the Gershgorin disks is changed so that the radius of the kth 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. Let $A = \mathcal{M}(T)$ be the matrix of T with respect to a basis $v_1, \ldots, v_n$ for V. Let λ be an eigenvalue of T. By Exercise 5B28, we know that T and T' have the same minimal polynomial, so they have the same eigenvalues. Moreover, we know that the matrix of T' with respect to the dual basis is just A^{t}. Using Theorem 5.67 on T', there must be some $k \in \{1, \ldots, n\}$ such that

$$|\lambda - A^{\text{t}}_{k,k}| \leq \sum_{\substack{j=1 \\ j \neq k}}^{n} |A^{\text{t}}_{k,j}|.$$

Since $A^{\text{t}}_{k,k} = A_{k,k}$ and $A^{\text{t}}_{k,j} = A_{j,k}$, we have

$$|\lambda - A_{k,k}| \leq \sum_{\substack{j=1 \\ j \neq k}}^{n} |A_{j,k}|,$$

so λ lies in the kth Gershgorin disk with the changed definition. ■

5E Commuting Operators

Exercise 5E1

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 e_1, e_2, e_3, e_4 be the standard basis of $\mathbf{F}^4$. Consider the operators $S, T \in \mathcal{L}(\mathbf{F}^4)$ defined by

$$S(x, y, z, w) = (0, x, 0, 0) \quad \text{and} \quad T(x, y, z, w) = (0, 0, x, 0).$$

Then

$$ST(x, y, z, w) = S(0, 0, x, 0) = (0, 0, 0, 0),$$
$$TS(x, y, z, w) = T(0, x, 0, 0) = (0, 0, 0, 0),$$

so $ST = TS$. The subspace $U = \operatorname{span}(e_1, e_2)$ is invariant under S, since $Se_1 = e_2 \in U$ and $Se_2 = 0 \in U$, but not under T, since $Te_1 = e_3 \notin U$. The subspace $W = \operatorname{span}(e_1, e_3)$ is invariant under T, since $Te_1 = e_3 \in W$ and $Te_3 = 0 \in W$, but not under S, since $Se_1 = e_2 \notin W$. ∎

Exercise 5E2

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.
Axler Note. *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 a basis of V with respect to which every element of $\mathcal{E}$ has a diagonal matrix. Then every pair of elements of $\mathcal{E}$ commutes, since diagonal matrices commute.

Conversely, suppose every pair of elements of $\mathcal{E}$ commutes. We proceed by induction on $\dim V$. If $\dim V = 0$, then $\varnothing$ is a basis, every element of $\mathcal{E}$ has empty matrix representation, and every pair of elements of $\mathcal{E}$ commutes. If $\dim V = 1$, then every operator on V is diagonalizable and has a diagonal matrix with respect to any basis of V. Now suppose $\dim V > 1$ and the statement holds for all vector spaces of dimension less than $\dim V$. If every element of $\mathcal{E}$ is a scalar multiple of the identity operator, then every element of $\mathcal{E}$ has a diagonal matrix with respect to any basis of V. Otherwise, there exists an operator $S \in \mathcal{E}$ that is not a scalar multiple of the identity operator. Decompose V into a direct sum of eigenspaces of S:

$$V = E(\lambda_1, S) \oplus \cdots \oplus E(\lambda_m, S),$$

where $\lambda_1, \ldots, \lambda_m$ are the distinct eigenvalues of S. Since every element of $\mathcal{E}$ commutes with S, each eigenspace $E(\lambda_i, S)$ is invariant under every element of $\mathcal{E}$. Since S is not a scalar multiple of the identity, $m \geq 2$, so $\dim E(\lambda_i, S) < \dim V$ for each i. For any $T \in \mathcal{E}$, Theorem 5.62 implies that the minimal polynomial of T is a product of distinct linear factors. Since the minimal polynomial of $T|_{E(\lambda_i, S)}$ divides the minimal polynomial of T, it follows that $T|_{E(\lambda_i, S)}$ is diagonalizable. Furthermore, any pair of elements of $\mathcal{E}$ commutes on $E(\lambda_i, S)$ since they commute on V and $E(\lambda_i, S)$ is invariant. By the induction hypothesis, for each i there exists a basis of $E(\lambda_i, S)$ with respect to which every restricted element of $\mathcal{E}$ has a diagonal matrix. Concatenating these bases gives a basis of V with respect to which every element of $\mathcal{E}$ has a diagonal matrix. ∎

Exercise 5E3

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.

Axler Note. *See 5.18 for the special case $S = T$.*

Solution. We first show that $ST = TS$ implies $S^n T = TS^n$ for every nonnegative integer n by induction on n. The base case $n = 0$ is trivial since $S^0 T = IT = T = TI = TS^0$. Now suppose $S^n T = TS^n$ for some nonnegative integer n. Then

$$S^{n+1}T = S(S^nT) = S(TS^n) = (ST)S^n = (TS)S^n = T(S^{n+1}),$$

so $S^{n+1}T = TS^{n+1}$. This completes the induction, so $S^n T = TS^n$ for any nonnegative integer n. Using this result, it follows that $p(S)T = Tp(S)$ for any polynomial $p \in \mathcal{P}(\mathbf{F})$.

(a) Suppose $v \in \operatorname{null} p(S)$. Then $p(S)v = 0$, so

$$p(S)(Tv) = Tp(S)v = T(0) = 0,$$

which shows that $Tv \in \operatorname{null} p(S)$. Thus, $\operatorname{null} p(S)$ is invariant under T.

(b) Suppose $v \in \operatorname{range} p(S)$. Then there exists $u \in V$ such that $p(S)u = v$, so

$$p(S)(Tu) = Tp(S)u = T(v),$$

which shows that $Tv \in \operatorname{range} p(S)$. Thus, $\operatorname{range} p(S)$ is invariant under T. ∎

Exercise 5E4

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 statement is false. Consider the diagonal matrix

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$$

and the upper-triangular matrix

$$B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Then

$$AB = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix},$$

but

$$BA = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}.$$

Since $AB \neq BA$, A and B do not commute. ∎

Exercise 5E5

Prove that a pair of operators on a finite-dimensional vector space commute if and only if their dual operators commute.

Axler Note. *See 3.118 for the definition of the dual of an operator.*

Solution. Suppose $S, T \in \mathcal{L}(V)$ commute. Then for any $\varphi \in V'$ and $v \in V$,

$$(S'T'\varphi)(v) = (T'\varphi)(Sv) = \varphi(TSv) = \varphi(STv) = (S'\varphi)(Tv) = (T'S'\varphi)(v),$$

so $S'T' = T'S'$.

Conversely, suppose $S', T' \in \mathcal{L}(V')$ commute. Applying the result we just proved to S' and T', we have that $(S')'(T')' = (T')'(S')'$, or equivalently $S''T'' = T''S''$. By Exercise 3F32, we know that S and T are isomorphic to S'' and T'', respectively, via the isomorphism Λ. Then for any $v \in V$,

$$\Lambda(STv) = S''T''\Lambda(v) = T''S''\Lambda(v) = \Lambda(TSv),$$

and injectivity of Λ implies that $ST = TS$. ∎

Exercise 5E6

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. Assume $V \neq \{0\}$. Since V is finite-dimensional and over $\mathbf{C}$, S has an eigenvalue $\alpha \in \mathbf{C}$. Let $W = \operatorname{range}(S - \alpha I)$. Since $\operatorname{null}(S - \alpha I) \neq \{0\}$, the fundamental theorem of linear maps gives $W \neq V$, so V/W is a nonzero complex finite-dimensional space. Since $ST = TS$, the subspace W is invariant under T: for $w = (S - \alpha I)u \in W$,

$$Tw = T(S - \alpha I)u = (S - \alpha I)Tu \in W.$$

Thus, the quotient operator $T/W \in \mathcal{L}(V/W)$ is well-defined and has an eigenvalue $\lambda \in \mathbf{C}$.

Suppose for contradiction that $W + \operatorname{range}(T - \lambda I) = V$. For every $v \in V$, write $v = w + (T - \lambda I)u$ with $w \in W$ and $u \in V$. Then

$$v + W = (T - \lambda I)u + W = (T/W - \lambda I)(u + W),$$

so $T/W - \lambda I$ is surjective on V/W. Since V/W is finite-dimensional, $T/W - \lambda I$ is invertible, contradicting λ being an eigenvalue of T/W. ∎

Exercise 5E7

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. Since S is diagonalizable, we can decompose V into a direct sum of eigenspaces of S:

$$V = E(\lambda_1, S) \oplus \cdots \oplus E(\lambda_m, S),$$

where $\lambda_1, \ldots, \lambda_m$ are the distinct eigenvalues of S. Since $ST = TS$, each eigenspace $E(\lambda_i, S)$ is invariant under T. Since V is complex and each $E(\lambda_i, S)$ is invariant under T, there exists a basis of $E(\lambda_i, S)$ with respect to which $T|_{E(\lambda_i, S)}$ has an upper-triangular matrix by Theorem 5.44. Let β_i be such a basis of $E(\lambda_i, S)$, and let β be the union of the β_i. Since $V = E(\lambda_1, S) \oplus \cdots \oplus E(\lambda_m, S)$, β is a basis of V. Moreover, each basis vector in β_i is an eigenvector of S, so $\mathcal{M}(S, \beta)$ is diagonal. Lastly, for each $v \in \beta_i$, $Tv \in E(\lambda_i, S) = \operatorname{span}(\beta_i)$. By choice of β_i, $\mathcal{M}(T|_{E(\lambda_i, S)}, \beta_i)$ is upper-triangular, so $\mathcal{M}(T, \beta)$ is upper-triangular. ∎

Exercise 5E8

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 basis $1, x, y, x^2, xy, y^2, x^3, x^2y, xy^2, y^3$ of $\mathcal{P}_3(\mathbf{R}^2)$. Enumerate the basis vectors as $v_1, \ldots, v_{10}$ in the order given. Then $D_x v_j \in \operatorname{span}(v_1, \ldots, v_{j-1})$ and $D_y v_j \in \operatorname{span}(v_1, \ldots, v_{j-1})$ for each $j = 1, \ldots, 10$, so $\mathcal{M}(D_x)$ and $\mathcal{M}(D_y)$ are upper-triangular with respect to this basis. ∎

Exercise 5E9

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.

Axler Note. *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) Let $T \in \mathcal{E}$. We proceed by induction on $\dim V$. For the base case $\dim V = 1$, every nonzero vector in V is an eigenvector for every element of $\mathcal{E}$. Now suppose $\dim V > 1$ and the statement holds for all complex vector spaces of dimension less than $\dim V$. If $\mathcal{E} = \varnothing$, then any nonzero vector in V is an eigenvector for every element of $\mathcal{E}$. Since V is complex and finite-dimensional, T has an eigenvalue $\lambda \in \mathbf{C}$. By Exercise 5E3, $E(\lambda, T)$ is invariant under every element of $\mathcal{E}$.

If $E(\lambda, T) = V$, then $T = \lambda I$, so every nonzero vector is an eigenvector of T. If all elements of $\mathcal{E}$ are scalar multiples of the identity, then every nonzero vector is an eigenvector for every element of $\mathcal{E}$. Otherwise, there exists an operator $S \in \mathcal{E}$ that is not a scalar multiple of the identity. Let μ be an eigenvalue of S. Then $E(\mu, S)$ is a nonzero proper subspace of V that is invariant under every element of $\mathcal{E}$ by Exercise 5E3, so $\dim E(\mu, S) < \dim V$. The restrictions $\{R|_{E(\mu,S)} : R \in \mathcal{E}\}$ pairwise commute because for any $R_1, R_2 \in \mathcal{E}$,

$$R_1|_{E(\mu,S)}\, R_2|_{E(\mu,S)} = (R_1R_2)|_{E(\mu,S)} = (R_2R_1)|_{E(\mu,S)} = R_2|_{E(\mu,S)}\, R_1|_{E(\mu,S)}.$$

By the inductive hypothesis, there exists a nonzero $v \in E(\mu, S)$ that is an eigenvector for every restriction, hence for every element of $\mathcal{E}$.

(b) We proceed by induction on $\dim V$. For the base case $\dim V = 1$, every operator on V has an upper-triangular matrix with respect to any basis of V. Now suppose $\dim V > 1$ and the statement holds for all complex vector spaces of dimension less than $\dim V$. By part (a), there exists a nonzero $v_1 \in V$ that is an eigenvector for every element of $\mathcal{E}$. Let $U = \operatorname{span}(v_1)$. Since v_1 is an eigenvector for every element of $\mathcal{E}$, U is invariant under every element of $\mathcal{E}$, so the quotient operators $\{T/U : T \in \mathcal{E}\}$ on V/U are well-defined. For any $S, T \in \mathcal{E}$,

$$(S/U)(T/U) = (ST)/U = (TS)/U = (T/U)(S/U),$$

so the quotient operators pairwise commute. Since $\dim(V/U) = \dim V - 1$, the inductive hypothesis gives a basis $v_2 + U, \ldots, v_n + U$ of V/U with respect to which every T/U has an upper-triangular matrix. Then $v_1, v_2, \ldots, v_n$ is a basis of V. For any $T \in \mathcal{E}$, $Tv_1 \in \operatorname{span}(v_1)$ since v_1 is an eigenvector of T. For $k \geq 2$, T/U being upper-triangular gives $(T/U)(v_k + U) \in \operatorname{span}(v_2 + U, \ldots, v_k + U)$, so $Tv_k \in \operatorname{span}(v_1, v_2, \ldots, v_k)$. Hence, every element of $\mathcal{E}$ has an upper-triangular matrix with respect to $v_1, \ldots, v_n$. ∎

Exercise 5E10

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.

Axler Note. *This exercise shows that 5.81 does not hold on real vector spaces.*

Solution. Consider the operators $S, T \in \mathcal{L}(\mathbf{R}^2)$ defined by

$$S(x, y) = (-y, x) \quad \text{and} \quad T(x, y) = (y, -x).$$

Since both are rotation operators, they commute. Neither S nor T has any real eigenvalues since no nonzero vector in $\mathbf{R}^2$ is a scalar multiple of its image under either operator. However, $S + T$ is the zero operator, which has eigenvalue 0, and $ST = I$, which has eigenvalue 1. ■

6 Inner Product Spaces

6A Inner Products and Norms

Exercise 6A1

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. By the additivity of inner products in both slots, we have

$$\sum_{j=1}^{m} \sum_{k=1}^{m} \langle v_j, v_k \rangle = \left\langle \sum_{j=1}^{m} v_j, \sum_{k=1}^{m} v_k \right\rangle = \left\| \sum_{j=1}^{m} v_j \right\|^2 \geq 0. \qquad \blacksquare$$

Exercise 6A2

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. Define $\langle u, v \rangle_1 = \langle Su, Sv \rangle$ for all $u, v \in V$. From the linearity of S and the properties of the regular inner product $\langle \cdot, \cdot \rangle$, we see that $\langle \cdot, \cdot \rangle_1$ satisfies the properties:

- $\langle u, v \rangle_1 = \langle Su, Sv \rangle = \overline{\langle Sv, Su \rangle} = \overline{\langle v, u \rangle_1}$ for all $u, v \in V$.
- $\langle u + w, v \rangle_1 = \langle S(u + w), Sv \rangle = \langle Su + Sw, Sv \rangle = \langle Su, Sv \rangle + \langle Sw, Sv \rangle = \langle u, v \rangle_1 + \langle w, v \rangle_1$ for all $u, v, w \in V$.
- $\langle \alpha u, v \rangle_1 = \langle S(\alpha u), Sv \rangle = \langle \alpha Su, Sv \rangle = \alpha \langle Su, Sv \rangle = \alpha \langle u, v \rangle_1$ for all $u, v \in V$ and $\alpha \in \mathbf{F}$.
- $\langle u, u \rangle_1 = \langle Su, Su \rangle = \|Su\|^2 \geq 0$ for all $u \in V$.

The only property that may not hold is definiteness.

If S is injective, then $\langle u, u \rangle_1 = 0 \implies \|Su\|^2 = 0 \implies Su = 0 \implies u = 0$, so definiteness holds. Conversely, if $\langle \cdot, \cdot \rangle_1$ is an inner product, then definiteness holds, so $Su = 0 \implies \langle u, u \rangle_1 = 0 \implies u = 0$, so S is injective. $\blacksquare$

Exercise 6A3

(a) Show that the function taking an ordered pair $((x_1, x_2), (y_1, y_2))$ of elements of $\mathbf{R}^2$ to $|x_1 y_1| + |x_2 y_2|$ 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 $x_1 y_1 + x_3 y_3$ is not an inner product on $\mathbf{R}^3$.

Solution.

(a) It is not linear in the first slot, since

$$\langle (1, 0) + (-1, 0), (1, 0) \rangle = \langle (0, 0), (1, 0) \rangle = |0 \cdot 1| + |0 \cdot 0| = 0,$$

but

$$\langle (1, 0), (1, 0) \rangle + \langle (-1, 0), (1, 0) \rangle = |1 \cdot 1| + |0 \cdot 0| + |-1 \cdot 1| + |0 \cdot 0| = 2.$$

(b) It is not positive definite, since $(0,1,0) \neq (0,0,0)$ but

$$\langle (0,1,0),(0,1,0)\rangle = 0 \cdot 0 + 0 \cdot 0 = 0.$$

■

Exercise 6A4

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. Suppose $v \in V$ is such that $(T - \sqrt{2}I)v = 0$. Then $Tv = \sqrt{2}v$, so $\|Tv\| = \sqrt{2}\|v\|$. But $\|Tv\| \leq \|v\|$, so $\sqrt{2}\|v\| \leq \|v\|$, which implies $\|v\| = 0$, so $v = 0$. Thus, $T - \sqrt{2}I$ is injective. ■

Exercise 6A5

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) We have

$$\begin{aligned}\langle u+v, u-v\rangle &= \langle u, u-v\rangle + \langle v, u-v\rangle \\ &= \langle u,u\rangle - \langle u,v\rangle + \langle v,u\rangle - \langle v,v\rangle \\ &= \|u\|^2 - \|v\|^2.\end{aligned}$$

(b) If $\|u\| = \|v\|$, then $\langle u+v, u-v\rangle = \|u\|^2 - \|v\|^2 = 0$, so $u+v$ is orthogonal to $u-v$.

(c) Let u and v be the vectors corresponding to two adjacent sides of a rhombus. Then $\|u\| = \|v\|$, so $u+v$ is orthogonal to $u-v$, where $u+v$ and $u-v$ represent the diagonals of the rhombus. ■

Exercise 6A6

Suppose $u, v \in V$. Prove that $\langle u, v\rangle = 0 \iff \|u\| \leq \|u + av\|$ for all $a \in \mathbf{F}$.

Solution. Suppose $\langle u, v\rangle = 0$. Then for all $a \in \mathbf{F}$, we have $\langle u, av\rangle = 0$. By the Pythagorean theorem, we have

$$\|u + av\|^2 = \|u\|^2 + \|av\|^2 \geq \|u\|^2,$$

Conversely, suppose that $\|u\| \leq \|u + av\|$ for all $a \in \mathbf{F}$. If $v = 0$, then $\langle u, v\rangle = 0$. Otherwise, consider $v \neq 0$, and let $a = -\frac{\langle u,v\rangle}{\|v\|^2}$. We have $\langle u + av, v\rangle = 0$, so $\langle u + av, av\rangle = 0$. Then by the Pythagorean theorem, we have

$$\|u\|^2 = \|(u + av) - av\|^2 = \|u + av\|^2 + \|av\|^2.$$

Note that $\|u\| \leq \|u + av\|$ implies $\|u\|^2 \leq \|u + av\|^2$. Hence, $\|av\|^2 \leq 0$, which implies $a = 0$ and $\langle u, v\rangle = 0$. ■

Exercise 6A7

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 all $a, b \in \mathbf{R}$, we have

$$\begin{aligned} \|au + bv\|^2 = \|bu + av\|^2 &\iff \langle au + bv, au + bv\rangle = \langle bu + av, bu + av\rangle \\ &\iff a^2\|u\|^2 + 2ab\,\mathrm{Re}\langle u, v\rangle + b^2\|v\|^2 = b^2\|u\|^2 + 2ab\,\mathrm{Re}\langle u, v\rangle + a^2\|v\|^2 \\ &\iff (a^2 - b^2)(\|u\|^2 - \|v\|^2) = 0. \end{aligned}$$

Since there exist $a, b \in \mathbf{R}$ such that $a^2 - b^2 \neq 0$, we have $\|au + bv\| = \|bu + av\|$ for all $a, b \in \mathbf{R}$ if and only if $\|u\|^2 - \|v\|^2 = 0$, which is equivalent to $\|u\| = \|v\|$. ∎

Exercise 6A8

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 $u = (a, b, c, x, y)$ and $v = (1, 1, 1, 4, 9)$. Then $\|u\|^2 = a^2+b^2+c^2+x^2+y^2 \leq 1$ and $\|v\|^2 = 1^2+1^2+1^2+4^2+9^2 = 100$. By the Cauchy–Schwarz inequality, we have

$$a + b + c + 4x + 9y = \langle u, v\rangle \leq |\langle u, v\rangle| \leq \|u\|\,\|v\| \leq \sqrt{100} = 10.$$ ∎

Exercise 6A9

Suppose $u, v \in V$ and $\|u\| = \|v\| = 1$ and $\langle u, v\rangle = 1$. Prove that $u = v$.

Solution. We have

$$\|u - v\|^2 = \langle u - v, u - v\rangle = \langle u, u\rangle - \langle u, v\rangle - \langle v, u\rangle + \langle v, v\rangle = 1 - 1 - 1 + 1 = 0,$$

where $\langle v, u\rangle = \overline{\langle u, v\rangle} = \overline{1} = 1$. Therefore, $\|u - v\| = 0$, which implies $u - v = 0$, so $u = v$. ∎

Exercise 6A10

Suppose $u, v \in V$ and $\|u\| \leq 1$ and $\|v\| \leq 1$. Prove that

$$\sqrt{1 - \|u\|^2}\sqrt{1 - \|v\|^2} \leq 1 - |\langle u, v\rangle|.$$

Solution. Consider the vectors $(\|u\|, \sqrt{1 - \|u\|^2})$ and $(\|v\|, \sqrt{1 - \|v\|^2})$ in $\mathbf{R}^2$. By the Cauchy–Schwarz inequality, we have

$$\|u\|\,\|v\| + \sqrt{1 - \|u\|^2}\sqrt{1 - \|v\|^2} \leq \sqrt{\|u\|^2 + 1 - \|u\|^2}\sqrt{\|v\|^2 + 1 - \|v\|^2} = 1.$$

By the Cauchy–Schwarz inequality, we also have $|\langle u, v\rangle| \leq \|u\|\,\|v\|$. Therefore,

$$\sqrt{1 - \|u\|^2}\sqrt{1 - \|v\|^2} \leq 1 - \|u\|\,\|v\| \leq 1 - |\langle u, v\rangle|.$$ ∎

Exercise 6A11

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 $u = a(1, 3) = (a, 3a)$ for some $a \in \mathbf{R}$. Then $v = (1, 2) - u = (1 - a, 2 - 3a)$. Since v is orthogonal to $(1, 3)$, we have

$$\langle v, (1, 3)\rangle = (1 - a) + 3(2 - 3a) = 7 - 10a = 0,$$

which implies $a = \frac{7}{10}$. Therefore, $u = \left(\frac{7}{10}, \frac{21}{10}\right)$ and $v = \left(\frac{3}{10}, -\frac{1}{10}\right)$. ∎

Exercise 6A12

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) Consider the vectors in $\mathbf{R}^4$:

$$u = \left(\sqrt{a}, \sqrt{b}, \sqrt{c}, \sqrt{d}\right) \implies \|u\|^2 = a+b+c+d,$$
$$v = \left(\frac{1}{\sqrt{a}}, \frac{1}{\sqrt{b}}, \frac{1}{\sqrt{c}}, \frac{1}{\sqrt{d}}\right) \implies \|v\|^2 = \frac{1}{a}+\frac{1}{b}+\frac{1}{c}+\frac{1}{d}.$$

Note that

$$\langle u, v\rangle = \sqrt{a}\cdot\frac{1}{\sqrt{a}} + \sqrt{b}\cdot\frac{1}{\sqrt{b}} + \sqrt{c}\cdot\frac{1}{\sqrt{c}} + \sqrt{d}\cdot\frac{1}{\sqrt{d}} = 4.$$

By the Cauchy–Schwarz inequality, we have $|\langle u, v\rangle| \leq \|u\|\,\|v\|$, so

$$16 = |\langle u, v\rangle|^2 \leq \|u\|^2\,\|v\|^2 = (a+b+c+d)\left(\frac{1}{a}+\frac{1}{b}+\frac{1}{c}+\frac{1}{d}\right).$$

(b) By the equality condition of the Cauchy–Schwarz inequality, we have equality if and only if u is a scalar multiple of v. If $u = \alpha v$ for some $\alpha \in \mathbf{R}$, then

$$\begin{cases} \sqrt{a} = \alpha\frac{1}{\sqrt{a}} \implies a = \alpha, \\ \sqrt{b} = \alpha\frac{1}{\sqrt{b}} \implies b = \alpha, \\ \sqrt{c} = \alpha\frac{1}{\sqrt{c}} \implies c = \alpha, \\ \sqrt{d} = \alpha\frac{1}{\sqrt{d}} \implies d = \alpha, \end{cases}$$

so $a = b = c = d$. ∎

Exercise 6A13

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 $a_1^2, \ldots, a_n^2$.

Manual Remark. *This is known as the QM–AM inequality, which is a special case of the power mean inequality.*

Solution. Let $u = (a_1, \ldots, a_n)$ and $v = (1, \ldots, 1) \in \mathbf{R}^n$. Then $\|u\|^2 = a_1^2 + \cdots + a_n^2$ and $\|v\|^2 = n$. By the Cauchy–Schwarz inequality, we have

$$|a_1 + \cdots + a_n|^2 = |\langle u, v\rangle|^2 \leq \|u\|^2\,\|v\|^2 = n(a_1^2 + \cdots + a_n^2),$$

so

$$\left(\frac{a_1 + \cdots + a_n}{n}\right)^2 = \frac{|a_1 + \cdots + a_n|^2}{n^2} \leq \frac{n(a_1^2 + \cdots + a_n^2)}{n^2} = \frac{a_1^2 + \cdots + a_n^2}{n}.$$

∎

Exercise 6A14

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. Let $u \in V$ with $\|u\| = 1$. Then

$$\begin{aligned} \|v - u\|^2 &= \|v\|^2 - 2\,\mathrm{Re}\langle v, u\rangle + \|u\|^2 \\ &\geq \|v\|^2 - 2|\langle v, u\rangle| + 1 \\ &\geq \|v\|^2 - 2\|v\|\,\|u\| + 1 \\ &= \|v\|^2 - 2\|v\| + 1 \\ &= (\|v\| - 1)^2 \end{aligned}$$

Equality holds if and only if $\mathrm{Re}\langle v, u\rangle = |\langle v, u\rangle| = \|v\|$. The second equality implies $v = \alpha u$ for some $\alpha \in \mathbf{F}$. Since $\alpha = \langle v, u\rangle$ and $\mathrm{Re}\langle v, u\rangle = |\langle v, u\rangle|$, we have $\alpha = \|v\| > 0$. Therefore, $u = v/\|v\|$. ∎

Exercise 6A15

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).
Axler Hint. *Use the law of cosines on the triangle formed by u, v, and $u - v$.*

Solution. Let θ be the angle between u and v. By the law of cosines, we have

$$\|u - v\|^2 = \|u\|^2 + \|v\|^2 - 2\|u\|\,\|v\| \cos\theta.$$

On the other hand, we have

$$\|u - v\|^2 = \langle u - v, u - v\rangle = \langle u, u\rangle - 2\langle u, v\rangle + \langle v, v\rangle = \|u\|^2 + \|v\|^2 - 2\langle u, v\rangle.$$

Therefore,

$$\|u\|^2 + \|v\|^2 - 2\|u\|\,\|v\| \cos\theta = \|u\|^2 + \|v\|^2 - 2\langle u, v\rangle,$$

which implies $\langle u, v\rangle = \|u\|\,\|v\| \cos\theta$. ∎

Exercise 6A16

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 6A15. Explain why the Cauchy–Schwarz inequality is needed to show that this definition makes sense.

Solution. The Cauchy–Schwarz inequality states that $|\langle x, y \rangle| \le \|x\| \, \|y\|$ for all $x, y \in \mathbf{R}^n$. Therefore, if x and y are nonzero vectors in $\mathbf{R}^n$, then

$$-1 \le -\frac{|\langle x, y \rangle|}{\|x\| \, \|y\|} \le \frac{\langle x, y \rangle}{\|x\| \, \|y\|} \le \frac{|\langle x, y \rangle|}{\|x\| \, \|y\|} \le 1.$$

Since the domain of arccos is $[-1, 1]$, the Cauchy–Schwarz inequality ensures that arccos is defined for the given expression. ∎

Exercise 6A17

Prove that

$$\left(\sum_{k=1}^{n} a_k b_k \right)^2 \le \left(\sum_{k=1}^{n} k a_k^2 \right) \left(\sum_{k=1}^{n} \frac{b_k^2}{k} \right)$$

for all real numbers $a_1, \ldots, a_n$ and $b_1, \ldots, b_n$.

Solution. Define the vectors

$$u = \left(\sqrt{1} a_1, \sqrt{2} a_2, \ldots, \sqrt{n} a_n \right), \quad v = \left(\frac{b_1}{\sqrt{1}}, \frac{b_2}{\sqrt{2}}, \ldots, \frac{b_n}{\sqrt{n}} \right) \in \mathbf{R}^n.$$

Then

$$\|u\|^2 = \sum_{k=1}^{n} k a_k^2, \quad \|v\|^2 = \sum_{k=1}^{n} \frac{b_k^2}{k}, \quad \langle u, v \rangle = \sum_{k=1}^{n} a_k b_k.$$

By the Cauchy–Schwarz inequality, we have

$$\left(\sum_{k=1}^{n} a_k b_k \right)^2 = |\langle u, v \rangle|^2 \le \|u\|^2 \, \|v\|^2 = \left(\sum_{k=1}^{n} k a_k^2 \right) \left(\sum_{k=1}^{n} \frac{b_k^2}{k} \right).$$ ∎

Exercise 6A18

(a) Suppose $f : [1, \infty) \to [0, \infty)$ is continuous. Show that

$$\left(\int_1^\infty f \right)^2 \le \int_1^\infty x^2 (f(x))^2 \, dx.$$

(b) For which continuous functions $f : [1, \infty) \to [0, \infty)$ is the inequality in (a) an equality with both sides finite?

Solution.

(a) Consider the functions over $[1, \infty)$:

$$u(x) = x f(x), \quad v(x) = \frac{1}{x}.$$

Then

$$\langle u, v\rangle = \int_1^\infty u(x)v(x)\,dx = \int_1^\infty f(x)\,dx, \quad \|u\|^2 = \int_1^\infty (xf(x))^2\,dx, \quad \|v\|^2 = \int_1^\infty \frac{1}{x^2}\,dx = 1.$$

By the Cauchy–Schwarz inequality, we have

$$\left(\int_1^\infty f\right)^2 = |\langle u, v\rangle|^2 \le \|u\|^2\,\|v\|^2 = \int_1^\infty x^2(f(x))^2\,dx.$$

(b) The equality condition of the Cauchy–Schwarz inequality states that we have equality if and only if u is a scalar multiple of v. Note that $f(x) \ge 0$ for all $x \in [1, \infty)$, so the scalar multiple must be nonnegative. Therefore, we have equality if and only if there exists a nonnegative constant c such that $xf(x) = c/x$ for all $x \in [1, \infty)$, which is equivalent to $f(x) = c/x^2$ for all $x \in [1, \infty)$. Moreover, both sides are finite as follows:

$$\int_1^\infty f(x)\,dx = \int_1^\infty \frac{c}{x^2}\,dx = c, \quad \int_1^\infty x^2(f(x))^2\,dx = \int_1^\infty x^2\left(\frac{c}{x^2}\right)^2 dx = c^2. \qquad \blacksquare$$

Exercise 6A19

Suppose $v_1, \dots, v_n$ is a basis of V and $T \in \mathcal{L}(V)$. Prove that if λ is an eigenvalue of T, then

$$|\lambda|^2 \le \sum_{j=1}^n \sum_{k=1}^n |\mathcal{M}(T)_{j,k}|^2,$$

where $\mathcal{M}(T)_{j,k}$ denotes the entry in row j, column k of the matrix of T with respect to the basis $v_1, \dots, v_n$.

Solution. Let $u \in V$ be an eigenvector of T corresponding to the eigenvalue λ, so $Tu = \lambda u$. Let $A = \mathcal{M}(T)$ be the matrix of T with respect to the basis $v_1, \dots, v_n$. From $Tu = \lambda u$, the coordinate vector $x \in \mathbf{F}^n$ of u with respect to the basis $v_1, \dots, v_n$ satisfies $Ax = \lambda x$. Note that $x \ne 0$ since $u \ne 0$.

Applying the Cauchy–Schwarz inequality to the rows of A, we have

$$|\lambda|^2\|x\|^2 = \|Ax\|^2 = \sum_{j=1}^n \left|\sum_{k=1}^n A_{j,k}x_k\right|^2 \le \left(\sum_{j=1}^n \sum_{k=1}^n |A_{j,k}|^2\right)\|x\|^2.$$

Dividing both sides by $\|x\|^2$ gives the desired inequality. $\blacksquare$

Exercise 6A20

Prove that if $u, v \in V$, then $\big|\,\|u\| - \|v\|\,\big| \le \|u - v\|$.

Axler Hint. *The inequality above is called the* **reverse triangle inequality**. *For the reverse triangle inequality when* $V = \mathbf{C}$, *see* Exercise 4-2.

Solution. By the triangle inequality, we have

$$\|u\| = \|(u - v) + v\| \le \|u - v\| + \|v\|,$$

which implies $\|u\| - \|v\| \le \|u - v\|$. Similarly, we have

$$\|v\| = \|(v - u) + u\| \le \|v - u\| + \|u\| = \|u - v\| + \|u\|,$$

which implies $\|v\| - \|u\| \leq \|u - v\|$. Therefore, we have

$$-\|u - v\| \leq \|u\| - \|v\| \leq \|u - v\|,$$

which is equivalent to $\big|\|u\| - \|v\|\big| \leq \|u - v\|$. ■

Exercise 6A21

Suppose $u, v \in V$ are such that

$$\|u\| = 3, \quad \|u + v\| = 4, \quad \|u - v\| = 6.$$

What number does $\|v\|$ equal?

Solution. By the parallelogram equality,

$$\|u + v\|^2 + \|u - v\|^2 = 2\|u\|^2 + 2\|v\|^2 = 16 + 36 = 52,$$

so $\|v\|^2 = 26 - 9 = 17$, which implies $\|v\| = \sqrt{17}$. ■

Exercise 6A22

Show that if $u, v \in V$, then

$$\|u + v\|\,\|u - v\| \leq \|u\|^2 + \|v\|^2.$$

Solution. By the triangle inequality, we have

$$\begin{aligned}\|u + v\|^2\|u - v\|^2 &\leq (\|u\|^2 + \|v\|^2)(\|u\|^2 + \|-v\|^2) \\ &= (\|u\|^2 + \|v\|^2)(\|u\|^2 + \|v\|^2) \\ &= (\|u\|^2 + \|v\|^2)^2.\end{aligned}$$

Taking the square root of both sides gives the desired inequality. ■

Exercise 6A23 ✔

Suppose $v_1, \ldots, v_m \in V$ are such that $\|v_k\| \leq 1$ for each $k = 1, \ldots, m$. Show that there exist $a_1, \ldots, a_m \in \{1, -1\}$ such that

$$\|a_1 v_1 + \cdots + a_m v_m\| \leq \sqrt{m}.$$

Manual Remark. *Note that* $\min\{A, B\}^2 \leq \frac{A^2+B^2}{2}$ *so long as* $A, B \geq 0$. *This follows from the fact that either* $A \leq B$ *or* $B \leq A$ *forces* $\min\{A, B\}^2 \leq A^2$ *and* $\min\{A, B\}^2 \leq B^2$.

Solution. We proceed by induction on m. For the base case $m = 1$, we have $\|v_1\| \leq 1 = \sqrt{1}$, so the statement holds.

Now suppose the statement holds for some $m \geq 1$. Let $v_1, \ldots, v_{m+1} \in V$ be such that $\|v_k\| \leq 1$ for each $k = 1, \ldots, m+1$. By the induction hypothesis, there exist $a_1, \ldots, a_m \in \{1, -1\}$ such that

$$\|a_1 v_1 + \cdots + a_m v_m\| \leq \sqrt{m}.$$

Let $u = a_1 v_1 + \cdots + a_m v_m$. Then

$$\|u + v_{m+1}\|^2 = \|u\|^2 + 2\,\mathrm{Re}\langle u, v_{m+1}\rangle + \|v_{m+1}\|^2,$$

and

$$\|u - v_{m+1}\|^2 = \|u\|^2 - 2\,\mathrm{Re}\langle u, v_{m+1}\rangle + \|v_{m+1}\|^2.$$

Adding these two equations gives

$$\|u + v_{m+1}\|^2 + \|u - v_{m+1}\|^2 = 2\|u\|^2 + 2\|v_{m+1}\|^2.$$

Since $\|v_{m+1}\|^2 \le 1$, we have

$$2\|u\|^2 + 2\|v_{m+1}\|^2 \le 2\|u\|^2 + 2.$$

Therefore,

$$\min\{\|u + v_{m+1}\|, \|u - v_{m+1}\|\}^2 \le \frac{\|u + v_{m+1}\|^2 + \|u - v_{m+1}\|^2}{2} \le \|u\|^2 + 1 \le m + 1,$$

where the last inequality follows from the induction hypothesis. Taking the square root of both sides gives

$$\min\{\|u + v_{m+1}\|, \|u - v_{m+1}\|\} \le \sqrt{m+1}.$$

Therefore, there exists $a_{m+1} \in \{1, -1\}$ such that

$$\|a_1 v_1 + \cdots + a_m v_m + a_{m+1} v_{m+1}\| = \|u + a_{m+1} v_{m+1}\| \le \sqrt{m+1}.$$ ■

Exercise 6A24

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 $\|(x, y)\| \ne \max\{|x|, |y|\}$.

Solution. If we define the norm as $\|(x, y)\| = \max\{|x|, |y|\}$ for all $(x, y) \in \mathbf{R}^2$, then take $u = (1, 0)$ and $v = (0, 1)$. Then

$$\|u + v\|^2 + \|u - v\|^2 = 1 + 1 = 2,$$

but

$$2\|u\|^2 + 2\|v\|^2 = 2 + 2 = 4,$$

contradicting the parallelogram equality. ■

Exercise 6A25

Suppose $p > 0$. Prove that there is an inner product on $\mathbf{R}^2$ such that the associated norm is given by

$$\|(x, y)\| = (|x|^p + |y|^p)^{1/p}$$

for all $(x, y) \in \mathbf{R}^2$ if and only if $p = 2$.

Solution. If $p = 2$, then the standard inner product on $\mathbf{R}^2$ gives the associated norm

$$\|(x, y)\| = (|x|^2 + |y|^2)^{1/2}.$$

Now suppose there is an inner product on $\mathbf{R}^2$ such that the associated norm is given by

$$\|(x, y)\| = (|x|^p + |y|^p)^{1/p}$$

for all $(x, y) \in \mathbf{R}^2$. Then the parallelogram equality must hold for this norm. Let $u = (1, 0)$ and $v = (0, 1)$. Then

$$\|u + v\|^2 + \|u - v\|^2 = 2^{2/p} + 2^{2/p} = 2 \cdot 2^{2/p} = 2^{1+2/p},$$

and

$$2\|u\|^2 + 2\|v\|^2 = 4.$$

Therefore, we must have

$$2^{1+2/p} = 4,$$

which implies $1 + 2/p = 2$, so $p = 2$. ■

Exercise 6A26

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. We have

$$\begin{aligned}\|u + v\|^2 - \|u - v\|^2 &= \langle u + v, u + v\rangle - \langle u - v, u - v\rangle \\ &= (\|u\|^2 + 2\langle u, v\rangle + \|v\|^2) - (\|u\|^2 - 2\langle u, v\rangle + \|v\|^2) \\ &= 4\langle u, v\rangle.\end{aligned}$$

Dividing both sides by 4 gives the desired equality. ■

Exercise 6A27

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. We have

$$\begin{aligned}\|u + v\|^2 - \|u - v\|^2 &= \langle u + v, u + v\rangle - \langle u - v, u - v\rangle \\ &= (\|u\|^2 + 2\,\mathrm{Re}\langle u, v\rangle + \|v\|^2) - (\|u\|^2 - 2\,\mathrm{Re}\langle u, v\rangle + \|v\|^2) \\ &= 4\,\mathrm{Re}\langle u, v\rangle,\end{aligned}$$

and

$$\begin{aligned}\|u + iv\|^2 - \|u - iv\|^2 &= \langle u + iv, u + iv\rangle - \langle u - iv, u - iv\rangle \\ &= (\|u\|^2 + 2\,\mathrm{Re}\langle u, iv\rangle + \|v\|^2) - (\|u\|^2 - 2\,\mathrm{Re}\langle u, iv\rangle + \|v\|^2) \\ &= 4\,\mathrm{Re}\langle u, iv\rangle.\end{aligned}$$

Let $\langle u, v\rangle = a + bi$ for some $a, b \in \mathbf{R}$. Then

$$\langle u, iv\rangle = \bar{i}\langle u, v\rangle = -i(a + bi) = b - ai,$$

so $\operatorname{Re}\langle u, iv\rangle = \operatorname{Im}\langle u, v\rangle$. Therefore, we have

$$\|u+v\|^2 - \|u-v\|^2 + \|u+iv\|^2 i - \|u-iv\|^2 i = 4\operatorname{Re}\langle u, v\rangle + 4\operatorname{Im}\langle u, v\rangle i = 4\langle u, v\rangle.$$

Dividing both sides by 4 gives the desired equality. ∎

Exercise 6A28

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 $u \in U$).

Manual Remark. *This is a highly non-trivial result known as the* **Jordan–von Neumann theorem**. *The proof is not easy since it requires density of* $\mathbf{Q}$ *in* $\mathbf{R}$ *to ensure homogeneity from the rational numbers extends to the real numbers via continuity.*

Solution. Suppose $\|\cdot\|$ is a norm on U satisfying the parallelogram equality. We first work in the case that $\mathbf{F} = \mathbf{R}$. Define

$$\langle u, v\rangle = \frac{\|u+v\|^2 - \|u-v\|^2}{4}$$

for all $u, v \in U$. We now check that $\langle\cdot,\cdot\rangle$ is an inner product on U.

- **Positive-Definiteness**: For any $u \in U$, we have

$$\langle u, u\rangle = \frac{\|u+u\|^2 - \|u-u\|^2}{4} = \frac{\|2u\|^2 - 0}{4} = \frac{4\|u\|^2}{4} = \|u\|^2 \ge 0,$$

  since $\|u\| \ge 0$. This establishes that $\|u\| = \langle u, u\rangle^{1/2}$ for all $u \in U$. Moreover, $\langle u, u\rangle = 0$ if and only if $\|u\|^2 = 0$, which is equivalent to $u = 0$.
- **Symmetry**: Note that $\|a\| = \|-a\|$ for any $a \in U$. For any $u, v \in U$, we have

$$\langle u, v\rangle = \frac{\|u+v\|^2 - \|u-v\|^2}{4} = \frac{\|v+u\|^2 - \|v-u\|^2}{4} = \langle v, u\rangle.$$

- **Additivity in the First Slot**: Let $x, y, z \in U$. Applying the parallelogram equality to the pairs $(x+z, y)$ and $(x-z, y)$ gives

$$\begin{aligned}\|(x+z)+y\|^2 + \|(x+z)-y\|^2 &= 2\|x+z\|^2 + 2\|y\|^2,\\ \|(x-z)+y\|^2 + \|(x-z)-y\|^2 &= 2\|x-z\|^2 + 2\|y\|^2.\end{aligned}$$

  Subtracting the second equation from the first gives

$$\|(x+z)+y\|^2 - \|(x-z)+y\|^2 + \|(x+z)-y\|^2 - \|(x-z)-y\|^2 = 2\|x+z\|^2 - 2\|x-z\|^2.$$

  Rearranging terms inside the norms gives

$$\|(x+y)+z\|^2 - \|(x+y)-z\|^2 + \|(x-y)+z\|^2 - \|(x-y)-z\|^2 = 2\|x+z\|^2 - 2\|x-z\|^2.$$

  Dividing both sides by 4 gives the relation $\langle x+y, z\rangle + \langle x-y, z\rangle = 2\langle x, z\rangle$. Note that $x = y$ yields $\langle 2x, z\rangle = 2\langle x, z\rangle$.

Then for any $u, v, w \in V$, define

$$x = \frac{u+v}{2}, \quad y = \frac{u-v}{2}$$

so that

$$\langle u+v, w\rangle = \langle 2x, w\rangle = 2\langle x, w\rangle = \langle x+y, w\rangle + \langle x-y, w\rangle = \langle u, w\rangle + \langle v, w\rangle.$$

- **Homogeneity in the First Slot**: Let $u, v \in U$ and $\alpha \in \mathbf{R}$. We show that $\langle \alpha u, v\rangle = \alpha\langle u, v\rangle$ for all $\alpha \in \mathbf{R}$. If $\alpha = 0$, then

$$\langle 0, v\rangle = \frac{\|0+v\|^2 - \|0-v\|^2}{4} = \frac{\|v\|^2 - \|v\|^2}{4} = 0 = 0\langle u, v\rangle.$$

We first consider when α is a positive integer. We proceed by induction on α. For the base case $\alpha = 1$, we have $\langle 1 \cdot u, v\rangle = \langle u, v\rangle$. Now suppose the statement holds for some positive integer α. Then

$$\langle (\alpha+1)u, v\rangle = \langle \alpha u + u, v\rangle = \langle \alpha u, v\rangle + \langle u, v\rangle = \alpha\langle u, v\rangle + \langle u, v\rangle = (\alpha+1)\langle u, v\rangle.$$

Therefore, the statement holds for all positive integers α. Next, we consider when α is a negative integer. By direct computation, we have for $\alpha = -1$ that

$$\langle -u, v\rangle = \frac{\|-u+v\|^2 - \|-u-v\|^2}{4} = \frac{\|v-u\|^2 - \|-(u+v)\|^2}{4} = \frac{\|v-u\|^2 - \|u+v\|^2}{4} = -\langle u, v\rangle.$$

We may now consider $\alpha = -\beta$ for some positive integer β. Then

$$\langle \alpha u, v\rangle = \langle -\beta u, v\rangle = \langle -(\beta u), v\rangle = -\langle \beta u, v\rangle = -\beta\langle u, v\rangle = \alpha\langle u, v\rangle.$$

We now consider when α is a rational number. Let $\alpha = p/q$ for some integers p and q with $q > 0$. Then

$$q\langle \alpha u, v\rangle = \langle q(\alpha u), v\rangle = \langle pu, v\rangle = p\langle u, v\rangle,$$

so $\langle \alpha u, v\rangle = \alpha\langle u, v\rangle$.

Finally, we consider when α is an arbitrary real number. By the density of the rational numbers in the real numbers, there exists a sequence $(\alpha_n)_{n=1}^{\infty}$ of rational numbers such that $\alpha_n \to \alpha$ as $n \to \infty$. By the reverse triangle inequality, we have

$$\big|\, \|\alpha_n u + v\| - \|\alpha u + v\| \,\big| \le \|(\alpha_n - \alpha)u\| = |\alpha_n - \alpha|\,\|u\|,$$

where $|\alpha_n - \alpha| \to 0$ as $n \to \infty$, so $\|\alpha_n u + v\| \to \|\alpha u + v\|$ as $n \to \infty$. Similarly, we have $\|\alpha_n u - v\| \to \|\alpha u - v\|$ as $n \to \infty$. Then $\langle \alpha_n u, v\rangle \to \langle \alpha u, v\rangle$ as $n \to \infty$. It follows that $\langle \alpha_n u, v\rangle = \alpha_n\langle u, v\rangle \to \alpha\langle u, v\rangle$ as $n \to \infty$, so $\langle \alpha u, v\rangle = \alpha\langle u, v\rangle$.

Now suppose $\mathbf{F} = \mathbf{C}$. Define

$$\langle u, v\rangle = \frac{\|u+v\|^2 - \|u-v\|^2}{4} + \frac{\|u+iv\|^2 - \|u-iv\|^2}{4}i$$

for all $u, v \in U$. We again check that $\langle \cdot, \cdot \rangle$ is an inner product on U.

- **Positive-Definiteness**: For any $u \in U$, $\mathrm{Re}\langle u, u\rangle$ follows from the real case. Considering $\mathrm{Im}\langle u, u\rangle$, we have that

$$\|u+iu\|^2 = \|(1+i)u\|^2 = |1+i|^2\|u\|^2 = 2\|u\|^2,$$
$$\|u-iu\|^2 = \|(1-i)u\|^2 = |1-i|^2\|u\|^2 = 2\|u\|^2,$$

so $\mathrm{Im}\langle u, u\rangle = 0$. Therefore, $\langle u, u\rangle = \mathrm{Re}\langle u, u\rangle + \mathrm{Im}\langle u, u\rangle i = \mathrm{Re}\langle u, u\rangle \ge 0$. This again establishes that $\|u\| = \langle u, u\rangle^{1/2}$ for all $u \in U$. Moreover, $\langle u, u\rangle = 0$ if and only if $\mathrm{Re}\langle u, u\rangle = 0$, which is equivalent to $u = 0$.

- **Conjugate Symmetry**: For any $u, v \in U$, $\mathrm{Re}\langle u, v\rangle$ follows from the real case. Moreover, note that $\|u+iv\| = \|i(v-iu)\| = |i|\,\|v-iu\| = \|v-iu\|$. Similarly, $\|u-iv\| = \|-i(v+iu)\| = |-i|\,\|v+iu\| = \|v+iu\|$. Thus, for $\mathrm{Im}\langle u, v\rangle$,

we have

$$\begin{aligned}\operatorname{Im}\langle u, v\rangle &= \frac{\|u + iv\|^2 - \|u - iv\|^2}{4} \\ &= \frac{\|v - iu\|^2 - \|v + iu\|^2}{4} \\ &= -\operatorname{Im}\langle v, u\rangle.\end{aligned}$$

Therefore, $\langle u, v\rangle = \operatorname{Re}\langle u, v\rangle + \operatorname{Im}\langle u, v\rangle i = \operatorname{Re}\langle v, u\rangle - \operatorname{Im}\langle v, u\rangle i = \overline{\langle v, u\rangle}$.

- **Additivity in the First Slot**: Let $u, v, w \in U$. Then $\operatorname{Re}\langle u + v, w\rangle$ follows from the real case. For $\operatorname{Im}\langle u + v, w\rangle$, the real-case additivity identity applied with w replaced by iw gives

$$\operatorname{Im}\langle u + v, w\rangle = \operatorname{Im}\langle u, w\rangle + \operatorname{Im}\langle v, w\rangle.$$

Therefore,

$$\begin{aligned}\langle u + v, w\rangle &= \operatorname{Re}\langle u + v, w\rangle + \operatorname{Im}\langle u + v, w\rangle i \\ &= \operatorname{Re}\langle u, w\rangle + \operatorname{Re}\langle v, w\rangle + (\operatorname{Im}\langle u, w\rangle + \operatorname{Im}\langle v, w\rangle)i \\ &= (\operatorname{Re}\langle u, w\rangle + \operatorname{Im}\langle u, w\rangle i) + (\operatorname{Re}\langle v, w\rangle + \operatorname{Im}\langle v, w\rangle i) \\ &= \langle u, w\rangle + \langle v, w\rangle.\end{aligned}$$

- **Homogeneity in the First Slot**: Let $u, v \in U$ and $\alpha \in \mathbf{C}$. We first consider $\alpha = i$. Then

$$\begin{aligned}\langle iu, v\rangle &= \frac{\|iu + v\|^2 - \|iu - v\|^2}{4} + \frac{\|iu + iv\|^2 - \|iu - iv\|^2}{4} i \\ &= \frac{\|v - iu\|^2 - \|v + iu\|^2}{4} + \frac{\|i(u + v)\|^2 - \|i(u - v)\|^2}{4} i \\ &= -\operatorname{Im}\langle v, u\rangle + \operatorname{Re}\langle u, v\rangle i = i\langle u, v\rangle.\end{aligned}$$

Now let $\alpha = a + bi$ for some $a, b \in \mathbf{R}$. From the real case, we have

$$\langle au, v\rangle = a\langle u, v\rangle \quad \text{and} \quad \langle ibu, v\rangle = b\langle iu, v\rangle = bi\langle u, v\rangle.$$

Therefore,

$$\langle \alpha u, v\rangle = \langle au + ibu, v\rangle = \langle au, v\rangle + \langle ibu, v\rangle = a\langle u, v\rangle + bi\langle u, v\rangle = \alpha\langle u, v\rangle. \qquad \blacksquare$$

Exercise 6A29

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 $V_1 \times \cdots \times V_m$.

Axler Note. *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. Let $\langle \cdot, \cdot\rangle_k$ denote the inner product on V_k for each $k = 1, \ldots, m$. To see that $\langle \cdot, \cdot\rangle$ is an inner product on $V_1 \times \cdots \times V_m$, note that the properties of conjugate symmetry, additivity in the first slot, and homogeneity in the first slot follow from the corresponding properties of each $\langle \cdot, \cdot\rangle_k$.

For positive-definiteness, let $(u_1, \ldots, u_m) \in V_1 \times \cdots \times V_m$. Then

$$\langle (u_1, \ldots, u_m), (u_1, \ldots, u_m)\rangle = \langle u_1, u_1\rangle_1 + \cdots + \langle u_m, u_m\rangle_m \geq 0.$$

Moreover, $\langle (u_1, \dots, u_m), (u_1, \dots, u_m) \rangle = 0$ if and only if $\langle u_k, u_k \rangle_k = 0$ for each $k = 1, \dots, m$, which is equivalent to $u_k = 0$ for each $k = 1, \dots, m$, so $(u_1, \dots, u_m) = 0$. ∎

Exercise 6A30

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

$$\langle u, v \rangle_{\mathbf{C}} = \langle u, v \rangle \quad \text{and} \quad \|u + iv\|_{\mathbf{C}}^2 = \|u\|^2 + \|v\|^2.$$

Axler Hint. *See Exercise 1B8 for the definition of the complexification* $V_{\mathbf{C}}$.

Solution.

(a) We first check that $\langle \cdot, \cdot \rangle_{\mathbf{C}}$ is an inner product on $V_{\mathbf{C}}$.

- **Positive-Definiteness**: For any $u + iv \in V_{\mathbf{C}}$, we have

$$\langle u + iv, u + iv \rangle_{\mathbf{C}} = \langle u, u \rangle + \langle v, v \rangle + (\langle v, u \rangle - \langle u, v \rangle)i = \|u\|^2 + \|v\|^2 \geq 0,$$

since $\|u\|, \|v\| \geq 0$. Moreover, $\langle u + iv, u + iv \rangle_{\mathbf{C}} = 0$ if and only if $\|u\|^2 + \|v\|^2 = 0$, which is equivalent to $u = v = 0$, so $u + iv = 0$.

- **Conjugate Symmetry**: For any $u + iv, w + ix \in V_{\mathbf{C}}$, we have

$$\begin{aligned} \langle w + ix, u + iv \rangle_{\mathbf{C}} &= \langle w, u \rangle + \langle x, v \rangle + (\langle x, u \rangle - \langle w, v \rangle)i \\ &= \overline{\langle u, w \rangle} + \overline{\langle v, x \rangle} + (\overline{\langle u, x \rangle} - \overline{\langle v, w \rangle})i \\ &= \overline{\langle u, w \rangle + \langle v, x \rangle + (\langle v, w \rangle - \langle u, x \rangle)i} = \overline{\langle u + iv, w + ix \rangle_{\mathbf{C}}}. \end{aligned}$$

- **Additivity in the First Slot**: Let $u + iv, w + ix, y + iz \in V_{\mathbf{C}}$. Then

$$\begin{aligned} \langle (u + iv) + (y + iz), w + ix \rangle_{\mathbf{C}} &= \langle (u + y) + i(v + z), w + ix \rangle_{\mathbf{C}} \\ &= \langle u + y, w \rangle + \langle v + z, x \rangle + (\langle v + z, w \rangle - \langle u + y, x \rangle)i \end{aligned}$$

Expanding the inner products gives

$$\begin{aligned} &(\langle u, w \rangle + \langle y, w \rangle) + (\langle v, x \rangle + \langle z, x \rangle) + ((\langle v, w \rangle + \langle z, w \rangle) - (\langle u, x \rangle + \langle y, x \rangle))i \\ &= (\langle u, w \rangle + \langle v, x \rangle + (\langle v, w \rangle - \langle u, x \rangle)i) + (\langle y, w \rangle + \langle z, x \rangle + (\langle z, w \rangle - \langle y, x \rangle)i) \\ &= \langle u + iv, w + ix \rangle_{\mathbf{C}} + \langle y + iz, w + ix \rangle_{\mathbf{C}}. \end{aligned}$$

- **Homogeneity in the First Slot**: Let $u + iv, w + ix \in V_{\mathbf{C}}$ and $\lambda \in \mathbf{R}$. Then

$$\begin{aligned} \langle \lambda(u + iv), w + ix \rangle_{\mathbf{C}} &= \langle (\lambda u) + i(\lambda v), w + ix \rangle_{\mathbf{C}} \\ &= \langle \lambda u, w \rangle + \langle \lambda v, x \rangle + (\langle \lambda v, w \rangle - \langle \lambda u, x \rangle)i \\ &= \lambda \langle u, w \rangle + \lambda \langle v, x \rangle + (\lambda \langle v, w \rangle - \lambda \langle u, x \rangle)i \\ &= \lambda(\langle u, w \rangle + \langle v, x \rangle + (\langle v, w \rangle - \langle u, x \rangle)i) = \lambda \langle u + iv, w + ix \rangle_{\mathbf{C}}. \end{aligned}$$

Moreover, note that $i(u+iv) = -v+iu$. Then

$$\begin{aligned}\langle i(u+iv), w+ix\rangle_{\mathbf{C}} &= \langle -v+iu, w+ix\rangle_{\mathbf{C}} \\ &= \langle -v, w\rangle + \langle u, x\rangle + (\langle u, w\rangle - \langle -v, x\rangle)i \\ &= -\langle v, w\rangle + \langle u, x\rangle + (\langle u, w\rangle + \langle v, x\rangle)i \\ &= i(\langle u, w\rangle + \langle v, x\rangle + (\langle v, w\rangle - \langle u, x\rangle)i) = i\langle u+iv, w+ix\rangle_{\mathbf{C}}.\end{aligned}$$

Hence, homogeneity in the first slot holds for all $\alpha \in \mathbf{C}$ by using additivity in the first slot and the fact that $\alpha = a+bi$ for some $a, b \in \mathbf{R}$.

Hence, $\langle \cdot, \cdot \rangle_{\mathbf{C}}$ is an inner product on $V_{\mathbf{C}}$, so $V_{\mathbf{C}}$ is a complex inner product space.

(b) We have

$$\langle u, v\rangle_{\mathbf{C}} = \langle u+i0, v+i0\rangle_{\mathbf{C}} = \langle u, v\rangle + \langle 0, 0\rangle + (\langle 0, v\rangle - \langle u, 0\rangle)i = \langle u, v\rangle,$$

and

$$\|u+iv\|_{\mathbf{C}}^2 = \langle u+iv, u+iv\rangle_{\mathbf{C}} = \langle u, u\rangle + \langle v, v\rangle + (\langle v, u\rangle - \langle u, v\rangle)i = \|u\|^2 + \|v\|^2. \qquad \blacksquare$$

Exercise 6A31

Suppose $u, v, w \in V$. Prove that

$$\|w - \tfrac{1}{2}(u+v)\|^2 = \frac{\|w-u\|^2 + \|w-v\|^2}{2} - \frac{\|u-v\|^2}{4}.$$

Solution. Set $a = w-u$ and $b = w-v$. Then the left-hand side is

$$\|w - \tfrac{1}{2}(u+v)\|^2 = \|\tfrac{1}{2}(a+b)\|^2 = \frac{1}{4}\|a+b\|^2.$$

Moreover, the right-hand side is

$$\frac{\|w-u\|^2 + \|w-v\|^2}{2} - \frac{\|u-v\|^2}{4} = \frac{\|a\|^2 + \|b\|^2}{2} - \frac{\|a-b\|^2}{4} = \frac{2\|a\|^2 + 2\|b\|^2 - \|a-b\|^2}{4}.$$

From the parallelogram equality, we have

$$\|a+b\|^2 + \|a-b\|^2 = 2\|a\|^2 + 2\|b\|^2 \implies \|a+b\|^2 = 2\|a\|^2 + 2\|b\|^2 - \|a-b\|^2,$$

so division by 4 gives our desired equality. $\blacksquare$

Exercise 6A32

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 points to w. Then $\|w-u\| = \|w-v\|$. By the property of E, we have $\frac{1}{2}(u+v) \in E$. By Exercise 6A31, we have

$$\|w - \tfrac{1}{2}(u+v)\|^2 = \frac{\|w-u\|^2 + \|w-v\|^2}{2} - \frac{\|u-v\|^2}{4} = \|w-u\|^2 - \frac{\|u-v\|^2}{4}.$$

Since $\frac{1}{2}(u+v) \in E$, we have

$$\|w-u\|^2 \le \|w-\tfrac{1}{2}(u+v)\|^2 = \|w-u\|^2 - \frac{\|u-v\|^2}{4},$$

so $\|u-v\|^2 = 0$, which implies that $u = v$. ■

Exercise 6A33

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.

Axler Note. *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) We have

$$\langle f(t), g(t)\rangle' = \left(\sum_{k=1}^{n} f_k(t)g_k(t)\right)' = \sum_{k=1}^{n}\left(f_k'(t)g_k(t) + f_k(t)g_k'(t)\right) = \langle f'(t), g(t)\rangle + \langle f(t), g'(t)\rangle.$$

(b) Suppose $\|f(t)\| = c$ for every $t \in \mathbf{R}$. Then $\langle f(t), f(t)\rangle = c^2$ for every $t \in \mathbf{R}$. Differentiating both sides with respect to t, we get

$$\langle f'(t), f(t)\rangle + \langle f(t), f'(t)\rangle = 2\langle f'(t), f(t)\rangle = 0,$$

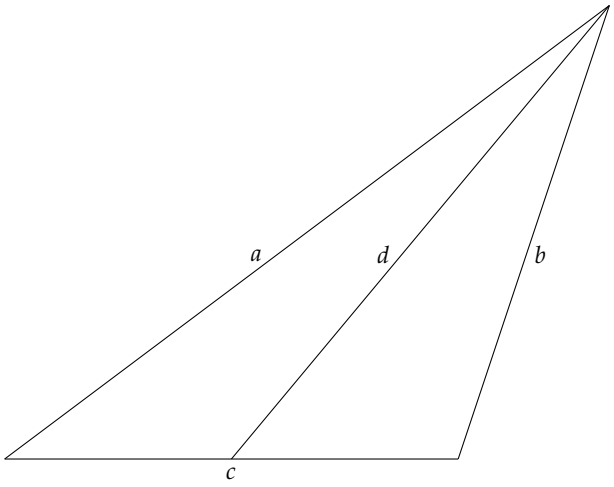
so $\langle f'(t), f(t)\rangle = 0$ for every $t \in \mathbf{R}$.

(c) Geometrically, the result in (b) says that the tangent vector to a curve lying on a sphere in $\mathbf{R}^n$ centered at the origin is orthogonal to the position vector of the point on the curve. ■

Exercise 6A34

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. Consider $\triangle ABC$ with $\|C-A\| = a$, $\|C-B\| = b$, and $\|B-A\| = c$. Let M be the midpoint of $\overline{AB}$, so $M = \frac{1}{2}(A+B)$. Then $d = \|C - M\|$. By Exercise 6A31, we have

$$\|C - M\|^2 = \frac{\|C-A\|^2 + \|C-B\|^2}{2} - \frac{\|A-B\|^2}{4} = \frac{a^2+b^2}{2} - \frac{c^2}{4}.$$

Rearranging terms gives $a^2 + b^2 = \frac{1}{2}c^2 + 2d^2$. ■

Exercise 6A35

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 $m_1, \ldots, m_n$ 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 $\|x\| = 1$.

Axler Note. *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 $p = 0$.*

Axler Hint. *A reasonable guess is that the desired harmonic polynomial p is of the form $q + (1 - \|x\|^2)r$ for some polynomial r. Prove that there is a polynomial r on $\mathbf{R}^n$ such that $q + (1 - \|x\|^2)r$ is harmonic by defining an operator T on a suitable vector space by*

$$Tr = \Delta((1 - \|x\|^2)r)$$

and then showing that T is injective and hence surjective.

Solution. Consider $T \in \mathcal{L}(\mathcal{P}_m(\mathbf{R}^n))$ defined by $Tr = \Delta((1 - \|x\|^2)r)$, where $\deg q = m$. Note that T is well-defined, since multiplication by $1 - \|x\|^2$ increases the degree of a polynomial by 2, and Δ decreases the degree by at most 2. Moreover, linearity of T follows from both linearity of Δ and linearity of multiplication by $1 - \|x\|^2$.

We show that T is injective. Suppose $Tr = 0$ for some $r \in \mathcal{P}_m(\mathbf{R}^n)$. Then $\Delta((1 - \|x\|^2)r) = 0$, so $(1 - \|x\|^2)r$ is harmonic. Since $(1 - \|x\|^2)r(x) = 0$ for all $x \in \mathbf{R}^n$ with $\|x\| = 1$, we have $(1 - \|x\|^2)r = 0$ by the fact about harmonic functions given in the exercise.

Now fix $y \in \mathbf{R}^n$. Evaluating at $x = ty$ for $t \in \mathbf{R}$, we have $(1-t^2\|y\|^2)r(ty) = 0$ for all $t \in \mathbf{R}$. If $y = 0$, then $r(ty) = r(0) = 0$ for all $t \in \mathbf{R}$. Otherwise, assume $y \neq 0$. Since $1 - t^2\|y\|^2$ is a polynomial in t of degree at most 2, it has at most two roots. Therefore, $r(ty) = 0$ for all $t \in \mathbf{R}$ except for at most two values of t. Since $r(ty)$ is a polynomial in t, it must be identically zero, so $r(ty) = 0$ for all $t \in \mathbf{R}$. Since $y \in \mathbf{R}^n$ was arbitrary, we have $r = 0$, so T is injective. Moreover, finite-dimensionality of $\mathcal{P}_m(\mathbf{R}^n)$ implies that T is surjective. Therefore, there exists $r \in \mathcal{P}_m(\mathbf{R}^n)$ such that $Tr = \Delta((1 - \|x\|^2)r) = -\Delta q$. We then have

$$\Delta(q + (1 - \|x\|^2)r) = \Delta q + \Delta((1 - \|x\|^2)r) = \Delta q - \Delta q = 0,$$

so $p = q + (1 - \|x\|^2)r$ is harmonic. ∎

6B Orthonormal Bases

Exercise 6B1

Suppose $e_1, \dots, e_m$ 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 $a_1, \dots, a_m \in \mathbf{F}$. Show that $e_1, \dots, e_m$ is an orthonormal list.

Axler Note. *This exercise provides a converse to 6.24.*

Solution. Let $j, k \in \{1, \dots, m\}$ with $j \neq k$. Then for $a_j = 1$ and $a_i = 0$ for all $i \neq j, k$, we have

$$\|e_j + a_k e_k\|^2 = |1|^2 + |a_k|^2 = 1 + |a_k|^2$$

by the hypothesis. If $a_k = 0$, then $\langle e_j, e_j \rangle = \|e_j\|^2 = 1$ for all j. Otherwise, let $|a_k| = 1$ so that $\|e_j + a_k e_k\|^2 = 1 + |a_k|^2 = 2$.

If $\mathbf{F} = \mathbf{R}$, then

$$\langle e_j, e_k \rangle = \frac{1}{4}(\|e_j + e_k\|^2 - \|e_j - e_k\|^2) = \frac{1}{4}(2 - 2) = 0$$

by Exercise 6A26. If $\mathbf{F} = \mathbf{C}$, then

$$\langle e_j, e_k \rangle = \frac{1}{4}(\|e_j + e_k\|^2 - \|e_j - e_k\|^2 + \|e_j + ie_k\|^2 i - \|e_j - ie_k\|^2 i) = \frac{1}{4}(2 - 2 + 2i - 2i) = 0$$

by Exercise 6A27. Therefore, the list $e_1, \dots, e_m$ is orthonormal. ∎

Exercise 6B2

(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) Let $v_1 = (\cos\theta, \sin\theta)$ and $v_2 = (-\sin\theta, \cos\theta)$. Then $\|v_1\|^2 = \cos^2\theta + \sin^2\theta = 1$ and $\|v_2\|^2 = \sin^2\theta + \cos^2\theta = 1$. Moreover,

$$\langle v_1, v_2 \rangle = -\cos\theta\sin\theta + \sin\theta\cos\theta = 0.$$

Since v_1, v_2 is an orthonormal list of two vectors in $\mathbf{R}^2$, it is an orthonormal basis of $\mathbf{R}^2$. Similarly, let $w_1 = (\cos\theta, \sin\theta)$ and $w_2 = (\sin\theta, -\cos\theta)$. Then $\|w_1\|^2 = \cos^2\theta + \sin^2\theta = 1$ and $\|w_2\|^2 = \sin^2\theta + \cos^2\theta = 1$. Again,

$$\langle w_1, w_2 \rangle = \cos\theta\sin\theta - \sin\theta\cos\theta = 0.$$

Since w_1, w_2 is an orthonormal list of two vectors in $\mathbf{R}^2$, it is an orthonormal basis of $\mathbf{R}^2$.

(b) Suppose u_1, u_2 is an orthonormal basis of $\mathbf{R}^2$. If $u_1 = (x_1, y_1)$, then $\|u_1\|^2 = x_1^2 + y_1^2 = 1$, so there exists $\theta \in \mathbf{R}$ such that $x_1 = \cos\theta$ and $y_1 = \sin\theta$. If $u_2 = (x_2, y_2)$, then $\|u_2\|^2 = x_2^2 + y_2^2 = 1$ and $\langle u_1, u_2 \rangle = x_1 x_2 + y_1 y_2 = 0$. We now solve the system of equations

$$\begin{cases} x_2^2 + y_2^2 = 1, \\ \cos\theta x_2 + \sin\theta y_2 = 0. \end{cases}$$

If $\sin\theta = 0$, then $\theta = 0$ or $\theta = \pi$, so $\cos\theta = \pm 1$. Then $x_2 = 0$, so $y_2^2 = 1$, and thus $y_2 = \pm 1$. Then $u_2 = (0, \pm 1)$, which is of the form given in (a).

Otherwise, $y_2 = -\frac{\cos\theta}{\sin\theta} x_2$, so

$$x_2^2 + \left(-\frac{\cos\theta}{\sin\theta} x_2\right)^2 = 1 \implies x_2^2 + \frac{\cos^2\theta}{\sin^2\theta} x_2^2 = 1 \implies x_2^2 \left(1 + \frac{\cos^2\theta}{\sin^2\theta}\right) = 1.$$

We then have $x_2^2 = \sin^2\theta$, so $x_2 = \pm\sin\theta$. Then $y_2 = -\frac{\cos\theta}{\sin\theta}(\pm\sin\theta) = \mp\cos\theta$. Thus, $u_2 = (\pm\sin\theta, \mp\cos\theta)$, which is of the form given in (a). ■

Exercise 6B3

Suppose $e_1, \dots, e_m$ is an orthonormal list in V and $v \in V$. Prove that

$$\|v\|^2 = |\langle v, e_1\rangle|^2 + \cdots + |\langle v, e_m\rangle|^2 \iff v \in \operatorname{span}(e_1, \dots, e_m).$$

Solution. Since $e_1, \dots, e_m$ is an orthonormal basis for $\operatorname{span}(e_1, \dots, e_m)$, if $v \in \operatorname{span}(e_1, \dots, e_m)$, then we have $\|v\|^2 = |\langle v, e_1\rangle|^2 + \cdots + |\langle v, e_m\rangle|^2$.

Conversely, suppose $\|v\|^2 = |\langle v, e_1\rangle|^2 + \cdots + |\langle v, e_m\rangle|^2$. Let $w = v - (\langle v, e_1\rangle e_1 + \cdots + \langle v, e_m\rangle e_m)$. Note that $\langle w, e_k\rangle = \langle v, e_k\rangle - \langle v, e_k\rangle = 0$ for each $k = 1, \dots, m$, so w is orthogonal to each e_k. Then

$$\|v\|^2 = \left\| w + \sum_{k=1}^{m} \langle v, e_k\rangle e_k \right\|^2 = \|w\|^2 + \sum_{k=1}^{m} |\langle v, e_k\rangle|^2.$$

The assumption of $\|v\|^2 = |\langle v, e_1\rangle|^2 + \cdots + |\langle v, e_m\rangle|^2$ gives $\|w\|^2 = 0$, so $w = 0$, and $v \in \operatorname{span}(e_1, \dots, e_m)$. ■

Exercise 6B4

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

$$\langle f, g\rangle = \int_{-\pi}^{\pi} fg.$$

Axler 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. Define the vectors

$$f_0 = \frac{1}{\sqrt{2\pi}}, \quad f_k = \frac{\cos kx}{\sqrt{\pi}}, \quad g_k = \frac{\sin kx}{\sqrt{\pi}} \quad \text{for } k = 1, \dots, n.$$

Then

$$\|f_0\|^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} 1 \, dx = 1.$$

Moreover, note the following for $k = 1, \ldots, n$,

$$\langle f_0, f_k \rangle = \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{\pi}} \int_{-\pi}^{\pi} \cos kx \, \mathrm{d}x = \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{\pi}} \frac{\sin kx}{k} \Big|_{-\pi}^{\pi} = 0,$$

and $\langle f_0, g_k \rangle = 0$ because $\sin kx$ is an odd function. Hence, f_0 is orthogonal to each of the other vectors in the list. Using the $\cos kx$ and $\sin kx$ computations in the above integrals, we may simplify the calculations of the norms of f_k and g_k to get

$$\begin{aligned} \|f_k\|^2 &= \frac{1}{\pi} \int_{-\pi}^{\pi} \cos^2 kx \, \mathrm{d}x = \frac{1}{2\pi} \int_{-\pi}^{\pi} (1 + \cos 2kx) \, \mathrm{d}x = 1, \\ \|g_k\|^2 &= \frac{1}{\pi} \int_{-\pi}^{\pi} \sin^2 kx \, \mathrm{d}x = \frac{1}{2\pi} \int_{-\pi}^{\pi} (1 - \cos 2kx) \, \mathrm{d}x = 1. \end{aligned}$$

We now show that the remaining vectors are pairwise orthogonal. For $j, k \in \{1, \ldots, n\}$ with $j \neq k$, we have that $j - k$ and $j + k$ are nonzero integers. Note that $\sin mx = 0$ for all integers m when $x = \pm\pi$. Then

$$\begin{aligned} \langle f_j, f_k \rangle &= \frac{1}{\pi} \int_{-\pi}^{\pi} \cos jx \cos kx \, \mathrm{d}x \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} (\cos(j-k)x + \cos(j+k)x) \, \mathrm{d}x \\ &= \frac{1}{2\pi} \left(\frac{\sin(j-k)x}{j-k} + \frac{\sin(j+k)x}{j+k} \right) \Big|_{-\pi}^{\pi} = 0, \\ \langle g_j, g_k \rangle &= \frac{1}{\pi} \int_{-\pi}^{\pi} \sin jx \sin kx \, \mathrm{d}x \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} (\cos(j-k)x - \cos(j+k)x) \, \mathrm{d}x \\ &= \frac{1}{2\pi} \left(\frac{\sin(j-k)x}{j-k} - \frac{\sin(j+k)x}{j+k} \right) \Big|_{-\pi}^{\pi} = 0. \end{aligned}$$

For $j, k = 1, \ldots, n$, note that $f_j g_k$ is a product of an even function and an odd function, so $\langle f_j, g_k \rangle = 0$. Thus, the list of vectors is orthonormal. ■

Exercise 6B5

Suppose $f : [-\pi, \pi] \to \mathbf{R}$ is continuous. For each nonnegative integer k, define

$$a_k = \frac{1}{\sqrt{\pi}} \int_{-\pi}^{\pi} f(x) \cos(kx) \, \mathrm{d}x \quad \text{and} \quad b_k = \frac{1}{\sqrt{\pi}} \int_{-\pi}^{\pi} f(x) \sin(kx) \, \mathrm{d}x.$$

Prove that

$$\frac{a_0^2}{2} + \sum_{k=1}^{\infty} (a_k^2 + b_k^2) \leq \int_{-\pi}^{\pi} f^2.$$

Axler Note. *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. The list of vectors in Exercise 6B4 is an orthonormal list in $C[-\pi, \pi]$. Using the same naming conventions of f_0, f_k, g_k as in Exercise 6B4, we have that

$$\langle f, f_0 \rangle = \frac{a_0}{\sqrt{2}}, \quad \langle f, f_k \rangle = a_k, \quad \langle f, g_k \rangle = b_k.$$

By Bessel's inequality, for every $n \geq 1$,

$$\frac{a_0^2}{2} + \sum_{k=1}^{n}(a_k^2 + b_k^2) = |\langle f, f_0\rangle|^2 + \sum_{k=1}^{n}(|\langle f, f_k\rangle|^2 + |\langle f, g_k\rangle|^2) \leq \|f\|^2 = \int_{-\pi}^{\pi} f^2.$$

Since the right side is independent of n, taking $n \to \infty$ gives the result. ■

Exercise 6B6

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.

Axler Note. *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) Suppose $v_1, \ldots, v_n$ are linearly dependent. Then there exist scalars $a_1, \ldots, a_n$, not all zero, such that

$$a_1 v_1 + \cdots + a_n v_n = 0.$$

Then

$$\begin{aligned}
\sqrt{\sum_{k=1}^{n}|a_k|^2} &= \left\|\sum_{k=1}^{n} a_k e_k\right\| \\
&= \left\|\sum_{k=1}^{n} a_k (e_k - v_k)\right\| \\
&\leq \sum_{k=1}^{n} |a_k|\|e_k - v_k\| \\
&< \frac{1}{\sqrt{n}}\sum_{k=1}^{n}|a_k| \\
&\leq \frac{1}{\sqrt{n}}\sqrt{n}\sqrt{\sum_{k=1}^{n}|a_k|^2} \\
&= \sqrt{\sum_{k=1}^{n}|a_k|^2}.
\end{aligned}$$

This results in $1 < 1$, which is a contradiction, so $v_1, \ldots, v_n$ must be linearly independent. Since $\dim V = n$, it follows that $v_1, \ldots, v_n$ is a basis of V.

(b) Consider the vector

$$u = \frac{1}{\sqrt{n}}(e_1 + \cdots + e_n), \quad \|u\|^2 = \frac{1}{n}\|e_1 + \cdots + e_n\|^2 = \frac{1}{n}(n) = 1.$$

For each $k = 1, \dots, n$, define

$$v_k = e_k - \frac{1}{\sqrt{n}}u \implies \|e_k - v_k\| = \|\frac{1}{\sqrt{n}}u\| = \frac{1}{\sqrt{n}}\|u\| = \frac{1}{\sqrt{n}}.$$

Then

$$v_1 + \cdots + v_n = (e_1 - \frac{1}{\sqrt{n}}u) + \cdots + (e_n - \frac{1}{\sqrt{n}}u) = (e_1 + \cdots + e_n) - \frac{n}{\sqrt{n}}u = (e_1 + \cdots + e_n) - \sqrt{n}u = 0,$$

so $v_1, \dots, v_n$ is linearly dependent. ■

Exercise 6B7

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. Using the Gram–Schmidt procedure, we can produce an orthonormal basis of $\mathbf{R}^3$ from the given basis. Let $v_1 = (1,0,0), v_2 = (1,1,1), v_3 = (1,1,2)$. Set $f_1 = v_1 = (1,0,0)$, so $\|f_1\|^2 = 1^2 + 0^2 + 0^2 = 1$. Then

$$f_2 = v_2 - \frac{\langle v_2, f_1\rangle}{\|f_1\|^2} f_1 = (1,1,1) - \frac{1}{1}(1,0,0) = (0,1,1)$$

with $\|f_2\|^2 = 0^2 + 1^2 + 1^2 = 2$. Finally,

$$f_3 = v_3 - \frac{\langle v_3, f_1\rangle}{\|f_1\|^2} f_1 - \frac{\langle v_3, f_2\rangle}{\|f_2\|^2} f_2 = (1,1,2) - (1,0,0) - \frac{3}{2}(0,1,1) = \left(0, -\frac{1}{2}, \frac{1}{2}\right)$$

with $\|f_3\|^2 = 0^2 + (-\frac{1}{2})^2 + (\frac{1}{2})^2 = \frac{1}{2}$. Then the orthonormal basis of $\mathbf{R}^3$ is

$$e_1 = \frac{f_1}{\|f_1\|} = (1,0,0), \quad e_2 = \frac{f_2}{\|f_2\|} = \frac{1}{\sqrt{2}}(0,1,1), \quad e_3 = \frac{f_3}{\|f_3\|} = \sqrt{2}\left(0, -\frac{1}{2}, \frac{1}{2}\right) = \left(0, -\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right).$$

By the Gram–Schmidt procedure, we know $\operatorname{span}(e_1, \dots, e_k) = \operatorname{span}(v_1, \dots, v_k)$ for each $k = 1, 2, 3$. Since $\mathcal{M}(T, (v_1, v_2, v_3))$ is upper-triangular, $\operatorname{span}(v_1, \dots, v_k)$ is invariant under T for each $k = 1, 2, 3$. Hence, $\operatorname{span}(e_1, \dots, e_k)$ is invariant under T for each $k = 1, 2, 3$, so $\mathcal{M}(T, (e_1, e_2, e_3))$ is upper-triangular. ■

Exercise 6B8

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) Let $v_1 = 1, v_2 = x, v_3 = x^2$. Then $f_1 = v_1 = 1$, so $\|f_1\|^2 = \int_0^1 1^2 = 1$. Next,

$$f_2 = v_2 - \frac{\langle v_2, f_1\rangle}{\|f_1\|^2} f_1 = x - \int_0^1 x \cdot 1 = x - \frac{1}{2}$$

with $\|f_2\|^2 = \int_0^1 (x - \frac{1}{2})^2 \, dx = \frac{1}{12}$. Finally,

$$f_3 = v_3 - \frac{\langle v_3, f_1 \rangle}{\|f_1\|^2} f_1 - \frac{\langle v_3, f_2 \rangle}{\|f_2\|^2} f_2 = x^2 - \int_0^1 x^2 \cdot 1 \, dx - 12 \left(\int_0^1 x^2 \left(x - \frac{1}{2} \right) dx \right) \left(x - \frac{1}{2} \right) = x^2 - x + \frac{1}{6}$$

with $\|f_3\|^2 = \int_0^1 (x^2 - x + \frac{1}{6})^2 \, dx = \frac{1}{180}$. Then the orthonormal basis of $\mathcal{P}_2(\mathbf{R})$ is

$$e_1 = \frac{f_1}{\|f_1\|} = 1, \quad e_2 = \frac{f_2}{\|f_2\|} = 2\sqrt{3}\left(x - \frac{1}{2} \right), \quad e_3 = \frac{f_3}{\|f_3\|} = 6\sqrt{5}\left(x^2 - x + \frac{1}{6} \right).$$

(b) The differentiation operator D on $\mathcal{P}_2(\mathbf{R})$ has the following matrix with respect to the basis $1, x, x^2$:

$$\mathcal{M}(D, (1, x, x^2)) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix}.$$

Then

$$\begin{aligned} De_1 &= D(1) = 0, \\ De_2 &= D\left(2\sqrt{3}\left(x - \frac{1}{2} \right) \right) = 2\sqrt{3} = 2\sqrt{3} e_1, \\ De_3 &= D\left(6\sqrt{5}\left(x^2 - x + \frac{1}{6} \right) \right) = 6\sqrt{5}(2x - 1) = 6\sqrt{5} \cdot \frac{1}{\sqrt{3}} e_2 = 2\sqrt{15} e_2 \end{aligned}$$

Hence, the matrix of D with respect to the orthonormal basis e_1, e_2, e_3 is

$$\mathcal{M}(D, (e_1, e_2, e_3)) = \begin{pmatrix} 0 & 2\sqrt{3} & 0 \\ 0 & 0 & 2\sqrt{15} \\ 0 & 0 & 0 \end{pmatrix},$$

which is upper triangular, as expected. ■

Exercise 6B9

Suppose $e_1, \ldots, e_m$ is the result of applying the Gram–Schmidt procedure to a linearly independent list $v_1, \ldots, v_m$ in V. Prove that $\langle v_k, e_k \rangle > 0$ for each $k = 1, \ldots, m$.

Solution. By the Gram–Schmidt procedure, we have

$$e_k = \frac{f_k}{\|f_k\|}, \quad f_k = v_k - \sum_{j=1}^{k-1} \langle v_k, e_j \rangle e_j.$$

From the linear independence of $v_1, \ldots, v_m$, $v_k \notin \operatorname{span}(v_1, \ldots, v_{k-1}) = \operatorname{span}(e_1, \ldots, e_{k-1})$, so $f_k \neq 0$ and $\|f_k\| > 0$. Then by orthogonality of f_k and e_j for $j < k$, we have

$$\langle v_k, e_k \rangle = \frac{1}{\|f_k\|} \langle v_k, f_k \rangle = \frac{1}{\|f_k\|} \left\langle f_k + \sum_{j=1}^{k-1} \langle v_k, e_j \rangle e_j, f_k \right\rangle = \frac{1}{\|f_k\|} \langle f_k, f_k \rangle = \|f_k\| > 0.$$

■

Exercise 6B10

Suppose $v_1, \dots, v_m$ 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 $e_1, \dots, e_m$ in V such that $\langle v_k, e_k \rangle > 0$ and $\operatorname{span}(v_1, \dots, v_k) = \operatorname{span}(e_1, \dots, e_k)$ for each $k = 1, \dots, m$.

Axler Note. *The result in this exercise is used in the proof of 7.58.*

Solution. Suppose $e_1, \dots, e_m$ is an orthonormal list in V such that $\langle v_k, e_k \rangle > 0$ and $\operatorname{span}(v_1, \dots, v_k) = \operatorname{span}(e_1, \dots, e_k)$ for each $k = 1, \dots, m$. We induce on k to show that e_k agrees with the Gram–Schmidt output.

For $k = 1$, we have $\operatorname{span}(v_1) = \operatorname{span}(e_1)$ and $\|e_1\| = 1$, so $v_1 = \langle v_1, e_1 \rangle e_1$. Taking norms and using $\langle v_1, e_1 \rangle > 0$ gives $\langle v_1, e_1 \rangle = \|v_1\|$, so $e_1 = v_1 / \|v_1\|$, which is the Gram–Schmidt output.

Now suppose the claim holds for $e_1, \dots, e_k$. Since $v_{k+1} \in \operatorname{span}(e_1, \dots, e_{k+1})$ and $e_1, \dots, e_{k+1}$ is orthonormal,

$$v_{k+1} = \sum_{j=1}^{k+1} \langle v_{k+1}, e_j \rangle e_j \implies \langle v_{k+1}, e_{k+1} \rangle e_{k+1} = v_{k+1} - \sum_{j=1}^{k} \langle v_{k+1}, e_j \rangle e_j.$$

By the induction hypothesis, the right-hand side is the Gram–Schmidt vector f_{k+1}. Taking norms and using $\langle v_{k+1}, e_{k+1} \rangle > 0$ gives $\langle v_{k+1}, e_{k+1} \rangle = \|f_{k+1}\|$, so $e_{k+1} = f_{k+1} / \|f_{k+1}\|$, which is the Gram–Schmidt output. ∎

Exercise 6B11

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. The map $p \mapsto p(\frac{1}{2})$ is a linear functional on $\mathcal{P}_2(\mathbf{R})$. By the Riesz representation theorem, there exists a unique $q \in \mathcal{P}_2(\mathbf{R})$ such that $p(\frac{1}{2}) = \langle p, q \rangle$ for every $p \in \mathcal{P}_2(\mathbf{R})$, given by

$$q = e_1\left(\tfrac{1}{2}\right) e_1 + e_2\left(\tfrac{1}{2}\right) e_2 + e_3\left(\tfrac{1}{2}\right) e_3,$$

where e_1, e_2, e_3 is the orthonormal basis from Exercise 6B8. Evaluating e_1, e_2, e_3 at $\frac{1}{2}$ gives

$$e_1(\tfrac{1}{2}) = 1, \quad e_2(\tfrac{1}{2}) = 2\sqrt{3}\left(\frac{1}{2} - \frac{1}{2}\right) = 0, \quad e_3(\tfrac{1}{2}) = 6\sqrt{5}\left(\frac{1}{4} - \frac{1}{2} + \frac{1}{6}\right) = -\frac{\sqrt{5}}{2}.$$

Hence, the polynomial q is

$$q = e_1 - \frac{\sqrt{5}}{2} e_3 = 1 - 15\left(x^2 - x + \frac{1}{6}\right) = -15x^2 + 15x - \frac{3}{2}.$$ ∎

Exercise 6B12

Find a polynomial $q \in \mathcal{P}_2(\mathbf{R})$ such that

$$\int_0^1 p(x) \cos(\pi x)\, \mathrm{d}x = \int_0^1 pq$$

for every $p \in \mathcal{P}_2(\mathbf{R})$.

Solution. By the Riesz representation theorem, there exists a unique $q \in \mathcal{P}_2(\mathbf{R})$ such that

$$\int_0^1 p(x) \cos(\pi x)\, \mathrm{d}x = \langle p, q \rangle = \int_0^1 pq$$

for every $p \in \mathcal{P}_2(\mathbf{R})$, given by

$$q = \langle e_1, \cos(\pi x)\rangle e_1 + \langle e_2, \cos(\pi x)\rangle e_2 + \langle e_3, \cos(\pi x)\rangle e_3,$$

where e_1, e_2, e_3 is the orthonormal basis from Exercise 6B8. Evaluating the inner products gives

$$\begin{aligned}
\langle e_1, \cos(\pi x)\rangle &= \int_0^1 e_1 \cos(\pi x) = \int_0^1 \cos(\pi x) = 0,\\
\langle e_2, \cos(\pi x)\rangle &= \int_0^1 e_2 \cos(\pi x) = \int_0^1 2\sqrt{3}\left(x - \frac{1}{2}\right)\cos(\pi x) = \frac{-4\sqrt{3}}{\pi^2},\\
\langle e_3, \cos(\pi x)\rangle &= \int_0^1 e_3 \cos(\pi x) = \int_0^1 6\sqrt{5}\left(x^2 - x + \frac{1}{6}\right)\cos(\pi x) = 0.
\end{aligned}$$

Hence,

$$q = \frac{-4\sqrt{3}}{\pi^2} e_2 = \frac{-4\sqrt{3}}{\pi^2} \cdot 2\sqrt{3}\left(x - \frac{1}{2}\right) = \frac{-24}{\pi^2}\left(x - \frac{1}{2}\right). \qquad \blacksquare$$

Exercise 6B13

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 $k \in \{1, \ldots, m\}$.

Axler Note. *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. Let f_k, e_k be the vectors obtained from the Gram–Schmidt procedure, with the process continuing as long as $f_k \neq 0$. If $v_1, \ldots, v_m$ is linearly dependent, take the smallest k with $v_k \in \operatorname{span}(v_1, \ldots, v_{k-1})$. Then $v_1, \ldots, v_{k-1}$ is linearly independent, so $e_1, \ldots, e_{k-1}$ is an orthonormal basis of that span. Hence,

$$v_k = \sum_{j=1}^{k-1} \langle v_k, e_j\rangle e_j \implies f_k = v_k - \sum_{j=1}^{k-1} \langle v_k, e_j\rangle e_j = 0.$$

Conversely, if the procedure gives $f_k = 0$ for some minimal k, then

$$v_k = \sum_{j=1}^{k-1} \langle v_k, e_j\rangle e_j \in \operatorname{span}(e_1, \ldots, e_{k-1}) = \operatorname{span}(v_1, \ldots, v_{k-1}),$$

so $v_1, \ldots, v_k$ is a linearly dependent list. $\blacksquare$

Exercise 6B14

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 all $k \in \{1, \ldots, m\}$.

Solution. By Exercise 6B10, there is a unique orthonormal list $e_1, \ldots, e_m$ such that $\operatorname{span}(v_1, \ldots, v_k) = \operatorname{span}(e_1, \ldots, e_k)$ and $\langle v_k, e_k\rangle > 0$ for all k. Let $e_1', \ldots, e_m'$ be any orthonormal list satisfying the same span condition. For each k, the vector

e'_k lies in $\operatorname{span}(e_1, \ldots, e_k)$ and is orthogonal to $\operatorname{span}(e'_1, \ldots, e'_{k-1}) = \operatorname{span}(e_1, \ldots, e_{k-1})$. Now, write e'_k as

$$e'_k = \sum_{j=1}^{k} c_j e_j$$

for some scalars $c_1, \ldots, c_k$. Then $\langle e'_k, e_j \rangle = 0$ for $j < k$ gives $c_j = 0$ for $j < k$, so $e'_k = c_k e_k$. Since $\|e'_k\| = 1$ and V is real, we have $e'_k = \pm e_k$. These m sign choices are independent, giving exactly 2^m such lists. ∎

Exercise 6B15

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 $\langle u, v \rangle_1 = c \langle u, v \rangle_2$ for every $u, v \in V$.

Axler Note. *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 the result holds trivially. Otherwise, assume $V \neq \{0\}$, and let $e_1 \in V$ such that $\|e_1\|_2 = 1$. Let $c = \|e_1\|_1^2$. For any $u, v \in V$, $U = \operatorname{span}(u, v, e_1)$ is finite-dimensional. Extend e_1 to an orthonormal basis $e_1, \ldots, e_n$ of U with respect to $\langle \cdot, \cdot \rangle_2$. Since $\langle e_j, e_k \rangle_2 = 0$ for $j \neq k$, the same holds for $\langle \cdot, \cdot \rangle_1$.

For each $j > 1$, $\langle e_1 + e_j, e_1 - e_j \rangle_2 = 0$, so $\langle e_1 + e_j, e_1 - e_j \rangle_1 = 0$. Expanding gives

$$0 = \langle e_1 + e_j, e_1 - e_j \rangle_1 = \langle e_1, e_1 \rangle_1 - \langle e_j, e_j \rangle_1 = c - \langle e_j, e_j \rangle_1,$$

so $\langle e_j, e_j \rangle_1 = c > 0$ for all j. Write $u = a_1 e_1 + \cdots + a_n e_n$ and $v = b_1 e_1 + \cdots + b_n e_n$ for some scalars $a_1, \ldots, a_n, b_1, \ldots, b_n$. Then

$$\langle u, v \rangle_1 = \sum_{j=1}^{n} a_j \overline{b_j} \langle e_j, e_j \rangle_1 = c \sum_{j=1}^{n} a_j \overline{b_j} = c \langle u, v \rangle_2.$$ ∎

Exercise 6B16

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 $\|v\|_1 \leq c \|v\|_2$ for every $v \in V$.

Solution. If $V = \{0\}$, then the result holds trivially. Otherwise, assume $V \neq \{0\}$, and let $e_1, \ldots, e_n$ be an orthonormal basis of V with respect to $\langle \cdot, \cdot \rangle_2$. For any $v \in V$, write $v = a_1 e_1 + \cdots + a_n e_n$ for some scalars $a_1, \ldots, a_n$. Then

$$\|v\|_1 \leq \sum_{j=1}^{n} |a_j| \|e_j\|_1 \leq \left(\sum_{j=1}^{n} |a_j|^2 \right)^{1/2} \left(\sum_{j=1}^{n} \|e_j\|_1^2 \right)^{1/2} = \|v\|_2 \left(\sum_{j=1}^{n} \|e_j\|_1^2 \right)^{1/2}.$$

Taking $c = \left(\sum_{j=1}^{n} \|e_j\|_1^2 \right)^{1/2} > 0$ gives the desired inequality. ∎

Exercise 6B17

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 $\|Tv\| \leq \|v\|$ for all $v \in V$, then T is the identity operator.

Solution. By Schur's theorem, there exists an orthonormal basis $e_1, \ldots, e_n$ of V such that $\mathcal{M}(T)$ is upper-triangular.

Since 1 is the only eigenvalue of T, the diagonal entries are all 1, so

$$Te_k = e_k + \sum_{i=1}^{k-1} a_{i,k} e_i$$

for each k. By orthonormality,

$$\|Te_k\|^2 = \left\| e_k + \sum_{i=1}^{k-1} a_{i,k} e_i \right\|^2 = 1 + \sum_{i=1}^{k-1} |a_{i,k}|^2.$$

Since $\|Te_k\| \le \|e_k\| = 1$ by hypothesis, we have $\sum_{i=1}^{k-1} |a_{i,k}|^2 = 0$, forcing each $a_{i,k} = 0$. Thus, $Te_k = e_k$ for all k, so $T = I$. ■

Exercise 6B18

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 $U = \operatorname{span}(u_1, \dots, u_m)$. Since $u_1, \dots, u_m$ is a basis of U, there exists a unique linear functional $f : U \to \mathbf{F}$ satisfying $f(u_k) = 1$ for all $k \in \{1, \dots, m\}$. Because U is a finite-dimensional inner product space (with the inner product inherited from V), the Riesz representation theorem provides a unique $v \in U$ such that $f(u) = \langle u, v \rangle$ for all $u \in U$. In particular, $\langle u_k, v \rangle = f(u_k) = 1$ for all k. ■

Exercise 6B19

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. For each $k \in \{1, \dots, n\}$, since $v_1, \dots, v_n$ is a basis of V, there is a unique linear functional $f_k : V \to \mathbf{F}$ satisfying $f_k(v_j) = \delta_{jk}$ for all $j \in \{1, \dots, n\}$. By the Riesz representation theorem, there exists a unique $u_k \in V$ such that $f_k(v) = \langle v, u_k \rangle$ for all $v \in V$. In particular, $\langle v_j, u_k \rangle = f_k(v_j) = \delta_{jk}$. The list $u_1, \dots, u_n$ is linearly independent because if $c_1 u_1 + \cdots + c_n u_n = 0$, then for each j,

$$0 = \left\langle v_j, \sum_{k=1}^{n} c_k u_k \right\rangle = \overline{c_j},$$

which implies $c_j = 0$ for all j. Hence, $u_1, \dots, u_n$ is a basis of V with the desired properties. ■

Exercise 6B20

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.

Axler Note. *This exercise strengthens Exercise 5E9(b) (in the context of inner product spaces) by asserting that the basis in that exercise can be chosen to be orthonormal.*

Solution. By Exercise 5E9(b), there exists a basis $v_1, \dots, v_n$ of V such that every element of $\mathcal{E}$ has an upper-triangular matrix with respect to this basis. Equivalently, every $T \in \mathcal{E}$ satisfies $T(\operatorname{span}(v_1, \dots, v_k)) \subseteq \operatorname{span}(v_1, \dots, v_k)$ for each $k \in \{1, \dots, n\}$.

Apply the Gram–Schmidt procedure to $v_1, \dots, v_n$ to obtain an orthonormal basis $e_1, \dots, e_n$ of V. By the key property of Gram–Schmidt,

$$\operatorname{span}(e_1, \dots, e_k) = \operatorname{span}(v_1, \dots, v_k)$$

for each k. Hence, every $T \in \mathcal{E}$ still satisfies $T(\text{span}(e_1, \ldots, e_k)) \subseteq \text{span}(e_1, \ldots, e_k)$ for each k, which means every element of $\mathcal{E}$ has an upper-triangular matrix with respect to the orthonormal basis $e_1, \ldots, e_n$. ■

Exercise 6B21

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 $v \in V$.

Manual Remark. *This exercise is quite difficult with a long solution. A high level overview is this: Schur's theorem gives an orthonormal basis of V with respect to which $\mathcal{M}(T)$ is upper-triangular. This reduces the problem to showing each basis vector is contracted by T^m for sufficiently large m, which is done by induction on the index of the basis vector. The final step is to use the triangle inequality to show that T^m contracts every vector in V.*

Solution. If $V = \{0\}$, then the result holds trivially. Otherwise, assume $V \neq \{0\}$, and let $\dim V = n \geq 1$. By Schur's theorem, there exists an orthonormal basis $e_1, \ldots, e_n$ of V such that $\mathcal{M}(T)$ is upper-triangular whose diagonal entries are the eigenvalues of T. Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of T, so $|\lambda_j| < 1$ for each j. Write $Te_j = \lambda_j e_j + v_j$ for each j, where $v_j \in \text{span}(e_1, \ldots, e_{j-1})$ and $v_1 = 0$. Let

$$\lambda = \max_j |\lambda_j| < 1,$$

and choose α such that $\lambda < \alpha < 1$.

We prove by induction on j that there exists some $C_j > 0$ such that $\|T^k e_j\| \leq C_j \alpha^k$ for all $k \geq 0$. For the base case $j = 1$, we have $Te_1 = \lambda_1 e_1$, which gives

$$\|T^k e_1\| = |\lambda_1|^k \leq \lambda^k < \alpha^k,$$

so we can take $C_1 = 1$.

Assume the inductive hypothesis holds for all indices less than j. Consider

$$v_j = \sum_{i=1}^{j-1} a_i e_i.$$

By the inductive hypothesis, $\|T^k e_i\| \leq C_i \alpha^k$ for all $i < j$, so

$$\|T^k v_j\| \leq \sum_{i=1}^{j-1} |a_i| \, \|T^k e_i\| \leq B\alpha^k, \quad B = \sum_{i=1}^{j-1} |a_i| C_i.$$

Repeatedly applying T where $Te_j = \lambda_j e_j + v_j$ gives

$$T^k e_j = \lambda_j^k e_j + \sum_{i=0}^{k-1} \lambda_j^{k-1-i} T^i v_j.$$

Taking norms and using the triangle inequality gives

$$\|T^k e_j\| \leq \lambda^k + B \sum_{i=0}^{k-1} \lambda^{k-1-i} \alpha^i = \lambda^k + B \frac{\alpha^k - \lambda^k}{\alpha - \lambda} < \left(1 + \frac{B}{\alpha - \lambda}\right) \alpha^k,$$

so we can take $C_j = 1 + \frac{B}{\alpha - \lambda}$ to complete the inductive step.

Let $C = \max\{C_1, \ldots, C_n\}$. For any $v \in V$, write $v = a_1 e_1 + \cdots + a_n e_n$ for some scalars $a_1, \ldots, a_n$. Then $|a_j| \leq \|v\|$ by orthonormality, so

$$\|T^k v\| \leq \sum_{j=1}^{n} |a_j| \, \|T^k e_j\| \leq nC\alpha^k \|v\|.$$

Choosing such m that $nC\alpha^m \le \varepsilon$ gives $\|T^m v\| \le \varepsilon\|v\|$ for every $v \in V$. ■

Exercise 6B22

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 $f \in C[-1,1]$.

Axler Note. *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 for contradiction that there exists $g \in C[-1,1]$ such that $f(0) = \int_{-1}^{1} fg$ for every $f \in C[-1,1]$. Take $f(x) = x^2 g(x)$, which is continuous. Then

$$0 = f(0) = \int_{-1}^{1} x^2 (g(x))^2 \, dx.$$

The integrand is continuous and nonnegative, so $x^2(g(x))^2 = 0$ for all $x \in [-1,1]$. Hence, $g(x) = 0$ for all $x \in [-1,1] \setminus \{0\}$, and continuity forces $g(0) = 0$, so $g \equiv 0$. But then $f(0) = 0$ for every $f \in C[-1,1]$, which is false (take $f = 1$), a contradiction. Hence, no such g exists. ■

Exercise 6B23

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).

Axler Note. *This exercise requires familiarity with metric spaces.*

Solution.

(a) We verify the metric axioms for d:

- **Non-negativity and Identity**: For all $u, v \in V$, $\|u - v\| \ge 0$ and $\|u - v\| = 0$ if and only if $u = v$.
- **Symmetry**: For all $u, v \in V$, we have

$$d(u,v) = \|u - v\| = \|-(v - u)\| = \|v - u\| = d(v,u).$$

- **Triangle Inequality**: For all $u, v, w \in V$, we have

$$d(u,w) = \|u - w\| = \|(u - v) + (v - w)\| \le \|u - v\| + \|v - w\| = d(u,v) + d(v,w).$$

Therefore, (V, d) is a metric space.

(b) If V is finite-dimensional, let $e_1, \ldots, e_n$ be an orthonormal basis of V. A Cauchy sequence $(v_k)_{k\ge1}$ in V can be

written as $v_k = a_{k,1}e_1 + \cdots + a_{k,n}e_n$ for scalars $a_{k,j}$. By the Cauchy–Schwarz inequality, for each j,

$$|a_{k,j} - a_{m,j}| = |\langle v_k - v_m, e_j \rangle| \le \|v_k - v_m\|,$$

so each coordinate sequence $(a_{k,j})_{k\ge1}$ is Cauchy in $\mathbf{F}$. Since $\mathbf{F}$ is complete, each $(a_{k,j})$ converges to some $a_j \in \mathbf{F}$. Define $v = a_1e_1 + \cdots + a_ne_n \in V$. Since $e_1, \ldots, e_n$ is orthonormal,

$$\|v_k - v\|^2 = \sum_{j=1}^{n} |a_{k,j} - a_j|^2 \to 0,$$

so $v_k \to v$ in V, and (V, d) is complete.

(c) Let U be a finite-dimensional subspace of V, and let $(u_k)_{k\ge1}$ be a sequence in U converging to some $v \in V$. Since (u_k) converges, it is Cauchy. Restricting the inner product on V to U, we see that (u_k) is also a Cauchy sequence in U. By part (b), U is complete, so (u_k) converges to some $u \in U$. Since (u_k) converges to v in V and to u in U, we have $v = u \in U$. Therefore, U is closed in V. ∎

6C Orthogonal Complements and Minimization Problems

Exercise 6C1

Suppose $v_1, \dots, v_m \in V$. Prove that

$$\{v_1, \dots, v_m\}^\perp = (\operatorname{span}(v_1, \dots, v_m))^\perp.$$

Solution. Suppose $v \in \{v_1, \dots, v_m\}^\perp$. Then $\langle v, v_k \rangle = 0$ for each $k = 1, \dots, m$. For any $u \in \operatorname{span}(v_1, \dots, v_m)$, we can write $u = a_1 v_1 + \cdots + a_m v_m$ for some scalars $a_1, \dots, a_m$. Then

$$\langle v, u \rangle = \sum_{k=1}^{m} \overline{a_k} \langle v, v_k \rangle = 0,$$

so $v \in (\operatorname{span}(v_1, \dots, v_m))^\perp$.

Conversely, suppose $v \in (\operatorname{span}(v_1, \dots, v_m))^\perp$. Since $\{v_1, \dots, v_m\} \subseteq \operatorname{span}(v_1, \dots, v_m)$, Theorem 6.48(e) shows that $\operatorname{span}(v_1, \dots, v_m)^\perp \subseteq \{v_1, \dots, v_m\}^\perp$, so $v \in \{v_1, \dots, v_m\}^\perp$. ∎

Exercise 6C2

Suppose U is a subspace of V with basis $u_1, \dots, u_m$ and

$$u_1, \dots, u_m, v_1, \dots, v_n$$

is a basis of V. Prove that if the Gram–Schmidt procedure is applied to the basis of V above, producing a list $e_1, \dots, e_m, f_1, \dots, f_n$, then $e_1, \dots, e_m$ is an orthonormal basis of U and $f_1, \dots, f_n$ is an orthonormal basis of $U^\perp$.

Solution. Let $e_1, \dots, e_m, f_1, \dots, f_n$ be the result of applying the Gram–Schmidt procedure to the basis $u_1, \dots, u_m, v_1, \dots, v_n$ of V. By the properties of the Gram–Schmidt procedure, $\operatorname{span}(e_1, \dots, e_k) = \operatorname{span}(u_1, \dots, u_k)$ for each $k \le m$. In particular, $\operatorname{span}(e_1, \dots, e_m) = \operatorname{span}(u_1, \dots, u_m) = U$, so $e_1, \dots, e_m$ is an orthonormal basis of U.

Since each f_j is constructed to be orthogonal to all previous vectors in the Gram–Schmidt process, it follows that each f_j is orthogonal to every e_k for $k = 1, \dots, m$. Hence, $f_1, \dots, f_n \in U^\perp$. Additionally, $f_1, \dots, f_n$ is an orthonormal list in $U^\perp$ because the Gram–Schmidt procedure produces orthonormal vectors. By Theorem 6.49, $V = U \oplus U^\perp$, so $\dim U + \dim U^\perp = \dim V$. Since $\dim U = m$ and $\dim V = m + n$, we have $\dim U^\perp = n$. Therefore, $f_1, \dots, f_n$ is an orthonormal basis of $U^\perp$. ∎

Exercise 6C3

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. Let $v_1 = (1, 2, 3, -4)$ and $v_2 = (-5, 4, 3, 2)$. Applying the Gram–Schmidt procedure to v_1 and v_2, we first set $f_1 = v_1$ with $\|f_1\|^2 = 30$. Then

$$f_2 = v_2 - \frac{\langle v_2, f_1 \rangle}{\|f_1\|^2} f_1 = (-5, 4, 3, 2) - \frac{4}{30}(1, 2, 3, -4) = \frac{1}{15}(-77, 56, 39, 38),$$

where

$$\|f_2\|^2 = \frac{1}{225}((-77)^2 + 56^2 + 39^2 + 38^2) = \frac{12030}{225}.$$

Then an orthonormal basis of U is given by

$$e_1 = \frac{1}{\sqrt{30}}(1,2,3,-4), \quad e_2 = \frac{1}{\sqrt{12030}}(-77,56,39,38).$$

An orthonormal basis for $U^\perp$ consists of vectors orthogonal to both v_1 and v_2. To find such vectors, consider the system of equations

$$\begin{cases} x + 2y + 3z - 4w = 0 \\ -5x + 4y + 3z + 2w = 0 \end{cases}$$

The first equation yields $x = -2y - 3z + 4w$, and substituting into the second equation gives

$$-5(-2y - 3z + 4w) + 4y + 3z + 2w = 0 \implies 14y + 18z - 18w = 0 \implies 7y = -9z + 9w.$$

Choosing $z = w = 1$ gives $y = 0$, and then $x = -2(0) - 3(1) + 4(1) = 1$. Choosing $z = 0$ and $w = 7$ gives $y = 9$ and then $x = -2(9) - 3(0) + 4(7) = -18 + 28 = 10$. Hence, we have two vectors $u_1 = (1,0,1,1)$ and $u_2 = (10,9,0,7)$ in $U^\perp$. Applying the Gram–Schmidt procedure to u_1 and u_2, we first set $g_1 = u_1$ with $\|g_1\|^2 = 3$. Then

$$g_2 = u_2 - \frac{\langle u_2, g_1\rangle}{\|g_1\|^2} g_1 = (10,9,0,7) - \frac{17}{3}(1,0,1,1) = \frac{1}{3}(13,27,-17,4),$$

where

$$\|g_2\|^2 = \frac{1}{9}(13^2 + 27^2 + (-17)^2 + 4^2) = \frac{1203}{9}.$$

Then an orthonormal basis of $U^\perp$ is given by

$$e_1' = \frac{1}{\sqrt{3}}(1,0,1,1), \quad e_2' = \frac{1}{\sqrt{1203}}(13,27,-17,4). \qquad \blacksquare$$

Exercise 6C4

Suppose $e_1, \ldots, e_n$ is a list of vectors in V with $\|e_k\| = 1$ for each $k = 1, \ldots, n$ and

$$\|v\|^2 = |\langle v, e_1\rangle|^2 + \cdots + |\langle v, e_n\rangle|^2$$

for all $v \in V$. Prove that $e_1, \ldots, e_n$ is an orthonormal basis of V.
Axler Note. *This exercise provides a converse to 6.30(b).*

Solution. Setting $v = e_j$, we have

$$1 = \|e_j\|^2 = |\langle e_j, e_1\rangle|^2 + \cdots + |\langle e_j, e_n\rangle|^2 = |\langle e_j, e_j\rangle|^2 + \sum_{k \neq j} |\langle e_j, e_k\rangle|^2.$$

Since $\|e_j\| = 1$, we have $|\langle e_j, e_j\rangle|^2 = 1$, so

$$1 = 1 + \sum_{k \neq j} |\langle e_j, e_k\rangle|^2 \implies \sum_{k \neq j} |\langle e_j, e_k\rangle|^2 = 0,$$

which forces $|\langle e_j, e_k\rangle|^2 = 0$ for each $k \neq j$, so $\langle e_j, e_k\rangle = 0$ for each $j \neq k$. Hence, $e_1, \ldots, e_n$ is an orthonormal list. To show that $e_1, \ldots, e_n$ is a basis of V, let $v \in V$ be arbitrary. Consider

$$u = v - \sum_{k=1}^{n} \langle v, e_k\rangle e_k.$$

Then for each j:

$$\langle u, e_j \rangle = \langle v, e_j \rangle - \sum_{k=1}^{n} \langle v, e_k \rangle \langle e_k, e_j \rangle = \langle v, e_j \rangle - \langle v, e_j \rangle = 0$$

by orthonormality. Thus, u is orthogonal to each e_j, so by the given condition,

$$\|u\|^2 = |\langle u, e_1 \rangle|^2 + \cdots + |\langle u, e_n \rangle|^2 = 0,$$

which implies $u = 0$. Therefore, $v \in \operatorname{span}(e_1, \ldots, e_n)$. Since $e_1, \ldots, e_n$ is an orthonormal list and $\operatorname{span}(e_1, \ldots, e_n) = V$, it follows that $e_1, \ldots, e_n$ is an orthonormal basis of V. ■

Exercise 6C5

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. For every $v \in V$, $v = P_U v + P_{U^\perp} v$ by the definition of orthogonal projection. Then

$$P_{U^\perp} v = v - P_U v = Iv - P_U v = (I - P_U)v.$$

Hence, $P_{U^\perp} = I - P_U$. ■

Exercise 6C6

Suppose V is finite-dimensional and $T \in \mathcal{L}(V, W)$. Show that

$$T = T P_{(\operatorname{null} T)^\perp} = P_{\operatorname{range} T} T.$$

Solution. We first show that $T = T P_{(\operatorname{null} T)^\perp}$. For any $v \in V$, write $v = u + w$ for some $u \in \operatorname{null} T$ and $w \in (\operatorname{null} T)^\perp$. Then

$$T P_{(\operatorname{null} T)^\perp} v = T P_{(\operatorname{null} T)^\perp}(u + w) = T P_{(\operatorname{null} T)^\perp} u + T P_{(\operatorname{null} T)^\perp} w = T0 + Tw = Tw.$$

Since $u \in \operatorname{null} T$, we have $Tu = 0$, so $Tv = T(u + w) = Tu + Tw = 0 + Tw = Tw$. Hence, $T = T P_{(\operatorname{null} T)^\perp}$.

Next, we show that $T = P_{\operatorname{range} T} T$. For any $v \in V$, we have $Tv \in \operatorname{range} T$, so $P_{\operatorname{range} T} Tv = Tv$. Therefore, $T = P_{\operatorname{range} T} T$. ■

Exercise 6C7

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. $\langle x, y \rangle = 0$ for all $x \in X, y \in Y \iff Y \subseteq X^\perp \iff \operatorname{range} P_Y \subseteq \operatorname{null} P_X \iff P_X P_Y = 0$. ■

Exercise 6C8

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. By the Riesz representation theorem, we have $w \in U$. For any $u \in U$, we have

$$\langle u, v - w \rangle = \langle u, v \rangle - \langle u, w \rangle = \varphi(u) - \varphi(u) = 0.$$

Hence, $v - w \in U^\perp$, so $w = P_U v$. ∎

Exercise 6C9

Suppose V is finite-dimensional. Suppose $P \in \mathcal{L}(V)$ is such that $P^2 = P$ and every vector in null P is orthogonal to every vector in range P. Prove that there exists a subspace U of V such that $P = P_U$.

Solution. Let $U = \operatorname{range} P$. By Exercise 3B27, $P^2 = P$ implies $V = U \oplus \operatorname{null} P$. Since every $v \in \operatorname{null} P$ is orthogonal to every $u \in U$, we have $\operatorname{null} P \subseteq U^\perp$. From the fundamental theorem of linear maps,

$$\dim \operatorname{null} P = \dim V - \dim U = \dim U^\perp.$$

Hence, $\operatorname{null} P = U^\perp$, so $V = U \oplus U^\perp$. Then for any $x \in V$, we can write $x = u + w$ for some $u \in U$ and $w \in U^\perp$. Since $u \in \operatorname{range} P$ and $P^2 = P$, we have $Pu = u$. Then $Px = P(u + w) = Pu + Pw = u + 0 = u = P_U x$, so $P = P_U$. ∎

Exercise 6C10

Suppose V is finite-dimensional and $P \in \mathcal{L}(V)$ is such that $P^2 = P$ and

$$\|Pv\| \le \|v\|$$

for every $v \in V$. Prove that there exists a subspace U of V such that $P = P_U$.

Solution. Let $U = \operatorname{range} P$. From Exercise 3B27, $P^2 = P$ implies $V = U \oplus \operatorname{null} P$. Let $w \in \operatorname{null} P$ and $u \in U$. For every $\alpha \in \mathbf{F}$,

$$\|u\| = \|P(u + \alpha w)\| \le \|u + \alpha w\|.$$

By Exercise 6A6, $\langle u, w \rangle = 0$. Therefore, $\operatorname{null} P \subseteq U^\perp$. From the fundamental theorem of linear maps,

$$\dim \operatorname{null} P = \dim V - \dim U = \dim U^\perp.$$

Hence, $\operatorname{null} P = U^\perp$, so $V = U \oplus U^\perp$. For any $v \in V$, write $v = u + w$ with $u \in U$ and $w \in U^\perp = \operatorname{null} P$. Then $Pv = Pu + Pw = Pu = u = P_U v$, where $Pu = u$ since $u \in \operatorname{range} P$ and $P^2 = P$. Hence, $P = P_U$. ∎

Exercise 6C11

Suppose $T \in \mathcal{L}(V)$ and U is a finite-dimensional subspace of V. Prove that

$$U \text{ is invariant under } T \iff P_U T P_U = T P_U.$$

Solution. Suppose U is invariant under T. For any $v \in V$, we have $P_U v \in U$, so $T P_U v \in U$ by invariance. Since P_U acts as the identity on U, it follows that $P_U T P_U v = T P_U v$.

Conversely, suppose $P_U T P_U = T P_U$. For any $u \in U$, we have $P_U u = u$, so

$$Tu = T P_U u = P_U T P_U u \in \operatorname{range} P_U = U,$$

so U is invariant under T. ∎

Exercise 6C12

Suppose V is finite-dimensional, $T \in \mathcal{L}(V)$, and U is a subspace of V. Prove that

$$U \text{ and } U^\perp \text{ are both invariant under } T \iff P_U T = T P_U.$$

Solution. Suppose U and $U^\perp$ are both invariant under T. For any $v \in V$, write $v = u + w$ with $u \in U$ and $w \in U^\perp$. Then

$$P_U T v = P_U(Tu + Tw) = P_U Tu + P_U Tw = Tu + 0 = Tu,$$

since $Tu \in U$ and $Tw \in U^\perp$. On the other hand,

$$T P_U v = T P_U(u + w) = Tu,$$

so $P_U T v = T P_U v$ for all $v \in V$, and hence $P_U T = T P_U$.

Conversely, suppose $P_U T = T P_U$. For any $u \in U$, we have $P_U(Tu) = T(P_U u) = Tu$, so $Tu \in U$, and U is invariant under T. For any $w \in U^\perp$, we have $P_U(Tw) = T(P_U w) = T(0) = 0$, so $Tw \in U^\perp$, and $U^\perp$ is invariant under T. ∎

Exercise 6C13

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'.

Axler Note. *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) Define the map $\Phi : V \to V'$ by $\Phi(v) = \varphi_v$ for all $v \in V$. To show that Φ is linear, let $v_1, v_2 \in V$ and $a, b \in \mathbf{R}$. Then

for any $u \in V$,

$$\begin{aligned}\Phi(av_1 + bv_2)(u) &= \varphi_{av_1+bv_2}(u)\\ &= \langle u, av_1 + bv_2\rangle\\ &= a\langle u, v_1\rangle + b\langle u, v_2\rangle\\ &= a\varphi_{v_1}(u) + b\varphi_{v_2}(u)\\ &= a\Phi(v_1)(u) + b\Phi(v_2)(u).\end{aligned}$$

Since this holds for all $u \in V$, we have $\Phi(av_1 + bv_2) = a\Phi(v_1) + b\Phi(v_2)$, so Φ is linear. Injectivity follows from the fact that if $\Phi(v) = 0$, then $\varphi_v(u) = 0$ for all $u \in V$, which implies $\langle u, v\rangle = 0$ for all $u \in V$. In particular, taking $u = v$ gives $\langle v, v\rangle = 0$, so $v = 0$.

(b) By part (a), Φ is an injective linear map from V to V'. Since V is finite-dimensional, we have $\dim V = \dim V'$. Hence, Φ is also surjective, and therefore an isomorphism from V onto V'. ∎

Exercise 6C14

Suppose that $e_1, \ldots, e_n$ is an orthonormal basis of V. Explain why the dual basis (see 3.112) of $e_1, \ldots, e_n$ is $e_1, \ldots, e_n$ under the identification of V' with V provided by the Riesz representation theorem (6.58).

Solution. Let $\varphi_1, \ldots, \varphi_n$ be the dual basis of $e_1, \ldots, e_n$. By the Riesz representation theorem, for each j there exists a unique $v_j \in V$ such that $\varphi_j(u) = \langle u, v_j\rangle$ for all $u \in V$. Define $\psi_j \in V'$ by $\psi_j(u) = \langle u, e_j\rangle$. Then for any k,

$$\psi_j(e_k) = \langle e_k, e_j\rangle = \delta_{jk} = \varphi_j(e_k).$$

Since ψ_j and φ_j agree on a basis of V, they are equal. By uniqueness in the Riesz representation theorem, $v_j = e_j$. Hence, the dual basis $\varphi_1, \ldots, \varphi_n$ corresponds to $e_1, \ldots, e_n$ under the identification of V' with V provided by the Riesz representation theorem. ∎

Exercise 6C15

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. Applying the Gram–Schmidt procedure to the basis $v_1 = (1, 1, 0, 0)$ and $v_2 = (1, 1, 1, 2)$ of U, we first set $f_1 = v_1$ with $\|f_1\|^2 = 2$. Then

$$f_2 = v_2 - \frac{\langle v_2, f_1\rangle}{\|f_1\|^2} f_1 = (1, 1, 1, 2) - \frac{2}{2}(1, 1, 0, 0) = (0, 0, 1, 2),$$

with $\|f_2\|^2 = 5$. Projecting $(1, 2, 3, 4)$ onto U gives

$$P_U(1, 2, 3, 4) = \frac{\langle (1, 2, 3, 4), f_1\rangle}{\|f_1\|^2} f_1 + \frac{\langle (1, 2, 3, 4), f_2\rangle}{\|f_2\|^2} f_2 = \frac{3}{2}(1, 1, 0, 0) + \frac{11}{5}(0, 0, 1, 2) = \left(\frac{3}{2}, \frac{3}{2}, \frac{11}{5}, \frac{22}{5}\right).$$ ∎

Exercise 6C16 ✔

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) Let $g \in U^\perp$, and let $f(x) = x^2 g(x)$. Then f is continuous on $[-1,1]$ and $f(0) = 0$, so $f \in U$. Hence,

$$0 = \langle f, g\rangle = \int_{-1}^{1} x^2 g(x)^2 \, dx.$$

Since $x^2 g(x)^2 \geq 0$ for all $x \in [-1,1]$, it follows that $x^2 g(x)^2 = 0$ for all $x \in [-1,1]$. Therefore, $g(x) = 0$ for all $x \in [-1,1] \setminus \{0\}$, and by continuity, $g(0) = 0$ as well. Hence, $g = 0$, so $U^\perp = \{0\}$.

(b) 6.49 asserts that $V = U \oplus U^\perp$, but $U^\perp = \{0\}$, so $U \oplus U^\perp = U \neq C[-1,1]$ since the constant function 1 is not in U. 6.52 asserts that $(U^\perp)^\perp = U$, but $(U^\perp)^\perp = \{0\}^\perp = C[-1,1] \neq U$. ■

Exercise 6C17

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. The conditions $p(0) = 0$ and $p'(0) = 0$ imply that $p(x) = ax^2 + bx^3$ for some scalars $a, b \in \mathbf{R}$. Let $f(x) = 2 + 3x$ and $g(x) = ax^2 + bx^3$. To minimize the integral, we project f onto $\operatorname{span}(x^2, x^3)$ in $C[0,1]$ to get the resulting system of equations:

$$0 = \langle f - g, x^2\rangle = \int_0^1 (2 + 3x - ax^2 - bx^3)x^2 \, dx = \frac{2}{3} + \frac{3}{4} - \frac{a}{5} - \frac{b}{6},$$
$$0 = \langle f - g, x^3\rangle = \int_0^1 (2 + 3x - ax^2 - bx^3)x^3 \, dx = \frac{1}{2} + \frac{3}{5} - \frac{a}{6} - \frac{b}{7}.$$

Solving this system yields $a = 24$ and $b = -\frac{203}{10}$, so the polynomial $p(x) = 24x^2 - \frac{203}{10}x^3$ minimizes the integral. ■

Exercise 6C18

Find $p \in \mathcal{P}_5(\mathbf{R})$ that makes $\int_{-\pi}^{\pi} |\sin x - p(x)|^2 \, dx$ as small as possible.

Axler Note. *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. Similar to Exercise 6C17, we project $\sin x$ onto $\operatorname{span}(1, x, x^2, x^3, x^4, x^5)$ in $C[-\pi, \pi]$. Note that $\sin x$ is orthogonal to 1, x^2, and x^4 on $[-\pi, \pi]$ because it is an odd function. Hence, we only need to consider the projection onto $\operatorname{span}(x, x^3, x^5)$. Using the Gram–Schmidt procedure, we can find an orthonormal basis of $\operatorname{span}(x, x^3, x^5)$ and then compute the projection of $\sin x$ onto this subspace. Calculating the orthogonal basis, we begin with $f_1 = x$ with

$\|f_1\|^2 = \frac{2\pi^3}{3}$. Then

$$f_2 = x^3 - \frac{\langle x^3, f_1\rangle}{\|f_1\|^2} f_1 = x^3 - \frac{3\pi^2}{5}x, \quad \|f_2\|^2 = \frac{8\pi^7}{175}.$$

Finally,

$$f_3 = x^5 - \frac{\langle x^5, f_1\rangle}{\|f_1\|^2} f_1 - \frac{\langle x^5, f_2\rangle}{\|f_2\|^2} f_2 = x^5 - \frac{10\pi^2}{9}x^3 + \frac{5\pi^4}{21}x, \quad \|f_3\|^2 = \frac{128\pi^{11}}{43659}.$$

Projecting $\sin x$ onto $U = \operatorname{span}(x, x^3, x^5)$ gives

$$P_U \sin x = \frac{\langle \sin x, f_1\rangle}{\|f_1\|^2} f_1 + \frac{\langle \sin x, f_2\rangle}{\|f_2\|^2} f_2 + \frac{\langle \sin x, f_3\rangle}{\|f_3\|^2} f_3,$$

where the inner products are computed as follows:

$$\begin{aligned}
\langle \sin x, f_1\rangle &= \int_{-\pi}^{\pi} x \sin x \, dx = 2\pi, \\
\langle \sin x, f_2\rangle &= \int_{-\pi}^{\pi} \left(x^3 - \frac{3\pi^2}{5}x\right) \sin x \, dx = \frac{4\pi}{5}(\pi^2 - 15), \\
\langle \sin x, f_3\rangle &= \int_{-\pi}^{\pi} \left(x^5 - \frac{10\pi^2}{9}x^3 + \frac{5\pi^4}{21}x\right) \sin x \, dx = \frac{2}{63}(8\pi^5 - 840\pi^3 + 7560\pi).
\end{aligned}$$

The coefficients of f_1, f_2, and f_3 in the projection are then

$$\frac{\langle \sin x, f_1\rangle}{\|f_1\|^2} = \frac{3}{\pi^2}, \quad \frac{\langle \sin x, f_2\rangle}{\|f_2\|^2} = \frac{35(\pi^2 - 15)}{2\pi^6}, \quad \frac{\langle \sin x, f_3\rangle}{\|f_3\|^2} = \frac{693(\pi^4 - 105\pi^2 + 945)}{8\pi^{10}}.$$

Hence, the polynomial $p(x) = P_U \sin x$ that minimizes the integral is $p(x) = \alpha x + \beta x^3 + \gamma x^5$, where

$$\begin{aligned}
\alpha &= \frac{105}{8\pi^6}(\pi^4 - 153\pi^2 + 1485), \\
\beta &= -\frac{315}{4\pi^8}(\pi^4 - 125\pi^2 + 1155), \\
\gamma &= \frac{693}{8\pi^{10}}(\pi^4 - 105\pi^2 + 945).
\end{aligned}$$

■

Exercise 6C19

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. For $u, v \in V$, write $u = u_U + u_{U^\perp}$ and $v = v_U + v_{U^\perp}$ with $u_U, v_U \in U = \operatorname{range} P$ and $u_{U^\perp}, v_{U^\perp} \in U^\perp = \operatorname{null} P$. Then $Pu = u_U$ and $Pv = v_U$, so

$$\langle Pu, v\rangle = \langle u_U, v_U + v_{U^\perp}\rangle = \langle u_U, v_U\rangle$$

since $u_U \in U$ and $v_{U^\perp} \in U^\perp$. On the other hand,

$$\langle u, Pv\rangle = \langle u_U + u_{U^\perp}, v_U\rangle = \langle u_U, v_U\rangle$$

since $u_{U^\perp} \in U^\perp$ and $v_U \in U$. Hence, $\langle Pu, v\rangle = \langle u, Pv\rangle$ for all $u, v \in V$, so $P^\dagger = P$. ■

Exercise 6C20

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. Let P be the orthogonal projection of W onto $\operatorname{range} T$ and let $\tilde{T} : (\operatorname{null} T)^\perp \to \operatorname{range} T$ denote the restriction of T to $(\operatorname{null} T)^\perp$, so that $T^\dagger = \tilde{T}^{-1}P$. Since $\tilde{T}^{-1} : \operatorname{range} T \to (\operatorname{null} T)^\perp$ is a bijection, and P maps W onto $\operatorname{range} T$,

$$\operatorname{null} T^\dagger = \operatorname{null}(\tilde{T}^{-1}P) = \operatorname{null} P = (\operatorname{range} T)^\perp,$$

where the middle equality uses injectivity of $\tilde{T}^{-1}$, and

$$\operatorname{range} T^\dagger = \tilde{T}^{-1}(P(W)) = \tilde{T}^{-1}(\operatorname{range} T) = (\operatorname{null} T)^\perp,$$

where the second equality uses surjectivity of P onto $\operatorname{range} T$. ∎

Exercise 6C21

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) From $a + b + c = 2b + 3c = 0$, we see that $\operatorname{null} T = \operatorname{span}(1, -3, 2)$, so

$$(\operatorname{null} T)^\perp = \{(a, b, c) : a - 3b + 2c = 0\}.$$

Moreover, $\operatorname{range} T = \mathbf{F}^2$, so $P_{\operatorname{range} T} = I$. Let $T^\dagger(x, y) = (a, b, c)$. Then $T(a, b, c) = (x, y)$ and $(a, b, c) \in (\operatorname{null} T)^\perp$. Hence, we have the system of equations

$$\begin{cases} x = a + b + c, \\ y = 2b + 3c, \\ 0 = a - 3b + 2c. \end{cases}$$

This yields the solution

$$T^\dagger(x, y) = \left(\frac{13x - 5y}{14}, \frac{3x + y}{14}, \frac{-2x + 4y}{14}\right).$$

(b) For $(x, y) \in \mathbf{F}^2$, we have

$$TT^\dagger(x, y) = T\left(\frac{13x - 5y}{14}, \frac{3x + y}{14}, \frac{-2x + 4y}{14}\right) = (x, y),$$

as required.

(c) We have $P_{(\operatorname{null} T)^\perp}$ being the projection onto the plane $a - 3b + 2c = 0$. Hence,

$$P_{(\operatorname{null} T)^\perp}(a, b, c) = (a, b, c) - \frac{a - 3b + 2c}{14}(1, -3, 2) = \left(\frac{13a + 3b - 2c}{14}, \frac{3a + 5b + 6c}{14}, \frac{-2a + 6b + 10c}{14}\right).$$

For $(a, b, c) \in \mathbf{F}^3$, we have

$$T^\dagger T(a, b, c) = T^\dagger(a + b + c, 2b + 3c) = \left(\frac{13a + 3b - 2c}{14}, \frac{3a + 5b + 6c}{14}, \frac{-2a + 6b + 10c}{14}\right),$$

as required. ∎

Exercise 6C22

Suppose V is finite-dimensional and $T \in \mathcal{L}(V, W)$. Prove that

$$TT^\dagger T = T \quad \text{and} \quad T^\dagger T T^\dagger = T^\dagger.$$

Axler Note. *Both formulas above clearly hold if T is invertible because in that case we can replace $T^\dagger$ with T^{-1}.*

Solution. Let $V = \operatorname{null} T \oplus (\operatorname{null} T)^\perp$ so that any $v \in V$ can be written as $v = u + w$ with $u \in \operatorname{null} T$ and $w \in (\operatorname{null} T)^\perp$. For the first equality, we have $Tv = Tw \in \operatorname{range} T$. Then $T^\dagger T w = w$ since $T^\dagger$ restricted to $\operatorname{range} T$ is the inverse of T restricted to $(\operatorname{null} T)^\perp$. Hence,

$$TT^\dagger T v = TT^\dagger T w = Tw = Tv$$

for all $v \in V$, so $TT^\dagger T = T$.

For the second equality, let $w \in W$ and write $w = x + y$ with $x \in \operatorname{range} T$ and $y \in (\operatorname{range} T)^\perp$. Then $T^\dagger w = T^\dagger x \in (\operatorname{null} T)^\perp$, so

$$T^\dagger T T^\dagger w = T^\dagger T (T^\dagger x) = T^\dagger x = T^\dagger w.$$

Hence, $T^\dagger T T^\dagger = T^\dagger$. ∎

Exercise 6C23

Suppose V and W are finite-dimensional and $T \in \mathcal{L}(V, W)$. Prove that

$$(T^\dagger)^\dagger = T.$$

Axler Note. *The equation above is analogous to the equation $(T^{-1})^{-1} = T$ that holds if T is invertible.*

Solution. By Exercise 6C20, $\operatorname{null} T^\dagger = (\operatorname{range} T)^\perp$ and $\operatorname{range} T^\dagger = (\operatorname{null} T)^\perp$. Applying 6.68 to $T^\dagger$ in place of T and substituting via Exercise 6C20, for $v \in V$,

$$(T^\dagger)^\dagger v = (T^\dagger|_{(\operatorname{null} T^\dagger)^\perp})^{-1} P_{\operatorname{range} T^\dagger} v = (T^\dagger|_{\operatorname{range} T})^{-1} P_{(\operatorname{null} T)^\perp} v.$$

By 6.67, $T|_{(\operatorname{null} T)^\perp}$ is an isomorphism onto $\operatorname{range} T$ whose inverse is $T^\dagger|_{\operatorname{range} T}$, so $(T^\dagger|_{\operatorname{range} T})^{-1} = T|_{(\operatorname{null} T)^\perp}$. Because $v - P_{(\operatorname{null} T)^\perp} v \in ((\operatorname{null} T)^\perp)^\perp = \operatorname{null} T$, we have $T(v - P_{(\operatorname{null} T)^\perp} v) = 0$, so

$$(T^\dagger)^\dagger v = T(P_{(\operatorname{null} T)^\perp} v) = Tv.$$ ∎

7 Operators on Inner Product Spaces

7A Self-Adjoint and Normal Operators

Exercise 7A1

Suppose n is a positive integer. Define $T \in \mathcal{L}(\mathbf{F}^n)$ by

$$T(z_1, \dots, z_n) = (0, z_1, \dots, z_{n-1}).$$

Find a formula for $T^*(z_1, \dots, z_n)$.

Solution. If $n = 1$, then $T(z_1) = 0$, so $T^*(z_1) = 0$. Let $n > 1$. Then for all $(z_1, \dots, z_n), (w_1, \dots, w_n) \in \mathbf{F}^n$, we have

$$\langle T(z_1, \dots, z_n), (w_1, \dots, w_n) \rangle = \langle (0, z_1, \dots, z_{n-1}), (w_1, \dots, w_n) \rangle = z_1 \overline{w_2} + \cdots + z_{n-1} \overline{w_n}.$$

We need

$$\langle (z_1, \dots, z_n), T^*(w_1, \dots, w_n) \rangle = z_1 \overline{w_2} + \cdots + z_{n-1} \overline{w_n}.$$

By comparison, we see that $T^*(w_1, \dots, w_n) = (w_2, \dots, w_n, 0)$, so $T^*(z_1, \dots, z_n) = (z_2, \dots, z_n, 0)$ for all $(z_1, \dots, z_n) \in \mathbf{F}^n$. ∎

Exercise 7A2

Suppose $T \in \mathcal{L}(V, W)$. Prove that

$$T = 0 \iff T^* = 0 \iff T^*T = 0 \iff TT^* = 0.$$

Solution. Since $(T^*)^* = T$, then $T = 0 \iff T^* = 0$. Moreover, if $T^* = 0$, then $T^*T = 0$. Conversely, if $T^*T = 0$, then for all $v \in V$, we have

$$\|Tv\|^2 = \langle Tv, Tv \rangle = \langle v, T^*Tv \rangle = \langle v, 0 \rangle = 0.$$

Hence, $Tv = 0$ for all $v \in V$, so $T = 0$. Therefore,

$$T = 0 \iff T^* = 0 \iff T^*T = 0.$$

Replacing T with T^* in the equivalence $T = 0 \iff T^*T = 0$ gives $T^* = 0 \iff TT^* = 0$. We then have

$$T = 0 \iff T^* = 0 \iff T^*T = 0 \iff TT^* = 0.$$ ∎

Exercise 7A3

Suppose $T \in \mathcal{L}(V)$ and $\lambda \in \mathbf{F}$. Prove that

$$\lambda \text{ is an eigenvalue of } T \iff \overline{\lambda} \text{ is an eigenvalue of } T^*.$$

Solution. Since V is finite-dimensional,

$$\dim \operatorname{null}(T - \lambda I) = \dim V - \dim \operatorname{range}(T - \lambda I) = \dim(\operatorname{range}(T - \lambda I)^\perp) = \dim \operatorname{null}((T - \lambda I)^*).$$

Hence,

$$\operatorname{null}(T - \lambda I) \neq \{0\} \iff \operatorname{null}((T - \lambda I)^*) \neq \{0\}.$$

Since $(T-\lambda I)^* = T^* - \overline{\lambda} I$, we have

$$\lambda \text{ is an eigenvalue of } T \iff \overline{\lambda} \text{ is an eigenvalue of } T^*.$$

■

Exercise 7A4

Suppose $T \in \mathcal{L}(V)$ and U is a subspace of V. Prove that

$$U \text{ is invariant under } T \iff U^\perp \text{ is invariant under } T^*.$$

Solution.

$$\begin{aligned}
U^\perp \text{ is invariant under } T^* &\iff T^* v \in U^\perp \text{ for all } v \in U^\perp \\
&\iff \langle u, T^* v\rangle = 0 \text{ for all } u \in U, v \in U^\perp \\
&\iff \langle Tu, v\rangle = 0 \text{ for all } u \in U, v \in U^\perp \\
&\iff Tu \in (U^\perp)^\perp \text{ for all } u \in U \\
&\iff Tu \in U \text{ for all } u \in U \\
&\iff U \text{ is invariant under } T
\end{aligned}$$

■

Exercise 7A5

Suppose $T \in \mathcal{L}(V, W)$. Suppose $e_1, \dots, e_n$ is an orthonormal basis of V and $f_1, \dots, f_m$ is an orthonormal basis of W. Prove that

$$\|Te_1\|^2 + \cdots + \|Te_n\|^2 = \|T^* f_1\|^2 + \cdots + \|T^* f_m\|^2.$$

Axler Note. *The numbers $\|Te_1\|^2, \dots, \|Te_n\|^2$ in the equation above depend on the orthonormal basis $e_1, \dots, e_n$, but the right side of the equation does not depend on $e_1, \dots, e_n$. Thus, the equation above shows that the sum on the left side does not depend on which orthonormal basis $e_1, \dots, e_n$ is used.*

Manual Remark. *For students familiar with this topic already, note that the left-hand side is the squared Frobenius norm of the matrix representation of T with respect to the orthonormal bases $e_1, \dots, e_n$ and $f_1, \dots, f_m$. This exercise shows that $\sum_i \|Te_i\|^2 = \sum_j \|T^* f_j\|^2$, which is independent of the choice of orthonormal basis $e_1, \dots, e_n$. Moreover, noting that $\operatorname{tr}(T^*T) = \operatorname{tr}(TT^*)$ by the cyclic property of the trace, we see that the sum of the squared norms of the columns of the matrix representation of T equals the sum of the squared norms of the rows, which aligns with the result of this exercise.*

Solution. Let

$$Te_j = \sum_{k=1}^m \langle Te_j, f_k\rangle f_k \implies \|Te_j\|^2 = \sum_{k=1}^m |\langle Te_j, f_k\rangle|^2.$$

Similarly, we may write

$$T^* f_i = \sum_{\ell=1}^n \langle T^* f_i, e_\ell\rangle e_\ell \implies \|T^* f_i\|^2 = \sum_{\ell=1}^n |\langle T^* f_i, e_\ell\rangle|^2.$$

Note that

$$\langle T^* f_i, e_j\rangle = \langle f_i, Te_j\rangle = \overline{\langle Te_j, f_i\rangle} \implies |\langle T^* f_i, e_j\rangle|^2 = |\langle Te_j, f_i\rangle|^2.$$

Therefore, we have

$$\sum_{k=1}^m \|T^* f_k\|^2 = \sum_{k=1}^m \sum_{\ell=1}^n |\langle Te_\ell, f_k\rangle|^2 = \sum_{\ell=1}^n \sum_{k=1}^m |\langle Te_\ell, f_k\rangle|^2 = \sum_{\ell=1}^n \|Te_\ell\|^2.$$

■

Exercise 7A6

Suppose $T \in \mathcal{L}(V, W)$. Prove that

(a) T is injective $\iff$ T^* is surjective;

(b) T is surjective $\iff$ T^* is injective.

Solution.

(a) T injective $\iff$ $\operatorname{null} T = \{0\}$ $\iff$ $(\operatorname{range} T^*)^\perp = \{0\}$ $\iff$ $\operatorname{range} T^* = V$ $\iff$ T^* surjective.

(b) Using part (a) on $T^* \in \mathcal{L}(W, V)$ and the fact that $(T^*)^* = T$, we have

$$T^* \text{ is injective} \iff (T^*)^* \text{ is surjective} \iff T \text{ is surjective.}$$

■

Exercise 7A7

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) By Theorem 7.6(a), we have $\operatorname{null} T^* = (\operatorname{range} T)^\perp$, so $\dim \operatorname{null} T^* = \dim W - \dim \operatorname{range} T$. Using the fundamental theorem of linear maps, we have $\dim \operatorname{range} T = \dim V - \dim \operatorname{null} T$, so

$$\dim \operatorname{null} T^* = \dim W - \dim V + \dim \operatorname{null} T.$$

(b) We have $\operatorname{range} T^* = (\operatorname{null} T)^\perp$, so $\dim \operatorname{range} T^* = \dim V - \dim \operatorname{null} T = \dim \operatorname{range} T$. ■

Exercise 7A8

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.

Axler Note. *This exercise asks for yet another alternative proof of a result that was previously proved in 3.57 and 3.133.*

Solution. Let $T \in \mathcal{L}(\mathbf{F}^n, \mathbf{F}^m)$ have the matrix $A = \mathcal{M}(T)$ with respect to the standard bases of $\mathbf{F}^n$ and $\mathbf{F}^m$. Since the standard bases of $\mathbf{F}^n$ and $\mathbf{F}^m$ are orthonormal, $A^* = \mathcal{M}(T^*)$ are the conjugate-transposed rows of A. Since complex conjugation preserves linear independence, the column rank of A^* equals the row rank of A. By Exercise 7A7(b), $\dim \operatorname{range} T^* = \dim \operatorname{range} T$, so the column rank of A^* equals the column rank of A, and the result follows. ■

Exercise 7A9

Prove that the product of two self-adjoint operators on V is self-adjoint if and only if the two operators commute.

Solution. Let $S, T \in \mathcal{L}(V)$ be self-adjoint, so $S^* = S$ and $T^* = T$. Note that $(ST)^* = T^*S^* = TS$. Hence,

$$ST \text{ is self-adjoint} \iff (ST)^* = ST \iff TS = ST.$$

■

Exercise 7A10

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.

$\boxed{\Rightarrow}$ Suppose $\langle Tv, v\rangle = \langle T^*v, v\rangle$ for all $v \in V$. Then for all $v \in V$,

$$\langle (T - T^*)v, v\rangle = 0.$$

Since $\mathbf{F} = \mathbf{C}$, we may use 7.13 with $T - T^*$ in place of T to conclude that $T - T^* = 0$, so $T = T^*$, i.e., T is self-adjoint.

$\boxed{\Leftarrow}$ If T is self-adjoint, then $T = T^*$, so for all $v \in V$,

$$\langle Tv, v\rangle = \langle T^*v, v\rangle. \qquad \blacksquare$$

Exercise 7A11

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.

Axler Note. *If $\mathbf{F} = \mathbf{R}$, then S is the operator on $\mathbf{R}^2$ of counterclockwise rotation by $90°$.*

Solution.

(a) For any $(w, z), (x, y) \in \mathbf{F}^2$,

$$\langle S(w, z), (x, y)\rangle = \langle (-z, w), (x, y)\rangle = -z\bar{x} + w\bar{y}.$$

Suppose $S^*(x, y) = (u, v)$. Then

$$\langle S(w, z), (x, y)\rangle = \langle (w, z), S^*(x, y)\rangle = \langle (w, z), (u, v)\rangle = w\bar{u} + z\bar{v}.$$

Comparing with $-z\bar{x} + w\bar{y}$, we get $w\bar{u} + z\bar{v} = -z\bar{x} + w\bar{y}$ for all $w, z \in \mathbf{F}$. Hence, $u = y$ and $v = -x$, so $S^*(x, y) = (y, -x)$.

(b) Observe that

$$SS^*(w, z) = S(S^*(w, z)) = S(z, -w) = (w, z),$$

and

$$S^*S(w, z) = S^*(S(w, z)) = S^*(-z, w) = (w, z).$$

Hence, $SS^* = S^*S$, so S is normal. However,

$$S(1, 0) = (0, 1) \neq (0, -1) = S^*(1, 0).$$

Hence, S is not self-adjoint.

(c) If $S(w, z) = \lambda(w, z)$ for some $\lambda \in \mathbf{F}$ and $(w, z) \neq (0, 0)$, then $-z = \lambda w$ and $w = \lambda z$. From this, we get $w = -\lambda^2 w$, so $\lambda^2 = -1$. If $\mathbf{F} = \mathbf{R}$, then S has no real eigenvalues. If $\mathbf{F} = \mathbf{C}$, then the eigenvalues of S are $\lambda = i$ and $\lambda = -i$. $\blacksquare$

Exercise 7A12

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.

$\boxed{\Rightarrow}$ Suppose there exist commuting operators A and B such that A is self-adjoint, B is skew, and $T = A + B$. Then

$$T^* = (A + B)^* = A^* + B^* = A - B.$$

Hence,

$$TT^* = (A + B)(A - B) = A^2 - AB + BA - B^2 = A^2 - B^2,$$

and

$$T^*T = (A - B)(A + B) = A^2 + AB - BA - B^2 = A^2 - B^2.$$

Therefore, $TT^* = T^*T$, so T is normal.

$\boxed{\Leftarrow}$ If T is normal, then $TT^* = T^*T$. Consider the operators

$$A = \frac{T + T^*}{2}, \quad B = \frac{T - T^*}{2}.$$

Then A is self-adjoint, B is skew, and $T = A + B$. Moreover,

$$\begin{aligned} AB &= \frac{T + T^*}{2} \cdot \frac{T - T^*}{2} \\ &= \frac{1}{4}(TT - TT^* + T^*T - T^*T^*) \\ &= \frac{1}{4}(TT + TT^* - T^*T - T^*T^*) \\ &= \frac{T - T^*}{2} \cdot \frac{T + T^*}{2} = BA, \end{aligned}$$

hence A and B commute. ∎

Exercise 7A13

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) If $V = \{0\}$, then $\mathcal{L}(V) = \{0\}$ and $\mathcal{A}$ has no eigenvalues. Otherwise, assume $\dim V \geq 1$. If $\mathcal{A}T = \lambda T$ for some nonzero $T \in \mathcal{L}(V)$ and $\lambda \in \mathbf{R}$, then $T^* = \lambda T$. Taking the adjoint of both sides gives

$$T = \lambda T^* = \lambda(\lambda T) = \lambda^2 T.$$

Hence, $\lambda^2 = 1$, so $\lambda = 1$ or $\lambda = -1$. To see that $\lambda = 1$ always occurs as an eigenvalue for $\dim V \geq 1$, note that $I^* = I$, so $\mathcal{A}I = I = 1 \cdot I$. If $\dim V = 1$, then $T = \mu I$ for some $\mu \in \mathbf{R}$. In this case, $T^* = T$, so $\mathcal{A}T = T = 1 \cdot T$, and the only eigenvalue of $\mathcal{A}$ is 1.

If $\dim V \geq 2$, let $e_1, \ldots, e_n$ be an orthonormal basis of V with $n = \dim V$. Define $S \in \mathcal{L}(V)$ by $Se_1 = e_2$, $Se_2 = -e_1$, and $Se_j = 0$ for $j \geq 3$. Then $S^*e_1 = -e_2$, $S^*e_2 = e_1$, and $S^*e_j = 0$ for $j \geq 3$, so $S^* = -S$. Hence, -1 is also an eigenvalue of $\mathcal{A}$. Therefore, the eigenvalues of $\mathcal{A}$ are 1 and -1 when $\dim V \geq 2$.

(b) Observe that $\mathcal{A}(\mathcal{A}T) = \mathcal{A}T^* = (T^*)^* = T$ for all $T \in \mathcal{L}(V)$. Hence, $\mathcal{A}^2 - I = 0$, so the minimal polynomial of $\mathcal{A}$ must divide $z^2 - 1$. If $V = \{0\}$, then $\mathcal{L}(V) = \{0\}$, and the minimal polynomial of $\mathcal{A}$ is 1. If $\dim V = 1$, then $\mathcal{A}$ has only one eigenvalue, 1, so the minimal polynomial of $\mathcal{A}$ is $z - 1$. Otherwise, both eigenvalues 1 and -1 occur when $\dim V \geq 2$, so the minimal polynomial of $\mathcal{A}$ is $z^2 - 1$. ■

Exercise 7A14

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) Consider $p = 1$ and $q = x$. Then

$$\langle Tp, q \rangle = \langle 0, x \rangle = 0,$$

but

$$\langle p, Tq \rangle = \langle 1, x \rangle = \int_0^1 x \, dx = \frac{1}{2} \neq 0.$$

Hence, T is not self-adjoint.

(b) Observe that

$$\|x\|^2 = \langle x, x \rangle = \int_0^1 x^2 \, dx = \frac{1}{3}.$$

By 7.9, given an orthonormal basis, T is self-adjoint if and only if the matrix of T with respect to that basis equals its conjugate transpose. However, the basis $1, x, x^2$ is not orthonormal with respect to the given inner product, so the fact that the matrix of T with respect to this basis equals its conjugate transpose does not imply that T is self-adjoint. ■

Exercise 7A15

Suppose $T \in \mathcal{L}(V)$ is invertible. Prove that

(a) T is self-adjoint $\iff$ T^{-1} is self-adjoint;

(b) T is normal $\iff$ T^{-1} is normal.

Solution. We first note that

$$TT^{-1} = I \implies (T^{-1})^* T^* = I^* = I \implies (T^{-1})^* = (T^*)^{-1}.$$

(a)

$$\begin{aligned} T^{-1} \text{ is self-adjoint} &\iff (T^{-1})^* = T^{-1} \\ &\iff (T^*)^{-1} = T^{-1} \\ &\iff T^* = T \\ &\iff T \text{ is self-adjoint.} \end{aligned}$$

(b)

$$\begin{aligned} T^{-1} \text{ is normal} &\iff (T^{-1})^* T^{-1} = T^{-1}(T^{-1})^* \\ &\iff (T^*)^{-1} T^{-1} = T^{-1}(T^*)^{-1} \\ &\iff (T^* T)^{-1} = (TT^*)^{-1} \\ &\iff T^* T = TT^* \\ &\iff T \text{ is normal.} \end{aligned}$$

■

Exercise 7A16

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) Let $\mathcal{S}$ be the set of self-adjoint operators on V. Since $0^* = 0$, we have $0 \in \mathcal{S}$. For $S, T \in \mathcal{S}$ and $\lambda \in \mathbf{R}$, we have

$$(S+T)^* = S^* + T^* = S + T \in \mathcal{S},$$

and

$$(\lambda T)^* = \bar{\lambda} T^* = \lambda T \in \mathcal{S}.$$

Therefore, $\mathcal{S}$ is closed under addition and scalar multiplication, and contains the zero operator, so it is a subspace of $\mathcal{L}(V)$.

(b) Let $n = \dim V$, and let $e_1, \ldots, e_n$ be an orthonormal basis of V. By 7.9, it follows that T is self-adjoint if and only if $\mathcal{M}(T)$ is a symmetric matrix, i.e., $\mathcal{M}(T) = \mathcal{M}(T)^t$. Since a symmetric $n \times n$ matrix is determined by the entries on and above the main diagonal, the dimension of the subspace of symmetric matrices is

$$\frac{n(n+1)}{2}.$$

■

Exercise 7A17

Suppose $\mathbf{F} = \mathbf{C}$. Show that the set of self-adjoint operators on V is not a subspace of $\mathcal{L}(V)$.

Solution. Note the result is not true when $V = \{0\}$, since the only operator on V is the zero operator, which is self-adjoint. Let $\mathcal{S}$ be the set of self-adjoint operators on V. If $\mathcal{S}$ were a subspace of $\mathcal{L}(V)$, consider iI, where $I \in \mathcal{S}$ is the identity operator on V. Then

$$(iI)^* = \bar{i} I = -iI \neq iI.$$

Therefore, iI is not self-adjoint, which shows that the set of self-adjoint operators on V is not a subspace of $\mathcal{L}(V)$ when $V \neq \{0\}$. ■

Exercise 7A18

Suppose $\dim V \geq 2$. Show that the set of normal operators on V is not a subspace of $\mathcal{L}(V)$.

Solution. Fix an orthonormal pair $e_1, e_2 \in V$. Extend this to an orthonormal basis $e_1, \ldots, e_n$ of V. Consider the operators S and T defined by

$$Se_1 = e_2, Se_2 = e_1, Se_i = 0 \text{ for } i > 2, \quad Te_1 = -e_2, Te_2 = e_1, Te_i = 0 \text{ for } i > 2.$$

Then

$$\mathcal{M}(S) = \begin{pmatrix} 0 & 1 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}.$$

From this, we see that $\mathcal{M}(S) = \mathcal{M}(S)^{\mathrm{t}}$, so S is normal. Moreover,

$$\mathcal{M}(T) = \begin{pmatrix} 0 & 1 & \cdots & 0 \\ -1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}.$$

Direct computation yields $\mathcal{M}(T)\mathcal{M}(T)^* = \mathcal{M}(T)^*\mathcal{M}(T)$, so T is normal. However, $S + T$ is not normal, as can be seen by computing

$$\mathcal{M}(S+T) = \begin{pmatrix} 0 & 2 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}, \quad \mathcal{M}(S+T)^* = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 2 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}.$$

Direct computation shows that $\mathcal{M}(S+T)\mathcal{M}(S+T)^* \neq \mathcal{M}(S+T)^*\mathcal{M}(S+T)$, so $S + T$ is not normal. ■

Exercise 7A19

Suppose $T \in \mathcal{L}(V)$ and $\|T^*v\| \leq \|Tv\|$ for every $v \in V$. Prove that T is normal.

Axler Note. *This exercise fails on infinite-dimensional inner product spaces, leading to what are called* **hyponormal operators**, *which have a well-developed theory.*

Solution. If $\dim V = 0$, then T is trivially normal. So assume $\dim V > 0$. Note that for all $v \in V$, we have

$$\|T^*v\|^2 = \langle T^*v, T^*v\rangle = \langle TT^*v, v\rangle, \quad \|Tv\|^2 = \langle Tv, Tv\rangle = \langle T^*Tv, v\rangle.$$

From the hypothesis, we have $\|T^*v\|^2 \leq \|Tv\|^2$ for all $v \in V$, which gives

$$\langle (T^*T - TT^*)v, v\rangle \geq 0 \quad \text{for all } v \in V.$$

Let $S = T^*T - TT^*$. Then $S^* = (T^*T - TT^*)^* = T^*T - TT^* = S$, so S is self-adjoint. Moreover, $\langle Sv, v\rangle \geq 0$ for all $v \in V$. Let $e_1, \ldots, e_n$ be an orthonormal basis of V. From Exercise 7A5, then

$$\sum_{i=1}^{n} \|Te_i\|^2 = \sum_{i=1}^{n} \langle T^*Te_i, e_i\rangle = \sum_{i=1}^{n} \langle TT^*e_i, e_i\rangle = \sum_{i=1}^{n} \|T^*e_i\|^2.$$

Then

$$\sum_{i=1}^{n}\langle Se_i, e_i\rangle = \sum_{i=1}^{n}\langle (T^*T - TT^*)e_i, e_i\rangle = \sum_{i=1}^{n}\langle T^*Te_i, e_i\rangle - \sum_{i=1}^{n}\langle TT^*e_i, e_i\rangle = 0.$$

Since each $\langle Se_i, e_i\rangle \geq 0$ and their sum is 0, it follows that $\langle Se_i, e_i\rangle = 0$ for all i. For arbitrary $u \in V$ with $\|u\| = 1$, we may extend it to an orthonormal basis $u, f_2, \ldots, f_n$ of V. Then

$$\langle Su, u\rangle + \sum_{i=2}^{n}\langle Sf_i, f_i\rangle = \sum_{i=1}^{n}\langle Se_i, e_i\rangle = 0.$$

Since each term is nonnegative, it follows that $\langle Su, u\rangle = 0$ for all unit vectors $u \in V$. By homogeneity, $\langle Sv, v\rangle = 0$ for all $v \in V$. From 7.16, if S is self-adjoint and $\langle Sv, v\rangle = 0$ for all $v \in V$, then $S = 0$. Hence, $TT^* - T^*T = 0$, so $TT^* = T^*T$, and therefore T is normal. ∎

Exercise 7A20

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.

(a) $\implies$ (b): If P is self-adjoint, then $P = P^*$. Hence,

$$PP^* = PP = P^2 = P = P^*P,$$

so P is normal.

(b) $\implies$ (c): Since P is normal,

$$(\operatorname{range} P)^{\perp} = \operatorname{null} P^* = \operatorname{null} P.$$

Since $P^2 = P$, $V = \operatorname{null} P \oplus \operatorname{range} P$ is an orthogonal direct sum. Letting $U = \operatorname{range} P$, then any $v \in V$ can be written uniquely as $v = u + w$ with $u \in U$ and $w \in \operatorname{null} P$. Hence,

$$Pv = P(u + w) = Pu + Pw = u + 0 = u \in U.$$

Therefore, $P = P_U$.

(c) $\implies$ (a): Write $V = U \oplus U^{\perp}$. Then for $v = u_1 + w_1$ and $w = u_2 + w_2$ with $u_1, u_2 \in U$ and $w_1, w_2 \in U^{\perp}$, we have

$$\langle P_U v, w\rangle = \langle u_1, u_2 + w_2\rangle = \langle u_1, u_2\rangle = \langle u_1 + w_1, u_2\rangle = \langle v, P_U w\rangle.$$

Hence, P_U is self-adjoint. ∎

Exercise 7A21

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 there did exist an inner product on $\mathcal{P}_8(\mathbf{R})$ that makes D a normal operator, then $\|Dp\| = \|D^*p\|$ for all $p \in \mathcal{P}_8(\mathbf{R})$, by 7.20. Consider $p = x$ so that $Dp = 1 \neq 0$, but $D^2p = 0$. Then applying 7.20 to Dp, we get

$$\|D^*(Dp)\| = \|D(Dp)\| = \|D^2p\| = \|0\| = 0$$

so that $D^*(Dp) = 0$. Then

$$\langle D^*(Dp), p\rangle = \langle Dp, Dp\rangle = \|Dp\|^2.$$

Since $D^*(Dp) = 0$, the left-hand side of the above equation is 0, while the right-hand side is $\|Dp\|^2 \neq 0$. This is a contradiction, so no such inner product can exist. ■

Exercise 7A22

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 have the matrix

$$\begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

with respect to the standard basis of $\mathbf{R}^3$, which is orthonormal. Then T is normal because

$$\mathcal{M}(T^*) = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \mathcal{M}(T)^*.$$

By direct computation, we have

$$TT^* = I, \quad T^*T = I.$$

Hence, T is normal but not self-adjoint, since $T \neq T^*$. ■

Exercise 7A23

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. Since v, w are eigenvectors of the normal operator T corresponding to distinct eigenvalues 3 and 4, they must be orthogonal. Therefore, $\langle v, w\rangle = 0$. Then

$$\begin{aligned} \|T(v+w)\|^2 &= \langle Tv + Tw, Tv + Tw\rangle \\ &= \langle 3v + 4w, 3v + 4w\rangle \\ &= \langle 3v, 3v\rangle + \langle 3v, 4w\rangle + \langle 4w, 3v\rangle + \langle 4w, 4w\rangle \\ &= 9\langle v, v\rangle + 12\langle v, w\rangle + 12\langle w, v\rangle + 16\langle w, w\rangle \\ &= 9\|v\|^2 + 16\|w\|^2 \\ &= 9 \cdot 4 + 16 \cdot 4 \\ &= 100. \end{aligned}$$

Therefore, $\|T(v + w)\| = 10$. ■

Exercise 7A24

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

$$\overline{a_0} + \overline{a_1} z + \overline{a_2} z^2 + \cdots + \overline{a_{m-1}} z^{m-1} + z^m.$$

Axler Note. *This exercise shows that the minimal polynomial of T^* equals the minimal polynomial of T if $\mathbf{F} = \mathbf{R}$.*

Solution. Define $p(z)$ to be the minimal polynomial of T. Let $q(z) = \sum_{i=0}^{m} \overline{a_i} z^i$. Then

$$q(T^*) = \sum_{i=0}^{m} \overline{a_i} (T^*)^i = \sum_{i=0}^{m} \overline{a_i} (T^i)^* = \left(\sum_{i=0}^{m} a_i T^i \right)^* = p(T)^* = 0.$$

Hence, $q(z)$ is a polynomial that annihilates T^*. Using a similar argument with $r(z)$ being the minimal polynomial of T^*, we can show that $\overline{r}(T) = 0$. Since $q(z)$ annihilates T^* and $r(z)$ is the minimal polynomial of T^*, it follows that $r(z)$ divides $q(z)$. Similarly, since $\overline{r}(T) = 0$, $p(z)$ divides $\overline{r}(z)$. Then $\deg r \le \deg q$ from the fact that $r(z)$ divides $q(z)$, and $\deg q \le \deg r$ from the fact that $p(z)$ divides $\overline{r}(z)$ and $\deg p = \deg q$. Hence, $\deg r = \deg q$, and since $r(z)$ is monic and divides $q(z)$, we must have $r(z) = q(z)$. Therefore, the minimal polynomial of T^* is $q(z)$. ∎

Exercise 7A25

Suppose $T \in \mathcal{L}(V)$. Prove that T is diagonalizable if and only if T^* is diagonalizable.

Solution.

$\boxed{\Rightarrow}$ Suppose T is diagonalizable. Then the minimal polynomial of T splits into distinct linear factors by 5.62. By Exercise 7A24, the minimal polynomial of T^* also splits into distinct linear factors whose roots are the complex conjugates of the roots of the minimal polynomial of T. Hence, T^* is diagonalizable.

$\boxed{\Leftarrow}$ Conversely, suppose T^* is diagonalizable. Apply the same argument as above, using the fact that $(T^*)^* = T$. Therefore, T is diagonalizable. ∎

Exercise 7A26

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

$$\langle Tv, w \rangle = \langle \langle v, u \rangle x, w \rangle = \langle v, u \rangle \langle x, w \rangle = \langle v, \overline{\langle x, w \rangle} u \rangle = \langle v, T^* w \rangle.$$

Since $\langle v, u_1 \rangle = \langle v, u_2 \rangle$ for all $v \in V$ implies $u_1 = u_2$, we conclude that $T^* w = \overline{\langle x, w \rangle} u = \langle w, x \rangle u$.

(a) $\boxed{\Rightarrow}$ If T is self-adjoint, then $T = T^*$, which means $\langle v, u \rangle x = \langle v, x \rangle u$ for all $v \in V$. If $u = 0$ or $x = 0$, then the list u, x is linearly dependent. If both are nonzero, then letting $v = u$ gives $\|u\|^2 x = \langle u, x \rangle u$, hence

$$x = \frac{\langle u, x \rangle}{\|u\|^2} u.$$

Hence, x and u are linearly dependent.

$\boxed{\Leftarrow}$ Conversely, suppose one of u or x is zero. Then $T = 0 = T^*$, so T is self-adjoint. If both are nonzero and linearly dependent, let $x = \lambda u$ for some $\lambda \in \mathbf{R}$. Then for all $v \in V$,

$$Tv = \langle v, u\rangle x = \langle v, u\rangle \lambda u = \lambda \langle v, u\rangle u = T^* v,$$

hence T is self-adjoint.

(b) $\boxed{\Rightarrow}$ Suppose T is normal. Then for any $y \in V$,

$$TT^* y = T(\langle y, x\rangle u) = \langle y, x\rangle Tu = \langle y, x\rangle \langle u, u\rangle x = \|u\|^2 \langle y, x\rangle x,$$

and

$$T^* T y = T^*(\langle y, u\rangle x) = \langle y, u\rangle T^* x = \langle y, u\rangle \langle x, x\rangle u = \|x\|^2 \langle y, u\rangle u.$$

If $u = 0$ or $x = 0$, then the list u, x is linearly dependent. Otherwise, for nonzero u and x, we must have $\|u\|^2 \langle y, x\rangle x = \|x\|^2 \langle y, u\rangle u$ for all $y \in V$. Let $y = u$. Then $\|u\|^2 \langle u, x\rangle x = \|x\|^2 \|u\|^2 u$, hence $\langle u, x\rangle x = \|x\|^2 u$. If $x \neq 0$, then

$$u = \frac{\langle u, x\rangle}{\|x\|^2} x,$$

so u and x are linearly dependent.

$\boxed{\Leftarrow}$ If one of u or x is zero, then $T = 0 = T^*$, so T is normal. If both are nonzero and linearly dependent, let $u = \mu x$ for some $\mu \in \mathbf{F}$. Then for all $y \in V$,

$$TT^* y = \|u\|^2 \langle y, x\rangle x = \|\mu x\|^2 \langle y, x\rangle x = |\mu|^2 \|x\|^2 \langle y, x\rangle x,$$

and

$$T^* T y = \|x\|^2 \langle y, u\rangle u = \|x\|^2 \langle y, \mu x\rangle \mu x = |\mu|^2 \|x\|^2 \langle y, x\rangle x.$$

Hence, $TT^* y = T^* T y$ for all $y \in V$, so T is normal. ■

Exercise 7A27

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. Let $v \in \operatorname{null} T$. Then $T^k v = 0$, so $v \in \operatorname{null} T^k$. Hence, $\operatorname{null} T \subseteq \operatorname{null} T^k$. Conversely, let $v \in \operatorname{null} T^k$. Since T is normal, $\|Tv\| = \|T^* v\|$ by 7.20. Hence, $\operatorname{null} T = \operatorname{null} T^*$. If $T^2 v = 0$, then $T(Tv) = 0$, so $Tv \in \operatorname{null} T = \operatorname{null} T^*$. Therefore, $T^* T v = 0$ so that $\|Tv\|^2 = \langle Tv, Tv\rangle = \langle T^* T v, v\rangle = 0$. Hence, $Tv = 0$, which shows that $v \in \operatorname{null} T$. By induction on k, assume that $\operatorname{null} T^{k-1} = \operatorname{null} T$. If $v \in \operatorname{null} T^k$, then $Tv \in \operatorname{null} T^{k-1} = \operatorname{null} T$, so $T(Tv) = T^2 v = 0$. By the base case earlier, $v \in \operatorname{null} T^2 = \operatorname{null} T$. Hence, by induction, $\operatorname{null} T^k = \operatorname{null} T$ for all positive integers k.

For the other equality, note that $v \in \operatorname{range} T^k$ implies that there exists $u \in V$ such that $v = T^k u$. Then $T(T^{k-1} u) = T^k u = v$, so $v \in \operatorname{range} T$. Hence, $\operatorname{range} T^k \subseteq \operatorname{range} T$. Conversely, note that

$$\dim \operatorname{range} T^k = \dim V - \dim \operatorname{null} T^k = \dim V - \dim \operatorname{null} T = \dim \operatorname{range} T$$

by the previous equality $\operatorname{null} T^k = \operatorname{null} T$. Since $\operatorname{range} T^k \subseteq \operatorname{range} T$ and they have the same dimension, we conclude that $\operatorname{range} T^k = \operatorname{range} T$. ■

Exercise 7A28

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. If the minimal polynomial $p(x)$ of T is a polynomial multiple of $(x-\lambda)^2$, then $p(x) = (x-\lambda)^2 q(x)$ for some polynomial $q(x)$. Then $(T-\lambda I)^2 q(T) = 0$, but $(T-\lambda I)q(T) \neq 0$ since $p(x)$ is the minimal polynomial. Pick v such that $(T-\lambda I)q(T)v \neq 0$. Then $(T-\lambda I)^2 q(T)v = 0$, so $q(T)v \in \operatorname{null}(T-\lambda I)^2$. By normality of T, we know that $(T-\lambda I)^* = T^* - \overline{\lambda} I$ which commutes with $T-\lambda I$. Hence, by Exercise 7A27, $\operatorname{null}(T-\lambda I)^2 = \operatorname{null}(T-\lambda I)$. Therefore, $q(T)v \in \operatorname{null}(T-\lambda I)$, which implies that $(T-\lambda I)q(T)v = 0$. This contradicts our choice of v. Hence, the minimal polynomial of T cannot be a polynomial multiple of $(x-\lambda)^2$. ∎

Exercise 7A29

Prove or give a counterexample: If $T \in \mathcal{L}(V)$ and there is an orthonormal basis $e_1, \dots, e_n$ of V such that $\|Te_k\| = \|T^* e_k\|$ for each $k = 1, \dots, n$, then T is normal.

Solution. Let $V = \mathbf{F}^2$, and consider the standard orthonormal basis e_1, e_2. Define $T \in \mathcal{L}(V)$ by $T(e_1) = e_1 - e_2$ and $T(e_2) = e_1$. Then

$$\mathcal{M}(T) = \begin{pmatrix} 1 & 1 \\ -1 & 0 \end{pmatrix}, \quad \mathcal{M}(T^*) = \begin{pmatrix} 1 & -1 \\ 1 & 0 \end{pmatrix}$$

so that $T^* e_1 = e_1 + e_2$ and $T^* e_2 = -e_1$. Moreover,

$$\|Te_1\|^2 = \|e_1 - e_2\|^2 = 2, \quad \|T^* e_1\|^2 = \|e_1 + e_2\|^2 = 2,$$

and

$$\|Te_2\|^2 = \|e_1\|^2 = 1, \quad \|T^* e_2\|^2 = \|-e_1\|^2 = 1.$$

Hence, $\|Te_k\| = \|T^* e_k\|$ for each $k = 1, 2$. However, T is not normal because

$$\mathcal{M}(TT^*) = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}, \quad \mathcal{M}(T^*T) = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}.$$

Therefore, $TT^* \neq T^*T$. ∎

Exercise 7A30

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. Since T is normal, we have $\operatorname{null} T = \operatorname{null} T^*$ by 7.20. Let $z = (z_1, z_2, z_3) \in \operatorname{null} T = \operatorname{null} T^*$. By $T(1,1,1) = 2(1,1,1)$ and the definition of the adjoint, we have

$$2\langle (1,1,1), z\rangle = \langle T(1,1,1), z\rangle = \langle (1,1,1), T^* z\rangle = \langle (1,1,1), 0\rangle = 0.$$

Then $\langle (1,1,1), z\rangle = 0$. Direct calculation shows that

$$\langle (1,1,1), z\rangle = \overline{z_1} + \overline{z_2} + \overline{z_3} = \overline{z_1 + z_2 + z_3}.$$

Since $\langle (1,1,1), z\rangle = 0$, we have $\overline{z_1 + z_2 + z_3} = 0$, which implies $z_1 + z_2 + z_3 = 0$. ∎

Exercise 7A31

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

$$V = \operatorname{span}(1, \cos x, \cos 2x, \dots, \cos nx, \sin x, \sin 2x, \dots, \sin nx).$$

(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) By a similar set of arguments as in Exercise 6B4, the list above is orthogonal with respect to the given inner product. Note that any $f, g \in V$ is a linear combination of $1, \cos kx, \sin kx$ for $k = 1, 2, \dots, n$, all of which are periodic with period 2π. Then $f(\pi) = f(-\pi)$ and $g(\pi) = g(-\pi)$. Integration by parts gives

$$\langle Df, g\rangle = \int_{-\pi}^{\pi} f'g = fg\,|_{-\pi}^{\pi} - \int_{-\pi}^{\pi} fg' = -\int_{-\pi}^{\pi} fg' = \langle f, -Dg\rangle.$$

Then $D^* = -D$, which shows that

$$DD^* = D(-D) = -D^2 \quad \text{and} \quad D^*D = (-D)D = -D^2.$$

Hence, $DD^* = D^*D$, which shows that D is normal. However, D is not self-adjoint, since $D\cos x = -\sin x \neq \sin x = D^* \cos x$.

(b) Note that $T = D^2$. By the previous part, $D^* = -D$, so $T^* = (D^2)^* = (D^*)^2 = (-D)^2 = D^2 = T$. Hence, T is self-adjoint. ■

Exercise 7A32

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 all $v \in V$ and $w \in W$, we have

$$(T'(\varphi_w))v = \varphi_w(Tv) = \langle Tv, w\rangle = \langle v, T^*w\rangle = \varphi_{T^*w}v.$$

Hence, $T'(\varphi_w) = \varphi_{T^*w}$ for all $w \in W$. ■

7B Spectral Theorem

Exercise 7B1

Prove that a normal operator on a complex inner product space is self-adjoint if and only if all its eigenvalues are real.

Axler Note. *This exercise strengthens the analogy (for normal operators) between self-adjoint operators and real numbers.*

Solution.

$\boxed{\Rightarrow}$ This follows from 7.12.

$\boxed{\Leftarrow}$ Let T be a normal operator on a complex inner product space with only real eigenvalues. By the complex spectral theorem, there exists an orthonormal basis $e_1, \dots, e_n$ of eigenvectors of T, with corresponding eigenvalues $\lambda_1, \dots, \lambda_n$. Because T is normal, each e_i is also an eigenvector of T^* with eigenvalue $\overline{\lambda_i}$ by 7.21(e). Since all the eigenvalues λ_i are real, we have $\overline{\lambda_i} = \lambda_i$ for all i. Hence,

$$Te_i = \lambda_i e_i = \overline{\lambda_i} e_i = T^* e_i$$

for all $i = 1, \dots, n$. Since T and T^* agree on a basis of V, we have $T = T^*$ by the linear map lemma. ■

Exercise 7B2

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. By the complex spectral theorem, T is diagonalizable. By 5.55, V is the direct sum of the eigenspaces of T. Since T has exactly one eigenvalue λ, $V = E(\lambda, T)$. Therefore, $Tv = \lambda v$ for all $v \in V$, which means $T = \lambda I$. ■

Exercise 7B3

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.

$\boxed{\Rightarrow}$ Suppose the set of eigenvalues of T is contained in $\{0, 1\}$. By the complex spectral theorem, T is diagonalizable. By 5.55, V is the direct sum of the eigenspaces of T. Thus, we can write $V = \operatorname{null}(T - I) \oplus \operatorname{null} T$. Let $U = \operatorname{null}(T - I)$. By 7.22, $E(1, T)$ is orthogonal to $E(0, T)$, so $\operatorname{null} T \subseteq U^\perp$. Since any $v \in V$ can be uniquely written as $v = u + w$ with $u \in U$ and $w \in U^\perp$, we have

$$Tv = T(u + w) = u + 0 = u = P_U v,$$

which shows that $T = P_U$.

$\boxed{\Leftarrow}$ If $T = P_U$, then by Exercise 7A20, we know $P_U^2 = P_U$. If $Tv = \lambda v$ for some $\lambda \in \mathbf{C}$ and $v \neq 0$, then $\lambda^2 v = \lambda v$. This yields $\lambda^2 = \lambda$, so $\lambda \in \{0, 1\}$. ■

Exercise 7B4

Prove that a normal operator on a complex inner product space is skew (meaning 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.

$\boxed{\Rightarrow}$ Let $v \in V$ be an eigenvector of T with eigenvalue λ. From 7.21(e), we know that $Tv = \lambda v$ implies $T^* v = \overline{\lambda} v$. If T is skew, then $T = -T^*$, so $Tv = -T^* v = -\overline{\lambda} v$. Therefore, $\lambda = -\overline{\lambda}$, which means λ is purely imaginary.

$\boxed{\Leftarrow}$ Suppose all eigenvalues of T are purely imaginary. By the complex spectral theorem, there exists an orthonormal basis $e_1, \ldots, e_n$ of eigenvectors of T with corresponding eigenvalues $\lambda_1, \ldots, \lambda_n$. Since T is normal, each e_i is also an eigenvector of T^* with eigenvalue $\overline{\lambda}_i$. Because each λ_i is purely imaginary, we have $\overline{\lambda_i} = -\lambda_i$ for each i. Therefore,

$$T^* e_i = \overline{\lambda}_i e_i = -\lambda_i e_i = -T e_i$$

for all $i = 1, \ldots, n$. Since T^* and $-T$ agree on the basis, $T^* = -T$ by the linear map lemma. ■

Exercise 7B5

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 $T \in \mathcal{L}(\mathbf{C}^3)$ with matrix

$$\mathcal{M}(T) = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

Since the diagonal entries are distinct, T is diagonalizable. However, $\|Te_1\| = 1$, while $\|T^* e_1\| = \sqrt{2}$, so T is not normal by 7.20. ■

Exercise 7B6

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. By the complex spectral theorem, there exists an orthonormal basis $e_1, \ldots, e_n$ of eigenvectors of T with corresponding eigenvalues $\lambda_1, \ldots, \lambda_n$. Define $Te_i = \lambda_i e_i$ for each $i = 1, \ldots, n$. From $T^9 = T^8$, we have $\lambda_i^9 = \lambda_i^8$, which implies $\lambda_i^8(\lambda_i - 1) = 0$. This gives $\lambda_i \in \{0, 1\}$. By Exercise 7B1, since all eigenvalues of T are real, T is self-adjoint. Moreover, since each eigenvalue λ_i satisfies $\lambda_i \in \{0, 1\}$, we have $\lambda_i^2 = \lambda_i$ for each i, so $T^2 e_i = \lambda_i^2 e_i = \lambda_i e_i = T e_i$ for each $i = 1, \ldots, n$. Since T^2 and T agree on the basis, $T^2 = T$ by the linear map lemma. ■

Exercise 7B7

Give an example of an operator T on a complex vector space such that $T^9 = T^8$ but $T^2 \neq T$.

Solution. Consider $T \in \mathcal{L}(\mathbf{C}^2)$ defined by $T(z_1, z_2) = (z_2, 0)$. Then $T^2 = 0$, so $T^9 = T^8 = 0$, but $T^2 = 0 \neq T$. ■

Exercise 7B8

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.

$\boxed{\Rightarrow}$ This follows directly from 7.21(e).

$\boxed{\Leftarrow}$ If $V = \{0\}$, then T is trivially normal. Otherwise, assume $\dim V > 0$. By Schur's theorem, there exists an orthonormal basis $e_1, \ldots, e_n$ of V with respect to which T has an upper-triangular matrix A. We show by induction on k that $A_{k,j} = 0$ for all $j \neq k$, so that A is diagonal. Suppose $A_{i,j} = 0$ for all $j \neq i$ and all $i < k$. Since A is upper-triangular,

$Te_k = \sum_{i=1}^{k} A_{i,k} e_i$. By the inductive hypothesis, $A_{i,k} = 0$ for all $i < k$. Hence, $Te_k = A_{k,k} e_k$, so e_k is an eigenvector of T. By hypothesis, e_k is an eigenvector of T^*, say $T^* e_k = \lambda e_k$. Since $\mathcal{M}(T^*) = A^*$, this gives

$$T^* e_k = \sum_{j=1}^{n} \overline{A_{k,j}} e_j = \lambda e_k,$$

and the uniqueness of the representation with respect to $e_1, \ldots, e_n$ forces $\overline{A_{k,j}} = 0$ for all $j \neq k$. Thus, A is diagonal, so A and A^* commute, and therefore $TT^* = T^*T$. ∎

Exercise 7B9

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.
$\boxed{\Rightarrow}$ If $V = \{0\}$, then $T^* = 0 = p(T)$ trivially with $p = 0$. Otherwise, suppose $V \neq \{0\}$. If T is normal, then by the complex spectral theorem, there exists an orthonormal basis of V consisting of eigenvectors of T. Let $e_1, \ldots, e_n$ be such a basis, and let $\lambda_1, \ldots, \lambda_n$ be the corresponding eigenvalues. Of the eigenvalues, let $\mu_1, \ldots, \mu_m$ be the distinct ones. By Exercise 4-7, there exists a polynomial $p \in \mathcal{P}(\mathbf{C})$ such that $p(\mu_i) = \overline{\mu_i}$ for all $i \in \{1, \ldots, m\}$. Since each eigenvalue λ_i is equal to some μ_j, we have $p(\lambda_i) = \overline{\lambda_i}$ for all i. Because T is normal and $Te_i = \lambda_i e_i$, we have $T^* e_i = \overline{\lambda_i} e_i$ by 7.21(e). Also, $p(T)e_i = p(\lambda_i)e_i = \overline{\lambda_i} e_i$. Hence, $T^* e_i = p(T)e_i$ for all i. Since T^* and $p(T)$ agree on a basis of V, we have $T^* = p(T)$ by the linear map lemma.
$\boxed{\Leftarrow}$ Suppose $T^* = p(T)$ for some polynomial $p \in \mathcal{P}(\mathbf{C})$. Since T commutes with $p(T)$, we have $TT^* = Tp(T) = p(T)T = T^*T$, so T is normal. ∎

Exercise 7B10

Suppose V is a complex inner product space. Prove that every normal operator on V has a square root.
Axler Note. *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. By the complex spectral theorem, there exists an orthonormal basis $e_1, \ldots, e_n$ of V consisting of eigenvectors of T with corresponding eigenvalues $\lambda_1, \ldots, \lambda_n$. For each i, the fundamental theorem of algebra applied to the polynomial $z^2 - \lambda_i$ asserts the existence of $\mu_i \in \mathbf{C}$ such that $\mu_i^2 = \lambda_i$. Let $S \in \mathcal{L}(V)$ be the operator defined by $Se_i = \mu_i e_i$ for all i. Then

$$S^2 e_i = S(\mu_i e_i) = \mu_i S e_i = \mu_i(\mu_i e_i) = \mu_i^2 e_i = \lambda_i e_i = Te_i$$

for all i. Since S^2 and T agree on a basis of V, we have $S^2 = T$. ∎

Exercise 7B11

Prove that every self-adjoint operator on V has a cube root.
Axler Note. *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. By 7.29 (for $\mathbf{F} = \mathbf{R}$) or 7.31 (for $\mathbf{F} = \mathbf{C}$ by normality of T), there exists an orthonormal basis $e_1, \ldots, e_n$ of V consisting of eigenvectors of T with corresponding eigenvalues $\lambda_1, \ldots, \lambda_n \in \mathbf{F}$. By 7.12, these eigenvalues $\lambda_1, \ldots, \lambda_n$ are real. For each i, the polynomial $z^3 - \lambda_i$ has a real root μ_i. Let $S \in \mathcal{L}(V)$ be the operator

defined by $Se_i = \mu_i e_i$ for all i. Then

$$S^3 e_i = S^2(\mu_i e_i) = S(\mu_i^2 e_i) = \mu_i^3 e_i = \lambda_i e_i = Te_i$$

for all i. Since S^3 and T agree on a basis of V, we have $S^3 = T$. ■

Exercise 7B12

Suppose V is a complex inner product 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^*.

Axler Note. *The result in this exercise is called* **Fuglede's theorem**.

Manual Remark. *The original exercise statement said "Suppose V is a complex vector space...". I believe this to be an oversight, since T^* is defined using the inner product.*

Solution. By Exercise 7B9, we may write $T^* = p(T)$ for some polynomial p. Since S commutes with T, it also commutes with any polynomial in T, and hence with T^*. ■

Exercise 7B13

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 6B20) 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. Normality of T implies that T and T^* commute. By Exercise 6B20, there exists an orthonormal basis of V with respect to which both T and T^* have upper-triangular matrices. By 7.9, $\mathcal{M}(T^*) = \mathcal{M}(T)^*$. Since $\mathcal{M}(T)^*$ is upper-triangular, $\mathcal{M}(T)$ must be lower-triangular. This forces $\mathcal{M}(T)$ to be diagonal. ■

Exercise 7B14 ✔

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 $\lambda_1, \ldots, \lambda_m$ denote the distinct eigenvalues of T.

Manual Remark. *This exercise adds another characterization to the real spectral theorem, where T is self-adjoint if and only if the eigenvectors corresponding to distinct eigenvalues are orthogonal and V decomposes as the direct sum of its eigenspaces. In particular, this is the orthogonal analogue of Theorem 5.55.*

Solution.

$\boxed{\Rightarrow}$ Since T is self-adjoint, it is normal. By 7.22, its eigenvectors corresponding to distinct eigenvalues are orthogonal. By the real spectral theorem, V has an orthonormal basis consisting of eigenvectors of T, so T is diagonalizable. By 5.55, this gives the desired decomposition of V.

$\boxed{\Leftarrow}$ In each $E(\lambda_i, T)$, apply the Gram–Schmidt process to a basis of that eigenspace to obtain an orthonormal basis for that eigenspace. By hypothesis, these orthonormal bases for the eigenspaces corresponding to distinct eigenvalues are mutually orthogonal. The combined list of these orthonormal bases forms an orthonormal list of eigenvectors of T, hence is linearly independent by 6.25. Moreover, this list spans V because $V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)$ since each eigenspace is the span of its orthonormal basis obtained via Gram–Schmidt. Thus, the list is an orthonormal basis of V consisting of eigenvectors of T, so T is self-adjoint by the real spectral theorem. ■

Exercise 7B15 ✔

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 $\lambda_1, \ldots, \lambda_m$ denote the distinct eigenvalues of T.

Manual Remark. *This has the same remark as Exercise 7B14 but in the context of the complex spectral theorem. It also uses normal operators instead of self-adjoint operators.*

Solution. The proof is similar to that of Exercise 7B14, but using the complex spectral theorem instead of the real spectral theorem. Moreover, the $\boxed{\Rightarrow}$ direction applies 7.22 directly since T is normal by hypothesis. ■

Exercise 7B16

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 $S, T \in \mathcal{E}$.

Axler Note. *This exercise extends the complex spectral theorem to the context of a collection of commuting normal operators.*

Solution.

$\boxed{\Rightarrow}$ By assumption, there exists an orthonormal basis of V with respect to which every element of $\mathcal{E}$ has a diagonal matrix. For each $T \in \mathcal{E}$, the diagonal matrix $\mathcal{M}(T)$ commutes with its conjugate transpose $\mathcal{M}(T)^*$. By 7.9, we know $\mathcal{M}(T^*) = \mathcal{M}(T)^*$, so $\mathcal{M}(T)$ commutes with $\mathcal{M}(T^*)$, which implies that T is normal. By 3.43, for all $S, T \in \mathcal{E}$, we have

$$\mathcal{M}(ST) = \mathcal{M}(S)\mathcal{M}(T) = \mathcal{M}(T)\mathcal{M}(S) = \mathcal{M}(TS),$$

so $ST = TS$.

$\boxed{\Leftarrow}$ Taking $S = T$ in the hypothesis gives that each $T \in \mathcal{E}$ is normal. By Exercise 7B12 (Fuglede's theorem), for all $S, T \in \mathcal{E}$, $ST^* = T^*S$. Then

$$S^*T^* = (TS)^* = (ST)^* = T^*S^*,$$

so S^* and T^* commute. Letting $\mathcal{E}^* = \{T^* : T \in \mathcal{E}\}$, we see that $\mathcal{E} \cup \mathcal{E}^*$ is a set of commuting operators. By Exercise 6B20, there exists an orthonormal basis of V with respect to which every element of $\mathcal{E} \cup \mathcal{E}^*$ has an upper-triangular matrix. For each $T \in \mathcal{E}$, both $\mathcal{M}(T)$ and $\mathcal{M}(T^*)$ are upper-triangular, and $\mathcal{M}(T^*) = \mathcal{M}(T)^*$ by 7.9. Hence, $\mathcal{M}(T)$ is both upper-triangular and lower-triangular, which implies that $\mathcal{M}(T)$ is diagonal. Therefore, there exists an orthonormal basis of V with respect to which every element of $\mathcal{E}$ has a diagonal matrix. ■

Exercise 7B17

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 $S, T \in \mathcal{E}$.

Axler Note. *This exercise extends the real spectral theorem to the context of a collection of commuting self-adjoint operators.*

Solution.

$\boxed{\Rightarrow}$ Fix an orthonormal basis of V with respect to which every element of $\mathcal{E}$ has a diagonal matrix. For any $T \in \mathcal{E}$, the matrix of T with respect to this orthonormal basis is diagonal. Moreover, 7.9 shows that $\mathcal{M}(T^*) = \mathcal{M}(T)^* = \mathcal{M}(T)$. Since the matrices are equal and over the same orthonormal basis, it follows that $T = T^*$, so T is self-adjoint. Additionally, for any $S, T \in \mathcal{E}$, the matrices of S and T with respect to this orthonormal basis are diagonal, and diagonal matrices commute. A similar argument from Exercise 7B16 shows that $ST = TS$.

$\boxed{\Leftarrow}$ We proceed by strong induction on $\dim V = n$. If $n = 0$, the empty list is an orthonormal basis of V, and the statement holds trivially. Now assume that the result holds for any real inner product space of dimension less than n and every

collection of commuting self-adjoint operators on it. If every $T \in \mathcal{E}$ is some scalar multiple of the identity, then any orthonormal basis of V will diagonalize every element of $\mathcal{E}$, and the statement holds. In particular, this holds even with $\mathcal{E} = \varnothing$.

Otherwise, choose $T \in \mathcal{E}$ that is not a scalar multiple of the identity. Let $\lambda_1, \dots, \lambda_m$ be the distinct eigenvalues of T. By the real spectral theorem and Exercise 7B14, we may write

$$V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T),$$

where the eigenspaces are pairwise orthogonal. Note that if $m = 1$, then $V = E(\lambda_1, T)$ so that $T = \lambda_1 I$, contradicting our choice of T as not a scalar multiple of the identity. Therefore, $m \geq 2$. For any i, nontriviality of each eigenspace implies that $\dim E(\lambda_i, T) = n - \sum_{j \neq i} \dim E(\lambda_j, T) \leq n - 1$.

We now claim that each eigenspace $E(\lambda_i, T)$ is invariant under every $S \in \mathcal{E}$. To see this, let $v \in E(\lambda_i, T)$ and $S \in \mathcal{E}$. Then

$$T(Sv) = S(Tv) = S(\lambda_i v) = \lambda_i (Sv),$$

so $Sv \in E(\lambda_i, T)$. Hence, each eigenspace $E(\lambda_i, T)$ is invariant under every $S \in \mathcal{E}$.

Now consider each eigenspace under the inner product inherited from V, and let $\mathcal{E}_i = \{S|_{E(\lambda_i,T)} : S \in \mathcal{E}\}$. For any $u, w \in E(\lambda_i, T)$,

$$\langle S|_{E(\lambda_i,T)} u, w \rangle = \langle Su, w \rangle = \langle u, Sw \rangle = \langle u, S|_{E(\lambda_i,T)} w \rangle$$

so that each $S|_{E(\lambda_i,T)}$ is a self-adjoint operator on $E(\lambda_i, T)$. Furthermore, for any $S|_{E(\lambda_i,T)}, R|_{E(\lambda_i,T)} \in \mathcal{E}_i$, we have

$$S|_{E(\lambda_i,T)} R|_{E(\lambda_i,T)} = (SR)|_{E(\lambda_i,T)} = (RS)|_{E(\lambda_i,T)} = R|_{E(\lambda_i,T)} S|_{E(\lambda_i,T)},$$

so $\mathcal{E}_i$ is a commuting family of self-adjoint operators on $E(\lambda_i, T)$. Then the induction hypothesis applies to $E(\lambda_i, T)$ and $\mathcal{E}_i$, so there exists an orthonormal basis of $E(\lambda_i, T)$ with respect to which every element of $\mathcal{E}_i$ has a diagonal matrix. By Exercise 7B14, concatenating these bases gives an orthonormal basis of V since the eigenspaces are pairwise orthogonal, and V is the direct sum of these eigenspaces. Now let $S \in \mathcal{E}$. For each basis vector e in this orthonormal basis, there exists an i such that e is one of the basis vectors of $E(\lambda_i, T)$. Since $\mathcal{M}(S|_{E(\lambda_i,T)})$ is diagonal with respect to the chosen orthonormal basis of $E(\lambda_i, T)$, it follows that $Se = \mu_e e$ for some $\mu_e \in \mathbf{R}$. Therefore, the matrix of S with respect to this orthonormal basis is diagonal. ∎

Exercise 7B18

Give an example of a real inner product space V, an operator $T \in \mathcal{L}(V)$, and real numbers b, c with $b^2 < 4c$ such that

$$T^2 + bT + cI$$

is not invertible.

Axler Note. *This exercise shows that the hypothesis that T is self-adjoint cannot be deleted in 7.26, even for real vector spaces.*

Solution. Consider the real inner product space $V = \mathbf{R}^2$ and the operator $T \in \mathcal{L}(V)$ whose matrix with respect to the standard basis is

$$\mathcal{M}(T) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

so that $T^2 = -I$. Now choose $b = 0$ and $c = 1$. Then $b^2 = 0 < 4c = 4$, and

$$T^2 + bT + cI = T^2 + I = -I + I = 0,$$

which is not invertible. ∎

Exercise 7B19

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) For any $u \in U$ and $w \in U^\perp$, we have

$$\langle Tw, u\rangle = \langle w, Tu\rangle = 0$$

since $Tu \in U$. Hence, $Tw \in U^\perp$.

(b) For any $u_1, u_2 \in U$, we have

$$\langle T|_U u_1, u_2\rangle = \langle Tu_1, u_2\rangle = \langle u_1, Tu_2\rangle = \langle u_1, T|_U u_2\rangle.$$

Hence, $T|_U$ is self-adjoint.

(c) Applying the same argument as in part (b) to $U^\perp$ gives the conclusion. ∎

Exercise 7B20

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.

Axler Note. *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.)*

Manual Remark. *The ordering of the parts in this exercise is strange. The solution below will prove (b) first as (a) falls out naturally as a consequence.*

Solution.

(b) If $U = \{0\}$, the result is immediate since the zero subspace is invariant under any operator. Otherwise, let $e_1, \ldots, e_n$ be an orthonormal basis of U. By definition of P_U, we have for all $v \in V$ that $P_U v = \sum_{i=1}^n \langle v, e_i\rangle e_i$. Moreover, the Pythagorean theorem gives $\|v\|^2 = \|P_U v\|^2 + \|v - P_U v\|^2$. Hence, $\|v\|^2 = \sum_{i=1}^n |\langle v, e_i\rangle|^2$ if and only if $v \in U$.

Now, for each $j = 1, \ldots, n$, the invariance of U under T gives $Te_j \in U$. Hence,

$$\|Te_j\|^2 = \sum_{i=1}^n |\langle Te_j, e_i\rangle|^2.$$

By Bessel's inequality, we have for all $j = 1, \ldots, n$ that

$$\|T^*e_j\|^2 \geq \sum_{i=1}^n |\langle T^*e_j, e_i\rangle|^2 = \sum_{i=1}^n |\langle e_j, Te_i\rangle|^2.$$

Since T is normal, we have $\|T^*e_j\| = \|Te_j\|$ for each $j = 1, \ldots, n$. Summing over $j = 1, \ldots, n$ gives

$$\sum_{j=1}^n \|Te_j\|^2 = \sum_{j=1}^n \|T^*e_j\|^2 \geq \sum_{j=1}^n \sum_{i=1}^n |\langle T^*e_j, e_i\rangle|^2 = \sum_{i=1}^n \sum_{j=1}^n |\langle e_j, Te_i\rangle|^2 = \sum_{i=1}^n \|Te_i\|^2,$$

where $Te_i \in U$ by the invariance of U under T. Since the first and last sums in the inequality are equal, equality must hold throughout. In particular,

$$\sum_{j=1}^{n}\left(\|T^*e_j\|^2 - \sum_{i=1}^{n}|\langle T^*e_j, e_i\rangle|^2\right) = 0.$$

By Bessel's inequality, each term in the sum is nonnegative. Hence, each term must be zero, and

$$\|T^*e_j\|^2 - \sum_{i=1}^{n}|\langle T^*e_j, e_i\rangle|^2 = 0 \quad \text{for each } j = 1, \ldots, n.$$

The above equality shows that $T^*e_j \in U$ for each $j = 1, \ldots, n$. The linearity of T^* then implies that $T^*u \in U$ for all $u \in U$, so U is invariant under T^*.

(a) Let $w \in U^\perp$ and $u \in U$. Using (b), $T^*u \in U$, so $\langle w, T^*u\rangle = \langle Tw, u\rangle = 0$. Hence, $Tw \in U^\perp$, and $U^\perp$ is invariant under T.

(c) By (b), we know $(T^*)|_U \in \mathcal{L}(U)$. Then for all $x, y \in U$, we have

$$\langle (T|_U)y, x\rangle = \langle Ty, x\rangle = \langle y, T^*x\rangle = \langle y, (T^*)|_U x\rangle.$$

Since the adjoint is unique, we conclude that $(T|_U)^* = (T^*)|_U$.

(d) For any $u' \in U$, (b) gives $T^*u' \in U$. Using (c), we have

$$(T|_U)(T|_U)^*u' = T(T^*u') = (TT^*)u' = (T^*T)u' = T^*(Tu') = (T|_U)^*(T|_U)u',$$

showing normality of $T|_U$.

Now, by (a), we have that $U^\perp$ is invariant under T. Then we may use the arguments in (b) and (c) with $U^\perp$ in place of U to conclude that $U^\perp$ is invariant under T^*. Additionally, the argument of the first paragraph with $U^\perp$ in place of U shows that $T|_{U^\perp}$ is normal. ■

Exercise 7B21

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. By either the real or complex spectral theorem, we may take an orthonormal basis of eigenvectors $e_1, \ldots, e_n$ of T with corresponding eigenvalues $\lambda_1, \ldots, \lambda_n$. For each $i \in \{1, \ldots, n\}$, we have $\lambda_i \in \{2, 3\}$. Then

$$(T^2 - 5T + 6I)e_i = (\lambda_i^2 - 5\lambda_i + 6)e_i = (\lambda_i - 2)(\lambda_i - 3)e_i = 0.$$

Since $T^2 - 5T + 6I$ evaluates to 0 on each basis vector e_i, the conclusion follows. ■

Exercise 7B22

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. Consider $T \in \mathcal{L}(\mathbf{C}^3)$ with matrix representation

$$\mathcal{M}(T) = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}.$$

Since T is upper-triangular with diagonal entries $2, 2, 3$, these are the eigenvalues of T. However,

$$(T - 2I)(T - 3I)e_2 = (T - 2I)(e_1 - e_2) = -e_1 \neq 0.$$ ■

Exercise 7B23

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 $|\lambda - \lambda'| < \varepsilon$.

Axler Note. *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. If $\mathbf{F} = \mathbf{R}$, the real spectral theorem guarantees an orthonormal basis of eigenvectors. If $\mathbf{F} = \mathbf{C}$, then T is normal, since T is self-adjoint, and the complex spectral theorem guarantees an orthonormal basis of eigenvectors. In either case, we obtain an orthonormal basis of eigenvectors $e_1, \dots, e_n$ of T with corresponding eigenvalues $\lambda_1, \dots, \lambda_n$. Hence, we may write $v = \sum_{i=1}^n \alpha_i e_i$ for scalars $\alpha_1, \dots, \alpha_n \in \mathbf{F}$. Since the basis is orthonormal and $\|v\| = 1$, we have $\sum_{i=1}^n |\alpha_i|^2 = 1$. Then

$$Tv - \lambda v = \sum_{i=1}^n \alpha_i \lambda_i e_i - \sum_{i=1}^n \alpha_i \lambda e_i = \sum_{i=1}^n \alpha_i (\lambda_i - \lambda) e_i.$$

By orthonormality of the basis, we have

$$\|Tv - \lambda v\|^2 = \sum_{i=1}^n |\alpha_i|^2 |\lambda - \lambda_i|^2.$$

Now, choose k such that

$$|\lambda - \lambda_k| = \min_{1 \leq i \leq n} |\lambda - \lambda_i|.$$

Let $\lambda' = \lambda_k$ so that λ' is an eigenvalue of T. Since $|\lambda - \lambda'| = |\lambda - \lambda_k| \leq |\lambda - \lambda_i|$ for all i, we have

$$|\lambda - \lambda'|^2 = |\lambda - \lambda_k|^2 \sum_{i=1}^n |\alpha_i|^2 \leq \sum_{i=1}^n |\alpha_i|^2 |\lambda - \lambda_i|^2 = \|Tv - \lambda v\|^2 < \varepsilon^2.$$

Taking square roots, we conclude that $|\lambda - \lambda'| < \varepsilon$. ■

Exercise 7B24

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.

Axler Note. *This exercise adds another equivalence to the list of conditions equivalent to diagonalizability in 5.55.*

Solution. Since U is only assumed to be a finite-dimensional vector space, we require an inner product to discuss orthonormality in the $\boxed{\Leftarrow}$ direction of both parts. Hence, we prove the following lemma: there exists an inner product on U that makes any basis $e_1, \dots, e_n$ orthonormal.

Let $f_1, \dots, f_n$ denote the standard basis of $\mathbf{F}^n$, and equip $\mathbf{F}^n$ with the Euclidean inner product. Define $\varphi \in \mathcal{L}(U, \mathbf{F}^n)$

to be the unique linear map

$$\varphi(e_i) = f_i \quad \text{for } i = 1, \dots, n.$$

Then for any $u \in U$ and scalars $\alpha_1, \dots, \alpha_n \in \mathbf{F}$ such that $u = \sum_{i=1}^n \alpha_i e_i$, we have $\varphi(u) = \sum_{i=1}^n \alpha_i f_i$. Additionally, if $\varphi(u) = 0$, then $\alpha_i = 0$ for each i so that $u = 0$. Hence, φ is injective. As in Exercise 6A2, applied verbatim to a linear map rather than a linear operator, we may define an inner product on U as

$$\langle u, v \rangle = \langle \varphi(u), \varphi(v) \rangle \quad \text{for all } u, v \in U.$$

Moreover, $\langle e_i, e_j \rangle = \langle f_i, f_j \rangle = \delta_{ij}$ for all i, j, so the basis $e_1, \dots, e_n$ is orthonormal with respect to this inner product. This completes the proof of the lemma.

(a) $\boxed{\Rightarrow}$ Suppose T is diagonalizable. Then the matrix of T with respect to some basis $u_1, \dots, u_n$ consisting of eigenvectors of T is a diagonal matrix whose diagonal entries are eigenvalues of T. Since a diagonal matrix is equal to its transpose, the matrix of T with respect to the basis $u_1, \dots, u_n$ equals its transpose.
$\boxed{\Leftarrow}$ Suppose there is a basis $e'_1, \dots, e'_n$ of U such that the matrix of T with respect to $e'_1, \dots, e'_n$ equals its transpose. By the lemma, there exists an inner product on U that makes the basis $e'_1, \dots, e'_n$ orthonormal. Let $\mathcal{M}(T)$ denote the matrix of T with respect to the basis $e'_1, \dots, e'_n$; similarly for $\mathcal{M}(T^*)$. By 7.9, $\mathcal{M}(T^*) = \mathcal{M}(T)^*$, and since all entries of $\mathcal{M}(T)$ are real, $\mathcal{M}(T)^* = \mathcal{M}(T)^t$. Moreover, the hypothesis gives $\mathcal{M}(T)^t = \mathcal{M}(T)$. Therefore, $\mathcal{M}(T^*) = \mathcal{M}(T)$ so that T is self-adjoint. By the real spectral theorem, T has a diagonal matrix with respect to some orthonormal basis, hence T is diagonalizable.

(b) $\boxed{\Rightarrow}$ Suppose T is diagonalizable. Then there exists a basis $v_1, \dots, v_n$ of U consisting of eigenvectors of T with corresponding eigenvalues $\lambda_1, \dots, \lambda_n$. By definition, $Tv_i = \lambda_i v_i$ for each $i = 1, \dots, n$. Let $A = \mathcal{M}(T, (v_1, \dots, v_n))$ so that A is a diagonal matrix with diagonal entries $\lambda_1, \dots, \lambda_n$. It follows that A^* is also a diagonal matrix with diagonal entries $\overline{\lambda_1}, \dots, \overline{\lambda_n}$. Since diagonal matrices commute, we have $AA^* = A^*A$.
$\boxed{\Leftarrow}$ Suppose there is a basis $e''_1, \dots, e''_n$ of U such that the matrix of T with respect to $e''_1, \dots, e''_n$ commutes with its conjugate transpose. By the lemma, there exists an inner product on U that makes the basis $e''_1, \dots, e''_n$ orthonormal. For any $S \in \mathcal{L}(U)$, let $\mathcal{M}(S)$ denote the matrix of S with respect to the basis $e''_1, \dots, e''_n$. By 7.9, $\mathcal{M}(T^*) = \mathcal{M}(T)^*$, and the hypothesis gives $\mathcal{M}(T)\mathcal{M}(T)^* = \mathcal{M}(T)^*\mathcal{M}(T)$. By 3.43, we have

$$\mathcal{M}(TT^*) = \mathcal{M}(T)\mathcal{M}(T^*) = \mathcal{M}(T)\mathcal{M}(T)^* = \mathcal{M}(T)^*\mathcal{M}(T) = \mathcal{M}(T^*)\mathcal{M}(T) = \mathcal{M}(T^*T),$$

so $TT^* = T^*T$, and hence T is normal. By the complex spectral theorem, T has a diagonal matrix with respect to some orthonormal basis of U, hence T is diagonalizable. ∎

Exercise 7B25

Suppose that $T \in \mathcal{L}(V)$ and there is an orthonormal basis $e_1, \dots, e_n$ of V consisting of eigenvectors of T, with corresponding eigenvalues $\lambda_1, \dots, \lambda_n$. Show that if $k \in \{1, \dots, n\}$, 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. Let $S = \{j : \lambda_j \neq 0\}$. For any $v \in V$, we can write

$$v = \sum_{i=1}^n \langle v, e_i \rangle e_i.$$

Then

$$Tv = \sum_{i=1}^{n} \langle v, e_i \rangle T e_i = \sum_{i=1}^{n} \langle v, e_i \rangle \lambda_i e_i.$$

It follows that $\operatorname{null} T = \operatorname{span}(e_j : j \notin S)$, and $\operatorname{range} T = \operatorname{span}(e_i : i \in S)$. If $k \notin S$, then e_k is orthogonal to $\operatorname{range} T$, so $P_{\operatorname{range} T} e_k = 0$, and hence $T^\dagger e_k = 0$. If $k \in S$, then $e_k \in \operatorname{range} T$ so that $P_{\operatorname{range} T} e_k = e_k$. Since $\langle e_k, e_j \rangle = 0$ for all $j \notin S$, and these vectors span $\operatorname{null} T$, we have $\frac{1}{\lambda_k} e_k \in (\operatorname{null} T)^\perp$. Since $T(\frac{1}{\lambda_k} e_k) = e_k$, we have that $(T|_{(\operatorname{null} T)^\perp})^{-1} e_k = \frac{1}{\lambda_k} e_k$, and hence $T^\dagger e_k = \frac{1}{\lambda_k} e_k$. This is exactly the formula we wanted to show. ∎

7C Positive Operators

Exercise 7C1

Suppose $T \in \mathcal{L}(V)$. Prove that if both T and $-T$ are positive operators, then $T = 0$.

Solution. For every $v \in V$, we have $\langle Tv, v\rangle \geq 0$ and $\langle -Tv, v\rangle = -\langle Tv, v\rangle \geq 0$. Both inequalities force $\langle Tv, v\rangle = 0$ for every $v \in V$. By 7.43, $Tv = 0$ for every $v \in V$, so $T = 0$. ∎

Exercise 7C2

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 standard basis is orthonormal, and $\mathcal{M}(T)$ as given is equal to its conjugate transpose, so T is self-adjoint. For any $x = (x_1, x_2, x_3, x_4) \in \mathbf{F}^4$, we have

$$Tx = (2x_1 - x_2, -x_1 + 2x_2 - x_3, -x_2 + 2x_3 - x_4, -x_3 + 2x_4).$$

Since $x_i\overline{x_j} + \overline{x_i}x_j = 2\operatorname{Re}(x_i\overline{x_j})$, we have

$$\begin{aligned} \langle Tx, x\rangle &= (2x_1 - x_2)\overline{x_1} + (-x_1 + 2x_2 - x_3)\overline{x_2} + (-x_2 + 2x_3 - x_4)\overline{x_3} + (-x_3 + 2x_4)\overline{x_4} \\ &= 2|x_1|^2 - \overline{x_1}x_2 - x_1\overline{x_2} + 2|x_2|^2 - \overline{x_2}x_3 - x_2\overline{x_3} + 2|x_3|^2 - \overline{x_3}x_4 - x_3\overline{x_4} + 2|x_4|^2 \\ &= 2|x_1|^2 - 2\operatorname{Re}(\overline{x_1}x_2) + 2|x_2|^2 - 2\operatorname{Re}(\overline{x_2}x_3) + 2|x_3|^2 - 2\operatorname{Re}(\overline{x_3}x_4) + 2|x_4|^2 \\ &= |x_1|^2 + |x_1 - x_2|^2 + |x_2 - x_3|^2 + |x_3 - x_4|^2 + |x_4|^2 \geq 0. \end{aligned}$$

Hence, T is a positive operator. Moreover, if $Tx = 0$ for some $x \in \mathbf{F}^4$, then

$$0 = \langle Tx, x\rangle = |x_1|^2 + |x_1 - x_2|^2 + |x_2 - x_3|^2 + |x_3 - x_4|^2 + |x_4|^2.$$

This forces $x_1 = x_4 = 0$ so that $x_2 = x_3 = 0$ as well. Hence, $x = 0$, and therefore T is invertible by 3.65. ∎

Exercise 7C3

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. Since the standard basis is orthonormal, $\mathcal{M}(T)$ is equal to its conjugate transpose, so T is self-adjoint by 7.9. Let $x = (1, \ldots, 1) \in \mathbf{F}^n$. For any $y \in \mathbf{F}^n$, observe that $Ty = \langle y, x\rangle x$. Therefore,

$$\langle Ty, y\rangle = \langle \langle y, x\rangle x, y\rangle = \langle y, x\rangle\langle x, y\rangle = |\langle y, x\rangle|^2 \geq 0$$

for all $y \in \mathbf{F}^n$. Hence, T is a positive operator. ∎

Exercise 7C4

Suppose n is an 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 be the n-by-n matrix whose entries are all 1 except $A_{1,2} = A_{2,1} = 2$. Then all entries of A are positive numbers, and $A = A^*$ by construction. Let $T \in \mathcal{L}(\mathbf{F}^n)$ be the operator whose matrix with respect to the standard basis is A. However, consider $x = (1, -1, 0, \ldots, 0) \in \mathbf{F}^n$. Then $(Tx)_1 = 1 + 2(-1) = -1$, $(Tx)_2 = 2(1) + (-1) = 1$, and $(Tx)_i = 1 - 1 = 0$ for $i > 2$. Hence,

$$\langle Tx, x \rangle = (-1)(1) + (1)(-1) + 0 = -2$$

so that T is not a positive operator. ■

Exercise 7C5

Suppose $T \in \mathcal{L}(V)$ is self-adjoint. Prove that T is a positive operator if and only if for every orthonormal basis $e_1, \ldots, e_n$ of V, all entries on the diagonal of $\mathcal{M}(T, (e_1, \ldots, e_n))$ are nonnegative numbers.

Solution. Note that the (j, k)th entry of $\mathcal{M}(T, (e_1, \ldots, e_n))$ is $\langle Te_k, e_j \rangle$, hence the diagonal entries are $\langle Te_j, e_j \rangle$ for $j = 1, \ldots, n$.

$\boxed{\Rightarrow}$ Since T is self-adjoint by hypothesis, T is positive if and only if $\langle Tv, v \rangle \geq 0$ for all $v \in V$. In particular, this must hold for each $v = e_j$ in any orthonormal basis $e_1, \ldots, e_n$ of V, which shows that all diagonal entries of $\mathcal{M}(T, (e_1, \ldots, e_n))$ are nonnegative.

$\boxed{\Leftarrow}$ If $v = 0$, then $\langle Tv, v \rangle = 0 \geq 0$. Suppose $v \neq 0$, and let $e_1 = v/\|v\|$. Then $\|e_1\| = 1$, so e_1 is an orthonormal list by itself. Extend e_1 to an orthonormal basis $e_1, \ldots, e_n$ of V. Then the first diagonal entry of $\mathcal{M}(T, (e_1, \ldots, e_n))$ is $\langle Te_1, e_1 \rangle = \langle T(v/\|v\|), v/\|v\| \rangle = \langle Tv, v \rangle / \|v\|^2 \geq 0$. Hence, $\langle Tv, v \rangle \geq 0$ for all $v \in V$, which shows that T is a positive operator. ■

Exercise 7C6

Prove that the sum of two positive operators on V is a positive operator.

Solution. Let $S, T \in \mathcal{L}(V)$ be positive operators. In particular, S and T are self-adjoint. By 7.5(a), $S + T$ is also self-adjoint. Positivity follows from the fact that for any $v \in V$,

$$\langle (S + T)v, v \rangle = \langle Sv, v \rangle + \langle Tv, v \rangle \geq 0 + 0 = 0,$$

since S and T are positive operators. Hence, $S + T$ is a positive operator. ■

Exercise 7C7

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. If $(S + T)v = 0$ for some $v \in V$, then $\langle Sv, v \rangle + \langle Tv, v \rangle = \langle (S + T)v, v \rangle = \langle 0, v \rangle = 0$. Since S and T are positive, $\langle Sv, v \rangle \geq 0$ and $\langle Tv, v \rangle \geq 0$. In particular, $\langle Sv, v \rangle = 0$. By 7.43, $Sv = 0$ so that $v = 0$ by the invertibility of S. Therefore, $(S + T)v = 0$ only if $v = 0$, which shows that $S + T$ is invertible by 3.65. ■

Exercise 7C8

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.
$\boxed{\Rightarrow}$ By 7.38(c), there is an orthonormal basis $e_1, \dots, e_n$ of V consisting of eigenvectors of T with respect to which the matrix of T is diagonal with nonnegative diagonal entries $\lambda_1, \dots, \lambda_n$. Hence, $Te_j = \lambda_j e_j$, where $\lambda_j \geq 0$ for each j. By Exercise 7B25,

$$T^\dagger e_j = \begin{cases} \frac{1}{\lambda_j} e_j & \text{if } \lambda_j \neq 0, \\ 0 & \text{if } \lambda_j = 0. \end{cases}$$

Hence, $\mathcal{M}(T^\dagger, (e_1, \dots, e_n))$ is a diagonal matrix with nonnegative entries, which shows that $T^\dagger$ is a positive operator by 7.38(c).
$\boxed{\Leftarrow}$ By Exercise 6C23, $(T^\dagger)^\dagger = T$. Applying the forward direction to $T^\dagger$ in place of T shows that T is positive. ∎

Exercise 7C9

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. Note that $S^* \in \mathcal{L}(V, W)$, so $S^*TS \in \mathcal{L}(W)$. By 7.5(c) and 7.5(d), we have $(S^*TS)^* = S^*T^*(S^*)^* = S^*TS$, where the last equality holds since T is self-adjoint. Moreover, for any $w \in W$,

$$\langle S^*TSw, w \rangle = \langle TSw, Sw \rangle \geq 0,$$

since T is positive. Hence, S^*TS is a positive operator on W. ∎

Exercise 7C10

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 $v = w$.
Manual Remark. *Alternatively, observe that $T(v - w) = -(v - w)$, so $v - w \in E(-1, T) = \{0\}$ by 7.38(b) and the fact that T is positive. Then $v - w = 0$, so $v = w$.*

Solution. Consider $v - w$. Then

$$\langle T(v - w), v - w \rangle = \langle Tv - Tw, v - w \rangle = \langle w - v, v - w \rangle = -\langle v - w, v - w \rangle = -\|v - w\|^2.$$

Since T is positive, $\langle T(v - w), v - w \rangle \geq 0$, so $-\|v - w\|^2 \geq 0$, which implies $\|v - w\|^2 = 0$ and hence $v - w = 0$, i.e., $v = w$. ∎

Exercise 7C11

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. Note that $T|_U \in \mathcal{L}(U)$ since U is invariant under T, and U carries the inner product inherited from V. Since T is self-adjoint on V, Exercise 7B19(b) gives that $T|_U$ is self-adjoint on U. Moreover, for any $u \in U$,

$$\langle T|_U u, u \rangle_U = \langle Tu, u \rangle_V \geq 0,$$

since T is positive on V. Hence, $T|_U$ is a positive operator on U. ■

Exercise 7C12

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 $e_1, \dots, e_n$ of V with respect to which $\mathcal{M}(T, (e_1, \dots, e_n))$ is diagonal with nonnegative entries $\lambda_1, \dots, \lambda_n$. By 3.43 and induction on k, $\mathcal{M}(T^k, (e_1, \dots, e_n)) = (\mathcal{M}(T, (e_1, \dots, e_n)))^k$ is diagonal with entries $\lambda_1^k, \dots, \lambda_n^k$, which are all nonnegative. By 7.38(c) again, this shows that T^k is positive. ■

Exercise 7C13

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) $\boxed{\Rightarrow}$ If $T - \alpha I$ is positive, then its eigenvalues are nonnegative by 7.38(b). Since $Tv = \lambda v$ if and only if $(T - \alpha I)v = (\lambda - \alpha)v$, the eigenvalues of $T - \alpha I$ are precisely $\lambda - \alpha$, where λ ranges over the eigenvalues of T. Hence, $\alpha \le \lambda$ for every eigenvalue λ of T.

$\boxed{\Leftarrow}$ By 7.5(a), (b), (e), self-adjointness of T, and $\alpha \in \mathbf{R}$,

$$(T - \alpha I)^* = T^* - \bar{\alpha} I^* = T - \alpha I,$$

so $T - \alpha I$ is self-adjoint. Since $\lambda - \alpha \ge 0$ for every eigenvalue λ of T, the eigenvalues of $T - \alpha I$ are nonnegative. Hence, $T - \alpha I$ is positive by 7.38(b).

(b) Note that $-T$ is self-adjoint by 7.5(b), and its eigenvalues are exactly $-\lambda$, where λ ranges over the eigenvalues of T. Applying (a) to $-T$ and $-\alpha$, the operator $-T - (-\alpha)I = \alpha I - T$ is positive if and only if $-\alpha \le -\lambda$, i.e., $\alpha \ge \lambda$, for every eigenvalue λ of T. ■

Exercise 7C14

Suppose T is a positive operator on V and $v_1, \dots, v_m \in V$. Prove that

$$\sum_{j=1}^{m} \sum_{k=1}^{m} \langle T v_k, v_j \rangle \ge 0.$$

Solution. Let $v = \sum_{k=1}^{m} v_k$. By linearity of T, additivity of the inner product, and positivity of T, we have

$$\sum_{j=1}^{m} \sum_{k=1}^{m} \langle T v_k, v_j \rangle = \sum_{j=1}^{m} \langle Tv, v_j \rangle = \langle Tv, v \rangle \ge 0.$$ ■

Exercise 7C15

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. If $\mathbf{F} = \mathbf{R}$, the real spectral theorem applies directly, and if $\mathbf{F} = \mathbf{C}$, self-adjointness implies normality, so the complex spectral theorem applies. In either case, there exists an orthonormal basis $e_1, \dots, e_n$ of V consisting of

eigenvectors of T with corresponding eigenvalues $\lambda_1, \dots, \lambda_n$. Since T is self-adjoint, $T^*T = T^2$. By the linear map lemma, there exists an operator C on V such that $Ce_k = |\lambda_k| e_k$ for each $k = 1, \dots, n$. Since $\mathcal{M}(C, (e_1, \dots, e_n))$ is diagonal with nonnegative entries, C is a positive operator by 7.38(c). Moreover, $C^2 e_k = |\lambda_k|^2 e_k = \lambda_k^2 e_k = T^2 e_k$ for each k, so $C^2 = T^2$ by the linear map lemma. By 7.39, $C = \sqrt{T^2} = \sqrt{T^*T}$.

Consider now the operators

$$A = \frac{1}{2}(C + T) \quad \text{and} \quad B = \frac{1}{2}(C - T).$$

By construction, $A - B = T$ and $A + B = C = \sqrt{T^*T}$. Then for each $k = 1, \dots, n$,

$$Ae_k = \frac{1}{2}(C + T)e_k = \frac{1}{2}(|\lambda_k| + \lambda_k)e_k \quad \text{and} \quad Be_k = \frac{1}{2}(C - T)e_k = \frac{1}{2}(|\lambda_k| - \lambda_k)e_k.$$

Then the diagonal entries of $\mathcal{M}(A, (e_1, \dots, e_n))$ and $\mathcal{M}(B, (e_1, \dots, e_n))$ are nonnegative, so A and B are positive operators by 7.38(c). Finally,

$$ABe_k = \frac{1}{4}((C + T)(C - T))e_k = \frac{1}{4}((|\lambda_k| + \lambda_k)(|\lambda_k| - \lambda_k))e_k = 0 \quad \text{and} \quad BAe_k = 0$$

so that $AB = BA = 0$ by the linear map lemma. ∎

Exercise 7C16

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. Since T is positive, $\sqrt{T}$ is also a positive operator. In particular, it is self-adjoint, hence normal. Let $S = \sqrt{T}$, so $S^2 = T$. By Exercise 7A27, $\operatorname{null} S = \operatorname{null} S^2 = \operatorname{null} T$ and $\operatorname{range} S = \operatorname{range} S^2 = \operatorname{range} T$. ∎

Exercise 7C17

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. If $V = \{0\}$, take $p = 0$. Otherwise, $V \neq \{0\}$. By 7.38(c), there exists an orthonormal basis $e_1, \dots, e_n$ of V such that $Te_k = \lambda_k e_k$ where $\lambda_k \geq 0$ for each $k = 1, \dots, n$. Since $n \geq 1$, T has at least one eigenvalue. Let $\mu_1, \dots, \mu_m$ be the distinct eigenvalues of T, so $m \geq 1$. Because $e_k \neq 0$, λ_k is an eigenvalue of T so that $\lambda_k \in \{\mu_1, \dots, \mu_m\}$ for each $k = 1, \dots, n$. By 7.38(b), each $\mu_j \geq 0$ so that $\sqrt{\mu_j} \in \mathbf{R}$ for each $j = 1, \dots, m$. By Exercise 4-7 applied over $\mathbf{R}$, there exists $p \in \mathcal{P}_{m-1}(\mathbf{R})$ such that $p(\mu_j) = \sqrt{\mu_j}$ for each $j = 1, \dots, m$. Since $T^i e_k = \lambda_k^i e_k$ for each nonnegative integer i, the definition of $p(T)$ gives $p(T)e_k = p(\lambda_k)e_k$ for each $k = 1, \dots, n$. Since $\lambda_k \in \{\mu_1, \dots, \mu_m\}$ and $p(\mu_j) = \sqrt{\mu_j}$ for each $j = 1, \dots, m$, we have $p(\lambda_k) = \sqrt{\lambda_k}$ for each $k = 1, \dots, n$. Then $\mathcal{M}(p(T), (e_1, \dots, e_n))$ is diagonal with diagonal entries $\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n}$. By 7.38(c), $p(T)$ is positive. Then

$$p(T)^2 e_k = p(\lambda_k)^2 e_k = (\sqrt{\lambda_k})^2 e_k = \lambda_k e_k = Te_k$$

for each $k = 1, \dots, n$. By the linear map lemma, $p(T)^2 = T$, and by 7.39, $p(T) = \sqrt{T}$. ∎

Exercise 7C18

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.

$\boxed{\Rightarrow}$ Note that S, T, and ST are all self-adjoint. By Exercise 7A9, ST is self-adjoint if and only if $ST = TS$.

$\boxed{\Leftarrow}$ By Exercise 7C17, there exists a polynomial p such that $\sqrt{S} = p(S)$. Since T commutes with S, it commutes with powers of S so that T commutes with $p(S) = \sqrt{S}$. Then $ST = \sqrt{S}\sqrt{S}T = \sqrt{S}^*\sqrt{S}T = \sqrt{S}^*T\sqrt{S}$, where the second equality uses the fact that $\sqrt{S}$ is self-adjoint. By Exercise 7C9, ST is positive. ■

Exercise 7C19

Show that the identity operator on $\mathbf{F}^2$ has infinitely many self-adjoint square roots.

Solution. Let $x, y \in \mathbf{R}$, and define the family of operators $T_{x,y}$ on $\mathbf{F}^2$ whose matrices with respect to the standard basis are

$$\mathcal{M}(T_{x,y}) = \begin{pmatrix} x & y \\ y & -x \end{pmatrix}.$$

Then $\mathcal{M}(T_{x,y}) = \mathcal{M}(T_{x,y})^*$. Since the standard basis is orthonormal, $T_{x,y}$ is self-adjoint by 7.9. Moreover,

$$\mathcal{M}(T_{x,y})^2 = \begin{pmatrix} x & y \\ y & -x \end{pmatrix}^2 = \begin{pmatrix} x^2+y^2 & 0 \\ 0 & x^2+y^2 \end{pmatrix} = (x^2+y^2)I.$$

Taking $x = \cos\theta$ and $y = \sin\theta$ for $\theta \in \mathbf{R}$, we get $x^2 + y^2 = 1$. Since $(\cos\theta, \sin\theta)$ takes on infinitely many distinct values and $T_{\cos\theta,\sin\theta}(1,0) = (\cos\theta, \sin\theta)$, the operators $T_{\cos\theta,\sin\theta}$ are all distinct. Therefore, the identity operator on $\mathbf{F}^2$ has infinitely many self-adjoint square roots. ■

Exercise 7C20

Suppose $T \in \mathcal{L}(V)$ and $e_1, \dots, e_n$ is an orthonormal basis of V. Prove that T is a positive operator if and only if there exist $v_1, \dots, v_n \in V$ such that

$$\langle Te_k, e_j \rangle = \langle v_k, v_j \rangle$$

for all $j, k = 1, \dots, n$.

Axler Note. *The numbers $\{\langle Te_k, e_j \rangle\}_{j,k=1,\dots,n}$ are the entries in the matrix of T with respect to the orthonormal basis $e_1, \dots, e_n$.*

Solution.

$\boxed{\Rightarrow}$ By 7.38(f), let $T = R^*R$ for some $R \in \mathcal{L}(V)$. Define $v_k = Re_k$ for $k = 1, \dots, n$. Then for all $j, k = 1, \dots, n$, we have

$$\langle v_k, v_j \rangle = \langle Re_k, Re_j \rangle = \langle R^*Re_k, e_j \rangle = \langle Te_k, e_j \rangle.$$

$\boxed{\Leftarrow}$ By the linear map lemma, we may define the unique operator $R \in \mathcal{L}(V)$ such that $Re_k = v_k$ for $k = 1, \dots, n$. Then

$$\langle Te_k, e_j \rangle = \langle v_k, v_j \rangle = \langle Re_k, Re_j \rangle = \langle R^*Re_k, e_j \rangle.$$

Let $S = T - R^*R$. Then

$$\langle Se_k, e_j \rangle = 0$$

for all $j, k = 1, \dots, n$. By 6.30(a), we have

$$Se_k = \sum_{j=1}^{n} \langle Se_k, e_j \rangle e_j = 0.$$

Since this holds for all $k = 1, \dots, n$, we have $S = 0$ by the linear map lemma so that $T = R^*R$. Therefore, T is positive by 7.38(f). ■

Exercise 7C21

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.

Axler Note. *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. Let $e_1, \dots, e_n$ be the orthonormal basis of V with respect to which the matrix of T is the n-by-n Hilbert matrix. Let $H_n = \mathcal{M}(T, (e_1, \dots, e_n))$. Since H_n has real entries, and $\frac{1}{j+k-1} = \frac{1}{k+j-1}$, it follows that $H_n = H_n^*$. By 7.9, it follows that T is self-adjoint.

Now, observe that

$$\int_0^1 x^{j-1} x^{k-1} \, dx = \frac{1}{j+k-1}$$

for all $j, k = 1, \dots, n$. Consider $v = \sum_{k=1}^n \alpha_k e_k \in V$ for $\alpha_1, \dots, \alpha_n \in \mathbf{F}$, and define the polynomial $p(x) = \sum_{k=1}^n \alpha_k x^{k-1}$. Note that $\overline{p(x)} = \sum_{k=1}^n \overline{\alpha_k} x^{k-1}$ since x is real. Then

$$\begin{aligned}
\langle Tv, v \rangle &= \sum_{j=1}^n \sum_{k=1}^n \alpha_k \overline{\alpha_j} \langle Te_k, e_j \rangle \\
&= \sum_{j=1}^n \sum_{k=1}^n \alpha_k \overline{\alpha_j} \frac{1}{j+k-1} \\
&= \sum_{j=1}^n \sum_{k=1}^n \alpha_k \overline{\alpha_j} \int_0^1 x^{j-1} x^{k-1} \, dx \\
&= \int_0^1 \sum_{j=1}^n \sum_{k=1}^n \alpha_k \overline{\alpha_j} x^{j-1} x^{k-1} \, dx \\
&= \int_0^1 \sum_{k=1}^n \alpha_k x^{k-1} \sum_{j=1}^n \overline{\alpha_j} x^{j-1} \, dx \\
&= \int_0^1 p(x) \overline{p(x)} \, dx = \int_0^1 |p(x)|^2 \, dx \geq 0.
\end{aligned}$$

Hence, T is positive. Moreover, suppose $Tv = 0$. Then $\langle Tv, v \rangle = 0$, which implies $\int_0^1 |p(x)|^2 \, dx = 0$. Note that $|p(x)|^2$ is a continuous and nonnegative function on $[0, 1]$, so $\int_0^1 |p(x)|^2 \, dx = 0$ implies that $|p(x)|^2 = 0$ for all $x \in [0, 1]$. Then p has infinitely many zeros in $[0, 1]$ while a nonzero polynomial of degree at most $n - 1$ can have at most $n - 1$ zeros. Therefore, p is the zero polynomial. Hence, $\alpha_1 = \cdots = \alpha_n = 0$, so $v = 0$. Since V is finite-dimensional, it follows that T is invertible by 3.65. ∎

Exercise 7C22

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. Since $u \in V$, $V \neq \{0\}$. By 7.38(c), there exists an orthonormal basis $e_1, \dots, e_n$ of V consisting of eigenvectors of

T with nonnegative eigenvalues $\lambda_1, \dots, \lambda_n$. Define

$$\lambda = \max_{1 \le k \le n} \lambda_k.$$

For any $v \in V$ with $\|v\| = 1$, write $v = \sum_{k=1}^{n} \alpha_k e_k$. Then

$$\|Tv\|^2 = \left\| \sum_{k=1}^{n} \alpha_k \lambda_k e_k \right\|^2 = \sum_{k=1}^{n} |\alpha_k|^2 \lambda_k^2 \le \lambda^2 \sum_{k=1}^{n} |\alpha_k|^2 = \lambda^2 \|v\|^2 = \lambda^2.$$

Since $\|Tu\| \ge \|Tv\|$ for all $v \in V$ with $\|v\| = 1$, the above shows that $\|Tu\| \le \lambda$. Moreover, if e_k is an eigenvector corresponding to the eigenvalue λ, then $\|Te_k\| = \lambda$. By hypothesis, $\|Tu\| \ge \|Te_k\| = \lambda$, so $\|Tu\| = \lambda$. Let $u = \sum_{k=1}^{n} \beta_k e_k$. Then $\sum_{k=1}^{n} |\beta_k|^2 = 1$, so

$$\|Tu\|^2 = \sum_{k=1}^{n} |\beta_k|^2 \lambda_k^2 = \lambda^2, \quad \text{and} \quad \lambda^2 \sum_{k=1}^{n} |\beta_k|^2 = \lambda^2.$$

Observe that

$$\sum_{k=1}^{n} |\beta_k|^2 (\lambda^2 - \lambda_k^2) = 0.$$

Since $\lambda_k \le \lambda$ for all k, equality holds if and only if $|\beta_k|^2 = 0$ for all k with $\lambda_k < \lambda$. Hence, u is a linear combination of eigenvectors corresponding to the eigenvalue λ. Therefore, u is an eigenvector of T corresponding to the largest eigenvalue λ. ∎

Exercise 7C23

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) $\boxed{\Rightarrow}$ Suppose $\langle \cdot, \cdot \rangle_T$ is an inner product on V. By positivity, $\langle Tv, v \rangle \ge 0$ for all $v \in V$. By conjugate symmetry,

$$\langle Tu, v \rangle = \langle u, v \rangle_T = \overline{\langle v, u \rangle_T} = \overline{\langle Tv, u \rangle} = \langle u, Tv \rangle$$

for all $u, v \in V$, so T is self-adjoint. Therefore, T is positive. If $Tv = 0$ for some $v \in V$, then $\langle Tv, v \rangle = 0$, which implies $v = 0$ by definiteness of the inner product $\langle \cdot, \cdot \rangle_T$. Hence, T is invertible by 3.65.

$\boxed{\Leftarrow}$ Suppose T is an invertible positive operator. By 7.38(f), there exists $R \in \mathcal{L}(V)$ such that $R^*R = T$, so

$$\langle u, v \rangle_T = \langle Tu, v \rangle = \langle R^*Ru, v \rangle = \langle Ru, Rv \rangle$$

for all $u, v \in V$. Since $\langle \cdot, \cdot \rangle$ is an inner product and R is linear, $\langle \cdot, \cdot \rangle_T$ satisfies all the inner product axioms, except possibly definiteness. If $\langle v, v \rangle_T = 0$, then $\|Rv\|^2 = 0$, which implies $Rv = 0$ and hence $Tv = R^*(Rv) = 0$. Since T is invertible, this implies $v = 0$. Conversely, $\langle 0, 0 \rangle_T = \langle T0, 0 \rangle = \langle 0, 0 \rangle = 0$. Therefore, $\langle \cdot, \cdot \rangle_T$ is an inner product on V.

(b) Let $\langle \cdot, \cdot \rangle'$ be an inner product on V. For each $v \in V$, define $\varphi_v : V \to \mathbf{F}$ by $\varphi_v(u) = \langle u, v \rangle'$. Since $\langle \cdot, \cdot \rangle'$ is an inner product on V, φ_v is a linear functional on V. By the Riesz representation theorem, there exists a unique $Sv \in V$ such that $\varphi_v(u) = \langle u, Sv \rangle$ for all $u \in V$. This defines a function $S : V \to V$. We claim that S is linear. To see this, let $v, w \in V$ and $\mu \in \mathbf{F}$. Then for all $u \in V$,

$$\langle u, S(v + w) \rangle = \langle u, v + w \rangle' = \langle u, v \rangle' + \langle u, w \rangle' = \langle u, Sv \rangle + \langle u, Sw \rangle = \langle u, Sv + Sw \rangle,$$

which implies $S(v+w) = Sv + Sw$ by the uniqueness part of the Riesz representation theorem. Similarly,

$$\langle u, S(\mu v)\rangle = \langle u, \mu v\rangle' = \bar{\mu}\langle u, v\rangle' = \bar{\mu}\langle u, Sv\rangle = \langle u, \mu Sv\rangle,$$

which implies $S(\mu v) = \mu Sv$ by the uniqueness part of the Riesz representation theorem. Hence, S is linear. By conjugate symmetry of both $\langle\cdot,\cdot\rangle'$ and $\langle\cdot,\cdot\rangle$,

$$\langle u, v\rangle' = \overline{\langle v, u\rangle'} = \overline{\langle v, Su\rangle} = \langle Su, v\rangle = \langle u, v\rangle_S$$

for all $u, v \in V$, so $\langle\cdot,\cdot\rangle' = \langle\cdot,\cdot\rangle_S$. Since $\langle\cdot,\cdot\rangle_S$ is an inner product, part (a) implies that S is positive and invertible. ■

Exercise 7C24

Suppose S and T are positive operators on V. Prove that

$$\operatorname{null}(S+T) = \operatorname{null} S \cap \operatorname{null} T.$$

Solution. $\boxed{\subseteq}$ Let $v \in \operatorname{null}(S+T)$. Then $(S+T)v = 0$, so

$$0 = \langle (S+T)v, v\rangle = \langle Sv, v\rangle + \langle Tv, v\rangle.$$

Since S and T are positive, $\langle Sv, v\rangle$ and $\langle Tv, v\rangle$ are both nonnegative. Because their sum is zero, we must have $\langle Sv, v\rangle = 0$ and $\langle Tv, v\rangle = 0$. By 7.43, $Sv = 0$ and $Tv = 0$, so $v \in \operatorname{null} S \cap \operatorname{null} T$.
$\boxed{\supseteq}$ Conversely, if $v \in \operatorname{null} S \cap \operatorname{null} T$, then $Sv = 0$ and $Tv = 0$, so $(S+T)v = Sv + Tv = 0$. Hence, $v \in \operatorname{null}(S+T)$. ■

Exercise 7C25

Let T be the second derivative operator in Exercise 7A31(b). Show that $-T$ is a positive operator.

Solution. Let $D \in \mathcal{L}(V)$ by $Df = f'$. By Exercise 7A31(a), $D^* = -D$ for $Df = f'$. Since $Tf = f''$, we have $T = D^2$. Therefore, $-T = -D^2 = (-D)(D) = D^*D$, which is positive by 7.38(f). ■

7D Isometries, Unitary Operators, and Matrix Factorization

Exercise 7D1

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.

$\boxed{\Rightarrow}$ Let S be an isometry, and let e_1, e_2 be an orthonormal list of length two in V. By 7.49(c), $\langle Se_i, Se_j \rangle = \langle e_i, e_j \rangle$ for all $i, j \in \{1, 2\}$, so Se_1, Se_2 is an orthonormal list in W.

$\boxed{\Leftarrow}$ Let $u \in V$ with $\|u\| = 1$. By 6.36 and $\dim V \geq 2$, extend u to an orthonormal list u, v of length two in V. By hypothesis, Su, Sv is an orthonormal list in W; in particular, $\|Su\| = 1$. Now let $x \in V$. If $x = 0$, then $\|Sx\| = 0 = \|x\|$. Otherwise, $\frac{x}{\|x\|}$ has norm 1. Then

$$\|Sx\| = \left\| S\left(\frac{x}{\|x\|}\right) \right\| \|x\| = \|x\|.$$

Hence, S preserves norms so that S is an isometry. ∎

Exercise 7D2

Suppose $T \in \mathcal{L}(V, W)$ and $T \neq 0$. Prove that T is a scalar multiple of an isometry if and only if T preserves orthogonality.

Axler Note. *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.

$\boxed{\Rightarrow}$ Let $T = \lambda S$, where $\lambda \in \mathbf{F}$ and $S \in \mathcal{L}(V, W)$ is an isometry. Then for any $u, v \in V$ such that $\langle u, v \rangle = 0$, we have

$$\langle Tu, Tv \rangle = \langle \lambda Su, \lambda Sv \rangle = \lambda \overline{\lambda} \langle Su, Sv \rangle = |\lambda|^2 \langle Su, Sv \rangle = |\lambda|^2 \langle u, v \rangle = 0$$

by 7.49(c). Hence, T preserves orthogonality.

$\boxed{\Leftarrow}$ Since $T \neq 0$, $V \neq \{0\}$. Let $e_1, \ldots, e_n$ be an orthonormal basis of V, where $n = \dim V \geq 1$. For $j, k \in \{1, \ldots, n\}$ with $j \neq k$, we have $\langle e_j, e_k \rangle = 0$, so by assumption $\langle Te_j, Te_k \rangle = 0$. Since $\|e_j\| = \|e_k\|$ and $\langle e_k, e_j \rangle = \langle e_j, e_k \rangle = 0$, it follows that $\langle e_j + e_k, e_j - e_k \rangle = 0$. Therefore,

$$0 = \langle T(e_j + e_k), T(e_j - e_k) \rangle = \langle Te_j + Te_k, Te_j - Te_k \rangle = \langle Te_j, Te_j \rangle - \langle Te_k, Te_k \rangle = \|Te_j\|^2 - \|Te_k\|^2,$$

where $\langle Te_j, Te_k \rangle = \langle Te_k, Te_j \rangle = 0$ for $j \neq k$. Hence, $\|Te_j\| = \|Te_k\|$ for all $j, k \in \{1, \ldots, n\}$. Let $\lambda = \|Te_1\|$. Then $\|Te_j\| = \lambda$ for all $j = 1, \ldots, n$. If $\lambda = 0$, then $Te_j = 0$ for all j, so $T = 0$ on a basis. This implies $T = 0$, which contradicts the assumption that $T \neq 0$. Hence, $\lambda > 0$. Define $S = \frac{1}{\lambda} T$. Then $\|Se_j\| = 1$ for all $j = 1, \ldots, n$. Moreover,

$$\langle Se_j, Se_k \rangle = \left\langle \frac{1}{\lambda} Te_j, \frac{1}{\lambda} Te_k \right\rangle = \frac{1}{\lambda^2} \langle Te_j, Te_k \rangle = 0$$

for $j \neq k$. Hence, $Se_1, \ldots, Se_n$ is an orthonormal list in W, and by 7.49(d), S is an isometry. Therefore, $T = \lambda S$ is a scalar multiple of an isometry. ∎

Exercise 7D3

(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.

Axler Note. *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) Let T_1 and T_2 be unitary operators on V. By 7.5(d) and 7.53(b), we have

$$(T_1T_2)^*(T_1T_2) = T_2^*T_1^*T_1T_2 = T_2^*T_2 = I, \quad (T_1T_2)(T_1T_2)^* = T_1T_2T_2^*T_1^* = T_1IT_1^* = T_1T_1^* = I.$$

Hence, T_1T_2 is a unitary operator by 7.53(b).

(b) Let T be a unitary operator on V. By 7.53(c), $T^{-1} = T^*$, and T^* is unitary by 7.53(f). ∎

Exercise 7D4

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. Let $T = A + iB$. By 7.5(a), (b), and the fact that A and B are self-adjoint, we have $T^* = A - iB$. Therefore,

$$TT^* = A^2 + B^2 + i(BA - AB), \quad T^*T = A^2 + B^2 + i(AB - BA).$$

$\boxed{\Rightarrow}$ Suppose T is unitary. By 7.53(b), $TT^* = T^*T = I$. Then

$$TT^* - T^*T = i(BA - AB) - i(AB - BA) = 2i(BA - AB) = 0,$$

so $AB = BA$. Moreover, $TT^* = I$ implies $A^2 + B^2 + i(BA - AB) = I$, and since $AB = BA$, this simplifies to $A^2 + B^2 = I$.
$\boxed{\Leftarrow}$ Suppose $AB = BA$ and $A^2 + B^2 = I$. Substituting these into either TT^* or T^*T, we get

$$TT^* = T^*T = A^2 + B^2 = I.$$

Hence, T is unitary by 7.53(b). ∎

Exercise 7D5

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.

$\boxed{\text{(a)} \Rightarrow \text{(b)}}$ Suppose S is a self-adjoint unitary operator. Then $S = S^*$. By 7.53(b), $S^*S = SS^* = I$, so $S^2 = I$. Let $P = \frac{1}{2}(S + I)$. By 7.5(a), (b), and (e), P is self-adjoint with

$$P^2 = \left(\frac{1}{2}(S + I)\right)^2 = \frac{1}{4}(S^2 + 2S + I) = \frac{1}{4}(I + 2S + I) = \frac{1}{2}(S + I) = P.$$

Since P is self-adjoint and $P^2 = P$, Exercise 7A20 gives that P is an orthogonal projection. Then $S = 2P - I$.

(b) $\Rightarrow$ (c) Suppose $S = 2P - I$ for some orthogonal projection P onto a subspace U of V. By 6.57, $Pu = u$ for $u \in U$, and $Pw = 0$ for $w \in U^\perp$. Hence,

$$Su = (2P - I)u = 2Pu - u = 2u - u = u, \quad Sw = (2P - I)w = 2Pw - w = 0 - w = -w$$

for $u \in U$ and $w \in U^\perp$.

(c) $\Rightarrow$ (a) Suppose there exists a subspace U of V such that $Su = u$ for every $u \in U$ and $Sw = -w$ for every $w \in U^\perp$. Since V is finite-dimensional, so is U. Let P be the orthogonal projection onto U. For any $v \in V$, write $v = u + w$ with $u \in U$ and $w \in U^\perp$ by 6.49. By 6.57(b) and (c),

$$(2P - I)v = (2P - I)(u + w) = 2P(u + w) - (u + w) = 2u - (u + w) = u - w = Su + Sw = Sv.$$

Hence, $S = 2P - I$. By Exercise 7A20, P is self-adjoint. It follows that $S^* = 2P^* - I = 2P - I = S$ by 7.5. By 6.57, $P^2 = P$, so

$$S^2 = (2P - I)^2 = 4P^2 - 4P + I = 4P - 4P + I = I$$

Therefore, $SS^* = S^*S = S^2 = I$, and S is a self-adjoint unitary operator by 7.53(b). ■

Exercise 7D6

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 $T_1 = S^*T_2S$.

Solution. Let e_2, e_5, e_7 and f_2, f_5, f_7 be unit eigenvectors of T_1 and T_2 corresponding to the eigenvalues $2, 5, 7$, respectively. By 7.22, each list is an orthonormal list in $\mathbf{F}^3$. By 6.28, each list is an orthonormal basis of $\mathbf{F}^3$. By the linear map lemma, there exists a linear operator $S \in \mathcal{L}(\mathbf{F}^3)$ such that $Se_2 = f_2$, $Se_5 = f_5$, and $Se_7 = f_7$. By 7.53(d), S is unitary, and $S^* = S^{-1}$ by 7.53(c). For each $k \in \{2, 5, 7\}$, we have $T_1e_k = ke_k$ and $T_2f_k = kf_k$, so

$$S^*T_2Se_k = S^*T_2f_k = S^*(kf_k) = kS^*f_k = kS^{-1}f_k = ke_k = T_1e_k.$$

Since T_1 and S^*T_2S agree on an orthonormal basis of $\mathbf{F}^3$, we have $T_1 = S^*T_2S$ by the linear map lemma. ■

Exercise 7D7

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 e_1, e_2, e_3, e_4 be the standard basis of $\mathbf{F}^4$, which is an orthonormal basis. Consider the operators $T_1, T_2 \in \mathcal{L}(\mathbf{F}^4)$ whose matrix representations with respect to e_1, e_2, e_3, e_4 are diagonal with diagonal entries $2, 2, 5, 7$ and $2, 5, 5, 7$, respectively. By 7.9, both operators are self-adjoint since real diagonal matrices equal their conjugate transposes. By 5.41, the eigenvalues of T_1 and T_2 are exactly $2, 5, 7$. Note that $E(2, T_1) = \operatorname{null}(T_1 - 2I) = \operatorname{span}(e_1, e_2)$ and $E(2, T_2) = \operatorname{null}(T_2 - 2I) = \operatorname{span}(e_1)$.

Now suppose $S \in \mathcal{L}(\mathbf{F}^4)$ is a unitary operator such that $T_1 = S^*T_2S$. By 7.53(c), $S^* = S^{-1}$. Then $ST_1 = S(S^*T_2S) = (SS^{-1})T_2S = T_2S$. For any $v \in E(2, T_1)$, we have $T_1v = 2v$. Thus,

$$T_2(Sv) = S(T_1v) = S(2v) = 2Sv,$$

which shows that S maps $E(2, T_1)$ into $E(2, T_2)$. Moreover, $\dim E(2, T_1) = 2$ and $\dim E(2, T_2) = 1$. By 3.22, $S|_{E(2,T_1)}$ cannot be injective, and therefore S cannot be injective. Since a unitary operator must be invertible by 7.53(c), this is a

contradiction. Hence, no such unitary operator exists. ■

Exercise 7D8

Prove or give a counterexample: If $S \in \mathcal{L}(V)$ and there exists an orthonormal basis $e_1, \dots, e_n$ of V such that $\|Se_k\| = 1$ for each e_k, then S is a unitary operator.

Solution. The statement is false in general. Consider $V = \mathbf{F}^2$ with the standard orthonormal basis e_1, e_2. Define $S \in \mathcal{L}(V)$ by $Se_1 = Se_2 = e_1$. Then $\|Se_1\| = \|Se_2\| = 1$, but S is not unitary since $S(e_1 - e_2) = 0$, which shows that S is not injective and therefore not invertible. ■

Exercise 7D9

Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$. Suppose every eigenvalue of T has absolute value 1 and $\|Tv\| \le \|v\|$ for every $v \in V$. Prove that T is a unitary operator.

Solution. If $V = \{0\}$, then T is trivially unitary. Otherwise, $V \neq \{0\}$, so by Schur's theorem, there exists an orthonormal basis $e_1, \dots, e_n$ of V such that $\mathcal{M}(T)$ is upper triangular with respect to this basis; call this matrix A. Then for each $k = 1, \dots, n$, $Te_k = \sum_{j=1}^{k} A_{j,k} e_j$. By 5.41, the diagonal entries $A_{k,k}$ are the eigenvalues of T, so $|A_{k,k}| = 1$ for each $k = 1, \dots, n$. By 6.24 and the hypothesis,

$$1 \ge \|Te_k\|^2 = \sum_{j=1}^{k} |A_{j,k}|^2 = |A_{k,k}|^2 + \sum_{j=1}^{k-1} |A_{j,k}|^2 = 1 + \sum_{j=1}^{k-1} |A_{j,k}|^2,$$

which implies $0 \ge \sum_{j=1}^{k-1} |A_{j,k}|^2$. Then $A_{j,k} = 0$ for all $j < k$, so A is diagonal, and $Te_k = A_{k,k} e_k$ for each $k = 1, \dots, n$. It follows that $e_1, \dots, e_n$ is an orthonormal basis of eigenvectors of T with eigenvalues of absolute value 1. By 7.55, T is unitary. ■

Exercise 7D10

Suppose $\mathbf{F} = \mathbf{C}$ and $T \in \mathcal{L}(V)$ is a self-adjoint operator such that $\|Tv\| \le \|v\|$ 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) By 7.5, $I - T^2$ is self-adjoint. Then for any $v \in V$,

$$\langle (I - T^2)v, v \rangle = \langle v, v \rangle - \langle T^2 v, v \rangle = \langle v, v \rangle - \langle Tv, T^* v \rangle = \langle v, v \rangle - \langle Tv, Tv \rangle = \|v\|^2 - \|Tv\|^2 \ge 0,$$

where the final inequality follows from $\|Tv\| \le \|v\|$ for all $v \in V$. Hence, $I - T^2$ is a positive operator.

(b) Let $B = \sqrt{I - T^2}$. By part (a) and 7.39, B is a positive self-adjoint operator. By Exercise 7C17, there exists a polynomial p such that $B = p(I - T^2)$, where p has real coefficients. Since T commutes with $I - T^2$, it commutes with B. Moreover,

$$T^2 + B^2 = T^2 + (I - T^2) = I.$$

By Exercise 7D4 $\boxed{\Leftarrow}$ with $A = T$ and B as above, $T + iB$ is a unitary operator. ■

Exercise 7D11

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. Consider the set $\mathcal{B} = \{v \in V : \|v\| \leq 1\}$. For any $T \in \mathcal{L}(V)$, let $T(\mathcal{B}) = \{Tv : v \in \mathcal{B}\}$. Now, if T is such that $T(\mathcal{B}) \subseteq \mathcal{B}$, then $\|Tv\| \leq 1$ for all $v \in \mathcal{B}$. If $v = 0$, then $\|T0\| = 0 = \|0\|$. If $v \neq 0$, then

$$\frac{\|Tv\|}{\|v\|} = \left\| T\left(\frac{v}{\|v\|}\right) \right\| \leq 1.$$

Hence, $\|Tv\| \leq \|v\|$ for all $v \in V$.

$\boxed{\Rightarrow}$ Suppose S is a unitary operator. Then S is also an isometry, so $S(\mathcal{B}) \subseteq \mathcal{B}$. By Exercise 7D3(b), S^{-1} is also unitary. Then for any $u \in \mathcal{B}$, $S^{-1}u \in \mathcal{B}$. Therefore, $u = S(S^{-1}u) \in S(\mathcal{B})$, so $S(\mathcal{B}) = \mathcal{B}$.

$\boxed{\Leftarrow}$ Now suppose $S(\mathcal{B}) = \mathcal{B}$. Let $w \in V$, and write $w = \|w\|w_0$ for some $w_0 \in \mathcal{B}$. Since $\mathcal{B} = S(\mathcal{B})$, there exists $v_0 \in \mathcal{B}$ such that $w_0 = Sv_0$. Hence, $w = \|w\|Sv_0 = S(\|w\|v_0)$, so $w \in \operatorname{range} S$. Therefore, S is surjective, and hence invertible by 3.65. Also, if $u \in \mathcal{B}$, then $u = Sv$ for some $v \in \mathcal{B}$. Hence, $v = S^{-1}u \in \mathcal{B}$, so $S^{-1}(\mathcal{B}) \subseteq \mathcal{B}$. Applying the preliminary result from the beginning of the solution to S and S^{-1}, we get $\|Sx\| \leq \|x\|$ and $\|x\| = \|S^{-1}(Sx)\| \leq \|Sx\|$, so $\|Sx\| = \|x\|$ for all $x \in V$. Therefore, S is an isometry, and since it is also invertible, it is unitary. ■

Exercise 7D12

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.

Manual Remark. *The solution below uses 2 and $\frac{1}{2}$, but replacing them with λ and $\frac{1}{\lambda}$ for any nonzero $\lambda \in \mathbf{F}$ with $|\lambda| \neq 1$ gives a similar counterexample.*

Solution. The statement is false. Consider $\mathbf{F}^2$ with the Euclidean inner product, and let $S \in \mathcal{L}(\mathbf{F}^2)$ be defined by $S(x, y) = (2y, \frac{1}{2}x)$. Then

$$S^2(x, y) = S\left(2y, \frac{1}{2}x\right) = \left(2 \cdot \frac{1}{2}x, \frac{1}{2} \cdot 2y\right) = (x, y).$$

Hence, S is invertible with $S^{-1} = S$. Then $\|S^{-1}(x, y)\| = \|S(x, y)\|$ for all $(x, y) \in \mathbf{F}^2$. However, $\|S(0, 1)\| = 2 \neq 1 = \|(0, 1)\|$. Therefore, S is not unitary. ■

Exercise 7D13

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. Let Q be an n-by-n matrix with complex entries. Define $T \in \mathcal{L}(\mathbf{C}^n)$ by $Tv = Qv$ for all $v \in \mathbf{C}^n$. Moreover, let $\mathbf{C}^n$ have the Euclidean inner product. If $e_1, \ldots, e_n$ denotes the standard (orthonormal) basis of $\mathbf{C}^n$, then Te_k is the kth column of Q. Hence, $\mathcal{M}(T) = Q$ with respect to $e_1, \ldots, e_n$. Then

$$\begin{aligned}
Q \text{ has orthonormal columns} &\iff Te_1, \ldots, Te_n \text{ is an orthonormal list} \\
&\iff Te_1, \ldots, Te_n \text{ is an orthonormal basis by 6.28} \\
&\iff T \text{ is unitary by 7.53(d)} \\
&\iff \text{the rows of } \mathcal{M}(T) = Q \text{ form an orthonormal basis by 7.53(e)} \\
&\iff \text{the rows of } Q \text{ form an orthonormal list by 6.28.}
\end{aligned}$$

■

Exercise 7D14

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.

$\boxed{\Rightarrow}$ Note that $\|Sv\| = \|v\| = 1$ since S is unitary. By the Cauchy–Schwarz inequality, $|b| = |\langle Sv, v\rangle| \leq \|Sv\|\,\|v\| = 1$.

$\boxed{\Leftarrow}$ Suppose $|b| \leq 1$. By 6.36, we may extend v to an orthonormal basis $v, u, e_3, \ldots, e_n$ of V, which has length $\dim V \geq 2$. Let $\beta = \sqrt{1 - |b|^2}$. By the linear map lemma, there exists an operator $S \in \mathcal{L}(V)$ such that

$$Sv = bv + \beta u, \quad Su = -\beta v + \bar{b}u, \quad Se_k = e_k \text{ for } k \geq 3.$$

Then

$$\|Sv\|^2 = \|bv + \beta u\|^2 = |b|^2\|v\|^2 + |\beta|^2\|u\|^2 = |b|^2 + (1 - |b|^2) = 1,$$

and

$$\|Su\|^2 = \|-\beta v + \bar{b}u\|^2 = |\beta|^2\|v\|^2 + |b|^2\|u\|^2 = (1 - |b|^2) + |b|^2 = 1.$$

Moreover, noting that $\bar{\bar{b}} = b$, we have

$$\langle Sv, Su\rangle = \langle bv + \beta u, -\beta v + \bar{b}u\rangle = b(-\beta)\langle v, v\rangle + b^2\langle v, u\rangle - \beta^2\langle u, v\rangle + \beta b\langle u, u\rangle = -b\beta + \beta b = 0.$$

Lastly, $Sv, Su \in \operatorname{span}(v, u)$, which is orthogonal to all e_k for $k \geq 3$. Then the list $Sv, Su, Se_3, \ldots, Se_n$ is an orthonormal basis of V by 6.28, so S is unitary by 7.53(d). Finally,

$$\langle Sv, v\rangle = \langle bv + \beta u, v\rangle = b\langle v, v\rangle + \beta\langle u, v\rangle = b. \qquad \blacksquare$$

Exercise 7D15

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.

Axler Note. *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. Define $A = (T + I)(T - I)^{-1}$. Observe that $T + I = (T - I) + 2I$. Then $A = I + 2(T - I)^{-1}$.

(a) By 7.5, $A^* = I + 2(T^* - I)^{-1}$. By 7.53(c), $T^* = T^{-1}$, so $T^* - I = -T^{-1}(T - I)$, and

$$(T^* - I)^{-1} = -(T - I)^{-1}T = -(T - I)^{-1}((T - I) + I) = -I - (T - I)^{-1}.$$

Therefore, $A^* = -I - 2(T - I)^{-1} = -A$.

(b) Using 7.5(b) and part (a), $(iA)^* = \bar{i}A^* = (-i)(-A) = iA$. $\blacksquare$

Exercise 7D16

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. Observe that every eigenvalue of T is real by 7.12. Therefore, neither i nor $-i$ is an eigenvalue of T so that $T - iI$ and $T + iI$ are invertible. Define $S = (T + iI)(T - iI)^{-1}$. By 7.5, $(T + iI)^* = T^* - iI = T - iI$ and

$((T - iI)^{-1})^* = ((T - iI)^*)^{-1} = (T + iI)^{-1}$. Thus,

$$S^* = (T + iI)^{-1}(T - iI).$$

Because $T - iI$ and $T + iI$ commute,

$$S^*S = (T + iI)^{-1}(T - iI)(T + iI)(T - iI)^{-1} = I,$$

showing that S is an isometry on V by 7.49(b), whence S is unitary.

Now, suppose $Sx = x$ for some $x \in V$. By definition of S, $(T + iI)(T - iI)^{-1}x = x$, so $(T + iI)y = x$ where $y = (T - iI)^{-1}x$. Since $x = (T - iI)y$, we have $(T + iI)y = (T - iI)y$, so $2iy = 0$, and hence $y = 0$. Therefore, $x = (T - iI)y = 0$, showing that 1 is not an eigenvalue of S. ∎

Exercise 7D17

Explain why the characterizations of unitary matrices given by 7.57 hold.

Solution. Let Q be an n-by-n matrix. Define $T \in \mathcal{L}(\mathbf{F}^n)$ by $Tv = Qv$ for all $v \in \mathbf{F}^n$, so $\mathcal{M}(T) = Q$ with respect to the standard basis of $\mathbf{F}^n$, which is orthonormal with respect to the Euclidean inner product.

(a) $\Leftrightarrow$ (c) Since $Tv = Qv$ for all $v \in \mathbf{F}^n$, (c) states precisely that T is an isometry. By the equivalence of 7.49(a) and (e), T is an isometry if and only if the columns of $\mathcal{M}(T) = Q$ are orthonormal, which is (a) by definition.

(a) $\Leftrightarrow$ $Q^*Q = I$, (b) $\Leftrightarrow$ $QQ^* = I$ For $j, k \in \{1, \dots, n\}$, consider the j, kth entry of both Q^*Q and QQ^*. By 3.46 and the definition of the conjugate transpose, we have

$$(Q^*Q)_{j,k} = \sum_{\ell=1}^{n} \overline{Q_{\ell,j}} Q_{\ell,k} = \langle Q_{\cdot,k}, Q_{\cdot,j} \rangle, \quad (QQ^*)_{j,k} = \sum_{\ell=1}^{n} Q_{j,\ell} \overline{Q_{k,\ell}} = \langle Q_{j,\cdot}, Q_{k,\cdot} \rangle.$$

Note that $Q_{j,\cdot}$ and $Q_{k,\cdot}$ are treated as vectors in $\mathbf{F}^n$. Therefore, $Q^*Q = I$ if and only if the columns of Q are orthonormal, which is (a), and $QQ^* = I$ if and only if the rows of Q are orthonormal, which is (b).

$Q^*Q = I \Leftrightarrow QQ^* = I$ By 7.9 and 3.43, $\mathcal{M}(T^*T) = Q^*Q$ and $\mathcal{M}(TT^*) = QQ^*$. Then $Q^*Q = I$ if and only if $T^*T = I$, and $QQ^* = I$ if and only if $TT^* = I$. By 3.68, $T^*T = I$ if and only if $TT^* = I$. Therefore, $Q^*Q = I$ if and only if $QQ^* = I$, so each condition is equivalent to (d).

This shows that (a) is equivalent to (c) and that (a) is equivalent to (d), which is equivalent to (b). ∎

Exercise 7D18

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. Let A be a symmetric n-by-n matrix with real entries. Define $T \in \mathcal{L}(\mathbf{R}^n)$ by $Tv = Av$ for all $v \in \mathbf{R}^n$, so $\mathcal{M}(T) = A$ with respect to the standard orthonormal basis of $\mathbf{R}^n$. Since A is real and symmetric, $A^* = A^t = A$. Then $\mathcal{M}(T^*) = A^* = A = \mathcal{M}(T)$, so T is self-adjoint by 7.9. By the real spectral theorem, there exists an orthonormal basis $f_1, \dots, f_n$ of $\mathbf{R}^n$ such that $Tf_k = \lambda_k f_k$ for some real numbers $\lambda_1, \dots, \lambda_n$.

Define Q to be the matrix whose columns are the vectors $f_1, \dots, f_n$. Moreover, let D be the diagonal matrix with entries $\lambda_1, \dots, \lambda_n$ on its diagonal. By definition, Q is a matrix with real entries and orthonormal columns, so it is unitary. Moreover, the kth column of AQ is $Af_k = Tf_k = \lambda_k f_k$, which is also the kth column of QD for each $k = 1, \dots, n$. Hence, $AQ = QD$. By 7.57(d), $Q^*Q = I$, so $Q^*AQ = Q^*QD = D$. ∎

Exercise 7D19

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 $z = (z_0, z_1, \ldots, z_{n-1})$. Define linear functionals $\omega_0, \omega_1, \ldots, \omega_{n-1}$ on $\mathbf{C}^n$ by

$$\omega_j(z_0, z_1, \ldots, z_{n-1}) = \frac{1}{\sqrt{n}} \sum_{m=0}^{n-1} z_m e^{-2\pi i j m/n}.$$

The **discrete Fourier transform** is the operator $\mathcal{F} : \mathbf{C}^n \to \mathbf{C}^n$ defined by

$$\mathcal{F}z = (\omega_0(z), \omega_1(z), \ldots, \omega_{n-1}(z)).$$

(a) Show that $\mathcal{F}$ is a unitary operator on $\mathbf{C}^n$.

(b) Show that if $(z_0, \ldots, z_{n-1}) \in \mathbf{C}^n$ and z_n is defined to equal z_0, then

$$\mathcal{F}^{-1}(z_0, z_1, \ldots, z_{n-1}) = \mathcal{F}(z_n, z_{n-1}, \ldots, z_1).$$

(c) Show that $\mathcal{F}^4 = I$.

Axler Note. *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. Let $e_0, e_1, \ldots, e_{n-1}$ be the standard orthonormal basis of $\mathbf{C}^n$, so the jth component of $\mathcal{F}e_k$ is

$$\omega_j(e_k) = \frac{1}{\sqrt{n}} \sum_{p=0}^{n-1} (e_k)_p e^{-2\pi i j p/n} = \frac{1}{\sqrt{n}} e^{-2\pi i j k/n},$$

where $(e_k)_p = \delta_{pk}$.

(a) Let $m, k \in \{0, 1, \ldots, n-1\}$. Then

$$\langle \mathcal{F}e_m, \mathcal{F}e_k \rangle = \sum_{j=0}^{n-1} \omega_j(e_m)\overline{\omega_j(e_k)} = \sum_{j=0}^{n-1} \frac{1}{\sqrt{n}} e^{-2\pi i j m/n} \cdot \frac{1}{\sqrt{n}} e^{2\pi i j k/n} = \frac{1}{n} \sum_{j=0}^{n-1} e^{2\pi i j (k-m)/n}.$$

If $m = k$, then $\langle \mathcal{F}e_m, \mathcal{F}e_k \rangle = 1$. Otherwise, let $\zeta = e^{2\pi i (k-m)/n}$. Then $\zeta^n = 1$ and $\zeta \neq 1$, since $0 < |k - m| < n$. Therefore,

$$\langle \mathcal{F}e_m, \mathcal{F}e_k \rangle = \frac{1}{n} \sum_{j=0}^{n-1} \zeta^j = \frac{1}{n} \cdot \frac{1 - \zeta^n}{1 - \zeta} = 0.$$

Hence, $\mathcal{F}e_0, \ldots, \mathcal{F}e_{n-1}$ is an orthonormal list in $\mathbf{C}^n$. By 6.28, $\mathcal{F}e_0, \ldots, \mathcal{F}e_{n-1}$ is an orthonormal basis of $\mathbf{C}^n$, so $\mathcal{F}$ is unitary on $\mathbf{C}^n$ by 7.53(d).

(b) Define $R \in \mathcal{L}(\mathbf{C}^n)$ by $R(z_0, z_1, \ldots, z_{n-1}) = (z_{n-1}, z_{n-2}, \ldots, z_0)$. Then for any $z \in \mathbf{C}^n$ and $j \in \{0, \ldots, n-1\}$,

$$(\mathcal{F}^2 z)_j = \frac{1}{\sqrt{n}} \sum_{m=0}^{n-1} (\mathcal{F}z)_m e^{-2\pi i j m/n} = \frac{1}{n} \sum_{m=0}^{n-1} \sum_{k=0}^{n-1} z_k e^{-2\pi i m(k+j)/n} = \sum_{k=0}^{n-1} z_k \left(\frac{1}{n} \sum_{m=0}^{n-1} e^{-2\pi i m(k+j)/n} \right)$$

where $r_k = e^{-2\pi i (k+j)/n}$. Observe that $0 \leq k + j \leq 2n - 2$, so $r_k \neq 1$ unless $k + j = 0$ or $k + j = n$ in which case the inner sum evaluates to 1. Otherwise, we have $r_k \neq 1$ and $r_k^n = 1$, so

$$\frac{1}{n} \sum_{m=0}^{n-1} r_k^m = \frac{1}{n} \cdot \frac{1 - r_k^n}{1 - r_k} = 0.$$

Observe that for $k \in \{0, \ldots, n-1\}$, one such k exists for which $k + j = 0$ or $k + j = n$, namely $k = 0$ if $j = 0$ or $k = n - j$ if $1 \le j \le n - 1$. Hence, $(\mathcal{F}^2 z)_0 = z_0$ and for $j \ge 1$, $(\mathcal{F}^2 z)_j = z_{n-j}$, whereupon $\mathcal{F}^2 z = Rz$. Since R fixes the 0th coordinate and reverses the jth and $(n-j)$th coordinates for $j \ge 1$, R is its own inverse, i.e., $R^2 = I$. From this, we get $\mathcal{F}(\mathcal{F}R) = \mathcal{F}^2 R = R^2 = I$. By 3.68, $(\mathcal{F}R)\mathcal{F} = I$, and hence $\mathcal{F}^{-1} = \mathcal{F}R$. Because $z_n = z_0$,

$$\mathcal{F}^{-1}(z_0, z_1, \ldots, z_{n-1}) = (\mathcal{F}R)(z_0, z_1, \ldots, z_{n-1}) = \mathcal{F}(z_0, z_{n-1}, \ldots, z_1).$$

(c) By part (b), $\mathcal{F}^2 = R$ and $R^2 = I$, so $\mathcal{F}^4 = R^2 = I$. ■

Exercise 7D20

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. Since A is n-by-n and has linearly independent columns, $\operatorname{column\,rank} A = n$. By 3.57, $\operatorname{row\,rank} A = \operatorname{column\,rank} A = n$, so the rows of A are linearly independent as well. Since the columns of A^* are the conjugate transposes of the rows of A, the columns of A^* are linearly independent. By 7.58, there exist unique matrices Q_1 and R_1 such that Q_1 is unitary, R_1 is upper triangular with positive diagonal entries, and $A^* = Q_1 R_1$. Taking conjugate transposes gives $A = R_1^* Q_1^*$, where R_1^* is lower triangular with the same positive diagonal entries as R_1, and Q_1^* is unitary by 7.57(d). Let $R = R_1^*$ and $Q = Q_1^*$, so $A = RQ$ where R and Q have the desired properties.

To see uniqueness of R and Q, suppose $A = RQ = R_0 Q_0$ where R_0 is lower triangular with positive diagonal entries and Q_0 is unitary. Taking conjugate transposes again yields $A^* = Q^* R^* = Q_0^* R_0^*$, where R^* and R_0^* are upper triangular with positive diagonal entries and Q^* and Q_0^* are unitary. By the uniqueness part of 7.58, we must have $R^* = R_0^*$ and $Q^* = Q_0^*$, which implies $R = R_0$ and $Q = Q_0$. ■

7E Singular Value Decomposition

Lemma 2

Suppose $T \in \mathcal{L}(V, W)$ and $v_1, \ldots, v_n$ is a basis of V with $T^*Tv_k = \mu_k v_k$ for $k = 1, \ldots, n$. Then the singular values of T are $\sqrt{\mu_1}, \ldots, \sqrt{\mu_n}$ sorted into decreasing order.

Proof. Each μ_k is an eigenvalue of T^*T since $v_k \neq 0$. By 7.64(a), T^*T is positive, so $\mu_k \geq 0$ for all $k = 1, \ldots, n$ by 7.38(b). Since $T^*Tv_k = \mu_k v_k$ for each k, the matrix $\mathcal{M}(T^*T, (v_1, \ldots, v_n))$ is diagonal with diagonal entries $\mu_1, \ldots, \mu_n$. By Exercise 5D2, each eigenvalue λ of T^*T appears in the list $\mu_1, \ldots, \mu_n$ exactly $\dim E(\lambda, T^*T)$ times. By the definition of singular values, the singular values of T are precisely $\sqrt{\mu_1}, \ldots, \sqrt{\mu_n}$ sorted into decreasing order. ■

Exercise 7E1

Suppose $T \in \mathcal{L}(V, W)$. Show that $T = 0$ if and only if all singular values of T are 0.

Solution.

$\boxed{\Rightarrow}$ If $T = 0$, then $\operatorname{range} T = \{0\}$. Hence, $\dim \operatorname{range} T = 0$. By 7.68(b), T has no positive singular values. Since all singular values are nonnegative, every singular value of T is 0.

$\boxed{\Leftarrow}$ Suppose all singular values of T are 0. Then T has no positive singular values, so $\dim \operatorname{range} T = 0$ by 7.68(b). Therefore, $\operatorname{range} T = \{0\}$, so $T = 0$. ■

Exercise 7E2

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

$$Tv = sw \quad \text{and} \quad T^*w = sv.$$

Axler Note. *The vectors v, w satisfying both equations above are called a* **Schmidt pair**. *Erhard Schmidt introduced the concept of singular values in 1907.*

Solution.

$\boxed{\Rightarrow}$ Suppose s is a singular value of T. Since s is a square root of an eigenvalue of T^*T, s^2 is an eigenvalue of T^*T. Let $v \in V$ be nonzero such that $T^*Tv = s^2 v$. Let $w = \frac{1}{s}Tv$. It follows that $Tv = sw$, and

$$T^*w = T^*\left(\frac{1}{s}Tv\right) = \frac{1}{s}T^*Tv = \frac{1}{s}s^2 v = sv.$$

Since $sv \neq 0$, we have $w \neq 0$.

$\boxed{\Leftarrow}$ Suppose there exist nonzero vectors $v \in V$ and $w \in W$ such that $Tv = sw$ and $T^*w = sv$. Observe that $T^*Tv = sT^*w = s^2 v$. Then s^2 is an eigenvalue of T^*T. Since $s > 0$, $\sqrt{s^2} = s$ is a singular value of T. ■

Exercise 7E3

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. Consider $T \in \mathcal{L}(\mathbf{C}^2)$ defined by $T(z, w) = (5w, 0)$. With respect to the standard orthonormal basis e_1, e_2 of $\mathbf{C}^2$, the matrix of T is

$$\begin{pmatrix} 0 & 5 \\ 0 & 0 \end{pmatrix}.$$

Since the matrix of T is upper triangular with all diagonal entries 0, 0 is the only eigenvalue of T by 5.41. Using 7.9 and 3.43,

$$\mathcal{M}(T^*T) = \begin{pmatrix} 0 & 0 \\ 0 & 25 \end{pmatrix}.$$

Then $T^*Te_1 = 0$ and $T^*Te_2 = 25e_2$. Therefore, by Lemma 2, the singular values of T are $\sqrt{25} = 5$ and $\sqrt{0} = 0$. ■

Exercise 7E4

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. Let $n = \dim V$. Let $\mathcal{S} = \{\|Tv\| : v \in V \text{ and } \|v\| = 1\}$. Since T has a largest and a smallest singular value, $V \neq \{0\}$, so $n \geq 1$. Since T^*T is positive by 7.64(a), 7.38(c) gives an orthonormal basis $e_1, \ldots, e_n$ of V consisting of eigenvectors of T^*T, say $T^*Te_k = \mu_k e_k$ for each $k = 1, \ldots, n$. After reordering, Lemma 2 shows that $\mu_k = s_k^2$ for each $k = 1, \ldots, n$.

$\boxed{\subseteq}$ Suppose $\|v\| = 1$. Let $c_k = \langle v, e_k \rangle$ for each $k = 1, \ldots, n$. By 6.30(a) and 6.30(b), $v = \sum_{k=1}^n c_k e_k$ and $\sum_{k=1}^n |c_k|^2 = \|v\|^2 = 1$. We then have

$$\|Tv\|^2 = \langle T^*Tv, v \rangle = \left\langle \sum_{k=1}^n c_k s_k^2 e_k, \sum_{j=1}^n c_j e_j \right\rangle = \sum_{k=1}^n s_k^2 |c_k|^2.$$

Therefore, $\|Tv\|^2$ is a weighted average of $s_1^2, \ldots, s_n^2$ with weights $|c_k|^2$, so $s_n^2 \leq \|Tv\|^2 \leq s_1^2$, and taking square roots gives $s_n \leq \|Tv\| \leq s_1$. Hence, $\mathcal{S} \subseteq [s_n, s_1]$.

$\boxed{\supseteq}$ Let $s \in [s_n, s_1]$. If $s_1 = s_n$, take $v = e_1$. Otherwise, $n \geq 2$ and $s_n < s_1$. Define

$$\theta = \frac{s^2 - s_n^2}{s_1^2 - s_n^2}, \quad v = \sqrt{\theta} e_1 + \sqrt{1 - \theta} e_n.$$

Since $s_n < s_1$ and $s \in [s_n, s_1]$, we have $\theta \in [0, 1]$. By orthonormality of e_1 and e_n, we have $\|v\|^2 = \theta + (1 - \theta) = 1$. Additionally, $\langle v, e_1 \rangle = \sqrt{\theta}$, $\langle v, e_n \rangle = \sqrt{1 - \theta}$, and $\langle v, e_k \rangle = 0$ for $1 < k < n$. Thus,

$$\|Tv\|^2 = s_1^2 \theta + s_n^2 (1 - \theta) = s^2.$$

Hence, $\|Tv\| = s$, so $s \in \mathcal{S}$. Therefore, $[s_n, s_1] \subseteq \mathcal{S}$. Combining with the previous inclusion gives $\mathcal{S} = [s_n, s_1]$. ■

Exercise 7E5

Suppose $T \in \mathcal{L}(\mathbf{C}^2)$ is defined by $T(x, y) = (-4y, x)$. Find the singular values of T.

Solution. With respect to the standard orthonormal basis e_1, e_2 of $\mathbf{C}^2$, the matrix of T is

$$\mathcal{M}(T) = \begin{pmatrix} 0 & -4 \\ 1 & 0 \end{pmatrix},$$

so by 7.9, the matrix of T^* with respect to the same basis is

$$\mathcal{M}(T^*) = \begin{pmatrix} 0 & 1 \\ -4 & 0 \end{pmatrix}.$$

By 3.43, $\mathcal{M}(T^*T)$ is diagonal with diagonal entries 1 and 16. Therefore, $T^*Te_1 = e_1$ and $T^*Te_2 = 16e_2$. By Lemma 2, the singular values of T are $\sqrt{16} = 4$ and $\sqrt{1} = 1$. ■

Exercise 7E6

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. Let e_1, e_2, e_3 be the orthonormal basis of $\mathcal{P}_2(\mathbf{R})$ given by

$$e_1 = \sqrt{\frac{1}{2}}, \quad e_2 = \sqrt{\frac{3}{2}}x, \quad e_3 = \sqrt{\frac{45}{8}}\left(x^2 - \frac{1}{3}\right)$$

as described in Example 6.34. Then

$$De_1 = 0, \quad De_2 = \sqrt{\frac{3}{2}} = \sqrt{3}e_1, \quad De_3 = \sqrt{\frac{45}{8}} \cdot 2x = \sqrt{15}e_2.$$

Therefore, the matrix of D with respect to the orthonormal basis e_1, e_2, e_3 is

$$\mathcal{M}(D) = \begin{pmatrix} 0 & \sqrt{3} & 0 \\ 0 & 0 & \sqrt{15} \\ 0 & 0 & 0 \end{pmatrix}.$$

By 7.9, the matrix of D^* with respect to the same basis is

$$\mathcal{M}(D^*) = \begin{pmatrix} 0 & 0 & 0 \\ \sqrt{3} & 0 & 0 \\ 0 & \sqrt{15} & 0 \end{pmatrix}.$$

By 3.43, $\mathcal{M}(D^*D)$ is diagonal with diagonal entries 0, 3, and 15. Therefore, $D^*De_1 = 0$, $D^*De_2 = 3e_2$, and $D^*De_3 = 15e_3$. By Lemma 2, the singular values of D are $\sqrt{15}$, $\sqrt{3}$, and 0. ∎

Exercise 7E7

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. If T is self-adjoint, then $T^*T = T^2 = TT^*$, so T is normal. Hence, T is normal in either case. If $\mathbf{F} = \mathbf{R}$, then T is self-adjoint, so the real spectral theorem applies. If $\mathbf{F} = \mathbf{C}$ and T is normal, then the complex spectral theorem applies. In either case, we obtain an orthonormal basis $e_1, \ldots, e_{\dim V}$ of V with $Te_k = \mu_k e_k$ for $k = 1, \ldots, \dim V$, where $\mu_k \in \mathbf{F}$. Then $\mathcal{M}(T, (e_1, \ldots, e_{\dim V}))$ is diagonal with diagonal entries $\mu_1, \ldots, \mu_{\dim V}$. Since $e_k \neq 0$, each μ_k is an eigenvalue of T. By Exercise 5D2, each eigenvalue λ of T appears in the list exactly $\dim E(\lambda, T)$ times, so $\mu_1, \ldots, \mu_{\dim V}$ is a rearrangement of $\lambda_1, \ldots, \lambda_n$. In particular, $\dim V = n$. Thus, we may reorder the basis so that $\mu_k = \lambda_k$ for $k = 1, \ldots, n$. By 7.21(e), $T^*e_k = \overline{\lambda_k}e_k$ for $k = 1, \ldots, n$, so

$$T^*Te_k = \lambda_k T^*e_k = \lambda_k \overline{\lambda_k} e_k = |\lambda_k|^2 e_k.$$

By Lemma 2, the singular values of T are $|\lambda_1|, \ldots, |\lambda_n|$, sorted into decreasing order. ∎

Exercise 7E8

Suppose $T \in \mathcal{L}(V, W)$. Suppose $s_1 \geq s_2 \geq \cdots \geq s_m > 0$ and $e_1, \ldots, e_m$ is an orthonormal list in V and $f_1, \ldots, f_m$ 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 $f_1, \ldots, f_m$ is an orthonormal basis of range T.

(b) Prove that $e_1, \ldots, e_m$ is an orthonormal basis of $(\operatorname{null} T)^\perp$.

(c) Prove that $s_1, \ldots, s_m$ are the positive singular values of T.

(d) Prove that if $k \in \{1, \ldots, m\}$, 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. Consider the maps $S_k \in \mathcal{L}(V, W)$ defined by $S_k v = \langle v, e_k\rangle f_k$ for $k = 1, \ldots, m$. Then $T = \sum_{k=1}^m s_k S_k$. Following the calculations in Example 7.3, $S_k^* w = \langle w, f_k\rangle e_k$ for $k = 1, \ldots, m$ and $w \in W$. Since $s_k \in \mathbf{R}$, by 7.5(a) and (b), we have

$$T^*w = \sum_{k=1}^m s_k S_k^* w = \sum_{k=1}^m s_k\langle w, f_k\rangle e_k.$$

Since $e_1, \ldots, e_m$ and $f_1, \ldots, f_m$ are orthonormal lists, we have $Te_k = s_k f_k$ and $T^* f_k = s_k e_k$ for $k = 1, \ldots, m$. Hence, $T^*Te_k = T^*(s_k f_k) = s_k^2 e_k$ for $k = 1, \ldots, m$.

(a) The computation of Tv for $v \in V$ gives $\operatorname{range} T \subseteq \operatorname{span}(f_1, \ldots, f_m)$. Since $f_k = Te_k/s_k$, we have $f_k \in \operatorname{range} T$ for $k = 1, \ldots, m$. Hence, $\operatorname{span}(f_1, \ldots, f_m) \subseteq \operatorname{range} T$, and equality follows. By 6.25, orthonormal lists are linearly independent, so $f_1, \ldots, f_m$ is an orthonormal basis of $\operatorname{range} T$.

(b) Using the same argument as in part (a) with T^* in place of T, we get $\operatorname{range} T^* = \operatorname{span}(e_1, \ldots, e_m)$. Since $\operatorname{range} T^* = (\operatorname{null} T)^\perp$ by 7.6(b), it follows that $e_1, \ldots, e_m$ is an orthonormal basis of $(\operatorname{null} T)^\perp$.

(c) By 6.36, we may extend $e_1, \ldots, e_m$ to an orthonormal basis $e_1, \ldots, e_n$ of V. For $k \leq m$, the beginning calculation shows that $T^*Te_k = s_k^2 e_k$. For $k > m$, we have $Te_k = 0$, so $T^*Te_k = 0$. By Lemma 2 and the hypothesis that $s_1 \geq \cdots \geq s_m > 0$, it follows that the singular values of T are $s_1, \ldots, s_m$ together with $n - m$ zeros. Thus, the positive singular values of T are exactly $s_1, \ldots, s_m$.

(d) By the beginning calculation, $T^*Te_k = s_k^2 e_k$ and $e_k \neq 0$ since $\|e_k\| = 1$ for $k = 1, \ldots, m$. Hence, e_k is indeed an eigenvector of T^*T corresponding to the eigenvalue s_k^2.

(e) For any $w \in W$, we have

$$TT^*w = \sum_{k=1}^m s_k\langle w, f_k\rangle Te_k = \sum_{k=1}^m s_k\langle w, f_k\rangle s_k f_k = \sum_{k=1}^m s_k^2\langle w, f_k\rangle f_k.$$

■

Exercise 7E9

Suppose $T \in \mathcal{L}(V, W)$. Show that T and T^* have the same positive singular values.

Solution. If $T = 0$, then $T^* = 0$ by 7.5(b), so neither T nor T^* has any positive singular values, and the statement holds trivially. Now assume $T \neq 0$. Let $s_1, \ldots, s_m$ denote the positive singular values of T. By definition, $s_1 \geq \cdots \geq s_m > 0$. Moreover, $m = \dim \operatorname{range} T \geq 1$ by 7.68(b), since $T \neq 0$. By the singular value decomposition, there exist orthonormal

lists $e_1, \dots, e_m$ in V and $f_1, \dots, f_m$ in W such that

$$Tv = \sum_{k=1}^{m} s_k \langle v, e_k \rangle f_k$$

for all $v \in V$. By 7.75,

$$T^* w = \sum_{k=1}^{m} s_k \langle w, f_k \rangle e_k$$

for all $w \in W$. Applying Exercise 7E8(c) to $T^* \in \mathcal{L}(W, V)$ with the orthonormal lists $f_1, \dots, f_m$ in W and $e_1, \dots, e_m$ in V, the positive singular values of T^* are exactly $s_1, \dots, s_m$. Hence, T and T^* have the same positive singular values. ■

Exercise 7E10

Suppose $T \in \mathcal{L}(V, W)$ has singular values $s_1, \dots, s_n$. Prove that if T is an invertible linear map, then T^{-1} has singular values

$$\frac{1}{s_n}, \dots, \frac{1}{s_1}.$$

Solution. If $\dim V = 0$, then $\dim W = 0$ by 3.70, so neither T nor T^{-1} has any singular values, and the statement holds trivially. Now assume $\dim V > 0$. Since T is invertible, T is injective. By 7.68(a) and the definition of singular values, $s_1 \geq \cdots \geq s_n > 0$. Since all singular values are positive, the positive singular values of T are exactly $s_1, \dots, s_n$. By the singular value decomposition, there exist orthonormal lists $e_1, \dots, e_n$ in V and $f_1, \dots, f_n$ in W such that

$$Tv = \sum_{k=1}^{n} s_k \langle v, e_k \rangle f_k$$

for all $v \in V$. By 6.69(a), $T^\dagger = T^{-1}$. Since $s_1 \geq \cdots \geq s_n > 0$, we have $\frac{1}{s_n} \geq \cdots \geq \frac{1}{s_1} > 0$. Reordering the sum, we have by 7.75 that

$$T^{-1} w = \sum_{k=1}^{n} \frac{1}{s_{n-k+1}} \langle w, f_{n-k+1} \rangle e_{n-k+1}$$

for all $w \in W$. Applying Exercise 7E8(c) to $T^{-1} \in \mathcal{L}(W, V)$ with the orthonormal lists $f_n, \dots, f_1$ in W and $e_n, \dots, e_1$ in V, the positive singular values of T^{-1} are exactly $\frac{1}{s_n}, \dots, \frac{1}{s_1}$. By 3.70, $\dim V = \dim W = n$, so T^{-1} has exactly n singular values. Since T^{-1} is injective, 0 is not a singular value of T^{-1} by 7.68(a). Hence, the singular values of T^{-1} are exactly $\frac{1}{s_n}, \dots, \frac{1}{s_1}$. ■

Exercise 7E11

Suppose that $T \in \mathcal{L}(V, W)$ and $v_1, \dots, v_n$ is an orthonormal basis of V. Let $s_1, \dots, s_n$ 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

$$\langle Tv_1, v_1 \rangle + \cdots + \langle Tv_n, v_n \rangle = s_1 + \cdots + s_n.$$

Axler Note. *See the comment after* Exercise 7A5.

Solution.

(a) Let m denote the number of positive singular values of T. By the singular value decomposition, there exist orthonormal lists $e_1, \dots, e_m$ in V and $f_1, \dots, f_m$ in W such that $Te_k = s_k f_k$ for $k = 1, \dots, m$. By 6.36, extend

$e_1, \dots, e_m$ to an orthonormal basis $e_1, \dots, e_n$ of V. By the formula in the singular value decomposition, $Te_k = 0$ for $k > m$ since $\langle e_k, e_j \rangle = 0$ for $j = 1, \dots, m$. Moreover, $s_k = 0$ for $k > m$, since T has exactly n singular values. Therefore,

$$\sum_{k=1}^{n} \|Te_k\|^2 = \sum_{k=1}^{m} \|s_k f_k\|^2 = \sum_{k=1}^{m} s_k^2 = \sum_{k=1}^{n} s_k^2.$$

By Exercise 7A5, the left-hand side is the same for any orthonormal basis of V. Therefore, the equality holds for the given orthonormal basis $v_1, \dots, v_n$ as well.

(b) Since $T \in \mathcal{L}(V)$ is a positive operator, it is self-adjoint. By Exercise 7E7, the singular values of T are $|\lambda_1|, \dots, |\lambda_n|$, where $\lambda_1, \dots, \lambda_n$ are the eigenvalues of T, and each eigenvalue λ is repeated exactly $\dim E(\lambda, T)$ times. By 7.38(b), the eigenvalues of T are nonnegative, so $s_1, \dots, s_n$ are exactly the eigenvalues of T. By 7.38(f), there exists an operator $R \in \mathcal{L}(V)$ such that $T = R^*R$. By part (a), the sum of the squares of the norms of Rv_k equals the sum of the squares of the singular values of R. Since $R^*R = T$, the singular values of R are, by definition, $\sqrt{s_1}, \dots, \sqrt{s_n}$. Hence,

$$\sum_{k=1}^{n} \langle Tv_k, v_k \rangle = \sum_{k=1}^{n} \langle R^*Rv_k, v_k \rangle = \sum_{k=1}^{n} \langle Rv_k, Rv_k \rangle = \sum_{k=1}^{n} \|Rv_k\|^2 = \sum_{k=1}^{n} s_k. \qquad \blacksquare$$

Exercise 7E12

(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 $T \in \mathcal{L}(\mathbf{C}^2)$ defined by $T(z, w) = (5w, 0)$. By Exercise 7E3, its singular values are 5 and 0. However, $T^2(z, w) = T(5w, 0) = (0, 0)$, so $T^2 = 0$ and its singular values are 0 and 0, not 25 and 0.

(b) Using 7.5(c), (d), and normality of T,

$$(T^2)^*T^2 = T^*(T^*T)T = T^*(TT^*)T = (T^*T)^2 = (T^*T)^*(T^*T).$$

By definition, T^2 and T^*T have the same singular values. Since T^*T is a positive operator by 7.64(a), its eigenvalues are nonnegative by 7.38(b). By Exercise 7E7 applied to T^*T, its singular values are its eigenvalues λ, each repeated $\dim E(\lambda, T^*T)$ times. The eigenvalues of T^*T are the squares of the singular values of T, so the singular values of T^2 are the squares of the singular values of T. $\blacksquare$

Exercise 7E13

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 $T_1 = S_1T_2S_2$.

Solution. Let $n = \dim V$, let $s_1 \geq \cdots \geq s_n \geq 0$ denote the singular values of T_2, and let m be the number of positive singular values of T_2. If $m = 0$, then T_2 has no positive singular values, so by the formula in the singular value decomposition, $T_2 = 0$. If T_1 and T_2 have the same singular values, the same reasoning yields $T_1 = 0$, so we may take $S_1 = S_2 = I$ and the result holds trivially. Conversely, if $T_1 = S_1T_2S_2$ for unitary $S_1, S_2 \in \mathcal{L}(V)$, then $T_1 = S_10S_2 = 0$, so $T_1^*T_1 = 0$. Since $E(0, T_1^*T_1) = V$, 0 is the only eigenvalue of $T_1^*T_1$, and $\dim E(0, T_1^*T_1) = n$. Therefore, the singular values of T_1 are all 0, which are the same as the singular values of T_2. Hence, we may assume $m > 0$ for the remainder of the proof.
$\boxed{\Rightarrow}$ Suppose T_1 and T_2 have the same singular values. By the singular value decomposition, there exist orthonormal lists

$e_1, \dots, e_m, f_1, \dots, f_m, e'_1, \dots, e'_m$, and $f'_1, \dots, f'_m$ in V such that

$$T_1 v = \sum_{i=1}^{m} s_i \langle v, e_i \rangle f_i, \quad T_2 v = \sum_{i=1}^{m} s_i \langle v, e'_i \rangle f'_i$$

for every $v \in V$. By 6.36, extend all orthonormal lists $e_1, \dots, e_m, f_1, \dots, f_m, e'_1, \dots, e'_m$, and $f'_1, \dots, f'_m$ to orthonormal bases $e_1, \dots, e_n, f_1, \dots, f_n, e'_1, \dots, e'_n$, and $f'_1, \dots, f'_n$ of V. By the orthonormality of these lists, $T_1 e_i = s_i f_i$ and $T_2 e'_i = s_i f'_i$ for $i \le m$, and $T_1 e_i = 0$ and $T_2 e'_i = 0$ for $i > m$. By the linear map lemma, we may define operators $S_1, S_2 \in \mathcal{L}(V)$ such that $S_1 f'_i = f_i$ and $S_2 e_i = e'_i$ for $i = 1, \dots, n$. By 7.53(d), S_1 and S_2 are unitary. Moreover, for $i = 1, \dots, m$, we have

$$S_1 T_2 S_2 e_i = S_1 T_2 e'_i = S_1(s_i f'_i) = s_i f_i = T_1 e_i,$$

and $S_1 T_2 S_2 e_i = T_1 e_i$ for $i > m$ as well, since both sides are 0. Since $S_1 T_2 S_2$ and T_1 agree on the basis $e_1, \dots, e_n$, we have $S_1 T_2 S_2 = T_1$ by the linear map lemma.

$\boxed{\Leftarrow}$ Suppose there exist unitary operators $S_1, S_2 \in \mathcal{L}(V)$ such that $T_1 = S_1 T_2 S_2$. By 7.5(d) and 7.53(b), we have

$$T_1^* T_1 = (S_1 T_2 S_2)^* (S_1 T_2 S_2) = S_2^* T_2^* S_1^* S_1 T_2 S_2 = S_2^* T_2^* T_2 S_2.$$

Moreover, 7.53(c) shows that $S_2^* = S_2^{-1}$. Therefore, $T_1^* T_1 v = \lambda v$ if and only if $T_2^* T_2 (S_2 v) = \lambda (S_2 v)$ for all $\lambda \in \mathbf{F}$ and $v \in V$. Hence, S_2 maps $E(\lambda, T_1^* T_1)$ onto $E(\lambda, T_2^* T_2)$ for all $\lambda \in \mathbf{F}$. Since S_2 is invertible, $\dim E(\lambda, T_1^* T_1) = \dim E(\lambda, T_2^* T_2)$ for all $\lambda \in \mathbf{F}$. By definition, the singular values of T_1 are the square roots of the eigenvalues of $T_1^* T_1$, each eigenvalue λ repeated $\dim E(\lambda, T_1^* T_1)$ times, and similarly for T_2. Therefore, T_1 and T_2 have the same singular values. ∎

Exercise 7E14

Suppose $T \in \mathcal{L}(V, W)$. Let s_n denote the smallest singular value of T. Prove that $s_n \|v\| \le \|Tv\|$ for every $v \in V$.

Solution. If $v = 0$, the inequality $s_n \|v\| \le \|Tv\|$ holds trivially. Now assume $v \neq 0$. Let $u = \frac{v}{\|v\|}$, where $\|u\| = 1$ by construction. By Exercise 7E4, $\|Tu\| \ge s_n$. Then

$$\|Tv\| = \|T(\|v\| u)\| = \|v\| \, \|Tu\| \ge \|v\| s_n = s_n \|v\|.$$ ∎

Exercise 7E15

Suppose $T \in \mathcal{L}(V)$ and $s_1 \ge \cdots \ge s_n$ are the singular values of T. Prove that if λ is an eigenvalue of T, then $s_1 \ge |\lambda| \ge s_n$.

Solution. Let λ be an eigenvalue of T with corresponding eigenvector $v \neq 0$, so $Tv = \lambda v$. Let $u = \frac{v}{\|v\|}$, so $\|u\| = 1$. Then $Tu = \frac{1}{\|v\|} Tv = \frac{1}{\|v\|} \lambda v = \lambda u$. Hence, $\|Tu\| = |\lambda|$. By Exercise 7E4, $s_n \le \|Tu\| \le s_1$. Therefore, $s_1 \ge |\lambda| \ge s_n$. ∎

Exercise 7E16

Suppose $T \in \mathcal{L}(V, W)$. Prove that $(T^*)^\dagger = (T^\dagger)^*$.

Axler Note. *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. By the singular value decomposition, there exist orthonormal lists $e_1, \dots, e_m$ in V and $f_1, \dots, f_m$ in W such that

$$Tv = \sum_{i=1}^{m} s_i \langle v, e_i \rangle f_i$$

for all $v \in V$. By 7.75,

$$T^* w = \sum_{i=1}^{m} s_i \langle w, f_i \rangle e_i, \quad T^\dagger w = \sum_{i=1}^{m} \frac{1}{s_i} \langle w, f_i \rangle e_i$$

for all $w \in W$. By Exercise 7E9, $s_1, \ldots, s_m$ are the positive singular values of T^* as well. Applying 7.75 to T^* with the orthonormal lists $f_1, \ldots, f_m$ in W and $e_1, \ldots, e_m$ in V, we get

$$(T^*)^\dagger v = \sum_{i=1}^{m} \frac{1}{s_i} \langle v, e_i \rangle f_i$$

for all $v \in V$. Since each $s_i \in \mathbf{R}$, for all $w \in W$ and $v \in V$,

$$\langle T^\dagger w, v \rangle = \sum_{i=1}^{m} \frac{1}{s_i} \langle w, f_i \rangle \langle e_i, v \rangle = \left\langle w, \sum_{i=1}^{m} \frac{1}{s_i} \langle v, e_i \rangle f_i \right\rangle.$$

By the definition of the adjoint, this shows that

$$(T^\dagger)^* v = \sum_{i=1}^{m} \frac{1}{s_i} \langle v, e_i \rangle f_i$$

for all $v \in V$. This matches the expression for $(T^*)^\dagger$ obtained earlier, so $(T^*)^\dagger = (T^\dagger)^*$. ■

Exercise 7E17

Suppose $T \in \mathcal{L}(V)$. Prove that T is self-adjoint if and only if $T^\dagger$ is self-adjoint.

Solution.

$\boxed{\Rightarrow}$ Suppose T is self-adjoint. Then $T = T^*$. By Exercise 7E16, $T^\dagger = (T^*)^\dagger = (T^\dagger)^*$, so $T^\dagger$ is self-adjoint.

$\boxed{\Leftarrow}$ Suppose $T^\dagger$ is self-adjoint. Applying $\boxed{\Rightarrow}$ to $T^\dagger$ shows that $(T^\dagger)^\dagger$ is self-adjoint. By Exercise 6C23, $(T^\dagger)^\dagger = T$, so T is self-adjoint. ■

7F Consequences of Singular Value Decomposition

Exercise 7F1

Prove that if $S, T \in \mathcal{L}(V, W)$, then $\big|\|S\| - \|T\|\big| \le \|S - T\|$.

Axler Note. *The inequality above is called the* **reverse triangle inequality**.

Solution. By 7.87(d), $\|S\| = \|(S - T) + T\| \le \|S - T\| + \|T\|$. Similarly, $\|T\| = \|(T - S) + S\| \le \|T - S\| + \|S\|$. By 7.87(c) with $\lambda = -1$, $\|T - S\| = \|-(S - T)\| = \|S - T\|$. Combining the inequalities yields the desired result. ■

Exercise 7F2 ✔

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. If $V = \{0\}$, then T has no eigenvalues, so we adopt the convention that the maximum over the empty set is 0. In this case, the only operator on V is $T = 0$, so $\|T\| = 0$ by 7.87(b), and the equality holds. Now assume $V \neq \{0\}$. Let $\lambda_1, \dots, \lambda_n$ be the eigenvalues of T, where each eigenvalue λ of T appears in the list as many times as $\dim E(\lambda, T)$. As in the solution to Exercise 7E7, $n = \dim V \ge 1$. By Exercise 7E7, the singular values of T are $|\lambda_1|, \dots, |\lambda_n|$ sorted in decreasing order. Each eigenvalue λ of T appears in the list $\lambda_1, \dots, \lambda_n$ at least once since $\dim E(\lambda, T) \ge 1$. We then have

$$\max_{1 \le k \le n} |\lambda_k| = \max\{|\lambda| : \lambda \text{ is an eigenvalue of } T\}.$$

By 7.88(a), $\|T\|$ is the largest singular value of T, which is the left-hand side of the equality in the previous line. ■

Exercise 7F3

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

$$\|Tv\| = \|T\| \, \|v\| \iff T^*Tv = \|T\|^2 v.$$

Solution.

$\boxed{\Rightarrow}$ Define $S = \|T\|^2 I - T^*T$. By 7.5, T^*T is self-adjoint. Since $\|T\|^2 \in \mathbf{R}$, $(\|T\|^2 I)^* = \|T\|^2 I$. Hence, S is self-adjoint. For every $u \in V$, 7.88(c) gives

$$\langle Su, u \rangle = \langle (\|T\|^2 I - T^*T)u, u \rangle = \|T\|^2 \|u\|^2 - \|Tu\|^2 \ge 0.$$

Hence, S is a positive operator. By hypothesis, $\langle Sv, v \rangle = \langle (\|T\|^2 I - T^*T)v, v \rangle = \|T\|^2 \|v\|^2 - \|Tv\|^2 = 0$. Since S is positive, $\langle Sv, v \rangle = 0$ implies $Sv = 0$ by 7.43. Therefore, $(\|T\|^2 I - T^*T)v = 0$, which gives $T^*Tv = \|T\|^2 v$.

$\boxed{\Leftarrow}$ Suppose $T^*Tv = \|T\|^2 v$. Then $\|Tv\|^2 = \langle Tv, Tv \rangle = \langle T^*Tv, v \rangle = \|T\|^2 \langle v, v \rangle = \|T\|^2 \|v\|^2$. Taking square roots gives $\|Tv\| = \|T\| \, \|v\|$. ■

Exercise 7F4

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. Suppose $u \in V$ with $\langle u, v \rangle = 0$. By Exercise 7F3, $T^*Tv = \|T\|^2 v$. By the definition of the adjoint, the previous line, and 6.6(e), since $\|T\|^2 \in \mathbf{R}$, we have

$$\langle Tu, Tv \rangle = \langle u, T^*Tv \rangle = \langle u, \|T\|^2 v \rangle = \|T\|^2 \langle u, v \rangle = 0.$$ ■

Exercise 7F5

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 7.87(a), $\|S\| \geq 0$. By 7.88(c) applied to S and T, for every $v \in V$, we have

$$\|STv\| \leq \|S\| \|Tv\| \leq \|S\| \|T\| \|v\|.$$

By 7.88(c), $\|ST\|$ is the smallest number c such that $\|STv\| \leq c\|v\|$ for all $v \in V$. Hence, $\|ST\| \leq \|S\| \|T\|$. ∎

Exercise 7F6

Prove or give a counterexample: If $S, T \in \mathcal{L}(V)$, then $\|ST\| = \|TS\|$.

Solution. The statement is false. Suppose $\dim V \geq 2$. By 6.35, there exists an orthonormal basis $e_1, e_2, \ldots, e_{\dim V}$ of V. Define the maps S and T by

$$Sv = \langle v, e_1 \rangle e_1, \quad Tv = \langle v, e_1 \rangle e_2$$

for all $v \in V$. By 6.6(a), $S, T \in \mathcal{L}(V)$. Then $STv = \langle v, e_1 \rangle Se_2 = \langle v, e_1 \rangle \langle e_2, e_1 \rangle e_1 = 0$ for all $v \in V$. On the other hand, $TSv = \langle v, e_1 \rangle Te_1 = \langle v, e_1 \rangle e_2 = Tv$ for all $v \in V$. Therefore, $ST = 0$ and $TS = T$. Moreover, $T \neq 0$ because $Te_1 = e_2 \neq 0$. By 7.87(a) and (b), we have $\|ST\| = 0 < \|TS\| = \|T\|$. ∎

Exercise 7F7

Show that defining $d(S, T) = \|S - T\|$ for $S, T \in \mathcal{L}(V, W)$ makes d a metric on $\mathcal{L}(V, W)$.

Axler Note. *This exercise is intended for readers who are familiar with metric spaces.*

Solution. We verify the metric axioms for d:

- **Nonnegativity and identity**: Nonnegativity follows from 7.87(a). By 7.87(b), $d(S, T) = \|S - T\| = 0$ if and only if $S - T = 0$, i.e., $S = T$.
- **Symmetry**: Since $T - S = (-1)(S - T)$, 7.87(c) implies $\|T - S\| = \|(-1)(S - T)\| = \|S - T\|$, so $d(S, T) = d(T, S)$.
- **Triangle inequality**: For any $R \in \mathcal{L}(V, W)$, $S - R = (S - T) + (T - R)$, so by 7.87(d),

$$d(S, R) = \|S - R\| \leq \|S - T\| + \|T - R\| = d(S, T) + d(T, R).$$

Therefore, $(\mathcal{L}(V, W), d)$ is a metric space. ∎

Exercise 7F8

(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.

Axler Note. *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 7F7.

Solution.

(a) By 3.65, it suffices to show that T is injective. Let $v \in V$ be such that $Tv = 0$. Then $(I - T)v = v$. By 7.88(c), $\|I - T\| \|v\| \geq \|(I - T)v\| = \|v\|$. If $v \neq 0$, then dividing both sides by $\|v\|$ gives $\|I - T\| \geq 1$, which contradicts the assumption that $\|I - T\| < 1$. Therefore, $v = 0$, and T is injective.

(b) Note that the hypothesis $\|S-T\| < 1/\|S^{-1}\|$ requires that $\|S^{-1}\| \neq 0$, so $\|S^{-1}\| > 0$ by 7.87(a). By 3.8, $I - S^{-1}T = S^{-1}(S-T)$. By Exercise 7F5,

$$\|I - S^{-1}T\| \leq \|S^{-1}\|\,\|S-T\|.$$

Since $\|S-T\| < 1/\|S^{-1}\|$ and $\|S^{-1}\| > 0$, $\|I - S^{-1}T\| < 1$. By part (a), $S^{-1}T$ is invertible, so $T = S(S^{-1}T)$ is invertible. ■

Exercise 7F9

Suppose $T \in \mathcal{L}(V)$. Prove that for every $\varepsilon > 0$, there exists an invertible operator $S \in \mathcal{L}(V)$ such that $0 < \|T-S\| < \varepsilon$.
Manual Remark. *Note that this exercise assumes $V \neq \{0\}$, for otherwise, $\|T-S\| = 0$ for every $S \in \mathcal{L}(V)$.*

Solution. Since $(0, \varepsilon)$ contains infinitely many real numbers, and T has at most $\dim V$ distinct eigenvalues by 5.12, there exists a real number $\lambda \in (0, \varepsilon)$ that is not an eigenvalue of T. Consider the operator $S = T - \lambda I$. By 5.7, S is invertible because λ is not an eigenvalue of T. Moreover, since I is an isometry, its singular values are all 1 by 7.69, so $\|I\| = 1$ by 7.88(a). Therefore, by 7.87(c), $\|T-S\| = \|\lambda I\| = |\lambda|\|I\| = \lambda$, where $|\lambda| = \lambda$ because $\lambda > 0$. Hence, $0 < \|T-S\| = \lambda < \varepsilon$. ■

Exercise 7F10

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. By 3.65, T is not injective since it is not invertible. Then there exists a nonzero $u \in \operatorname{null} T$. Let $U = (\operatorname{span}(u))^{\perp}$. Since $\dim \operatorname{span}(u) = 1$ and $\dim V > 1$, we have $\dim U = \dim V - 1 \geq 1$ by 6.51. Consider the operator $S \in \mathcal{L}(V)$ defined by $S = T + \frac{\varepsilon}{2}P_U$. Note that for every $w \in U$, $\langle w, u\rangle = 0$, so $u \in U^{\perp}$ and hence $P_U u = 0$ by 6.57(c). Then $Su = 0$, so S is not injective, and hence not invertible by 3.65. By 6.57(h), $\|P_U v\| \leq \|v\|$ for all $v \in V$, which yields $\|P_U\| \leq 1$ by 7.88(c). Since $\dim U \geq 1$, there exists a unit vector $x \in U$. By 6.57(b), $P_U x = x$, so $\|P_U\| \geq \|P_U x\| = \|x\| = 1$ by 7.88(b). Thus, $\|P_U\| = 1$. Finally, by 7.87(c),

$$\|T-S\| = \left\|-\frac{\varepsilon}{2}P_U\right\| = \frac{\varepsilon}{2}\|P_U\| = \frac{\varepsilon}{2},$$

so $0 < \|T-S\| < \varepsilon$. ■

Exercise 7F11

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 $0 < \|T-S\| < \varepsilon$.
Manual Remark. *Note that this exercise assumes $V \neq \{0\}$, for otherwise, $\|T-S\| = 0$ for every $S \in \mathcal{L}(V)$.*

Solution. Let $\dim V = n$. By Schur's theorem, since $\mathbf{F} = \mathbf{C}$, there exists an orthonormal basis $e_1, \ldots, e_n$ of V such that $\mathcal{M}(T)$ is upper triangular with respect to this basis with diagonal entries $\lambda_1, \ldots, \lambda_n$, which are the eigenvalues of T by 5.41, possibly with repetitions. Define $\delta_1 = \frac{\varepsilon}{2}$. For $k = 2, \ldots, n$, choose $\delta_k \in [0, \frac{\varepsilon}{2})$ such that

$$\lambda_k + \delta_k \notin \{\lambda_j + \delta_j : 1 \leq j < k\}.$$

Note that this construction is possible since $[0, \frac{\varepsilon}{2})$ contains infinitely many real numbers, and only $k-1$ values are excluded for each $k \geq 2$. By the linear map lemma, we may construct an operator $D \in \mathcal{L}(V)$ by $De_k = \delta_k e_k$. Then $\mathcal{M}(D, (e_1, \ldots, e_n))$ is diagonal with diagonal entries $\delta_1, \ldots, \delta_n$, which are its eigenvalues by 5.41.

Let $S = T + D$. By 3.35, $\mathcal{M}(S, (e_1, \ldots, e_n))$ is upper triangular with diagonal entries $\lambda_1 + \delta_1, \ldots, \lambda_n + \delta_n$, which are distinct by construction. Since these are n distinct eigenvalues by 5.41, S is diagonalizable by 5.58. By the complex spectral

theorem, D is normal since it has a diagonal matrix with respect to the orthonormal basis $e_1, \dots, e_n$. By Exercise 7F2, and since $\delta_k \geq 0$ for all k,

$$\|T - S\| = \|D\| = \max_{1 \leq k \leq n} |\delta_k| = \max_{1 \leq k \leq n} \delta_k = \frac{\varepsilon}{2} < \varepsilon.$$

Then $0 < \|T - S\| = \frac{\varepsilon}{2} < \varepsilon$. ∎

Exercise 7F12

Suppose $T \in \mathcal{L}(V)$ is a positive operator. Show that $\|\sqrt{T}\| = \sqrt{\|T\|}$.

Solution. By 7.39, $\sqrt{T}$ is the unique positive square root of T, so T and $\sqrt{T}$ are self-adjoint by 7.38(b). By Exercise 7F2 and the nonnegativity of the eigenvalues of T and $\sqrt{T}$ by 7.38(b), we have

$$\|T\| = \max_{\lambda} \lambda, \quad \left\|\sqrt{T}\right\| = \max_{\mu} \mu,$$

where λ and μ range over all eigenvalues of T and $\sqrt{T}$, respectively.

By the proof of 7.39, for every eigenvalue λ of T, $\sqrt{\lambda}$ is an eigenvalue of $\sqrt{T}$. Moreover, if $\sqrt{T}v = \mu v$ for some nonzero $v \in V$, then $Tv = \mu^2 v$. Since $\mu \geq 0$, we have $\mu = \sqrt{\mu^2}$, and μ^2 is an eigenvalue of T since $v \neq 0$. Therefore, the eigenvalues of $\sqrt{T}$ are precisely the square roots of the eigenvalues of T, so

$$\left\|\sqrt{T}\right\| = \max_{\mu} \mu = \max_{\lambda} \sqrt{\lambda} = \sqrt{\max_{\lambda} \lambda} = \sqrt{\|T\|},$$

where the third equality follows from the fact that $\lambda \mapsto \sqrt{\lambda}$ is an increasing function on $[0, \infty)$. ∎

Exercise 7F13

Suppose $S, T \in \mathcal{L}(V)$ are positive operators. Show that

$$\|S - T\| \leq \max\{\|S\|, \|T\|\} \leq \|S + T\|.$$

Solution. If $V = \{0\}$, then $S = T = 0$, and the inequality holds trivially. Now assume $V \neq \{0\}$. Since S and T are positive, the Cauchy–Schwarz inequality and 7.88(c) give

$$0 \leq \langle Sv, v \rangle \leq \|S\|, \quad 0 \leq \langle Tv, v \rangle \leq \|T\|$$

for every unit vector $v \in V$. By 7.38(b), S and T are self-adjoint, so $S - T$ is also self-adjoint by 7.5(a) and (b). By Exercise 7F2, there exists an eigenvalue λ of $S - T$ such that $\|S - T\| = |\lambda|$. Let $u \in V$ be a unit eigenvector corresponding to λ, so $(S - T)u = \lambda u$. Then

$$\lambda = \lambda \|u\|^2 = \langle (S - T)u, u \rangle = \langle Su, u \rangle - \langle Tu, u \rangle \in [-\|T\|, \|S\|].$$

Thus, $|\lambda| = \|S - T\| \leq \max\{\|S\|, \|T\|\}$.

By Exercise 7F2 again, there exists an eigenvalue μ of S such that $\|S\| = |\mu|$. By 7.38(b), $\mu \geq 0$, so $\|S\| = |\mu| = \mu$. Let $w \in V$ be a unit eigenvector corresponding to μ, so $Sw = \mu w$. Then $\langle Sw, w \rangle = \mu \|w\|^2 = \mu = \|S\|$. Consequently,

$$\|S\| \leq \|S\| + \langle Tw, w \rangle = \langle (S + T)w, w \rangle \leq \|S + T\|,$$

where the last inequality follows from the Cauchy–Schwarz inequality and 7.88(c). By symmetry, we also have $\|T\| \leq \|S + T\|$. Therefore, $\max\{\|S\|, \|T\|\} \leq \|S + T\|$, and the desired inequality is proved. ∎

Exercise 7F14

Suppose U and W are subspaces of V such that $\|P_U - P_W\| < 1$. Prove that $\dim U = \dim W$.

Solution. Suppose $\dim U > \dim W$. By 6.51, 2.43, and 2.37,

$$\dim(U \cap W^\perp) = \dim U + \dim W^\perp - \dim(U + W^\perp) \geq \dim U - \dim W > 0.$$

Let $u \in U \cap W^\perp$ be a unit vector. Then $(P_U - P_W)u = P_U u - P_W u = u - 0 = u$ by 6.57(b) and (c). By 7.88(c), $\|P_U - P_W\| \geq \|(P_U - P_W)u\| = \|u\| = 1$, contradicting the assumption that $\|P_U - P_W\| < 1$. Therefore, our assumption that $\dim U > \dim W$ must be false. By symmetry, $\dim W > \dim U$ is also impossible. Hence, $\dim U = \dim W$. ∎

Exercise 7F15

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. For any $(z_1, z_2, z_3), (w_1, w_2, w_3) \in \mathbf{F}^3$,

$$\begin{aligned}\langle T(z_1, z_2, z_3), (w_1, w_2, w_3)\rangle &= \langle (z_3, 2z_1, 3z_2), (w_1, w_2, w_3)\rangle \\ &= z_3\overline{w_1} + 2z_1\overline{w_2} + 3z_2\overline{w_3} \\ &= \langle (z_1, z_2, z_3), (2w_2, 3w_3, w_1)\rangle.\end{aligned}$$

Therefore,

$$T^*(w_1, w_2, w_3) = (2w_2, 3w_3, w_1), \quad T^*T(z_1, z_2, z_3) = (4z_1, 9z_2, z_3).$$

Let $R \in \mathcal{L}(\mathbf{F}^3)$ be defined by $R(z_1, z_2, z_3) = (2z_1, 3z_2, z_3)$. Then $\mathcal{M}(R)$ with respect to the standard orthonormal basis e_1, e_2, e_3 of $\mathbf{F}^3$ is diagonal with nonnegative diagonal entries $2, 3, 1$. By 7.38(c), R is positive. Additionally, $R^2(z_1, z_2, z_3) = (4z_1, 9z_2, z_3) = T^*T(z_1, z_2, z_3)$, so $R^2 = T^*T$ and hence $R = \sqrt{T^*T}$ by 7.39. Moreover, R^{-1} exists and is given by

$$R^{-1}(z_1, z_2, z_3) = \left(\frac{z_1}{2}, \frac{z_2}{3}, z_3\right).$$

Define $S \in \mathcal{L}(\mathbf{F}^3)$ by $S = TR^{-1}$. Then

$$S(z_1, z_2, z_3) = T(R^{-1}(z_1, z_2, z_3)) = T\left(\frac{z_1}{2}, \frac{z_2}{3}, z_3\right) = (z_3, z_1, z_2).$$

Observe that $Se_1 = e_2$, $Se_2 = e_3$, and $Se_3 = e_1$. Since e_2, e_3, e_1 is an orthonormal basis of $\mathbf{F}^3$, S is unitary by 7.53(d). Because $S = TR^{-1}$, we have $T = SR = S\sqrt{T^*T}$. ∎

Exercise 7F16

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. If $V = \{0\}$, every operator on V is positive, so any $\delta > 0$ works. Now assume $V \neq \{0\}$. By 7.38(c), there exists an orthonormal basis $e_1, \ldots, e_n$ of V such that $Se_k = \lambda_k e_k$ with $\lambda_k \geq 0$ for all $k = 1, \ldots, n$. Then each λ_k is an eigenvalue of S because $e_k \neq 0$, and since S is invertible, $\lambda_k > 0$ for all $k = 1, \ldots, n$ by 5.7. Let

$$\delta = \min_{1 \leq k \leq n} \lambda_k > 0.$$

Observe that for any $v \in V$ and $k = 1, \dots, n$, we have

$$\langle Sv, e_k \rangle = \langle v, Se_k \rangle = \overline{\lambda_k} \langle v, e_k \rangle = \lambda_k \langle v, e_k \rangle,$$

since S is self-adjoint and $\lambda_k \in \mathbf{R}$ for all k. By 6.30(b) and (c),

$$\langle Sv, v \rangle = \sum_{k=1}^{n} \langle Sv, e_k \rangle \overline{\langle v, e_k \rangle} = \sum_{k=1}^{n} \lambda_k |\langle v, e_k \rangle|^2 \geq \delta \sum_{k=1}^{n} |\langle v, e_k \rangle|^2 = \delta \|v\|^2.$$

Let $T \in \mathcal{L}(V)$ be self-adjoint with $\|S - T\| < \delta$. By the Cauchy–Schwarz inequality and 7.88(c),

$$\langle Tv, v \rangle = \langle Sv, v \rangle - \langle (S - T)v, v \rangle \geq \langle Sv, v \rangle - \|S - T\| \, \|v\|^2 \geq (\delta - \|S - T\|) \|v\|^2 \geq 0.$$

Since T is self-adjoint and $\langle Tv, v \rangle \geq 0$ for all $v \in V$, it follows that T is positive. ■

Exercise 7F17

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 $\|\varphi_u\| = \|u\|$.
Axler Note. *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. Observe that $\|\alpha\| = \sqrt{\langle \alpha, \alpha \rangle} = \sqrt{\alpha \overline{\alpha}} = |\alpha|$ for all $\alpha \in \mathbf{F}$. For any $v \in V$, the Cauchy–Schwarz inequality gives

$$|\varphi_u(v)| = |\langle v, u \rangle| \leq \|v\| \, \|u\|.$$

By 7.88(c), $\|\varphi_u\| \leq \|u\|$.

If $u = 0$, then $0 \leq \|\varphi_u\| \leq \|u\| = 0$ by 7.87(a). For $u \neq 0$, consider $w = \frac{u}{\|u\|}$. Since $\|w\| = 1$, 7.88(b) gives $\|\varphi_u\| \geq |\varphi_u(w)|$, so

$$\|\varphi_u\| \geq |\varphi_u(w)| = |\langle w, u \rangle| = \left| \left\langle \frac{u}{\|u\|}, u \right\rangle \right| = \frac{\|u\|^2}{\|u\|} = \|u\|.$$

Hence, $\|\varphi_u\| = \|u\|$. ■

Exercise 7F18

Suppose $e_1, \dots, e_n$ is an orthonormal basis of V and $T \in \mathcal{L}(V, W)$.
(a) Prove that $\max\{\|Te_1\|, \dots, \|Te_n\|\} \leq \|T\| \leq (\|Te_1\|^2 + \cdots + \|Te_n\|^2)^{1/2}$.
(b) Prove that $\|T\| = (\|Te_1\|^2 + \cdots + \|Te_n\|^2)^{1/2}$ if and only if $\dim \operatorname{range} T \leq 1$.
Axler Note. *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 7E11*(a)*.

Solution. Let $s_1 \geq \cdots \geq s_n$ be the singular values of T.

(a) Observe that $\max\{\|Te_1\|, \dots, \|Te_n\|\} \leq \|T\|$ by 7.88(b) applied to each e_k. By Exercise 7E11(a) and 7.88(a),

$$\sum_{k=1}^{n} \|Te_k\|^2 = \sum_{k=1}^{n} s_k^2 \geq s_1^2 = \|T\|^2.$$

Taking square roots gives the desired inequality.

(b) Note that $\|T\| = (\sum_{k=1}^{n} \|Te_k\|^2)^{1/2}$ if and only if $\|T\|^2 = \sum_{k=1}^{n} \|Te_k\|^2$, since both sides are nonnegative.

$$\begin{aligned}
\|T\|^2 = \sum_{k=1}^{n} \|Te_k\|^2 &\iff s_1^2 = \sum_{k=1}^{n} s_k^2 \\
&\iff \sum_{k=2}^{n} s_k^2 = 0 \\
&\iff s_2 = \cdots = s_n = 0 \\
&\iff T \text{ has at most one positive singular value} \\
&\iff \dim \operatorname{range} T \leq 1,
\end{aligned}$$

where the first equivalence follows from 7.88(a) and Exercise 7E11(a); the third equivalence from $s_k^2 \geq 0$ for all k; the fourth equivalence from the decreasing order of the singular values; and the fifth equivalence from 7.68(b). ■

Exercise 7F19

Prove that if $T \in \mathcal{L}(V, W)$, then $\|T^*T\| = \|T\|^2$.

Axler Note. *This formula for $\|T^*T\|$ leads to the important subject of C^*-algebras.*

Manual Remark. *Alternatively, if $T \neq 0$ and*

$$Tv = \sum_{k=1}^{m} s_k \langle v, e_k \rangle f_k$$

is a singular value decomposition of T by the singular value decomposition, then 7.75 gives

$$T^*Tv = \sum_{k=1}^{m} s_k^2 \langle v, e_k \rangle e_k,$$

so Exercise 7E8(c) and 7.88(a) give $\|T^*T\| = s_1^2 = \|T\|^2$.

Solution.

$\boxed{\leq}$ By 7.88(b), there exists $v \in V$ with $\|v\| = 1$ such that $\|Tv\| = \|T\|$. By the definition of the adjoint, the Cauchy–Schwarz inequality, and 7.88(c), we have

$$\|T\|^2 = \|Tv\|^2 = \langle Tv, Tv \rangle = \langle v, T^*Tv \rangle \leq \|v\| \, \|T^*Tv\| \leq \|T^*T\| \, \|v\|^2 = \|T^*T\|.$$

$\boxed{\geq}$ By Exercise 7F5 and 7.91, $\|T^*T\| \leq \|T^*\| \, \|T\| = \|T\|^2$. ■

Exercise 7F20

Suppose $T \in \mathcal{L}(V)$ is normal. Prove that $\|T^k\| = \|T\|^k$ for every positive integer k.

Solution. Let $n = \dim V$ and let $S = T^*T$. We first show that $(T^k)^*T^k = S^k$ for all positive integers k. The base case $k = 1$ is trivial since $(T^1)^*T^1 = T^*T = S$. Now suppose $(T^{k-1})^*T^{k-1} = S^{k-1}$ for $k \geq 2$. Since T is normal, S commutes with T, and hence with T^{k-1}. By 7.5(d) applied to T and T^{k-1} and the induction hypothesis, we have

$$(T^k)^*T^k = (TT^{k-1})^*T^k = (T^{k-1})^*(T^*T)T^{k-1} = (T^{k-1})^*ST^{k-1} = (T^{k-1})^*T^{k-1}S = S^{k-1}S = S^k,$$

so $(T^k)^*T^k = S^k$ for all positive integers k.

Observe that S is positive by 7.64(a), so by 7.38(c), $\mathcal{M}(S)$ is diagonal with entries $\lambda_1, \ldots, \lambda_n \geq 0$ with respect to some orthonormal basis of V. By 3.43, $\mathcal{M}(S^k)$ is also diagonal with entries $\lambda_1^k, \ldots, \lambda_n^k$ with respect to the same orthonormal

basis of V, which are exactly the eigenvalues of S^k by 5.41, so S^k is positive by 7.38(c). By Exercise 7F2 applied to S^k and S, the nonnegativity of the eigenvalues of S^k, and the fact that $x \mapsto x^k$ is increasing on $[0, \infty)$, we have

$$\|S^k\| = \max_{1 \le j \le n} \lambda_j^k = \left(\max_{1 \le j \le n} \lambda_j \right)^k = \|S\|^k.$$

The above computation and Exercise 7F19 applied to T^k and T then give $\|T^k\|^2 = \|S\|^k = \|T\|^{2k}$, and taking the nonnegative square roots of both sides yields $\|T^k\| = \|T\|^k$. ■

Exercise 7F21

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. Suppose the norm on $\mathcal{L}(V, W)$ comes from an inner product. By 6.35, there exist orthonormal bases $e_1, \dots, e_{\dim V}$ of V and $f_1, \dots, f_{\dim W}$ of W. Since $\dim V > 1$ and $\dim W > 1$, we may take the orthonormal lists e_1, e_2 in V and f_1, f_2 in W. Consider the maps S and R defined by

$$Sv = \langle v, e_1 \rangle f_1, \quad Rv = \langle v, e_2 \rangle f_2.$$

Note that $S, R \in \mathcal{L}(V, W)$ by 6.6(a). By the Cauchy–Schwarz inequality,

$$\|Sv\| = |\langle v, e_1 \rangle| \, \|f_1\| \le \|v\| \le 1$$

for $\|v\| \le 1$. Since $\|Se_1\| = \|f_1\| = 1$, we have $\|S\| = 1$. Similarly, $\|R\| = 1$, using e_2 and f_2. By 6.24 and Bessel's inequality, we have for any $v \in V$ that

$$\|(S \pm R)v\|^2 = |\langle v, e_1 \rangle|^2 + |\langle v, e_2 \rangle|^2 \le \|v\|^2.$$

It follows that $\|(S \pm R)v\| \le \|v\| \le 1$ for all $v \in V$ with $\|v\| \le 1$, so $\|S \pm R\| \le 1$. Because $(S \pm R)e_1 = f_1$, it follows that $\|S + R\| \ge 1$ and $\|S - R\| \ge 1$. Hence, $\|S + R\| = \|S - R\| = 1$. Consequently,

$$\|S + R\|^2 + \|S - R\|^2 = 1^2 + 1^2 = 2$$

but

$$2\|S\|^2 + 2\|R\|^2 = 2 \cdot 1^2 + 2 \cdot 1^2 = 4.$$

The parallelogram equality requires that $\|S + R\|^2 + \|S - R\|^2 = 2\|S\|^2 + 2\|R\|^2$, which is not the case here. Therefore, the norm on $\mathcal{L}(V, W)$ does not come from an inner product. ■

Exercise 7F22

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. Let m be the number of positive singular values of T, so $s_{m+1} = \cdots = s_n = 0$. By the singular value

decomposition, there exist orthonormal lists $e_1, \dots, e_m$ in V and $f_1, \dots, f_m$ in W such that

$$Tv = \sum_{j=1}^{m} s_j \langle v, e_j \rangle f_j$$

for all $v \in V$. Extend $e_1, \dots, e_m$ to an orthonormal basis $e_1, \dots, e_n$ of V by 6.35 if $m = 0$, or by 6.36 if $m > 0$. By 6.24,

$$\|Tv\|^2 = \sum_{j=1}^{m} s_j^2 |\langle v, e_j \rangle|^2 + \sum_{j=m+1}^{n} 0 \cdot |\langle v, e_j \rangle|^2 = \sum_{j=1}^{n} s_j^2 |\langle v, e_j \rangle|^2.$$

$\boxed{\leq}$ Consider the subspace $X = \operatorname{span}(e_{n-k+1}, \dots, e_n)$ of V. By 6.25, $\dim X = k$. For every $x \in X$ and $j \leq n - k$, we have $\langle x, e_j \rangle = 0$ by the orthonormality of $e_1, \dots, e_n$. By the above calculation,

$$\|Tx\|^2 = \sum_{j=n-k+1}^{n} s_j^2 |\langle x, e_j \rangle|^2 \leq s_{n-k+1}^2 \sum_{j=n-k+1}^{n} |\langle x, e_j \rangle|^2 = s_{n-k+1}^2 \|x\|^2,$$

where the last equality follows from 6.30(b). By 7.88(c), $\|T|_X\| \leq s_{n-k+1}$.
$\boxed{\geq}$ Let U be any subspace of V with $\dim U = k$. Let $Y = \operatorname{span}(e_1, \dots, e_{n-k+1})$, so $\dim Y = n - k + 1$ by 6.25. By 2.37, $\dim(U + Y) \leq n$. By 2.43,

$$\dim(U \cap Y) = \dim U + \dim Y - \dim(U + Y) \geq k + (n - k + 1) - n = 1.$$

Let $u \in U \cap Y$ be nonzero. For all $j > n - k + 1$, we have $\langle u, e_j \rangle = 0$ by the orthonormality of $e_1, \dots, e_n$. By the above calculation,

$$\|Tu\|^2 = \sum_{j=1}^{n-k+1} s_j^2 |\langle u, e_j \rangle|^2 \geq s_{n-k+1}^2 \sum_{j=1}^{n-k+1} |\langle u, e_j \rangle|^2 = s_{n-k+1}^2 \|u\|^2,$$

where the last equality follows from 6.30(b). Taking square roots yields $\|Tu\| \geq s_{n-k+1}\|u\|$. By 7.88(c), $\|T|_U\| \, \|u\| \geq \|Tu\| \geq s_{n-k+1}\|u\|$, so $\|T|_U\| \geq s_{n-k+1}$ by dividing both sides by $\|u\| > 0$ since $u \neq 0$. Since every U with $\dim U = k$ satisfies $\|T|_U\| \geq s_{n-k+1}$, and we have found an X with $\dim X = k$ such that $\|T|_X\| \leq s_{n-k+1}$, the minimum exists and is attained at X, and therefore equals s_{n-k+1}. ∎

Exercise 7F23

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 6B23).

Manual Remark. *For reference, the metrics on V and W that arise from the norms are $d_V(v_1, v_2) = \|v_1 - v_2\|$ and $d_W(w_1, w_2) = \|w_1 - w_2\|$. With these metrics, uniform continuity of T means that for every $\varepsilon > 0$, there exists a $\delta > 0$ such that for all $v_1, v_2 \in V$, if $\|v_1 - v_2\| < \delta$, then $\|Tv_1 - Tv_2\| < \varepsilon$.*

Solution. Let $\varepsilon > 0$ be given. Define

$$\delta = \frac{\varepsilon}{\|T\| + 1}.$$

Note that $\delta > 0$ since $\|T\| \geq 0$ by 7.87(a). Suppose $v_1, v_2 \in V$ such that $\|v_1 - v_2\| < \delta$. By 7.88(c),

$$\|Tv_1 - Tv_2\| = \|T(v_1 - v_2)\| \leq \|T\| \, \|v_1 - v_2\| \leq \|T\|\delta = \|T\| \frac{\varepsilon}{\|T\| + 1} < \varepsilon.$$

Hence, T is uniformly continuous. ∎

Exercise 7F24

Suppose $T \in \mathcal{L}(V)$ is invertible. Prove that

$$\|T^{-1}\| = \|T\|^{-1} \iff \frac{T}{\|T\|} \text{ is a unitary operator.}$$

Solution. Since T is invertible, $T \neq 0$ so $\|T\| \neq 0$ by 7.87(b). Let $S = \frac{T}{\|T\|}$, so $S^{-1} = \|T\|T^{-1}$. By 7.87(c), $\|S\| = 1$.
$\boxed{\Rightarrow}$ Suppose $\|T^{-1}\| = \|T\|^{-1}$. By 7.87(c), $\|S^{-1}\| = \|T^{-1}\|\|T\| = 1$. Then for every $v \in V$, we have by 7.88(c) that

$$\|v\| = \|S^{-1}Sv\| \leq \|S^{-1}\|\,\|Sv\| = \|Sv\| \leq \|S\|\,\|v\| = \|v\|,$$

so $\|Sv\| = \|v\|$. Therefore, S is an invertible isometry, so S is unitary.
$\boxed{\Leftarrow}$ Suppose S is unitary. Then $S^{-1} = S^*$ by 7.53(c), and hence is unitary by 7.53(f). Thus, $\|S^{-1}\| = 1$ by 7.88(b), so $\|T\|\,\|T^{-1}\| = \|S^{-1}\| = 1$ by 7.87(c). Hence, $\|T^{-1}\| = \|T\|^{-1}$. ∎

Exercise 7F25

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. Observe that $\langle Tv, w\rangle = \langle\langle v, u\rangle x, w\rangle = \langle v, u\rangle\langle x, w\rangle$ for all $v, w \in V$. Then $\langle v, u\rangle\langle x, w\rangle = \langle v, \overline{\langle x, w\rangle}u\rangle = \langle v, \langle w, x\rangle u\rangle$ for all $v, w \in V$, so $T^*w = \langle w, x\rangle u$ for all $w \in V$. Therefore,

$$T^*Tv = T^*(\langle v, u\rangle x) = \langle v, u\rangle T^*x = \langle v, u\rangle\langle x, x\rangle u = \|x\|^2\langle v, u\rangle u$$

for all $v \in V$. Define $R \in \mathcal{L}(V)$ by $Rv = \frac{\|x\|}{\|u\|}\langle v, u\rangle u$ for all $v \in V$. Since $\frac{\|x\|}{\|u\|} \geq 0$, $\langle Rv, w\rangle = \frac{\|x\|}{\|u\|}\langle v, u\rangle\langle u, w\rangle = \langle v, Rw\rangle$ for all $v, w \in V$, so R is self-adjoint. Moreover,

$$\langle Rv, v\rangle = \frac{\|x\|}{\|u\|}\langle v, u\rangle\langle u, v\rangle = \frac{\|x\|}{\|u\|}|\langle v, u\rangle|^2 \geq 0,$$

so R is positive. Finally,

$$R^2v = R\left(\frac{\|x\|}{\|u\|}\langle v, u\rangle u\right) = \frac{\|x\|}{\|u\|}\left\langle\frac{\|x\|}{\|u\|}\langle v, u\rangle u, u\right\rangle u = \frac{\|x\|^2}{\|u\|^2}\langle v, u\rangle\langle u, u\rangle u = \|x\|^2\langle v, u\rangle u = T^*Tv$$

for all $v \in V$. Then R is a positive square root of T^*T, so $R = \sqrt{T^*T}$ by 7.39. Hence,

$$\sqrt{T^*T}\,v = Rv = \frac{\|x\|}{\|u\|}\langle v, u\rangle u$$

for all $v \in V$. ∎

Exercise 7F26

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. Since $\sqrt{T^*T}$ is positive, it is self-adjoint. Then for any $v \in V$, we have

$$\left\|\sqrt{T^*T}\, v\right\|^2 = \langle T^*Tv, v\rangle = \|Tv\|^2.$$

Hence, $\operatorname{null}\sqrt{T^*T} = \operatorname{null} T$. By the polar decomposition of T, the existence of such S is guaranteed. As such, it remains to show that S is unique if and only if T is invertible.

$\boxed{\Rightarrow}$ Suppose T is invertible. By the above, $\operatorname{null}\sqrt{T^*T} = \operatorname{null} T = \{0\}$, so $\sqrt{T^*T}$ is injective, and hence invertible by 3.65. Moreover, if $S_1, S_2 \in \mathcal{L}(V)$ are unitary operators such that $T = S_1\sqrt{T^*T} = S_2\sqrt{T^*T}$, then right-multiplying by $(\sqrt{T^*T})^{-1}$ gives $S_1 = S_2$, so S is unique.

$\boxed{\Leftarrow}$ Suppose T is not invertible. By 3.65, $\operatorname{null} T \neq \{0\}$. By the polar decomposition of T, there exists a unitary operator S such that $T = S\sqrt{T^*T}$. By 6.35, there exists an orthonormal basis $e_1, \ldots, e_m$ of $\operatorname{null} T$, where $m \geq 1$. By 6.36, we may extend this to an orthonormal basis $e_1, \ldots, e_n$ of V. By the linear map lemma, there exists $R \in \mathcal{L}(V)$ such that $Re_j = -e_j$ for $j \leq m$ and $Re_j = e_j$ for $j > m$. Because $Re_1, \ldots, Re_n$ is an orthonormal basis of V, R is unitary by 7.53(d). Moreover, self-adjointness of $\sqrt{T^*T}$ gives

$$\operatorname{range}\sqrt{T^*T} = (\operatorname{null}\sqrt{T^*T})^\perp = (\operatorname{null} T)^\perp$$

by 7.6(b). Now, if $u \in (\operatorname{null} T)^\perp$, then $\langle u, e_j\rangle = 0$ for $j \leq m$. By 6.30(a), we may write

$$u = \sum_{j=m+1}^{n} \langle u, e_j\rangle e_j,$$

so $Ru = u$. Consequently, $R\sqrt{T^*T} = \sqrt{T^*T}$, so $SR\sqrt{T^*T} = S\sqrt{T^*T} = T$. By 7.5(d) and 7.53(b),

$$(SR)^*(SR) = R^*S^*SR = I, \quad (SR)(SR)^* = SRR^*S^* = I.$$

Hence, SR is unitary. Lastly, injectivity of S shows that $Se_1 \neq 0$, so $SRe_1 = -Se_1 \neq Se_1$. Therefore, S and SR are distinct unitary operators such that $T = S\sqrt{T^*T} = SR\sqrt{T^*T}$, and S is not unique. ■

Exercise 7F27

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 $\|T - E\| \geq \|T - S\|$.

Axler Note. *This exercise finds a unitary operator S that is as close as possible (among the unitary operators) to a given operator T.*

Solution.

(a) Since $Se_j = f_j$ for all j, S is unitary by 7.53(d). By assumption,

$$(T - S)v = \sum_{j=1}^{n}(s_j - 1)\langle v, e_j\rangle f_j$$

for any $v \in V$. Define

$$M = \max_{1 \le j \le n} |s_j - 1|.$$

By 6.24 and 6.30(b),

$$\|(T-S)v\|^2 = \sum_{j=1}^{n} |s_j - 1|^2 |\langle v, e_j \rangle|^2 \le M^2 \sum_{j=1}^{n} |\langle v, e_j \rangle|^2 = M^2 \|v\|^2$$

for all $v \in V$. Taking square roots and applying 7.88(c) gives $\|T - S\| \le M$. Choose k such that $|s_k - 1| = M$. Since $\|(T-S)e_k\| = |s_k - 1| = M$ and $\|e_k\| = 1$, 7.88(b) gives $\|T - S\| \ge M$. Therefore, $\|T - S\| = M$.

(b) Suppose $E \in \mathcal{L}(V)$ is unitary. Then it is an isometry, so $\|Ee_j\| = \|e_j\| = 1$ for all j. Moreover, since $Te_j = s_j f_j$ and $s_j \ge 0$ for all j, we have $\|Te_j\| = s_j$. Because $\|e_j\| = 1$, 7.88(b) and the reverse triangle inequality give

$$\|T - E\| \ge \|(T-E)e_j\| \ge \big|\|Te_j\| - \|Ee_j\|\big| = |s_j - 1|.$$

Since this holds for all j, taking the maximum over j gives

$$\|T - E\| \ge \max_{1 \le j \le n} |s_j - 1| = \|T - S\|. \qquad \blacksquare$$

Exercise 7F28

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. By the polar decomposition of T^*, there exists a unitary operator $U \in \mathcal{L}(V)$ such that $T^* = U\sqrt{(T^*)^* T^*}$. By 7.5(c), $(T^*)^* = T$, so $T^* = U\sqrt{TT^*}$. Since $\sqrt{TT^*}$ is positive, it is self-adjoint, so $\sqrt{TT^*} = (\sqrt{TT^*})^*$. By 7.5(c), (d), and taking adjoints, $T = (U\sqrt{TT^*})^* = \sqrt{TT^*}\, U^*$. By 7.53(f), U^* is unitary since U is unitary. Setting $S = U^*$ gives the desired result. $\blacksquare$

Exercise 7F29

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) By the polar decomposition of T, there exists a unitary operator $S \in \mathcal{L}(V)$ such that $T = S\sqrt{T^*T}$. By 7.5(d), $T^* = (\sqrt{T^*T})^* S^*$. Since $\sqrt{T^*T}$ is positive, it is self-adjoint, so $(\sqrt{T^*T})^* = \sqrt{T^*T}$, and hence $T^* = \sqrt{T^*T}\, S^*$. By the definition of the square root,

$$TT^* = S\sqrt{T^*T}\left(\sqrt{T^*T}\, S^*\right) = ST^*TS^*.$$

(b) Since T^*T is positive, 7.38(c) gives an orthonormal basis $e_1, \ldots, e_n$ of V and nonnegative numbers $\mu_1, \ldots, \mu_n$ such that $T^*Te_k = \mu_k e_k$ for each k. By 7.53(b), $S^*S = I$ since S is unitary. Therefore,

$$TT^*(Se_k) = ST^*TS^*(Se_k) = ST^*Te_k = S(\mu_k e_k) = \mu_k(Se_k).$$

Since S is invertible, $Se_1, \ldots, Se_n$ is also a basis of V. By Lemma 2 applied to T with the basis $e_1, \ldots, e_n$, the singular values of T are $\sqrt{\mu_1}, \ldots, \sqrt{\mu_n}$ sorted into decreasing order. By 7.5(c), $(T^*)^*T^* = TT^*$, so $(T^*)^*T^*(Se_k) = TT^*(Se_k) = \mu_k(Se_k)$ for each k. By Lemma 2 applied to T^* with the basis $Se_1, \ldots, Se_n$, the singular values of T^* are $\sqrt{\mu_1}, \ldots, \sqrt{\mu_n}$ sorted into decreasing order, which are the same as those of T. $\blacksquare$

Exercise 7F30

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 $R = \sqrt{T^*T}$.

Axler Note. *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. Since R is positive, it is self-adjoint. By 7.5(d) and 7.53(b), $T^*T = RS^*SR = R^2$. Then R is a positive square root of T^*T, so $R = \sqrt{T^*T}$ by 7.39. ∎

Exercise 7F31

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. By the complex spectral theorem, there exists an orthonormal basis $e_1, \dots, e_n$ of V consisting of eigenvectors of T, and scalars $\lambda_1, \dots, \lambda_n$ such that $Te_k = \lambda_k e_k$ for each k. By 7.21(e),

$$T^*Te_k = \lambda_k T^*e_k = \lambda_k \overline{\lambda_k} e_k = |\lambda_k|^2 e_k.$$

By the linear map lemma, there exists $R \in \mathcal{L}(V)$ such that $Re_k = |\lambda_k| e_k$ for each k. Then $\mathcal{M}(R, (e_1, \dots, e_n))$ is diagonal with entries $|\lambda_1|, \dots, |\lambda_n|$, so R is positive by 7.38(c) and (a). Moreover, $R^2 e_k = |\lambda_k|^2 e_k = T^*Te_k$ for each k, so $R^2 = T^*T$ by the linear map lemma. Since R is a positive square root of T^*T, we have $R = \sqrt{T^*T}$ by 7.39. By the linear map lemma again, there exists $S \in \mathcal{L}(V)$ such that $Se_k = \mu_k e_k$, where

$$\mu_k = \begin{cases} \frac{\lambda_k}{|\lambda_k|}, & \lambda_k \neq 0, \\ 1, & \lambda_k = 0. \end{cases}$$

Since each $Se_k = \mu_k e_k$ with $|\mu_k| = 1$, it follows that $Se_1, \dots, Se_n$ is an orthonormal basis of V, so S is unitary by 7.53(d). Moreover, $Te_k = \lambda_k e_k = \mu_k |\lambda_k| e_k = SRe_k$ for each k, so $T = SR = S\sqrt{T^*T}$ by the linear map lemma. It now follows that $\mathcal{M}(S, (e_1, \dots, e_n))$ is diagonal with entries $\mu_1, \dots, \mu_n$, and $\mathcal{M}(\sqrt{T^*T}, (e_1, \dots, e_n))$ is diagonal with entries $|\lambda_1|, \dots, |\lambda_n|$, so S and $\sqrt{T^*T}$ both have diagonal matrices with respect to the same orthonormal basis $e_1, \dots, e_n$. ∎

Exercise 7F32

Suppose that $T \in \mathcal{L}(V, W)$ and $T \neq 0$. Let $s_1, \dots, s_m$ denote the positive singular values of T. Show that there exists an orthonormal basis $e_1, \dots, e_m$ of $(\operatorname{null} T)^\perp$ such that

$$T\left(E\left(\frac{e_1}{s_1}, \dots, \frac{e_m}{s_m}\right)\right)$$

equals the ball in $\operatorname{range} T$ of radius 1 centered at 0.

Solution. Let $B = \{w \in \operatorname{range} T : \|w\| < 1\}$. Since $T \neq 0$, we have $\dim \operatorname{range} T = m \geq 1$ by 7.68(b). By the singular value decomposition, there exist orthonormal lists $e_1, \dots, e_m$ in V and $f_1, \dots, f_m$ in W such that

$$Tv = \sum_{k=1}^{m} s_k \langle v, e_k \rangle f_k$$

for all $v \in V$. Moreover, for any $u \in \operatorname{null} T$,

$$Tu = \sum_{k=1}^{m} s_k \langle u, e_k \rangle f_k = 0.$$

Since $f_1, \dots, f_m$ is linearly independent by 6.25 and $s_k > 0$ for all k, $\langle u, e_k \rangle = 0$ for all $k = 1, \dots, m$, so $e_k \in (\operatorname{null} T)^\perp$ for all k. By 6.51 and the fundamental theorem of linear maps, $\dim(\operatorname{null} T)^\perp = \dim \operatorname{range} T = m$, so $e_1, \dots, e_m$ form an orthonormal basis of $(\operatorname{null} T)^\perp$ by 6.28. By 7.96 in $(\operatorname{null} T)^\perp$,

$$E = E\left(\frac{e_1}{s_1}, \dots, \frac{e_m}{s_m}\right) = \left\{ v \in (\operatorname{null} T)^\perp : \sum_{k=1}^{m} s_k^2 |\langle v, e_k \rangle|^2 < 1 \right\}.$$

By 6.24, $\|Tv\|^2 = \sum_{k=1}^{m} s_k^2 |\langle v, e_k \rangle|^2$ for all $v \in V$, so

$$E = \{ v \in (\operatorname{null} T)^\perp : \|Tv\| < 1 \}.$$

We now show that $T(E) = B$.

$\boxed{\subseteq}$ This follows immediately from the definition of E.

$\boxed{\supseteq}$ Suppose $w \in B$. By 6.67, there exists $v \in (\operatorname{null} T)^\perp$ such that $Tv = w$. Then $\|Tv\| < 1$, so $v \in E$, and hence $w = Tv \in T(E)$. ∎

8 Operators on Complex Vector Spaces

8A Generalized Eigenvectors and Nilpotent Operators

Exercise 8A1

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, $\operatorname{null} T^4 \subseteq \operatorname{null} T^5 \subseteq \operatorname{null} T^6$. Then $8 \leq \dim \operatorname{null} T^5 \leq 9$ by 2.37. If $\dim \operatorname{null} T^5 = 8$, then 2.39 gives $\operatorname{null} T^4 = \operatorname{null} T^5$. 8.2 yields $\operatorname{null} T^4 = \operatorname{null} T^6$, contradicting $\dim \operatorname{null} T^6 = 9$. Therefore, $\dim \operatorname{null} T^5 = 9$, so $\operatorname{null} T^5 = \operatorname{null} T^6$ by 2.39. Hence, $\operatorname{null} T^5 = \operatorname{null} T^m$ for all $m \geq 5$ by 8.2, so $\dim \operatorname{null} T^m = 9$ for all $m \geq 5$. ■

Exercise 8A2

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.

Axler Note. *The result in this exercise is used in the proof of 8.45.*

Solution. By 3.10, $T^j v = T^{j-m}(T^m v) = T^{j-m}(0) = 0$ for all $j \geq m$. Suppose

$$\sum_{j=0}^{m-1} \lambda_j T^j v = 0$$

for some scalars $\lambda_0, \ldots, \lambda_{m-1} \in \mathbf{F}$, where not all λ_j are zero. Let k be the smallest index such that $\lambda_k \neq 0$. Applying T^{m-1-k} to both sides gives

$$\sum_{j=0}^{m-1} \lambda_j T^{j+m-1-k} v = 0.$$

If $j < k$, then $\lambda_j = 0$ by the choice of k. If $j > k$, then $j + m - 1 - k \geq m$, so $T^{j+m-1-k}v = 0$ by the observation above. Hence, the sum reduces to $\lambda_k T^{m-1} v = 0$. Since $\lambda_k \neq 0$, it must be that $T^{m-1}v = 0$, which contradicts the assumption that $T^{m-1}v \neq 0$. Therefore, all $\lambda_j = 0$, proving that $v, Tv, \ldots, T^{m-1}v$ is linearly independent. ■

Exercise 8A3

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

$$V = \operatorname{null} T \oplus \operatorname{range} T \iff \operatorname{null} T^2 = \operatorname{null} T.$$

Solution. By Exercise 5D4, it suffices to show that $\operatorname{null} T \cap \operatorname{range} T = \{0\}$ if and only if $\operatorname{null} T^2 = \operatorname{null} T$. By 8.1, $\operatorname{null} T \subseteq \operatorname{null} T^2$. Moreover, $0 \in \operatorname{null} T \cap \operatorname{range} T$. Consequently,

$$\begin{aligned}
\operatorname{null} T^2 = \operatorname{null} T &\iff \operatorname{null} T^2 \subseteq \operatorname{null} T \\
&\iff \text{for all } v \in V,\ T^2 v = 0 \text{ implies } Tv = 0 \\
&\iff \text{for all } w \in \operatorname{range} T,\ Tw = 0 \text{ implies } w = 0 \\
&\iff \operatorname{null} T \cap \operatorname{range} T = \{0\}.
\end{aligned}$$

■

Exercise 8A4

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, $p(z) = (z-\lambda)^m q(z)$ for some $q \in \mathcal{P}(\mathbf{F})$, which is monic since p and $(z-\lambda)^m$ are monic. We claim that for each $k = 1, \dots, m$,

$$\operatorname{null}(T-\lambda I)^{k-1} \subsetneq \operatorname{null}(T-\lambda I)^k.$$

Indeed, if equality did hold for some k, then 8.2 applied to $T-\lambda I$ yields $\operatorname{null}(T-\lambda I)^{k-1} = \operatorname{null}(T-\lambda I)^m$. Then for each $v \in V$,

$$(T-\lambda I)^m(q(T)v) = p(T)v = 0$$

by 5.17(a), so $q(T)v \in \operatorname{null}(T-\lambda I)^m = \operatorname{null}(T-\lambda I)^{k-1}$. Let $r(z) = (z-\lambda)^{k-1}q(z)$. Then r is monic, and

$$\deg r = \deg p - (m-k+1) < \deg p.$$

Additionally, $r(T)v = (T-\lambda I)^{k-1}q(T)v = 0$ for all $v \in V$ by 5.17(a), whence $r(T) = 0$. This contradicts the minimality of p by 5.22, so equality cannot hold for any $k = 1, \dots, m$. Thus,

$$\{0\} = \operatorname{null}(T-\lambda I)^0 \subsetneq \operatorname{null}(T-\lambda I)^1 \subsetneq \cdots \subsetneq \operatorname{null}(T-\lambda I)^m.$$

Since strict inclusion raises the dimension by at least 1 by 2.37 and 2.39, we have

$$\dim \operatorname{null}(T-\lambda I)^m \geq m.$$

■

Exercise 8A5

Suppose $T \in \mathcal{L}(V)$ and m is a positive integer. Prove that

$$\dim \operatorname{null} T^m \leq m \dim \operatorname{null} T.$$

Axler Hint. Exercise 3B21 *may be useful.*

Solution. We proceed by induction on m. The base case $m = 1$ is trivial since $\dim \operatorname{null} T^1 = \dim \operatorname{null} T \leq 1 \cdot \dim \operatorname{null} T$. Suppose $m \geq 2$ and the statement holds for $m-1$. Since $v \in \operatorname{null} T^m$ if and only if $T^{m-1}Tv = 0$, i.e., $Tv \in \operatorname{null} T^{m-1}$, we have

$$\operatorname{null} T^m = \{v \in V : Tv \in \operatorname{null} T^{m-1}\}.$$

By Exercise 3B21 with $W = V$ and $U = \operatorname{null} T^{m-1}$, we have

$$\begin{aligned}
\dim \operatorname{null} T^m &= \dim \operatorname{null} T + \dim(\operatorname{null} T^{m-1} \cap \operatorname{range} T) \\
&\leq \dim \operatorname{null} T + \dim \operatorname{null} T^{m-1} \\
&\leq \dim \operatorname{null} T + (m-1) \dim \operatorname{null} T \\
&= m \dim \operatorname{null} T,
\end{aligned}$$

where the first inequality follows from 2.37, and the second inequality follows from the induction hypothesis. This completes the induction step and proves the statement for all positive integers m. ■

Exercise 8A6

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. Note that $\operatorname{range} T^0 = \operatorname{range} I = V$. For each $k \geq 0$, every $v \in \operatorname{range} T^{k+1}$ has the form $v = T^{k+1}u$ for some $u \in V$, hence $v = T^k(Tu) \in \operatorname{range} T^k$. Therefore, $\operatorname{range} T^{k+1} \subseteq \operatorname{range} T^k$ for all $k \geq 0$. ■

Exercise 8A7

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. Observe that for each $j \geq 0$,

$$\operatorname{range} T^{j+1} = \{T(T^j v) : v \in V\} = \{Tw : w \in \operatorname{range} T^j\}.$$

We proceed by induction on $k \geq m+1$. The base case $k = m+1$ is given. Suppose $k \geq m+1$ and $\operatorname{range} T^k = \operatorname{range} T^m$. Then

$$\operatorname{range} T^{k+1} = \{Tw : w \in \operatorname{range} T^k\} = \{Tw : w \in \operatorname{range} T^m\} = \operatorname{range} T^{m+1} = \operatorname{range} T^m.$$

This completes the induction step and proves the statement for all $k > m$. ■

Exercise 8A8

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. By 8.3, $\operatorname{null} T^{\dim V+k} = \operatorname{null} T^{\dim V}$ for all $k \geq 0$. By the fundamental theorem of linear maps,

$$\dim \operatorname{range} T^{\dim V+k} = \dim V - \dim \operatorname{null} T^{\dim V+k} = \dim V - \dim \operatorname{null} T^{\dim V} = \dim \operatorname{range} T^{\dim V}$$

for all $k \geq 0$. By Exercise 8A6, $\operatorname{range} T^{\dim V+k} \subseteq \operatorname{range} T^{\dim V}$ for all $k \geq 0$. Equality follows from 2.39. ■

Exercise 8A9

Suppose $T \in \mathcal{L}(V)$ and m is a nonnegative integer. Prove that

$$\operatorname{null} T^m = \operatorname{null} T^{m+1} \iff \operatorname{range} T^m = \operatorname{range} T^{m+1}.$$

Solution. By 8.1 and Exercise 8A6, $\operatorname{null} T^m \subseteq \operatorname{null} T^{m+1}$ and $\operatorname{range} T^{m+1} \subseteq \operatorname{range} T^m$. By 2.39, each inclusion is an equality if and only if the dimensions are equal. By the fundamental theorem of linear maps, $\dim \operatorname{null} T^j = \dim V - \dim \operatorname{range} T^j$

for all $j \geq 0$. Therefore,

$$\begin{aligned} \operatorname{null} T^m = \operatorname{null} T^{m+1} &\iff \dim \operatorname{null} T^m = \dim \operatorname{null} T^{m+1} \\ &\iff \dim V - \dim \operatorname{range} T^m = \dim V - \dim \operatorname{range} T^{m+1} \\ &\iff \dim \operatorname{range} T^m = \dim \operatorname{range} T^{m+1} \\ &\iff \operatorname{range} T^m = \operatorname{range} T^{m+1}. \end{aligned}$$

■

Exercise 8A10

Define $T \in \mathcal{L}(\mathbf{C}^2)$ by $T(w, z) = (z, 0)$. Find all generalized eigenvectors of T.

Solution. Observe that $T^2(w, z) = T(z, 0) = (0, 0)$, so $T^2 = 0$, and hence T is nilpotent. By 8.17(a), the only eigenvalue of T is 0. Therefore, every generalized eigenvector of T corresponds to 0. For every nonzero $v \in \mathbf{C}^2$, we have $T^2 v = 0$, so v is a generalized eigenvector of T corresponding to 0. ■

Exercise 8A11

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 $\lambda_1, \ldots, \lambda_m \in \mathbf{F}$.

Axler Hint. *Assume* $\mathbf{F} = \mathbf{R}$ *because the case* $\mathbf{F} = \mathbf{C}$ *follows from 5.27(b) and 8.9.*

Axler Note. *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 $v_1, \ldots, v_n$ of V, then v_1 is an eigenvector of T, but it is not necessarily true that $v_2, \ldots, v_n$ are generalized eigenvectors of T.*

Solution. If $\mathbf{F} = \mathbf{C}$, the result follows from 5.27(b) and 8.9, so we may assume $\mathbf{F} = \mathbf{R}$.

We first claim that a monic $q_0 \in \mathcal{P}(\mathbf{F})$ that divides $q(z) = (z - \mu_1) \cdots (z - \mu_m)$ for some $\mu_1, \ldots, \mu_m \in \mathbf{F}$ is a product of linear factors, each of the form $z - \mu_k$ for some k. Write $q = q_0 r$, and factor q_0 and r by 4.16. By the uniqueness in 4.16, q_0 consists of linear factors of the form $z - \mu_k$.

Now, let $n = \dim V$ and let p be the minimal polynomial of T.

$\boxed{\Rightarrow}$ Suppose $v_1, \ldots, v_n$ is a basis of V consisting of generalized eigenvectors of T with corresponding eigenvalues $\nu_1, \ldots, \nu_n$. Then there exist positive integers $k_1, \ldots, k_n$ such that $(T - \nu_\ell I)^{k_\ell} v_\ell = 0$ for each $\ell = 1, \ldots, n$. Consider the polynomial

$$s(z) = \prod_{\ell=1}^{n} (z - \nu_\ell)^{k_\ell}.$$

By 5.17, the factors of $s(T) = \prod_\ell (T - \nu_\ell I)^{k_\ell}$ commute with each other. Moving $(T - \nu_\ell I)^{k_\ell}$ to the right gives $s(T) v_\ell = 0$ for each ℓ, so $s(T) = 0$ because $v_1, \ldots, v_n$ is a basis of V. By 5.29, p divides s, so the preliminary claim implies that p is a product of linear factors, each of the form $z - \nu_\ell$ for some ℓ.

$\boxed{\Leftarrow}$ We proceed by strong induction on n. Suppose $p(z) = (z - \lambda_1) \cdots (z - \lambda_m)$ for some $\lambda_1, \ldots, \lambda_m \in \mathbf{F}$. Suppose the result holds for all operators on all vector spaces of dimension less than n whose minimal polynomial is a product of linear factors. Let $U = \operatorname{range}(T - \lambda_1 I)^n$, so $V = \operatorname{null}(T - \lambda_1 I)^n \oplus U$ by 8.4.

If $U = \{0\}$, then $V = \operatorname{null}(T - \lambda_1 I)^n$ by 8.4, and any basis of V works. Otherwise, U is nontrivial and invariant under T by 5.18. Since λ_1 is an eigenvalue of T by 5.27(a), $\{0\} \neq \operatorname{null}(T - \lambda_1 I) \subseteq \operatorname{null}(T - \lambda_1 I)^n$, so $\dim U < n$. By 5.31 and the preliminary claim, the minimal polynomial of $T|_U$ is a product of linear factors. If $u \in U$ is a generalized eigenvector of $T|_U$ corresponding to an eigenvalue λ, then $(T - \lambda I)^k u = (T|_U - \lambda I)^k u = 0$ for some positive integer k. Additionally, λ is an eigenvalue of T because an eigenvector of $T|_U$ is also an eigenvector of T. Hence, u is a generalized eigenvector of T corresponding to λ. By the induction hypothesis and the preceding argument, U has a basis consisting of generalized

eigenvectors of T. Concatenating any basis of $\operatorname{null}(T-\lambda_1 I)^n$ with the basis of U from above gives a basis of V consisting of generalized eigenvectors of T by 8.4. ■

Exercise 8A12

Suppose $T \in \mathcal{L}(V)$ is such that every nonzero 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. Let $v \in V$ be a nonzero vector, and let $\dim V = n$. By 8.1 and 8.3, $\operatorname{null}(T-\alpha I)^k \subseteq \operatorname{null}(T-\alpha I)^n$ for all $\alpha \in \mathbf{F}$ and positive integers k. By hypothesis, v is a generalized eigenvector of T corresponding to some eigenvalue $\lambda \in \mathbf{F}$. We claim that every nonzero vector in V is a generalized eigenvector of T corresponding to the same eigenvalue λ.

Suppose some nonzero $u \in V$ is not a generalized eigenvector of T corresponding to λ. By hypothesis and 8.11, u corresponds to a unique eigenvalue $\mu \in \mathbf{F}$, where $\mu \neq \lambda$ by choice of u. By 8.12, the list v, u is linearly independent, so $v+u$ is a nonzero vector in V. By hypothesis and 8.11 again, $v+u$ is a generalized eigenvector of T corresponding to a unique eigenvalue $\nu \in \mathbf{F}$. If $\nu = \lambda$, then $u = (v+u) - v \in \operatorname{null}(T-\lambda I)^n$, which contradicts 8.11, since u corresponds to $\mu \neq \lambda$. If $\nu = \mu$, then $v = (v+u) - u \in \operatorname{null}(T-\mu I)^n$, which contradicts 8.11, since v corresponds to $\lambda \neq \mu$. Thus, $\nu \neq \lambda$ and $\nu \neq \mu$, whence the list $v, u, v+u$ is linearly independent by 8.12, which is a contradiction since $v+u \in \operatorname{span}(v,u)$. Therefore, every nonzero vector in V is a generalized eigenvector of T corresponding to the same eigenvalue λ. Hence, each nonzero vector of V lies in $\operatorname{null}(T-\lambda I)^k$ for some positive integer k, and thus lies in $\operatorname{null}(T-\lambda I)^n$. Moreover, $(T-\lambda I)^n 0 = 0$, so $(T-\lambda I)^n = 0$, which shows that $T - \lambda I$ is nilpotent. ■

Exercise 8A13

Suppose $S, T \in \mathcal{L}(V)$ and ST is nilpotent. Prove that TS is nilpotent.

Solution. Since ST is nilpotent, there exists a positive integer k such that $(ST)^k = 0$. Then

$$(TS)^{k+1} = T(ST)^k S = T0S = 0,$$

so TS is nilpotent. ■

Exercise 8A14

Suppose $T \in \mathcal{L}(V)$ is nilpotent and $T \neq 0$. Prove T is not diagonalizable.

Solution. Suppose T is diagonalizable. By 8.17(a), the only eigenvalue of T is 0. Then 5.55(c) gives $V = E(0,T) = \operatorname{null} T$, so $T = 0$, which contradicts the assumption that $T \neq 0$. ■

Exercise 8A15

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.

Axler Note. *For $\mathbf{F} = \mathbf{C}$, this exercise adds another equivalence to the list of conditions for diagonalizability in 5.55.*

Solution.

$\boxed{\Rightarrow}$ Suppose T is diagonalizable, and let $\lambda_1, \ldots, \lambda_m$ be the distinct eigenvalues of T. By 5.55(c), we have

$$V = \bigoplus_{k=1}^{m} E(\lambda_k, T).$$

Let $v \in V$ be a generalized eigenvector of T corresponding to an eigenvalue λ, so $\lambda = \lambda_j$ for some j, and $(T - \lambda I)^d v = 0$ for some positive integer d. Write $v = \sum_{k=1}^m v_k$ with $v_k \in E(\lambda_k, T)$. Then

$$\sum_{k=1}^m (T - \lambda I)^d v_k = \sum_{k=1}^m (\lambda_k - \lambda)^d v_k = 0.$$

Because the sum is direct, $(\lambda_k - \lambda)^d v_k = 0$ for each k. For $\lambda_k \neq \lambda$, this implies $v_k = 0$, so $v = v_j \in E(\lambda, T)$, which shows that v is an eigenvector of T corresponding to λ.

$\boxed{\Leftarrow}$ Suppose every generalized eigenvector of T is an eigenvector of T. By 8.9, V has a basis consisting of generalized eigenvectors of T, which are eigenvectors of T by assumption, so 5.55(b) implies that T is diagonalizable. ■

Exercise 8A16

(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 $V = \mathbf{F}^2$ and define $S, T \in \mathcal{L}(V)$ by $S(x, y) = (y, 0)$ and $T(x, y) = (0, x)$. Then $S^2(x, y) = S(y, 0) = (0, 0)$ and $T^2(x, y) = T(0, x) = (0, 0)$, so $S^2 = T^2 = 0$, which shows that S and T are nilpotent. Moreover,

$$(S + T)(1, 1) = (1, 1), \quad ST(1, 0) = (1, 0),$$

from $(S + T)(x, y) = (y, x)$ and $ST(x, y) = (x, 0)$. Then $(1, 1)$ and $(1, 0)$ are eigenvectors of $S + T$ and ST, respectively, corresponding to the eigenvalue 1, so neither $S + T$ nor ST is nilpotent by 8.17(a).

(b) By 8.16, $S^n = T^n = 0$, where $n = \dim V$. Since S and T commute, $(ST)^n = S^n T^n = 0$, so ST is nilpotent. By the commutativity of S and T and the distributive property of linear maps,

$$(S + T)^{2n-1} = \sum_{k=0}^{2n-1} \binom{2n-1}{k} S^k T^{2n-1-k}.$$

For $k \geq n$, $S^k = S^n S^{k-n} = 0$, and for $2n - 1 - k \geq n$, $T^{2n-1-k} = T^n T^{n-1-k} = 0$. If $k \leq n - 1$, then $2n - 1 - k \geq n$, so at least one of S^k or T^{2n-1-k} is 0. Hence, every term in the sum is 0, so $(S + T)^{2n-1} = 0$, which shows that $S + T$ is nilpotent. ■

Exercise 8A17

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) Let $R = I + T + \cdots + T^{m-1}$. Then

$$(I - T)R = (I + T + \cdots + T^{m-1}) - (T + T^2 + \cdots + T^m) = I - T^m = I.$$

By 5.17(b), $(I - T)R = R(I - T) = I$, so $I - T$ is invertible with inverse R.

(b) Recall the geometric series formula

$$\frac{1}{1 - x} = \sum_{k=0}^{\infty} x^k, \quad |x| < 1.$$

Replacing x with T and 1 with I, this suggests that

$$(I-T)^{-1} = \sum_{k=0}^{\infty} T^k.$$

Since $T^k = T^{k-m}T^m = 0$ for $k \geq m$, the series terminates at $k = m-1$, yielding

$$(I-T)^{-1} = I + T + \cdots + T^{m-1}. \qquad \blacksquare$$

Exercise 8A18

Suppose $T \in \mathcal{L}(V)$ is nilpotent. Prove that $T^{1+\dim \operatorname{range} T} = 0$.

Axler Note. *If* $\dim \operatorname{range} T < \dim V - 1$, *then this exercise improves 8.16.*

Solution. Given that T is nilpotent, there exists a positive integer k such that $T^k = 0$. Since $\operatorname{range} T$ is invariant under T, we can consider the restriction $S = T|_{\operatorname{range} T} \in \mathcal{L}(\operatorname{range} T)$. Because $S^k = T^k|_{\operatorname{range} T} = 0$, S is nilpotent. Applying 8.16 to S, we get $S^{\dim \operatorname{range} T} = 0$. Then for every $v \in V$,

$$T^{1+\dim \operatorname{range} T} v = S^{\dim \operatorname{range} T}(Tv) = 0. \qquad \blacksquare$$

Exercise 8A19

Suppose $T \in \mathcal{L}(V)$ is not nilpotent. Show that

$$V = \operatorname{null} T^{\dim V - 1} \oplus \operatorname{range} T^{\dim V - 1}.$$

Axler Note. *For operators that are not nilpotent, this exercise improves 8.4.*

Solution. If the first n inclusions in the chain of 8.1 were strict, then

$$\{0\} = \operatorname{null} T^0 \subsetneq \operatorname{null} T^1 \subsetneq \cdots \subsetneq \operatorname{null} T^n,$$

where the dimension of each successive null space increases by at least 1, so $\dim \operatorname{null} T^n \geq n = \dim V$. By 2.39, $\operatorname{null} T^n = V$, so $T^n = 0$, contradicting that T is not nilpotent. Hence, there must be some $k \leq n-1$ such that $\operatorname{null} T^k = \operatorname{null} T^{k+1}$, so by 8.2, $\operatorname{null} T^{n-1} = \operatorname{null} T^n$. Since $\operatorname{range} T^n \subseteq \operatorname{range} T^{n-1}$ and $\dim \operatorname{range} T^n = \dim \operatorname{range} T^{n-1}$ by the fundamental theorem of linear maps, it follows that $\operatorname{range} T^{n-1} = \operatorname{range} T^n$ by 2.39. By 8.4,

$$V = \operatorname{null} T^n \oplus \operatorname{range} T^n = \operatorname{null} T^{n-1} \oplus \operatorname{range} T^{n-1}. \qquad \blacksquare$$

Exercise 8A20 ✔

Suppose V is an inner product space and $T \in \mathcal{L}(V)$ is normal and nilpotent. Prove that $T = 0$.

Solution. Let $n = \dim V$. By 7.21(c), $V = \operatorname{null} T \oplus \operatorname{range} T$. By Exercise 8A3, $\operatorname{null} T^2 = \operatorname{null} T$. Applying 8.2 yields $\operatorname{null} T^k = \operatorname{null} T$ for all $k \geq 1$. By 8.16, $T^n = 0$, so $V = \operatorname{null} T^n = \operatorname{null} T$, and hence $T = 0$. $\blacksquare$

Exercise 8A21

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. Let $n = \dim V$. If $\operatorname{null} T^k = \operatorname{null} T^{k+1}$ for some $k \in \{0, \dots, n-1\}$, then 8.2 gives $\operatorname{null} T^{n-1} = \operatorname{null} T^n$, contradicting the hypothesis. Hence, every inclusion in the chain in 8.1 must be strict:

$$\{0\} = \operatorname{null} T^0 \subsetneq \operatorname{null} T^1 \subsetneq \cdots \subsetneq \operatorname{null} T^n.$$

It follows that the dimension of each successive null space increases by at least 1. Therefore, for $0 \le k \le n$,

$$\dim \operatorname{null} T^k \ge k \quad \text{and} \quad n \ge \dim \operatorname{null} T^n \ge \dim \operatorname{null} T^k + (n-k).$$

This forces $\dim \operatorname{null} T^k = k$ for all $0 \le k \le n$. By 2.39, $\operatorname{null} T^n = V$, so $T^n = 0$, and hence T is nilpotent. ■

Exercise 8A22

Suppose $T \in \mathcal{L}(\mathbf{C}^5)$ is such that $\operatorname{range} T^4 \ne \operatorname{range} T^5$. Prove that T is nilpotent.

Solution. By Exercise 8A9 with $m = 4$, $\operatorname{null} T^4 \ne \operatorname{null} T^5$. Since $\dim(\mathbf{C}^5) = 5$, Exercise 8A21 yields that T is nilpotent. ■

Exercise 8A23

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.

Axler Note. *This exercise shows that (b) in 8.17 does not hold without the hypothesis that* $\mathbf{F} = \mathbf{C}$.

Solution. Define $T \in \mathcal{L}(\mathbf{R}^3)$ by $T(x, y, z) = (0, -z, y)$. Since $T(1, 0, 0) = (0, 0, 0)$, it follows that 0 is an eigenvalue of T. Suppose (x, y, z) is an eigenvector of T with eigenvalue λ. Then $T(x, y, z) = \lambda(x, y, z)$ gives

$$\begin{cases} \lambda x = 0, \\ \lambda y = -z, \\ \lambda z = y. \end{cases}$$

The two equations $\lambda y = -z$ and $\lambda z = y$ imply $(\lambda^2 + 1)y = 0$. Since $\lambda \in \mathbf{R}$, we have $\lambda^2 + 1 \ne 0$, so $y = 0$, whence $z = -\lambda y = 0$. Since (x, y, z) is an eigenvector, it is nonzero, so $x \ne 0$, and the first equation $\lambda x = 0$ forces $\lambda = 0$. Thus, 0 is the only eigenvalue of T. Moreover, $T^3(0, 1, 0) = T^2(0, 0, 1) = T(0, -1, 0) = (0, 0, -1) \ne (0, 0, 0)$, so $T^3 \ne 0$, and hence T is not nilpotent by 8.16. ■

Exercise 8A24

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) Let $e_1, \dots, e_4$ be the standard basis of $\mathbf{F}^4$. Then $Te_1 = e_3$ and $Te_2 = e_4$, while $Te_3 = Te_4 = 0$, so a basis of $\operatorname{null} T$ is e_3, e_4. Since $\operatorname{null} T^2 = \mathbf{F}^4$ by the example, we can extend the basis of $\operatorname{null} T$ to a basis e_3, e_4, e_1, e_2 of $\mathbf{F}^4$. Therefore,

$$\mathcal{M}(T, (e_3, e_4, e_1, e_2)) = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

(b) Let $T \in \mathcal{L}(\mathbf{F}^3)$ be the operator whose matrix with respect to the standard basis is the matrix A given in the example, and consider $v = (0, 1, 0)$. Then $Tv = (9, 9, 0)$, the second column of A, and $T^2 v = T(9, 9, 0) = (54, 18, 36) \ne 0$. Lastly,

$T^3v = 0$ since $A^3 = 0$ as given in the example. By Exercise 8A2, the vectors v, Tv, T^2v are linearly independent, and hence form a basis of $\mathbf{F}^3$ by 2.38. Therefore,

$$\mathcal{M}(T, (T^2v, Tv, v)) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

(c) Consider the standard basis of $\mathcal{P}_m(\mathbf{R})$, and let D be the differentiation operator. Then $D1 = 0$ and $Dx^k = kx^{k-1}$ for $k = 1, \ldots, m$. Therefore, the matrix of D with respect to the standard basis has entries $1, 2, \ldots, m$ directly above the diagonal and 0 elsewhere. ■

Exercise 8A25

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. By 8.18(b), the minimal polynomial of T is $z^m = (z - 0)^m$ for some positive integer m. By 6.37, the matrix of T is upper triangular with respect to an orthonormal basis of V. By 5.41, the diagonal entries of the matrix are the eigenvalues of T, hence all 0 by 8.17(a). ■

8B Generalized Eigenspace Decomposition

Exercise 8B1

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. If $T(w,z) = \lambda(w,z)$ for $\lambda \in \mathbf{C}$ and nonzero $(w,z) \in \mathbf{C}^2$, then $-z = \lambda w$ and $w = \lambda z$. Together, these equations imply $(\lambda^2+1)z = 0$. Since $z \neq 0$, for otherwise $w = 0$, we must have $\lambda^2 + 1 = 0$, whence $\lambda = \pm i$. If $\lambda = i$, then $-z = iw$, so $E(i,T) = \operatorname{span}((i,1))$. A similar argument shows that $E(-i,T) = \operatorname{span}((-i,1))$. Since $E(\pm i,T) \subseteq G(\pm i,T)$, each generalized eigenspace must have dimension at least 1. Moreover, the multiplicities must sum to 2 by 8.25, so each generalized eigenspace has dimension exactly 1. By 2.39,

$$G(i,T) = \operatorname{span}((i,1)), \quad G(-i,T) = \operatorname{span}((-i,1)).$$

■

Exercise 8B2

Suppose $T \in \mathcal{L}(V)$ is invertible. Prove that $G(\lambda,T) = G\left(\frac{1}{\lambda}, T^{-1}\right)$ for every $\lambda \in \mathbf{F}$ with $\lambda \neq 0$.

Solution. Observe that $T^{-1} - \frac{1}{\lambda}I = -\frac{1}{\lambda}T^{-1}(T - \lambda I)$. Moreover, $T^{-1}(T-\lambda I) = (T-\lambda I)T^{-1} = I - \lambda T^{-1}$. Then for all positive integers k,

$$\left(T^{-1} - \frac{1}{\lambda}I\right)^k = \left(-\frac{1}{\lambda}T^{-1}(T-\lambda I)\right)^k = \left(-\frac{1}{\lambda}\right)^k T^{-k}(T-\lambda I)^k.$$

Since $\lambda \neq 0$, the operator $(-\frac{1}{\lambda})^k T^{-k}$ is injective, and therefore $\operatorname{null}(T^{-1} - \frac{1}{\lambda}I)^k = \operatorname{null}(T-\lambda I)^k$. By 8.19, the result follows. ■

Exercise 8B3

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. Note that

$$(S^{-1}TS - \lambda I)^k = (S^{-1}(T-\lambda I)S)^k = S^{-1}(T-\lambda I)^k S.$$

Because S^{-1} is injective, $S^{-1}(T-\lambda I)^k Sv = 0$ if and only if $(T-\lambda I)^k Sv = 0$, and therefore $\operatorname{null}(S^{-1}TS - \lambda I)^k = S^{-1}(\operatorname{null}(T-\lambda I)^k)$ for all positive integers k. For $k = 1$, we have $E(\lambda, S^{-1}TS) = S^{-1}(E(\lambda,T))$. Since S^{-1} is injective, $E(\lambda,S^{-1}TS) \neq \{0\}$ if and only if $E(\lambda,T) \neq \{0\}$, so T and $S^{-1}TS$ have the same eigenvalues. Taking $k = \dim V$ and using 8.20 yields $G(\lambda,S^{-1}TS) = S^{-1}(G(\lambda,T))$. Note that $S^{-1}|_{G(\lambda,T)} \in \mathcal{L}(G(\lambda,T),V)$ is an injective linear map whose range is $G(\lambda,S^{-1}TS)$. By the fundamental theorem of linear maps,

$$\dim G(\lambda,T) = \dim \operatorname{null} S^{-1}|_{G(\lambda,T)} + \dim \operatorname{range} S^{-1}|_{G(\lambda,T)} = \dim G(\lambda,S^{-1}TS),$$

so each eigenvalue has the same multiplicity for T as for $S^{-1}TS$. ■

Exercise 8B4

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 $n = \dim V$. Because $\operatorname{null} T^{n-2} \neq \operatorname{null} T^{n-1}$, the contrapositive of 8.2 gives $\operatorname{null} T^k \neq \operatorname{null} T^{k+1}$ for all $k \leq n-2$. Then all the inclusions in the chain up to $\operatorname{null} T^{n-1}$ in 8.1 must be strict:

$$\{0\} = \operatorname{null} T^0 \subsetneq \operatorname{null} T^1 \subsetneq \cdots \subsetneq \operatorname{null} T^{n-1}.$$

Consequently, the dimension of each successive null space increases by at least one, so $\dim \operatorname{null} T^{n-1} \geq n-1$. Because $\operatorname{null} T \neq \{0\}$, 5.7 gives that 0 is an eigenvalue of T. By 8.1 and 8.20, $\operatorname{null} T^{n-1} \subseteq \operatorname{null} T^n = G(0,T)$, so $\dim G(0,T) \geq \dim \operatorname{null} T^{n-1} \geq n-1$. Let $0, \lambda_2, \ldots, \lambda_m$ be the distinct eigenvalues of T. By 8.12 and 1.45, $G(0,T)+G(\lambda_2,T)+\cdots+G(\lambda_m,T)$ is a direct sum, so by 3.94 and 2.37,

$$\dim G(0,T) + \dim G(\lambda_2,T) + \cdots + \dim G(\lambda_m,T) \leq \dim V = n.$$

Since each dimension must be at least 1, $(n-1)+(m-1) \leq n$, so $m \leq 2$. ∎

Exercise 8B5

Suppose $T \in \mathcal{L}(V)$ and 3 and 8 are eigenvalues of T. Let $n = \dim V$. Prove that $V = (\operatorname{null} T^{n-2}) \oplus (\operatorname{range} T^{n-2})$.
Manual Remark. *In fact, if T has m distinct nonzero eigenvalues, then $V = \operatorname{null} T^{n-m} \oplus \operatorname{range} T^{n-m}$.*

Solution. By 5.12, $n \geq 2$ since T has at least two distinct eigenvalues, 3 and 8. We claim that $\operatorname{null} T^{n-2} = \operatorname{null} T^{n-1}$. If 0 is not an eigenvalue of T, then 5.7 yields $\operatorname{null} T^0 = \operatorname{null} T^1 = \{0\}$, so 8.2 gives $\operatorname{null} T^{n-2} = \operatorname{null} T^{n-1}$. If 0 is an eigenvalue of T, then $0, 3, 8$ are three distinct eigenvalues, and the contrapositive of Exercise 8B4 gives $\operatorname{null} T^{n-2} = \operatorname{null} T^{n-1}$. In either case, we have $\operatorname{null} T^{n-2} = \operatorname{null} T^n$ by 8.2. Applying Exercise 8A9 with $m = n-2$ yields $\operatorname{range} T^{n-2} = \operatorname{range} T^{n-1}$, so $\operatorname{range} T^{n-2} = \operatorname{range} T^n$ by Exercise 8A7. Therefore, 8.4 gives

$$V = \operatorname{null} T^n \oplus \operatorname{range} T^n = \operatorname{null} T^{n-2} \oplus \operatorname{range} T^{n-2}.$$

∎

Exercise 8B6

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 $n = \dim V$. Let p be the minimal polynomial of T. By 5.27(a), λ is a zero of p, so applying 4.6 repeatedly gives $p(z) = (z-\lambda)^k q(z)$ for some positive integer k and some $q \in \mathcal{P}(\mathbf{F})$ with $q(\lambda) \neq 0$. Let $G = G(\lambda,T)$ and $U = \operatorname{range}(T-\lambda I)^n$. By 8.4, $V = G \oplus U$, and both G and U are invariant under T by 5.18. Because $(T-\lambda I)^n|_G = 0$ by 8.20, there exists a smallest positive integer m such that $(T-\lambda I)^m|_G = 0$. We claim that $m = k$.
$\boxed{\leq}$ By 5.29, the minimal polynomial r of $T|_G$ divides $(z-\lambda)^m$, so by 4.13 if $\mathbf{F} = \mathbf{C}$ or 4.16 if $\mathbf{F} = \mathbf{R}$, we have $r(z) = (z-\lambda)^{m'}$ for some $m' \leq m$. Because λ is an eigenvalue of T, we have $\{0\} \neq E(\lambda,T) \subseteq G$. It follows that $r \neq 1$, so $m' \geq 1$. The minimality of m gives $m' = m$, so the minimal polynomial of $T|_G$ is $(z-\lambda)^m$. By 5.31, r also divides p. If $m > k$, then $(z-\lambda)^{m-k}$ would divide q, contradicting $q(\lambda) \neq 0$. Therefore, $m \leq k$.
$\boxed{\geq}$ Note that $\operatorname{null}(T-\lambda I) \subseteq G$ and $G \cap U = \{0\}$, so $(T-\lambda I)|_U$ is injective. For any $u \in U$, we have $(T-\lambda I)^k q(T)u = 0$, where $q(T)u \in U$. By injectivity, this implies $q(T)u = 0$. Let $s(z) = (z-\lambda)^m q(z)$. Then $s(T)u = (T-\lambda I)^m q(T)u = (T-\lambda I)^m 0 = 0$. On G, we have $(T-\lambda I)^m|_G = 0$, so $s(T)|_G = q(T)(T-\lambda I)^m|_G = 0$ as well. Therefore, $s(T) = 0$ on all of V. By 5.29, p divides s, so $m \geq k$, whence $m = k$. ∎

Exercise 8B7

Suppose $T \in \mathcal{L}(V)$ and λ is an eigenvalue of T with multiplicity d. Prove that $G(\lambda,T) = \operatorname{null}(T-\lambda I)^d$.
Axler Note. *If $d < \dim V$, then this exercise improves 8.20.*

Solution. Let $n = \dim V$ and let $G = G(\lambda, T)$. Define $N = (T - \lambda I)|_G$. By 8.23, $\dim G = d$.
$\boxed{\subseteq}$ By 8.20, $G = \operatorname{null}(T - \lambda I)^n$, which is invariant under T by 5.18 applied to $(z - \lambda)^n$, and $N^n = 0$. Thus, $N \in \mathcal{L}(G)$, so $N^d = 0$ by 8.16. Hence, $G \subseteq \operatorname{null}(T - \lambda I)^d$.
$\boxed{\supseteq}$ Since λ is an eigenvalue of T, we have $\{0\} \neq E(\lambda, T) \subseteq G$, so $d \geq 1$. Hence, $\operatorname{null}(T - \lambda I)^d \subseteq G$ by 8.19. Therefore, $G = \operatorname{null}(T - \lambda I)^d$. ∎

Exercise 8B8

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 $k_1, \ldots, k_m$.
Axler Note. *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.
$\boxed{\Rightarrow}$ By 3.94 and 2.42, taking bases of all the $G(\lambda_j, T)$ gives a basis of V consisting of generalized eigenvectors of T. By Exercise 8A11, the minimal polynomial p of T is $p(z) = (z - \mu_1) \cdots (z - \mu_M)$ for $\mu_1, \ldots, \mu_M \in \mathbf{F}$. By 5.27(a), $\{\mu_1, \ldots, \mu_M\} = \{\lambda_1, \ldots, \lambda_m\}$, so each λ_j appears with some positive exponent k_j. Grouping together the factors corresponding to the same λ_j gives the desired form $(z - \lambda_1)^{k_1} \cdots (z - \lambda_m)^{k_m}$.
$\boxed{\Leftarrow}$ By Exercise 8A11, V has a basis of generalized eigenvectors of T. Since each generalized eigenvector belongs to some $G(\lambda_j, T)$, the span of the union of the $G(\lambda_j, T)$ equals V, which gives

$$V = \sum_{j=1}^{m} G(\lambda_j, T).$$

By 8.12 and 1.45, this sum is direct. ∎

Exercise 8B9

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 $n = \dim V$ and let $\lambda_1, \ldots, \lambda_m$ be the distinct eigenvalues of T. By the generalized eigenspace decomposition, $V = G_1 \oplus \cdots \oplus G_m$, where $G_j = G(\lambda_j, T)$. Define D by $D(v_1 + \cdots + v_m) = \lambda_1 v_1 + \cdots + \lambda_m v_m$ for $v_k \in G_k$. Since the sum is direct, D is well-defined and linear. Define $N = T - D$. Since concatenating bases of all the G_j gives a basis of eigenvectors of D by 3.94 and 2.42, it follows that D is diagonalizable by 5.55.

Suppose $u \in G_k$ for some k. By 8.22(a), G_k is invariant under T. Then $Nu = (T - D)u = Tu - \lambda_k u \in G_k$. It follows that $N^n u = (T - \lambda_k I)^n u = 0$ by 8.20, so $N^n = 0$ on each G_k. Moreover, $Tu \in G_k$, so $DTu = \lambda_k Tu = T(\lambda_k u) = TDu$, which gives $DT = TD$ on each G_k. This gives $DN = D(T - D) = DT - D^2 = TD - D^2 = (T - D)D = ND$ on each G_k. Since V is the direct sum of the G_k, it follows that N is nilpotent and $DN = ND$ on V. ∎

Exercise 8B10

Suppose V is a complex inner product space, $e_1, \dots, e_n$ is an orthonormal basis of V, and $T \in \mathcal{L}(V)$. Let $\lambda_1, \dots, \lambda_n$ be the eigenvalues of T, each included as many times as its multiplicity. Prove that

$$|\lambda_1|^2 + \cdots + |\lambda_n|^2 \le \|Te_1\|^2 + \cdots + \|Te_n\|^2.$$

Axler Note. *See the comment after* Exercise 7A5.

Solution. By Schur's theorem, there exists an orthonormal basis $f_1, \dots, f_n$ of V such that the matrix A of T with respect to this basis is upper triangular. By 5.41, 8.31, and 8.25, the diagonal entries $A_{1,1}, \dots, A_{n,n}$ are the eigenvalues $\lambda_1, \dots, \lambda_n$ of T rearranged in some order. Observe that $Tf_k = \sum_{j \le k} A_{j,k} f_j$ for each k, so

$$\|Tf_k\|^2 = \sum_{j \le k} |A_{j,k}|^2 \ge |A_{k,k}|^2$$

by 6.24. By Exercise 7A5 applied to both bases $e_1, \dots, e_n$ and $f_1, \dots, f_n$, we have

$$\sum_{m=1}^{n} \|Te_m\|^2 = \sum_{m=1}^{n} \|Tf_m\|^2 \ge \sum_{m=1}^{n} |A_{m,m}|^2 = \sum_{m=1}^{n} |\lambda_m|^2. \qquad \blacksquare$$

Exercise 8B11

Give an example of an operator on $\mathbf{C}^4$ whose characteristic polynomial equals $(z-7)^2(z-8)^2$.

Solution. Define $T \in \mathcal{L}(\mathbf{C}^4)$ by $T(z_1, z_2, z_3, z_4) = (7z_1, 7z_2, 8z_3, 8z_4)$. Then $\mathcal{M}(T)$ with respect to the standard basis is diagonal with diagonal entries $7, 7, 8, 8$. By 5.41, the eigenvalues of T are $7, 8$, each with multiplicity 2 by 8.31. Hence, the characteristic polynomial of T is $(z-7)^2(z-8)^2$. $\blacksquare$

Exercise 8B12

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 e_1, e_2, e_3, e_4 be the standard basis of $\mathbf{C}^4$. Consider the operator $T \in \mathcal{L}(\mathbf{C}^4)$ defined by $T(z_1, z_2, z_3, z_4) = (z_1, 5z_2 + z_3, 5z_3, 5z_4)$. Then the matrix of T with respect to the standard basis is

$$\mathcal{M}(T) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 5 & 1 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 5 \end{pmatrix}.$$

By 5.41, the eigenvalues of T are $1, 5$, with 1 having multiplicity 1 and 5 having multiplicity 3 by 8.31. Therefore, the characteristic polynomial of T is $(z-1)(z-5)^3$. Let p be the minimal polynomial of T. By 8.30, 5.27(a), and uniqueness of the factorization of the minimal polynomial, we have $p(z) = (z-1)(z-5)^b$ for some $b \in \{1, 2, 3\}$. Observe that $(T-I)(T-5I)e_3 = 4e_2 \ne 0$. However, $(T-5I)^2 e_1 = 16e_1$ and $(T-5I)^2 e_k = 0$ for $k = 2, 3, 4$. Moreover, $(T-I)e_1 = 0$, so $(T-I)(T-5I)^2 = 0$. Hence, $b = 2$ and the minimal polynomial of T is $p(z) = (z-1)(z-5)^2$. $\blacksquare$

Exercise 8B13

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 e_1, e_2, e_3, e_4 be the standard basis of $\mathbf{C}^4$. Consider the operator $T \in \mathcal{L}(\mathbf{C}^4)$ defined by $T(z_1, z_2, z_3, z_4) = (0, z_2 + z_3, z_3, 3z_4)$. Then the matrix of T with respect to the standard basis is

$$\mathcal{M}(T) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

By 5.41, the eigenvalues of T are $0, 1, 3$, with 0 and 3 having multiplicity 1 and 1 having multiplicity 2 by 8.31. Therefore, the characteristic polynomial of T is $z(z-1)^2(z-3)$. Let p be the minimal polynomial of T. By 8.30, 5.27(a), and uniqueness of the factorization of the minimal polynomial, we have $p(z) = z(z-1)^b(z-3)$ for some $b \in \{1, 2\}$. Since $T(T-I)(T-3I)e_3 = -2e_2 \neq 0$, we have $b = 2$, so $p(z) = z(z-1)^2(z-3)$. ∎

Exercise 8B14

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 defined by $T(z_1, z_2, z_3, z_4) = (0, z_2, z_3, 3z_4)$. Then $\mathcal{M}(T)$ with respect to the standard basis is diagonal with diagonal entries $0, 1, 1, 3$. By 5.41, the eigenvalues of T are $0, 1, 3$, with 0 and 3 having multiplicity 1 and 1 having multiplicity 2 by 8.31. Therefore, the characteristic polynomial of T is $z(z-1)^2(z-3)$. Since $\mathcal{M}(T)$ is diagonal, T is diagonalizable, so the minimal polynomial of T has distinct linear factors by 5.62, and hence the minimal polynomial of T is $z(z-1)(z-3)$ by 5.27(a). ∎

Exercise 8B15

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 $\mathbf{C}^4$. Observe that $Te_k = e_{k+1}$ for $k = 1, 2, 3$ and $Te_4 = 0$, so $T^4 = 0$, and hence T is nilpotent. By 8.17(a), 0 is the only eigenvalue of T, and $\dim \operatorname{null} T^4 = \dim \mathbf{C}^4 = 4$, so the multiplicity of 0 is 4 by 8.23, and hence the characteristic polynomial of T is z^4. Moreover, 8.18(b) implies that the minimal polynomial of T is z^m for some positive integer m. By 5.29, $m \leq 4$, and since $T^3 e_1 = e_4 \neq 0$, we must have $m \geq 4$, so $m = 4$, and hence the minimal polynomial of T is z^4. ∎

Exercise 8B16

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 $e_1, \ldots, e_6$ be the standard basis of $\mathbf{C}^6$. Observe that $Te_k = e_{k+1}$ for $k = 1, 2, 4$ and $Te_j = 0$ for $j = 3, 5, 6$. Hence, $T^3 = 0$, and T is nilpotent. By 8.17(a), 0 is the only eigenvalue of T, and $\dim \operatorname{null} T^6 = \dim \mathbf{C}^6 = 6$, so the multiplicity of 0 is 6 by 8.23, and hence the characteristic polynomial of T is z^6. Moreover, 8.18(b) implies that the minimal polynomial of T is z^m for some positive integer m. By 5.29, $m \leq 3$, and since $T^2 e_1 = e_3 \neq 0$, we must have $m \geq 3$, so $m = 3$, and hence the minimal polynomial of T is z^3. ∎

Exercise 8B17

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. Let $d = \dim V$. By induction, $P^d = P$, so $V = \operatorname{null} P^d \oplus \operatorname{range} P^d = \operatorname{null} P \oplus \operatorname{range} P$ by 8.4. Choose a basis of V consisting of a basis of $\operatorname{null} P$ followed by a basis of $\operatorname{range} P$. For any $u \in \operatorname{range} P$, there exists some $w \in V$ such that $Pw = u$. Then $Pu = P(Pw) = P^2w = Pw = u$, and hence the matrix of P with respect to this basis is diagonal with m zeros followed by n ones. By 5.41, the eigenvalues of P are the diagonal entries, and the multiplicity of each eigenvalue is the number of times it appears on the diagonal by 8.31. Hence, the characteristic polynomial of P is $z^m(z-1)^n$. ∎

Exercise 8B18

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. Let $n = \dim V$. Define $S = T - \lambda I$. By 8.20, $G = G(\lambda, T) = \operatorname{null} S^n$. Let the integers in each part be denoted by a, b, c, d respectively. The integer a exists because of 5.27(a); b exists because $S^n|_G = 0$; c exists by 8.3; d exists by Exercise 8A9.

By Exercise 8B6, $a = b$. Applying Exercise 8A9 to S gives $c = d$. We close the proof by showing $b = c$.

$\boxed{\leq}$ By 8.1 and 8.2, $G \subseteq \operatorname{null} S^n \subseteq \operatorname{null} S^{c+n} = \operatorname{null} S^c$, so $b \leq c$.

$\boxed{\geq}$ By 8.1, $G \subseteq \operatorname{null} S^b \subseteq \operatorname{null} S^{b+1} \subseteq G$, so $c \leq b$. Therefore, $b = c$, and hence all four numbers are equal. ∎

Exercise 8B19

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. Let $\lambda_1, \ldots, \lambda_k$ be the distinct eigenvalues of S and let $m_j = \dim G(\lambda_j, S)$ for each j. The characteristic polynomial of S has constant term $(-\lambda_1)^{m_1} \cdots (-\lambda_k)^{m_k}$. Since S is unitary, $|\lambda_j| = 1$ for each j by 7.54, so the absolute value of the constant term is $|(-\lambda_1)^{m_1} \cdots (-\lambda_k)^{m_k}| = 1$. ∎

Exercise 8B20 ✔

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. Let q be the characteristic polynomial of T. For each k, let $n_k = \dim V_k$. Since each V_k is nonzero and complex, 5.47 yields a basis $v_{k,1}, \dots, v_{k,n_k}$ of V_k with respect to which $\mathcal{M}(T|_{V_k})$ is upper triangular. For each V_k, denote the diagonal entries of $\mathcal{M}(T|_{V_k})$ by $d_{k,j}$ for $j = 1, \dots, n_k$. Let $\mathcal{B}$ be the list obtained by concatenating the bases of the V_k. By 3.94, the sum of n_k over all k equals $\dim V$. Additionally, $\mathcal{B}$ spans V, so it is a basis of V by 2.42. Because $\mathcal{M}(T|_{V_k})$ is upper triangular for each k, 5.39(c) yields $Tv_{k,j} \in \operatorname{span}(v_{k,1}, \dots, v_{k,j})$ for each $j = 1, \dots, n_k$. Moreover, $\operatorname{span}(v_{k,1}, \dots, v_{k,j}) \subseteq \operatorname{span}(v_{1,1}, \dots, v_{1,n_1}, \dots, v_{k,1}, \dots, v_{k,j})$, so $\mathcal{M}(T, \mathcal{B})$ is upper triangular by 5.39(c), where the diagonal entries are precisely the $d_{k,j}$. By 5.41, 8.31, and 8.26 applied to T and $T|_{V_k}$ for each k, we obtain

$$q = \prod_{k=1}^{m} \prod_{j=1}^{n_k} (z - d_{k,j}) = \prod_{k=1}^{m} p_k. \qquad \blacksquare$$

Exercise 8B21

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.
Axler Note. *This exercise implies that every monic polynomial is the characteristic polynomial of some operator.*

Solution. Since $\mathbf{C}^{\deg q} \neq \{0\}$, we have $\deg q \geq 1$, so q has a zero by 4.12. Since p and q have the same zeros, p also has a zero. Write $q = pr$ for some $r \in \mathcal{P}(\mathbf{C})$, and factor r by the fundamental theorem of algebra. By uniqueness of factorization in $\mathcal{P}(\mathbf{C})$, we can write

$$q(z) = \prod_{k=1}^{n} (z - \lambda_k)^{m_k}, \quad p(z) = \prod_{k=1}^{n} (z - \lambda_k)^{\ell_k},$$

where $\lambda_1, \dots, \lambda_n$ are the distinct zeros of q. Moreover, all m_k and ℓ_k are positive integers since p and q have zeros at λ_k, and we have $\ell_k \leq m_k$ for each k since q is a polynomial multiple of p. Note that $\deg q = \sum_{k=1}^{n} m_k$. Let $e_1, \dots, e_{\deg q}$ be the standard basis of $\mathbf{C}^{\deg q}$, and split it into blocks of sizes $m_1, \dots, m_n$ corresponding to the distinct zeros $\lambda_1, \dots, \lambda_n$ of q. For each λ_k, let $v_{k,1}, \dots, v_{k,m_k}$ be the ordered subset of the standard basis of $\mathbf{C}^{\deg q}$ as described above. By the linear map lemma, there exists a unique linear map $T \in \mathcal{L}(\mathbf{C}^{\deg q})$ such that

$$Tv_{k,j} = \begin{cases} \lambda_k v_{k,j} + v_{k,j-1}, & \text{for } j = 2, \dots, \ell_k, \\ \lambda_k v_{k,j}, & \text{for } j = 1 \text{ and } j = \ell_k + 1, \dots, m_k. \end{cases}$$

Observe that the matrix of T with respect to the standard basis is upper triangular, where λ_k appears on the diagonal m_k times for each k. Therefore, the characteristic polynomial of T is q by 5.41 and 8.31.

Consider the subspaces $V_k = \operatorname{span}(v_{k,1}, \dots, v_{k,m_k})$ for each $k = 1, \dots, n$. Then $(T - \lambda_k I)^{\ell_k} v_{k,j} = 0$ for all $j = 1, \dots, m_k$. Therefore, $(T - \lambda_k I)^{\ell_k}|_{V_k} = 0$, but $(T - \lambda_k I)^{\ell_k - 1} v_{k,\ell_k} = v_{k,1} \neq 0$. Since $V_k \subseteq G(\lambda_k, T)$, and $\dim V_k = \dim G(\lambda_k, T) = m_k$ by 8.31 and 8.23, we have $V_k = G(\lambda_k, T)$. By 5.27(b), the minimal polynomial of T is a product of the factors $(z - \lambda_k)$, and by Exercise 8B6, the exponents of these factors are exactly ℓ_k. Therefore, the minimal polynomial of T is p. $\blacksquare$

Exercise 8B22

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

$$AB = \begin{pmatrix} A_1B_1 & & 0 \\ & \ddots & \\ 0 & & A_mB_m \end{pmatrix}.$$

Solution. Let A_k and B_k be n_k-by-n_k matrices for each $k = 1, \dots, m$. Let $N_k = n_1 + \cdots + n_{k-1}$ for each $k = 1, \dots, m$, with the convention that $N_1 = 0$. Then the kth block matrix has indices i, j with $N_k + 1 \le i, j \le N_k + n_k$. Fix ℓ and s with $1 \le s \le n_\ell$, and consider the $(N_\ell + s)$th column of B. The $(N_\ell + s)$th column of B has zeros outside the ℓth block and equals the sth column of B_ℓ within the ℓth block. By 3.48 and 3.50, the $(N_\ell + s)$th column of AB is

$$\sum_{t=1}^{n_\ell} (B_\ell)_{t,s} A_{\cdot, N_\ell + t}.$$

Since A is block diagonal, every column $A_{\cdot, N_\ell + t}$ has zeros outside the rows $N_\ell + 1 \le i \le N_\ell + n_\ell$, and equals $(A_\ell)_{\cdot, t}$ within those rows. Then the $(N_\ell + s)$th column of AB is zero outside the ℓth block and equals

$$\sum_{t=1}^{n_\ell} (B_\ell)_{t,s} (A_\ell)_{\cdot, t} = A_\ell (B_\ell)_{\cdot, s} = (A_\ell B_\ell)_{\cdot, s}$$

by 3.48 and 3.50. Since this holds for any column of any block, we conclude that AB is block diagonal with blocks $A_k B_k$. ∎

Exercise 8B23

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 the 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 5A17) to show that if $\dim V$ is an odd number, then T has an eigenvalue (thus giving an alternative proof of 5.34).

Axler Note. *See Exercise 3B33 for the definition of the complexification $T_{\mathbf{C}}$.*

Solution. Define $C : V_{\mathbf{C}} \to V_{\mathbf{C}}$ by $C(u + iv) = u - iv$ for $u, v \in V$. We claim that C is additive, conjugate linear, satisfies $C^2 = I$, and commutes with $T_{\mathbf{C}}$. To see additivity, let $u_1 + iv_1, u_2 + iv_2 \in V_{\mathbf{C}}$. Then

$$\begin{aligned} C((u_1 + iv_1) + (u_2 + iv_2)) &= C((u_1 + u_2) + i(v_1 + v_2)) \\ &= (u_1 + u_2) - i(v_1 + v_2) \\ &= (u_1 - iv_1) + (u_2 - iv_2) \\ &= C(u_1 + iv_1) + C(u_2 + iv_2). \end{aligned}$$

To see conjugate linearity, let $\alpha = a + bi \in \mathbf{C}$ and $u + iv \in V_{\mathbf{C}}$. Then

$$\begin{aligned} C(\alpha(u+iv)) &= C((a+bi)(u+iv)) \\ &= C((au-bv)+i(av+bu)) \\ &= (au-bv)-i(av+bu) \\ &= (a-bi)(u-iv) \\ &= \overline{\alpha}C(u+iv). \end{aligned}$$

Moreover, $C^2(u+iv) = C(u-iv) = u+iv$, so $C^2 = I$. Finally,

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

so C commutes with $T_{\mathbf{C}}$.

(a) Since C commutes with $T_{\mathbf{C}}$ and $C(\lambda(u+iv)) = \overline{\lambda}C(u+iv)$ for any $u+iv \in V_{\mathbf{C}}$, we have $C(T_{\mathbf{C}} - \lambda I) = (T_{\mathbf{C}} - \overline{\lambda}I)C$. Then for any positive integer k, it follows by induction that $C(T_{\mathbf{C}} - \lambda I)^k = (T_{\mathbf{C}} - \overline{\lambda}I)^k C$. By definition, $u+iv \in G(\lambda, T_{\mathbf{C}})$ if and only if $(T_{\mathbf{C}} - \lambda I)^k(u+iv) = 0$ for some positive integer k. Moreover, $C^2 = I$ implies that C is injective, and additivity gives $C0 = 0$, so for any positive integer k, $(T_{\mathbf{C}} - \lambda I)^k(u+iv) = 0$ if and only if $(T_{\mathbf{C}} - \overline{\lambda}I)^k C(u+iv) = 0$. Since $C(u+iv) = u - iv$, we have $(T_{\mathbf{C}} - \overline{\lambda}I)^k(u-iv) = 0$ for some k if and only if $u - iv \in G(\overline{\lambda}, T_{\mathbf{C}})$.

(b) Suppose $v_1, \ldots, v_m$ is a basis of $G(\lambda, T_{\mathbf{C}})$. By part (a), $Cv_j \in G(\overline{\lambda}, T_{\mathbf{C}})$ for each $j = 1, \ldots, m$. We claim that $Cv_1, \ldots, Cv_m$ is a basis of $G(\overline{\lambda}, T_{\mathbf{C}})$. To see this, suppose there are scalars $\alpha_1, \ldots, \alpha_m \in \mathbf{C}$ such that

$$\sum_{j=1}^{m} \alpha_j Cv_j = 0.$$

Applying C to both sides gives

$$C\left(\sum_{j=1}^{m} \alpha_j Cv_j\right) = \sum_{j=1}^{m} \overline{\alpha_j} C^2 v_j = \sum_{j=1}^{m} \overline{\alpha_j} v_j = C0 = 0.$$

By the linear independence of $v_1, \ldots, v_m$, we must have $\overline{\alpha_1} = \cdots = \overline{\alpha_m} = 0$, hence $\alpha_1 = \cdots = \alpha_m = 0$. Moreover, for any $w \in G(\overline{\lambda}, T_{\mathbf{C}})$, we have $Cw \in G(\lambda, T_{\mathbf{C}})$ by part (a), so $Cw = \sum_{j=1}^{m} \beta_j v_j$ for some $\beta_j \in \mathbf{C}$, hence

$$w = C(Cw) = C\left(\sum_{j=1}^{m} \beta_j v_j\right) = \sum_{j=1}^{m} \overline{\beta_j} Cv_j.$$

Therefore, $Cv_1, \ldots, Cv_m$ is a basis of $G(\overline{\lambda}, T_{\mathbf{C}})$, whence $\dim G(\overline{\lambda}, T_{\mathbf{C}}) = \dim G(\lambda, T_{\mathbf{C}})$. Thus, λ is an eigenvalue of $T_{\mathbf{C}}$ if and only if $\overline{\lambda}$ is an eigenvalue of $T_{\mathbf{C}}$.

(c) By Exercise 2B11, $\dim V_{\mathbf{C}} = \dim V$. Then $V_{\mathbf{C}}$ is nonzero with odd dimension, so we may apply 8.25 to $T_{\mathbf{C}}$ to conclude that $\dim V_{\mathbf{C}} = \sum_{\mu} d_\mu$, where μ ranges over the eigenvalues of $T_{\mathbf{C}}$ and d_μ is the multiplicity of μ. Note that for nonreal eigenvalues μ, we have $\mu \neq \overline{\mu}$. By part (b), $d_\mu = d_{\overline{\mu}}$ for each nonreal eigenvalue μ, so the sum of the multiplicities of nonreal eigenvalues is even. Since $\dim V_{\mathbf{C}}$ is odd, there must be at least one real eigenvalue of $T_{\mathbf{C}}$.

(d) By (c), $T_{\mathbf{C}}$ has a real eigenvalue $\nu \in \mathbf{R}$. By Exercise 5A17, ν is an eigenvalue of T. ■

8C Consequences of Generalized Eigenspace Decomposition

Manual Remark. *The terms* **superdiagonal**, **second superdiagonal**, *and* **subdiagonal** *refer to the entries immediately above, two above, and immediately below the main diagonal of a matrix, respectively. For* 1*-by-*1 *blocks, all three diagonals are empty, while for* 2*-by-*2 *blocks, the second superdiagonal is empty, so claims involving the entries of these empty diagonals hold vacuously.*

Exercise 8C1

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. Note that $T^2(z_1, z_2, z_3) = (z_3, 0, 0)$ and $T^3 = 0$. Suppose $S \in \mathcal{L}(\mathbf{C}^3)$ is such that $S^2 = T$. Then $S^6 = (S^2)^3 = T^3 = 0$, so S is nilpotent. By 8.16, $S^3 = 0$, but $T^2 = S^4 = SS^3 = 0$, which contradicts $T^2 \neq 0$. Therefore, T does not have a square root. ∎

Exercise 8C2

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) With respect to the standard basis of $\mathbf{F}^5$, the matrix $\mathcal{M}(T)$ of T on and below the diagonal is 0, so T is nilpotent by 8.18.

(b) By part (a) and 8.16, $T^5 = 0$. As in the proof of 8.39, we may construct the polynomial

$$p(z) = 1 + \frac{1}{2}z - \frac{1}{8}z^2 + \frac{1}{16}z^3 - \frac{5}{128}z^4.$$

Squaring $p(z)$ gives

$$p^2(z) = 1 + z + z^5 q(z)$$

for some $q \in \mathcal{P}(\mathbf{F})$. Let $S = p(T)$, so $S^2 = I + T + T^5 q(T) = I + T$ by 5.17(a) and $T^5 = 0$. ∎

Exercise 8C3

Suppose V is a complex vector space. Prove that every invertible operator on V has a cube root.

Manual Remark. *The coefficients in the recursion below are given by* $a_j = \binom{1/3}{j}$ *for* $j \geq 1$.

Solution. Let $T \in \mathcal{L}(V)$ be invertible with distinct eigenvalues $\lambda_1, \ldots, \lambda_m$ and let $n = \dim V$. By 5.7, $\lambda_k \neq 0$ for each $k = 1, \ldots, m$. Let $G_k = G(\lambda_k, T)$ for each $k = 1, \ldots, m$. Then the generalized eigenspace decomposition of V is

$$V = \bigoplus_{k=1}^{m} G_k.$$

By 8.22(a), we may define $N_k = (T - \lambda_k I)|_{G_k} \in \mathcal{L}(G_k)$ for each $k = 1, \ldots, m$. Since N_k is nilpotent on G_k by 8.22(b), and $\dim G_k \leq n$, we have $N_k^n = 0$ by 8.16. As in the proof of 8.39 with a cubic in place of a square, set $a_1 = \frac{1}{3}$. For each $2 \leq j \leq n-1$, the coefficient of z^j in the expansion $(1 + a_1 z + \cdots + a_j z^j)^3$ is $3a_j + f_j(a_1, \ldots, a_{j-1})$ for some polynomial f_j in $a_1, \ldots, a_{j-1}$. Choosing a_j such that $3a_j + f_j(a_1, \ldots, a_{j-1}) = 0$, we can recursively determine $a_2, \ldots, a_{n-1}$. We then obtain the polynomial

$$p(z) = 1 + a_1 z + \cdots + a_{n-1} z^{n-1}$$

such that $p^3(z) = 1 + z + z^n r(z)$ for some $r \in \mathcal{P}(\mathbf{F})$. By the fundamental theorem of algebra, the polynomial $z^3 - \lambda_k$ has a zero in $\mathbf{C}$ for each $k = 1, \dots, m$, say μ_k. Define the operators $S_k = \mu_k p(\lambda_k^{-1} N_k)$ on G_k for each k. By 5.17(a) and $N_k^n = 0$, it follows that $S_k^3 = (\mu_k p(\lambda_k^{-1} N_k))^3 = \lambda_k I + N_k = T|_{G_k}$. For each $v \in V$, by the generalized eigenspace decomposition, we may uniquely write $v = v_1 + \cdots + v_m$ with $v_k \in G_k$ for each $k = 1, \dots, m$. Define $S \in \mathcal{L}(V)$ by $Sv = S_1 v_1 + \cdots + S_m v_m$. Since $Sv_k = S_k v_k \in G_k$ for each k, we have $S^3 v_k = S_k^3 v_k = T|_{G_k} v_k = T v_k$ for each k, and thus $S^3 = T$ on all of V. ∎

Exercise 8C4

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.

$\boxed{\Rightarrow}$ Let $S \in \mathcal{L}(V)$ be such that $S^2 = -I$. Then $\operatorname{null}(S^2 + 0S + I) = \operatorname{null}(-I + I) = \operatorname{null}(0) = V$. Since $0^2 < 4$, 5.33 gives that $\dim V$ is even.

$\boxed{\Leftarrow}$ Let $\dim V = 2n$ for some positive integer n, and let V have a basis $v_1, \dots, v_{2n}$. By the linear map lemma, there exists a linear map $S \in \mathcal{L}(V)$ such that $Sv_{2k-1} = v_{2k}$ and $Sv_{2k} = -v_{2k-1}$ for each $k = 1, \dots, n$. Then $S^2 v_{2k-1} = Sv_{2k} = -v_{2k-1}$ and $S^2 v_{2k} = S(-v_{2k-1}) = -v_{2k}$ for each $k = 1, \dots, n$. Since S^2 and $-I$ agree on a basis of V, the uniqueness part of the linear map lemma implies that $S^2 = -I$. ∎

Exercise 8C5

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. Let $N = T - 2I$. Then $N(w, z) = (-3w - z, 9w + 3z) = (-3w - z)(1, -3)$, so

$$N^2(w, z) = (-3w - z)N(1, -3) = 0.$$

Let $v = (0, -1)$. Then $Nv = (1, -3) \neq 0$, so by Exercise 8A2, v, Nv is a linearly independent list of length 2, and thus is a basis of $\mathbf{C}^2$ by 2.38. We have $TNv = 2Nv$ and $Tv = Nv + 2v$, so

$$\mathcal{M}(T, (Nv, v)) = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}.$$

Therefore, the list $(1, -3), (0, -1)$ is a Jordan basis for T. ∎

Exercise 8C6

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. Define $v = \frac{x^4}{4!}$, so $D^4 v = 1$ and $D^5 v = 0$. By Exercise 8A2, the list $v, Dv, D^2 v, D^3 v, D^4 v$ is a linearly independent list of length 5 and thus forms a basis of $\mathcal{P}_4(\mathbf{R})$ by 2.38. Then

$$\mathcal{M}(D, (D^4 v, D^3 v, D^2 v, Dv, v)) = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

so the list $1, x, \frac{x^2}{2!}, \frac{x^3}{3!}, \frac{x^4}{4!}$ is a Jordan basis for D. ∎

Exercise 8C7

Suppose $T \in \mathcal{L}(V)$ is nilpotent and $v_1, \dots, v_n$ 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 $v_1, \dots, v_n$.

Solution. Let $A_1, \dots, A_p$ be the blocks as in 8.44, where each A_k is d_k-by-d_k. Let $v_{k,1}, \dots, v_{k,d_k}$ be the basis vectors of the block A_k for each $k = 1, \dots, p$. By 8.44,

$$Tv_{k,j} = \begin{cases} \lambda_k v_{k,1} & \text{if } j = 1, \\ \lambda_k v_{k,j} + v_{k,j-1} & \text{if } j > 1. \end{cases}$$

Note that $Tv_{k,1} = \lambda_k v_{k,1}$ with $v_{k,1} \neq 0$. Since T is nilpotent, 8.17(a) implies that $\lambda_k = 0$ for all k. Then $Tv_{k,1} = 0$ and $Tv_{k,j} = v_{k,j-1}$ for $j > 1$, so

$$T^{j-1} v_{k,j} = v_{k,1}, \quad T^j v_{k,j} = 0.$$

The entries above the diagonal between the blocks A_k and A_{k+1} are all 0, so the consecutive string of 1's cannot extend beyond a single block. Therefore, each block A_k contributes a consecutive string of 1's of length $d_k - 1$, so $m + 1 = \max_k d_k$. By 8.18, the minimal polynomial of T is z^r for some r. We now claim that $r = m + 1$.

$\boxed{\leq}$ Since each $v_{k,j}$ has $j \leq d_k \leq m + 1$, we have $T^{m+1} v_{k,j} = T^{m+1-j}(T^j v_{k,j}) = 0$. Therefore, $T^{m+1} = 0$, and 5.29 yields $r \leq m + 1$.

$\boxed{\geq}$ Consider some k such that $d_k = m + 1$. Then $T^m v_{k,m+1} = v_{k,1} \neq 0$, so $T^m \neq 0$. Therefore, $r \geq m + 1$ by 5.29. Combining this with the previous inequality gives $r = m + 1$. ∎

Exercise 8C8

Suppose $T \in \mathcal{L}(V)$ and $v_1, \dots, v_n$ is a basis of V that is a Jordan basis for T. Describe the matrix of T^2 with respect to this basis.

Solution. Let $A_1, \dots, A_p$ be the blocks of $\mathcal{M}(T)$ with respect to the basis $v_1, \dots, v_n$ and let λ_k be the eigenvalue corresponding to the block A_k. By 3.43 and Exercise 8B22, $\mathcal{M}(T^2) = \mathcal{M}(T)^2$ is block diagonal with blocks $A_1^2, \dots, A_p^2$. Write $A_k = \lambda_k I + N_k$, where N_k is the matrix with 1's on the superdiagonal and 0's elsewhere. Then $A_k^2 = \lambda_k^2 I + 2\lambda_k N_k + N_k^2$, and N_k^2 has 1's on the second superdiagonal and 0's elsewhere by 3.46. Therefore, each block A_k^2 has λ_k^2 on the diagonal, $2\lambda_k$ on the superdiagonal, 1's on the second superdiagonal, and 0's elsewhere. ∎

Exercise 8C9

Suppose $T \in \mathcal{L}(V)$ is nilpotent. Explain why there exist $v_1, \dots, v_n \in V$ and nonnegative integers $m_1, \dots, m_n$ such that (a) and (b) below both hold.

(a) $T^{m_1} v_1, \dots, Tv_1, v_1, \dots, T^{m_n} v_n, \dots, Tv_n, v_n$ is a basis of V.

(b) $T^{m_1+1} v_1 = \cdots = T^{m_n+1} v_n = 0$.

Solution. By 8.45, nilpotence of T gives a Jordan basis $u_1, \dots, u_{\dim V}$ for T. Let $A_1, \dots, A_n$ be the blocks of $\mathcal{M}(T, (u_1, \dots, u_{\dim V}))$ with diagonal entries $\lambda_1, \dots, \lambda_n$. Let $u_{k,1}, \dots, u_{k,d_k}$ be the basis vectors corresponding to the d_k-by-d_k block A_k. By 5.41 and 8.17(a), $\lambda_k = 0$, so $Tu_{k,1} = 0$ and $Tu_{k,j} = u_{k,j-1}$ for $2 \leq j \leq d_k$ by 8.44. Let $v_k = u_{k,d_k}$ and $m_k = d_k - 1$ for $1 \leq k \leq n$.

(a) is satisfied because $T^j v_k = u_{k,d_k-j}$ for $0 \leq j \leq m_k$, so the vectors $T^{m_k} v_k, \dots, Tv_k, v_k$ are exactly $u_{k,1}, \dots, u_{k,d_k}$ in order.

(b) is satisfied because $T^{m_k+1} v_k = T^{d_k} u_{k,d_k} = 0$ for each k, as $u_{k,1} \in \operatorname{null} T$. ∎

Exercise 8C10

Suppose $T \in \mathcal{L}(V)$ and $v_1, \dots, v_n$ is a basis of V that is a Jordan basis for T. Describe the matrix of T with respect to the basis $v_n, \dots, v_1$ obtained by reversing the order of the v's.

Solution. Let $A_1, \dots, A_p$ be the Jordan blocks of $\mathcal{M}(T, (v_1, \dots, v_n))$ with diagonal entries $\lambda_1, \dots, \lambda_p$. Let $v_{k,1}, \dots, v_{k,d_k}$ be the basis vectors among $v_1, \dots, v_n$ corresponding to A_k in order. By 8.44, $Tv_{k,1} = \lambda_k v_{k,1}$ and $Tv_{k,j} = v_{k,j-1} + \lambda_k v_{k,j}$ for $2 \le j \le d_k$. In the basis $v_n, \dots, v_1$, the basis vectors corresponding to A_k appear in the reverse order $v_{k,d_k}, \dots, v_{k,1}$, with the blocks appearing in the order $A_p, \dots, A_1$. Writing $w_\ell = v_{k,d_k+1-\ell}$ for $1 \le \ell \le d_k$, we have $Tw_\ell = \lambda_k w_\ell + w_{\ell+1}$ for $1 \le \ell \le d_k - 1$ and $Tw_{d_k} = \lambda_k w_{d_k}$. Therefore, $\mathcal{M}(T, (v_n, \dots, v_1))$ is block diagonal with the blocks $A_p^t, \dots, A_1^t$. Each block A_k^t has λ_k on its diagonal, 1's on its subdiagonal, and 0's elsewhere. ■

Exercise 8C11

Suppose $T \in \mathcal{L}(V)$. Explain why every vector in each Jordan basis for T is a generalized eigenvector of T.

Solution. Let $v_1, \dots, v_n$ be a Jordan basis for T. Let $A_1, \dots, A_p$ be the blocks of $\mathcal{M}(T, (v_1, \dots, v_n))$ with diagonal entries $\lambda_1, \dots, \lambda_p$. Let $v_{k,1}, \dots, v_{k,d_k}$ be the basis vectors corresponding to A_k in order. By 8.44, $(T - \lambda_k I)v_{k,1} = 0$ and $(T - \lambda_k I)v_{k,j} = v_{k,j-1}$ for $2 \le j \le d_k$. Hence, λ_k is an eigenvalue of T and $(T - \lambda_k I)^j v_{k,j} = 0$ for $1 \le j \le d_k$. Since $v_{k,j} \ne 0$, it follows that $v_{k,j}$ is a generalized eigenvector of T corresponding to λ_k. ■

Exercise 8C12

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 $v_1, \dots, v_n$ be a Jordan basis for T. Let $A_1, \dots, A_p$ be the blocks of $\mathcal{M}(T, (v_1, \dots, v_n))$ with diagonal entries $\lambda_1, \dots, \lambda_p$. Let $v_{k,1}, \dots, v_{k,d_k}$ be the basis vectors corresponding to the block A_k, and define $U_k = \operatorname{span}(v_{k,1}, \dots, v_{k,d_k})$. By 8.44, T maps the vectors $v_{k,1}, \dots, v_{k,d_k}$ into U_k, so U_k is invariant under T, and $\mathcal{M}(T|_{U_k}, (v_{k,1}, \dots, v_{k,d_k})) = A_k$. By 5.65, $T|_{U_k}$ is diagonalizable for each k. Since A_k is upper triangular with λ_k on the diagonal, 5.41 yields that λ_k is the only eigenvalue of $T|_{U_k}$. By 5.55, $U_k = E(\lambda_k, T|_{U_k})$, so $T|_{U_k} = \lambda_k I$, and hence $A_k = \lambda_k I$. Since this holds for each block A_k, it follows that $\mathcal{M}(T, (v_1, \dots, v_n))$ is diagonal. ■

Exercise 8C13

Suppose $T \in \mathcal{L}(V)$ is nilpotent. Prove that if $v_1, \dots, v_n$ are vectors in V and $m_1, \dots, m_n$ are nonnegative integers such that

$$T^{m_1}v_1, \dots, Tv_1, v_1, \dots, T^{m_n}v_n, \dots, Tv_n, v_n \text{ is a basis of } V$$

and

$$T^{m_1+1}v_1 = \cdots = T^{m_n+1}v_n = 0,$$

then $T^{m_1}v_1, \dots, T^{m_n}v_n$ is a basis of $\operatorname{null} T$.

Axler Note. *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. Since $T^{m_k+1}v_k = 0$ for each k, $T^{m_k}v_k \in \operatorname{null} T$ for each k. Observe that $T(T^j v_k) = T^{j+1}v_k$ for each k and $0 \le j \le m_k$, where $T^{m_k+1}v_k = 0$. Therefore, the list $Tv_1, \dots, T^{m_1}v_1, \dots, Tv_n, \dots, T^{m_n}v_n$ spans $\operatorname{range} T$ and is

linearly independent since it is part of the given basis. Consequently, the list forms a basis of range T, and hence $\dim \operatorname{range} T = \sum_{k=1}^{n} m_k$. Because $\dim V = \sum_{k=1}^{n}(m_k + 1)$, the fundamental theorem of linear maps gives

$$\dim \operatorname{null} T = \sum_{k=1}^{n}(m_k + 1) - \sum_{k=1}^{n} m_k = n.$$

Since $T^{m_1}v_1, \ldots, T^{m_n}v_n$ is part of the given basis, it is a linearly independent list of length n in $\operatorname{null} T$, wherefore it forms a basis of $\operatorname{null} T$ by 2.38. ∎

Exercise 8C14

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. Let p be the minimal polynomial of T and let $n = \dim V$.

$\boxed{\Rightarrow}$ Suppose no direct sum decomposition of V into two nonzero subspaces invariant under T exists. By 8.46, there exists a Jordan basis $v_1, \ldots, v_n$ of V with blocks $A_1, \ldots, A_q$. Define U to be the span of the basis vectors corresponding to the first block A_1 and W to be the span of the remaining basis vectors. By 8.44, T maps the basis vectors corresponding to each block into their own span, so U and W are invariant under T. Since $U + W$ contains the basis vectors and $\dim U + \dim W = n$, it follows that $V = U \oplus W$ by 3.94. Since $U \neq \{0\}$, the hypothesis implies that $W = \{0\}$. Hence, $\mathcal{M}(T, (v_1, \ldots, v_n)) = A_1$ is a single n-by-n block. Let $\lambda = \lambda_1$. By 8.44, $(T - \lambda I)v_1 = 0$ and $(T - \lambda I)v_j = v_{j-1}$ for $2 \leq j \leq n$, wherefore $(T - \lambda I)^n = 0$ and $(T - \lambda I)^{n-1}v_n = v_1 \neq 0$. By 5.29, p divides $(z - \lambda)^n$, so $p = (z - \lambda)^r$ for some positive integer $r \leq n$. On the other hand, if $r \leq n - 1$, then $(z - \lambda)^{n-1}$ is a multiple of p, so $(T - \lambda I)^{n-1} = 0$ by 5.29, contradicting $(T - \lambda I)^{n-1}v_n = v_1 \neq 0$. Hence, $r = n$, so $p = (z - \lambda)^n$.

$\boxed{\Leftarrow}$ Suppose $p = (z - \lambda)^n$ for some $\lambda \in \mathbf{C}$. Let $V = U \oplus W$ with U and W invariant under T. Since $(T - \lambda I)^n = 0$ and invariance of U under T gives invariance under $T - \lambda I$, it follows that $(T - \lambda I)^{\dim U} = 0$ on U by 8.16. A similar argument shows that $(T - \lambda I)^{\dim W} = 0$ on W. Let $M = \max(\dim U, \dim W)$. Then $(T - \lambda I)^M = 0$ on $U + W = V$, so $M \geq n$ by 5.29. By 3.94, $\dim U + \dim W = n$, so $M \geq \dim U + \dim W$, forcing $\min(\dim U, \dim W) = 0$. Hence, either $U = \{0\}$ or $W = \{0\}$, proving that no direct sum decomposition into two nonzero invariant subspaces exists. ∎

8D Trace: A Connection Between Matrices and Operators

Exercise 8D1

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 $n = \dim V$. By 6.35, there exists an orthonormal basis $e_1, \ldots, e_n$ of V. By 8.55, homogeneity in the first slot, conjugate symmetry of the inner product, and 6.30(c), we have

$$\operatorname{tr} T = \sum_{k=1}^{n} \langle Te_k, e_k\rangle = \sum_{k=1}^{n} \langle \langle e_k, v\rangle w, e_k\rangle = \sum_{k=1}^{n} \langle w, e_k\rangle \overline{\langle v, e_k\rangle} = \langle w, v\rangle. \qquad \blacksquare$$

Exercise 8D2

Suppose $P \in \mathcal{L}(V)$ satisfies $P^2 = P$. Prove that

$$\operatorname{tr} P = \dim \operatorname{range} P.$$

Solution. Let $n = \dim V$. By Exercise 3B27, $V = \operatorname{null} P \oplus \operatorname{range} P$. Let $v_1, \ldots, v_m$ be a basis of $\operatorname{range} P$ and $u_1, \ldots, u_{n-m}$ be a basis of $\operatorname{null} P$, where $n - m = \dim \operatorname{null} P$ from the fundamental theorem of linear maps. Since the list $v_1, \ldots, v_m, u_1, \ldots, u_{n-m}$ spans $\operatorname{null} P + \operatorname{range} P = V$ and is of length n, it is a basis of V by 2.42. For each $v_k \in \operatorname{range} P$, there is $w \in V$ such that $v_k = Pw$, and hence $Pv_k = P^2 w = Pw = v_k$. Since $Pu_j = 0$ for $j = 1, \ldots, n - m$, the matrix of P with respect to this basis is diagonal with m 1's and $n - m$ 0's on the diagonal. Therefore, $\operatorname{tr} P = m = \dim \operatorname{range} P$. $\blacksquare$

Exercise 8D3

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. Let $n = \dim V$. We show the existence of an operator $S \in \mathcal{L}(\mathbf{C}^n)$ such that $S^5 = S$ and $\operatorname{tr} S = \operatorname{tr} T$. Indeed, fix a basis $v_1, \ldots, v_n$ of V, and let $A = \mathcal{M}(T, (v_1, \ldots, v_n))$. Viewing $A \in \mathbf{C}^{n,n}$, the linear map lemma yields a unique $S \in \mathcal{L}(\mathbf{C}^n)$ such that $\mathcal{M}(S)$ with respect to the standard basis of $\mathbf{C}^n$ is A, so $\operatorname{tr} T = \operatorname{tr} A = \operatorname{tr} S$. 3.43 yields

$$\mathcal{M}(S^5) = A^5 = \mathcal{M}(T^5) = \mathcal{M}(T) = A = \mathcal{M}(S).$$

By 3.71, $S^5 = S$. Suppose $Sv = \lambda v$ for $v \neq 0$ and $\lambda \in \mathbf{C}$. Then $S^5 v = \lambda^5 v = Sv = \lambda v$, so $\lambda^5 = \lambda$. Because $z^5 - z = z(z-1)(z+1)(z-i)(z+i)$ has zeros $0, 1, -1, i, -i$, we see that the eigenvalues of S are all among these numbers. Letting d_λ be the multiplicity of the eigenvalue λ, where $d_\lambda = 0$ if λ is not an eigenvalue, we have by 8.52 that

$$\operatorname{tr} S = (d_1 - d_{-1}) + (d_i - d_{-i})i.$$

Therefore, both $\operatorname{Re}(\operatorname{tr} S) = d_1 - d_{-1}$ and $\operatorname{Im}(\operatorname{tr} S) = d_i - d_{-i}$ are integers. Since $\operatorname{tr} T = \operatorname{tr} S$, the real and imaginary parts of $\operatorname{tr} T$ are also integers. $\blacksquare$

Exercise 8D4

Suppose V is an inner product space and $T \in \mathcal{L}(V)$. Prove that

$$\operatorname{tr} T^* = \overline{\operatorname{tr} T}.$$

Solution. Let $n = \dim V$. By 6.35, there exists an orthonormal basis $e_1, \dots, e_n$ of V. By 8.55, we have

$$\operatorname{tr} T^* = \sum_{k=1}^{n} \langle T^* e_k, e_k \rangle = \sum_{k=1}^{n} \langle e_k, T e_k \rangle = \sum_{k=1}^{n} \overline{\langle T e_k, e_k \rangle} = \overline{\sum_{k=1}^{n} \langle T e_k, e_k \rangle} = \overline{\operatorname{tr} T}.$$

■

Exercise 8D5

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. Let $n = \dim V$. By 7.38(c), there exists an orthonormal basis $e_1, \dots, e_n$ of V such that $\mathcal{M}(T, (e_1, \dots, e_n))$ is diagonal with nonnegative diagonal entries $\lambda_1, \dots, \lambda_n$. Since $0 = \operatorname{tr} T = \sum_{k=1}^{n} \lambda_k$ and each $\lambda_k \geq 0$, we must have $\lambda_k = 0$ for all k. Therefore, $\mathcal{M}(T, (e_1, \dots, e_n)) = 0$, so $T = 0$. ■

Exercise 8D6

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. Let $n = \dim V$. By 6.57(g), $P^2 = P$ and $Q^2 = Q$, so P and Q are self-adjoint by Exercise 7A20. Since $Q^2 = Q$, we have $\operatorname{tr}(PQ) = \operatorname{tr}((PQ)Q)$, so 8.56 gives $\operatorname{tr}((PQ)Q) = \operatorname{tr}(Q(PQ)) = \operatorname{tr}(QPQ)$. By 7.5(d), $QPQ = QPPQ = (PQ)^*(PQ)$. 6.35 yields an orthonormal basis $e_1, \dots, e_n$ of V, so 8.55 gives

$$\operatorname{tr}(PQ) = \sum_{k=1}^{n} \langle (PQ)^*(PQ) e_k, e_k \rangle = \sum_{k=1}^{n} \|PQ e_k\|^2 \geq 0.$$

■

Exercise 8D7

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. Note that $\operatorname{tr} T = 51 + (-40) + 1 = 12$. By 8.25, the eigenvalues of T, repeated with multiplicity, form a list of length 3 that contains $-48, 24$, and a third eigenvalue λ. 8.52 gives $-48 + 24 + \lambda = 12$, so $\lambda = 36$. ■

Exercise 8D8

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. Take $V = \mathbf{F}^2$ with $S = T = I$, so $ST = I$. Then $\operatorname{tr}(ST) = 2$, but $(\operatorname{tr} S)(\operatorname{tr} T) = 4$. ■

Exercise 8D9

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. Let $v_1, \dots, v_n$ be a basis of V, and let $A = \mathcal{M}(T, (v_1, \dots, v_n))$. By the linear map lemma, there exist operators

$S_{j,k} \in \mathcal{L}(V)$ such that $E_{j,k} = \mathcal{M}(S_{j,k}, (v_1, \ldots, v_n))$ has a 1 in the (j,k)th entry and 0 elsewhere. By 3.46,

$$(E_{j,k}A)_{\ell,\ell} = \sum_{m=1}^{n} (E_{j,k})_{\ell,m} A_{m,\ell} = (E_{j,k})_{\ell,k} A_{k,\ell} = \begin{cases} A_{k,j} & \text{if } \ell = j, \\ 0 & \text{otherwise.} \end{cases}$$

By 3.43, $\operatorname{tr}(S_{j,k}T) = \operatorname{tr}(E_{j,k}A) = A_{k,j}$. By assumption, $\operatorname{tr}(S_{j,k}T) = 0$ for all j, k, so $A_{k,j} = 0$ for all j, k. Therefore, $A = 0$, so by the linear map lemma, $T = 0$. ∎

Exercise 8D10

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 $\tau(I) = \dim V$.

Axler 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. To show existence of such a linear functional, observe that 8.56 gives linearity of the trace and that $\operatorname{tr}(ST) = \operatorname{tr}(TS)$ for all $S, T \in \mathcal{L}(V)$. Since $\mathcal{M}(I) = I$, we have $\operatorname{tr} I = \dim V$, so the trace satisfies the same conditions as τ.

To show uniqueness, let $n = \dim V$ and let $v_1, \ldots, v_n$ be a basis of V. Let τ be a linear functional on $\mathcal{L}(V)$ such that $\tau(ST) = \tau(TS)$ for all $S, T \in \mathcal{L}(V)$ and $\tau(I) = n$. Let $P_{j,k}$ be as in the hint. Then for $a_1, \ldots, a_n \in \mathbf{F}$, we have

$$P_{j,k}P_{k,j}(a_1 v_1 + \cdots + a_n v_n) = P_{j,k}(a_j v_k) = a_j v_j = P_{j,j}(a_1 v_1 + \cdots + a_n v_n).$$

Therefore, $P_{j,k}P_{k,j} = P_{j,j}$. Swapping j and k gives $P_{k,j}P_{j,k} = P_{k,k}$. Since $\tau(ST) = \tau(TS)$ for all $S, T \in \mathcal{L}(V)$, we have

$$\tau(P_{j,j}) = \tau(P_{j,k}P_{k,j}) = \tau(P_{k,j}P_{j,k}) = \tau(P_{k,k})$$

for all j, k, so $\tau(P_{j,j})$ is the same for all j. Moreover,

$$I = \sum_{j=1}^{n} P_{j,j},$$

so $n = \tau(I) = n\tau(P_{1,1})$, whence $\tau(P_{1,1}) = 1$. Thus, $\tau(P_{j,j}) = 1$ for all j.

Fix $T \in \mathcal{L}(V)$, and let $A = \mathcal{M}(T)$ be the matrix of T with respect to the basis $v_1, \ldots, v_n$. Then for any $v = a_1 v_1 + \cdots + a_n v_n \in V$, we have

$$P_{j,j}TP_{j,j}v = a_j P_{j,j}\left(\sum_{\ell=1}^{n} A_{\ell,j} v_\ell\right) = a_j A_{j,j} v_j = A_{j,j} P_{j,j} v,$$

so $P_{j,j}TP_{j,j} = A_{j,j}P_{j,j}$. From $I = \sum_{k=1}^{n} P_{k,k}$, we get $T = \sum_{k=1}^{n} TP_{k,k}$. Then

$$\tau(T) = \sum_{k=1}^{n} \tau((TP_{k,k})P_{k,k}) = \sum_{k=1}^{n} \tau(P_{k,k}TP_{k,k}) = \sum_{k=1}^{n} A_{k,k}\tau(P_{k,k}) = \sum_{k=1}^{n} A_{k,k} = \operatorname{tr} T,$$

where the first equality follows from linearity and $P_{k,k}^2 = P_{k,k}$; the second equality follows from the property $\tau(ST) = \tau(TS)$; the third equality follows from $P_{k,k}TP_{k,k} = A_{k,k}P_{k,k}$ and linearity; the fourth equality follows from $\tau(P_{k,k}) = 1$; and the fifth equality is by the definition of the trace. ∎

Exercise 8D11

Suppose V and W are inner product spaces and $T \in \mathcal{L}(V, W)$. Prove that if $e_1, \dots, e_n$ is an orthonormal basis of V and $f_1, \dots, f_m$ is an orthonormal basis of W, then

$$\operatorname{tr}(T^*T) = \sum_{k=1}^{n} \sum_{j=1}^{m} |\langle Te_k, f_j \rangle|^2.$$

Axler Note. *The numbers $\langle Te_k, f_j \rangle$ are the entries of the matrix of T with respect to the orthonormal bases $e_1, \dots, e_n$ and $f_1, \dots, f_m$. 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. Apply 8.55 to $T^*T \in \mathcal{L}(V)$. Then the definition of the adjoint and 6.30(b) in W yield

$$\operatorname{tr}(T^*T) = \sum_{k=1}^{n} \langle T^*Te_k, e_k \rangle = \sum_{k=1}^{n} \|Te_k\|^2 = \sum_{k=1}^{n} \sum_{j=1}^{m} |\langle Te_k, f_j \rangle|^2.$$

∎

Exercise 8D12

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 $\mathbf{F}^{mn}$.

Axler Note. ***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. Let $n = \dim V$ and $m = \dim W$. By 6.35, there exist orthonormal bases $e_1, \dots, e_n$ of V and $f_1, \dots, f_m$ of W. Define the map $\varphi \colon \mathcal{L}(V, W) \to \mathbf{F}^{mn}$ by mapping T to $\mathcal{M}(T, (e_1, \dots, e_n), (f_1, \dots, f_m))$ as a vector in $\mathbf{F}^{mn}$, where the entries are listed column by column. 3.71 gives that φ is linear and injective. By 3.43 and 7.9, $\mathcal{M}(T^*S) = \mathcal{M}(T)^*\mathcal{M}(S)$ with respect to the fixed orthonormal bases, so

$$\operatorname{tr}(T^*S) = \sum_{k=1}^{n} \sum_{j=1}^{m} \mathcal{M}(S)_{j,k} \overline{\mathcal{M}(T)_{j,k}} = \langle \varphi(S), \varphi(T) \rangle,$$

where the right-hand side is the Euclidean inner product on $\mathbf{F}^{mn}$.

(a) Note that positivity, additivity, homogeneity in the first slot, and conjugate symmetry of the inner product on $\mathcal{L}(V, W)$ follow from the corresponding properties of the Euclidean inner product on $\mathbf{F}^{mn}$ via the map φ. To see definiteness, note that $\langle 0, 0 \rangle = 0$ and if $\langle T, T \rangle = 0$, then $\langle \varphi(T), \varphi(T) \rangle = 0$, which implies $\varphi(T) = 0$, and hence $T = 0$ by injectivity of φ.

(b) This is exactly the identity computed in the display above, which shows that the inner product on $\mathcal{L}(V, W)$ corresponds to the Euclidean inner product on $\mathbf{F}^{mn}$ under the map φ. ∎

Exercise 8D13

Find $S, T \in \mathcal{L}(\mathcal{P}(\mathbf{F}))$ such that $ST - TS = I$.

Axler Hint. *Make an appropriate modification of the operators in Example 3.9.*

Axler Note. *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 $S, T \in \mathcal{L}(\mathcal{P}(\mathbf{F}))$ be defined by $Sf = f'$ and $Tf = zf$ for all $f \in \mathcal{P}(\mathbf{F})$. Each f is a linear combination of monomials z^k, so it suffices to check the identity on $f = z^k$ for each $k \geq 0$. We have $STz^k = S(z^{k+1}) = (k+1)z^k$ for $k \geq 0$. Also, $TSz^k = T(kz^{k-1}) = kz^k$ holds for $k \geq 1$, and for $k = 0$, $TSz^0 = T(Sz^0) = T(0) = 0$. Therefore, for all $k \geq 0$, we have $(ST - TS)z^k = z^k$, which shows that $ST - TS = I$ on $\mathcal{P}(\mathbf{F})$. ■

9 Multilinear Algebra and Determinants

9A Bilinear Forms and Quadratic Forms

Exercise 9A1

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. Let $c = \beta(1, 1)$. Then for all $x, y \in \mathbf{F}$, homogeneity in the first and second slots gives

$$\beta(x, y) = \beta(x \cdot 1, y \cdot 1) = xy\beta(1, 1) = cxy. \qquad \blacksquare$$

Exercise 9A2

Let $n = \dim V$. Suppose β is a bilinear form on V. Prove that there exist $\varphi_1, \ldots, \varphi_n, \tau_1, \ldots, \tau_n \in V'$ such that

$$\beta(u, v) = \varphi_1(u) \cdot \tau_1(v) + \cdots + \varphi_n(u) \cdot \tau_n(v)$$

for all $u, v \in V$.

Axler Note. *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 $v_1, \ldots, v_n$ be a basis of V with dual basis $\varphi_1, \ldots, \varphi_n$ of V'. For any $u \in V$, 3.114 yields $u = \varphi_1(u)v_1 + \cdots + \varphi_n(u)v_n$. For each $k = 1, \ldots, n$, define $\tau_k \in V'$ by $\tau_k(v) = \beta(v_k, v)$. Then for any $v \in V$,

$$\beta(u, v) = \sum_{k=1}^{n} \varphi_k(u)\beta(v_k, v) = \sum_{k=1}^{n} \varphi_k(u)\tau_k(v). \qquad \blacksquare$$

Exercise 9A3

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. For any $u, v \in V$, we have $(u, v) = (u, 0) + (0, v)$, so linearity of β on $V \times V$ gives $\beta(u, v) = \beta(u, 0) + \beta(0, v)$. Applying 3.10 to $w \mapsto \beta(u, w)$ and $w \mapsto \beta(w, v)$ gives $\beta(u, 0) = 0$ and $\beta(0, v) = 0$, respectively. Hence, $\beta(u, v) = \beta(u, 0) + \beta(0, v) = 0$ for all $u, v \in V$, so $\beta = 0$. $\blacksquare$

Exercise 9A4 ✔

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 $u, v \in V$.

Axler Note. *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. For each $v \in V$, define $\varphi_v \in V'$ by $\varphi_v(u) = \langle u, v \rangle$. By Exercise 6C13, $\Phi : V \to V'$ defined by $\Phi(v) = \varphi_v$ is an isomorphism, so $\Phi^{-1} \in \mathcal{L}(V', V)$ exists. Define $\Psi : V \to V'$ by $\Psi(v)(u) = \beta(u, v)$ for all $u \in V$. By linearity of β in the

first slot, $\Psi(v) \in V'$. Moreover, for any $\lambda, \mu \in \mathbf{R}$ and $v, w \in V$, we have

$$\Psi(\lambda v + \mu w)(u) = \beta(u, \lambda v + \mu w) = \lambda \Psi(v)(u) + \mu \Psi(w)(u)$$

for all $u \in V$, so $\Psi \in \mathcal{L}(V, V')$. For any $S \in \mathcal{L}(V)$, we have

$$\begin{aligned} \beta(u,v) = \langle u, Sv \rangle \text{ for all } u, v \in V &\iff \Psi(v) = \varphi_{Sv} = \Phi(Sv) \text{ for all } v \in V \\ &\iff \Psi = \Phi S \\ &\iff S = \Phi^{-1}\Psi. \end{aligned}$$

Thus, the operator $T = \Phi^{-1}\Psi$ satisfies $\beta(u, v) = \langle u, Tv \rangle$ for all $u, v \in V$, and it is unique. ■

Exercise 9A5

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 9A4). Show that β is an inner product on V if and only if T is an invertible positive operator on V.

Solution. Let $\langle \cdot, \cdot \rangle_T$ be defined as in Exercise 7C23. Since $\mathbf{F} = \mathbf{R}$, for all $u, v \in V$, we have

$$\langle u, v \rangle_T = \langle Tu, v \rangle = \langle v, Tu \rangle = \beta(v, u),$$

so $\langle \cdot, \cdot \rangle_T$ is β with swapped arguments. Then β is an inner product if and only if $\langle \cdot, \cdot \rangle_T$ is an inner product, which holds if and only if T is an invertible positive operator, where the second equivalence follows from Exercise 7C23(a). ■

Exercise 9A6

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. The statement is false. Consider $V = \mathbf{F}^2$ and let e_1, e_2 be the standard basis of V. Define $\rho \colon V \times V \to \mathbf{F}$ by

$$\rho(ae_1 + be_2, ce_1 + de_2) = ad + bc.$$

Then ρ is symmetric and bilinear. However, $\rho(e_1, e_1) = \rho(e_2, e_2) = 0$, but $\rho(e_1 + e_2, e_1 + e_2) = 2 \neq 0$. Therefore, the set $\{v \in V : \rho(v, v) = 0\}$ is not a subspace. ■

Exercise 9A7

Explain why the proof of 9.13 (diagonalization of a symmetric bilinear form by an orthonormal basis on a real inner product space) fails if the hypothesis that $\mathbf{F} = \mathbf{R}$ is dropped.

Manual Remark. *I ultimately decided to point out one specific failure within the proof, but there are other places where the argument relies on the entries being real.*

Solution. The proof of 9.13 fails when it claims that T is self-adjoint: since $f_1, \ldots, f_n$ is an orthonormal basis of V, 7.9 gives $\mathcal{M}(T^*) = B^* = \overline{B}$ because $B^t = B$. Therefore, $T = T^*$ only holds when $B = \overline{B}$, i.e., when the entries of B are real, which fails when B is symmetric with non-real entries. ■

Exercise 9A8

Find formulas for $\dim V_{\text{sym}}^{(2)}$ and $\dim V_{\text{alt}}^{(2)}$ in terms of $\dim V$.

Solution. Let $n = \dim V$ and fix some basis of V. Let $\varphi\colon V^{(2)} \to \mathbf{F}^{n,n}$ be defined by $\varphi(\beta) = \mathcal{M}(\beta)$. Define $\mathcal{S}$ to be the set of symmetric n-by-n matrices over $\mathbf{F}$. By 9.5 and 9.12, φ restricts to an isomorphism from $V_{\text{sym}}^{(2)}$ to $\mathcal{S}$, so $\dim V_{\text{sym}}^{(2)} = \dim \mathcal{S}$. Consider the matrix $E_{j,k}$ with a 1 in the (j,k)th entry and 0 elsewhere. We claim that the set

$$\mathcal{E} = \{E_{j,j} : 1 \le j \le n\} \cup \{E_{j,k} + E_{k,j} : 1 \le j < k \le n\}$$

is a basis of $\mathcal{S}$. Indeed, each $A \in \mathcal{S}$ can be written as

$$A = \sum_{j=1}^{n} A_{j,j} E_{j,j} + \sum_{1 \le j < k \le n} A_{j,k}(E_{j,k} + E_{k,j}),$$

and comparing the (j,k)th entries, we see that the coefficients are uniquely determined, so 2.28 gives that $\mathcal{E}$ is a basis of $\mathcal{S}$. Counting the elements of $\mathcal{E}$, we have n matrices of the form $E_{j,j}$ and $\frac{n(n-1)}{2}$ matrices of the form $E_{j,k} + E_{k,j}$ with $j < k$. Hence,

$$\dim \mathcal{S} = n + \frac{n(n-1)}{2} = \frac{n(n+1)}{2} = \dim V_{\text{sym}}^{(2)}.$$

Finally, 9.5 gives that $\dim V^{(2)} = n^2$ and 9.17 gives that $V^{(2)} = V_{\text{sym}}^{(2)} \oplus V_{\text{alt}}^{(2)}$, so by 3.94,

$$\dim V_{\text{alt}}^{(2)} = \dim V^{(2)} - \dim V_{\text{sym}}^{(2)} = n^2 - \frac{n(n+1)}{2} = \frac{n(n-1)}{2}.$$

■

Exercise 9A9

Suppose that n is a positive integer and $V = \{p \in \mathcal{P}_n(\mathbf{R}) : p(0) = p(1)\}$. Define $\alpha\colon V \times V \to \mathbf{R}$ by

$$\alpha(p,q) = \int_0^1 pq'.$$

Show that α is an alternating bilinear form on V.

Solution. For fixed $q \in V$, the map $p \mapsto \alpha(p,q)$ is linear because the integral is linear. Similarly, for fixed $p \in V$, the map $q \mapsto \alpha(p,q)$ is linear because differentiation and the integral are linear, and thus α is bilinear. Finally, for any $r \in V$, we have

$$\alpha(r,r) = \int_0^1 rr' = \frac{1}{2}\int_0^1 (r^2)' = \frac{1}{2}(r(1)^2 - r(0)^2) = 0,$$

since $r(0) = r(1)$. Hence, α is alternating. ■

Exercise 9A10

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 $\rho\colon V \times V \to \mathbf{R}$ by

$$\rho(p,q) = \int_0^1 pq''.$$

Show that ρ is a symmetric bilinear form on V.

Solution. Let $U = \{p \in \mathcal{P}_n(\mathbf{R}) : p(0) = p(1)\}$ and let α be as in Exercise 9A9. For any $p, q \in V$, we have $p, q' \in U$ because $q' \in \mathcal{P}_{n-1}(\mathbf{R})$ and $q'(0) = q'(1)$. Therefore, $\rho(p, q) = \alpha(p, q')$. Moreover, ρ is linear in p because α is linear in its first slot. Also, ρ is linear in q since $q \mapsto q'$ is linear and α is linear in its second slot. Finally, for any $p, q \in V$, 9.16 gives

$$\rho(p, q) = \alpha(p, q') = -\alpha(q', p) = -\int_0^1 q'p',$$

and swapping p and q results in the same expression, so ρ is symmetric. ■

9B Alternating Linear Forms

Exercise 9B1

Suppose m is a positive integer. Show that $\dim V^{(m)} = (\dim V)^m$.

Solution. Let $n = \dim V$. Let $v_1, \dots, v_n$ be a basis of V and let $\varphi_1, \dots, \varphi_n$ be the dual basis. For each ordered m-tuple $(k_1, \dots, k_m) \in \{1, \dots, n\}^m$, define

$$\beta_{k_1,\dots,k_m}(u_1, \dots, u_m) = \varphi_{k_1}(u_1) \cdots \varphi_{k_m}(u_m).$$

The ℓth map $u_\ell \mapsto \beta_{k_1,\dots,k_m}(u_1, \dots, u_m)$ is linear because it is the product of $\varphi_{k_\ell}(u_\ell)$, which is linear in u_ℓ, and the other factors, which are independent of u_ℓ. Therefore, $\beta_{k_1,\dots,k_m}$ is m-linear. We claim that the list $\mathcal{B} = (\beta_{k_1,\dots,k_m})_{(k_1,\dots,k_m)\in\{1,\dots,n\}^m}$ forms a basis of $V^{(m)}$.

Indeed, suppose there are scalars $c_{k_1,\dots,k_m}$ such that

$$\sum_{(k_1,\dots,k_m)\in\{1,\dots,n\}^m} c_{k_1,\dots,k_m}\beta_{k_1,\dots,k_m} = 0.$$

Fix $(p_1, \dots, p_m) \in \{1, \dots, n\}^m$. Then for each $(k_1, \dots, k_m) \in \{1, \dots, n\}^m$, we have

$$\beta_{k_1,\dots,k_m}(v_{p_1}, \dots, v_{p_m}) = \delta_{k_1 p_1} \cdots \delta_{k_m p_m}.$$

Evaluating the sum at $(v_{p_1}, \dots, v_{p_m})$ gives

$$\sum_{(k_1,\dots,k_m)\in\{1,\dots,n\}^m} c_{k_1,\dots,k_m}\beta_{k_1,\dots,k_m}(v_{p_1}, \dots, v_{p_m}) = 0,$$

which simplifies to $c_{p_1,\dots,p_m} = 0$. Since $(p_1, \dots, p_m)$ was arbitrary, it follows that all $c_{k_1,\dots,k_m} = 0$, proving that $\mathcal{B}$ is a linearly independent list of length n^m. Moreover, let $\beta \in V^{(m)}$ and $u_1, \dots, u_m \in V$ be arbitrary. By 3.114, we may write each u_ℓ as $u_\ell = \sum_{t=1}^n \varphi_t(u_\ell)v_t$ for $\ell = 1, \dots, m$. Using the multilinearity of β and the fact that $\varphi_{t_1}(u_1) \cdots \varphi_{t_m}(u_m) = \beta_{t_1,\dots,t_m}(u_1, \dots, u_m)$, we have

$$\beta(u_1, \dots, u_m) = \sum_{t_1,\dots,t_m=1}^n \varphi_{t_1}(u_1) \cdots \varphi_{t_m}(u_m)\beta(v_{t_1}, \dots, v_{t_m}) = \sum_{t_1,\dots,t_m=1}^n \beta_{t_1,\dots,t_m}(u_1, \dots, u_m)\beta(v_{t_1}, \dots, v_{t_m}).$$

Since $u_1, \dots, u_m$ were arbitrary, this shows that

$$\beta = \sum_{t_1,\dots,t_m=1}^n \beta(v_{t_1}, \dots, v_{t_m})\beta_{t_1,\dots,t_m}.$$

Hence, $\mathcal{B}$ spans $V^{(m)}$, and thus $\mathcal{B}$ is a basis of $V^{(m)}$ with $\dim V^{(m)} = n^m = (\dim V)^m$. ∎

Exercise 9B2

Suppose $n \geq 3$ and $\alpha\colon \mathbf{F}^n \times \mathbf{F}^n \times \mathbf{F}^n \to \mathbf{F}$ is defined by

$$\begin{aligned}&\alpha((x_1, \dots, x_n), (y_1, \dots, y_n), (z_1, \dots, z_n))\\ &\quad = x_1y_2z_3 - x_2y_1z_3 - x_3y_2z_1 - x_1y_3z_2 + x_3y_1z_2 + x_2y_3z_1.\end{aligned}$$

Show that α is an alternating 3-linear form on $\mathbf{F}^n$.

Solution. Observe that the map $(x_1, \dots, x_n) \mapsto x_i$ is linear for each $i = 1, \dots, n$, so each term in the definition of α is

a product of three linear maps, one in each slot. Since each product is 3-linear and a linear combination of 3-linear maps is 3-linear, it follows that α is 3-linear. Let $x, y, z \in \mathbf{F}^n$ be arbitrary with $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n)$, and $z = (z_1, \ldots, z_n)$. Then

$$\begin{aligned}
\alpha(x, x, z) &= x_1x_2z_3 - x_2x_1z_3 - x_3x_2z_1 - x_1x_3z_2 + x_3x_1z_2 + x_2x_3z_1 = 0 \\
\alpha(x, y, x) &= x_1y_2x_3 - x_2y_1x_3 - x_3y_2x_1 - x_1y_3x_2 + x_3y_1x_2 + x_2y_3x_1 = 0 \\
\alpha(x, y, y) &= x_1y_2y_3 - x_2y_1y_3 - x_3y_2y_1 - x_1y_3y_2 + x_3y_1y_2 + x_2y_3y_1 = 0.
\end{aligned}$$

Hence, α is alternating. ∎

Exercise 9B3

Suppose m is a positive integer and α is an m-linear form on V such that $\alpha(v_1, \ldots, v_m) = 0$ whenever $v_1, \ldots, v_m$ is a list of vectors in V with $v_j = v_{j+1}$ for some $j \in \{1, \ldots, m-1\}$. Prove that α is an alternating m-linear form on V.

Solution. We claim that an adjacent swap of two vectors in the arguments of α changes the sign of α. Fix some $j \in \{1, \ldots, m-1\}$ and all slots except for the jth and $(j+1)$th slots. For arbitrary $u, w \in V$, consider $u + w$ in both the jth and $(j+1)$th slots. By hypothesis, $\alpha(\ldots, u, u, \ldots) = 0$ and $\alpha(\ldots, w, w, \ldots) = 0$, where the repeated entries are in the jth and $(j+1)$th slots. Expanding $\alpha(\ldots, u + w, u + w, \ldots)$ by linearity in the jth and $(j+1)$th slots, we get

$$0 = \alpha(\ldots, u + w, u + w, \ldots) = \alpha(\ldots, u, w, \ldots) + \alpha(\ldots, w, u, \ldots),$$

which shows that $\alpha(\ldots, u, w, \ldots) = -\alpha(\ldots, w, u, \ldots)$.

Now let $v_1, \ldots, v_m \in V$ be arbitrary and suppose $v_\ell = v_k$ for some $\ell < k$. Apply the swap relation $k - \ell - 1$ times to move v_k to the $(\ell+1)$th slot, changing the sign each time. Then

$$\alpha(v_1, \ldots, v_m) = (-1)^{k-\ell-1}\alpha(v_1, \ldots, v_\ell, v_k, v_{\ell+1}, \ldots, v_{k-1}, v_{k+1}, \ldots, v_m).$$

The ℓth and $(\ell+1)$th slots of the right side now contain the same vector $v_\ell = v_k$, so by the hypothesis, the right side is 0. Hence, $\alpha(v_1, \ldots, v_m) = 0$ whenever $v_\ell = v_k$ for some $\ell < k$, so α is alternating. ∎

Exercise 9B4

Prove or give a counterexample: If $\alpha \in V_{\text{alt}}^{(4)}$, 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. Take $V = \mathbf{F}^4$ and let U be the set in question. By 9.37, $\dim V_{\text{alt}}^{(4)} = 1$, so there exists a nonzero $\alpha \in V_{\text{alt}}^{(4)}$. Let $(x_1, x_2, x_3, x_4) \in V^4$ be such that $\alpha(x_1, x_2, x_3, x_4) \neq 0$. Applying 3.10 to the maps $v \mapsto \alpha(x_1, x_2, x_3, v)$ and $v \mapsto \alpha(v, 0, 0, x_4)$, we see that $\alpha(x_1, x_2, x_3, 0) = 0 = \alpha(0, 0, 0, x_4)$, so $(x_1, x_2, x_3, 0), (0, 0, 0, x_4) \in U$. However, $(x_1, x_2, x_3, 0) + (0, 0, 0, x_4) = (x_1, x_2, x_3, x_4) \notin U$. Hence, U is not closed under addition and therefore is not a subspace. ∎

Exercise 9B5

Suppose m is a positive integer and β is an m-linear form on V. Define an m-linear form α on V by

$$\alpha(v_1, \dots, v_m) = \sum_{(j_1,\dots,j_m)\in \operatorname{perm} m} (\operatorname{sign}(j_1, \dots, j_m))\beta(v_{j_1}, \dots, v_{j_m})$$

for $v_1, \dots, v_m \in V$. Explain why $\alpha \in V_{\text{alt}}^{(m)}$.

Solution. Suppose $v_\ell = v_k$ for some $\ell \neq k$. Let $\sigma = (j_1, \dots, j_m) \in \operatorname{perm} m$, and let $\sigma' = (j'_1, \dots, j'_m)$ be σ with the entries equal to ℓ and k swapped. The map $\sigma \mapsto \sigma'$ is a bijection on $\operatorname{perm} m$ and $\operatorname{sign}(\sigma') = -\operatorname{sign}(\sigma)$ by 9.34. Because $v_\ell = v_k$, the lists $(v_{j_1}, \dots, v_{j_m})$ and $(v_{j'_1}, \dots, v_{j'_m})$ are equal. Then

$$\alpha(v_1, \dots, v_m) = \sum_{\sigma\in \operatorname{perm} m} \operatorname{sign}(\sigma')\beta(v_{j'_1}, \dots, v_{j'_m}) = \sum_{\sigma\in \operatorname{perm} m} -\operatorname{sign}(\sigma)\beta(v_{j_1}, \dots, v_{j_m}) = -\alpha(v_1, \dots, v_m).$$

Hence, $\alpha(v_1, \dots, v_m) = 0$ whenever $v_\ell = v_k$ for some $\ell \neq k$, so α is alternating. ∎

Exercise 9B6

Suppose m is a positive integer and β is an m-linear form on V. Define an m-linear form α on V by

$$\alpha(v_1, \dots, v_m) = \sum_{(j_1,\dots,j_m)\in \operatorname{perm} m} \beta(v_{j_1}, \dots, v_{j_m})$$

for $v_1, \dots, v_m \in V$. Explain why

$$\alpha(v_{k_1}, \dots, v_{k_m}) = \alpha(v_1, \dots, v_m)$$

for all $v_1, \dots, v_m \in V$ and all $(k_1, \dots, k_m) \in \operatorname{perm} m$.

Solution. Fix some $(k_1, \dots, k_m) \in \operatorname{perm} m$ and $v_1, \dots, v_m \in V$. Let $u_p = v_{k_p}$ for $p = 1, \dots, m$. Distinctness of the entries in the lists $(j_1, \dots, j_m)$ and $(k_1, \dots, k_m)$ ensures that the map $(j_1, \dots, j_m) \mapsto (k_{j_1}, \dots, k_{j_m})$ lands in $\operatorname{perm} m$. Since the k_p are distinct, this map is injective, and because $\operatorname{perm} m$ is finite, it is a bijection. By reindexing the sum using this bijection, we have

$$\alpha(u_1, \dots, u_m) = \sum_{(j_1,\dots,j_m)\in \operatorname{perm} m} \beta(u_{j_1}, \dots, u_{j_m}) = \sum_{(j_1,\dots,j_m)\in \operatorname{perm} m} \beta(v_{k_{j_1}}, \dots, v_{k_{j_m}}) = \alpha(v_1, \dots, v_m),$$

which gives the desired equality. ∎

Exercise 9B7

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 $\alpha(v_1, v_2) = 0$.

Axler Note. *This exercise shows that 9.39 can fail if the hypothesis that $n = \dim V$ is deleted.*

Solution. Let $x = (x_1, x_2, x_3)$ and $y = (y_1, y_2, y_3)$ be vectors in $\mathbf{R}^3$, let e_1, e_2, e_3 be the standard basis of $\mathbf{R}^3$, and define

$$\alpha(x, y) = x_1y_2 - x_2y_1.$$

With y fixed, $\alpha(x, y)$ is a linear function of x, and with x fixed, $\alpha(x, y)$ is a linear function of y, so α is bilinear. Moreover, $\alpha(x, x) = x_1x_2 - x_2x_1 = 0$ for all $x \in \mathbf{R}^3$, so α is alternating. Because $\alpha(e_1, e_2) = 1 \cdot 1 - 0 \cdot 0 = 1 \neq 0$, we see that α is

nonzero. Finally, $v_1 = e_1$ and $v_2 = e_3$ form a linearly independent list in $\mathbf{R}^3$, and $\alpha(v_1, v_2) = \alpha(e_1, e_3) = 0$. ■

9C Determinants

Exercise 9C1

Prove or give a counterexample: $S, T \in \mathcal{L}(V) \implies \det(S+T) = \det S + \det T$.

Solution. This statement is false in general. Consider $S = T = I$ with $\dim V = 2$. For any $\alpha \in V_{\text{alt}}^{(2)}$, we have $\alpha_I = \alpha$, and

$$\alpha_{2I}(u, v) = \alpha(2u, 2v) = 4\alpha(u, v).$$

Hence, $\det(S+T) = \det(2I) = 4 \neq 1 + 1 = \det S + \det T$. ∎

Exercise 9C2

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. Let A be n-by-n with $n > 1$ so that B is well-defined. Let $\Lambda \subseteq \operatorname{perm} n$ denote the set of permutations with 1 in the first position. In 9.46, the factor $A_{j_1,1} = 0$ unless $j_1 = 1$. Therefore, the only permutations that contribute to the sum are of the form $\sigma = (1, j_2, \ldots, j_n)$, so

$$\det A = A_{1,1} \sum_{\sigma \in \Lambda} (\operatorname{sign} \sigma) \prod_{k=2}^{n} A_{j_k,k}.$$

Let $\ell_p = j_{p+1} - 1$ for $1 \leq p \leq n-1$, and let $\tau = (\ell_1, \ldots, \ell_{n-1})$. Then the map $\sigma \mapsto \tau$ is a bijection from Λ to $\operatorname{perm}(n-1)$. By 9.32, no pair $(1, j_p)$ with $p \geq 2$ contributes to the sign. Moreover, for $2 \leq q < r$, the value q appears after r in σ if and only if $q-1$ appears after $r-1$ in τ, which holds because τ is σ with the first element removed and each remaining element decreased by 1. Therefore, $\operatorname{sign} \sigma = \operatorname{sign} \tau$. Because $B_{p,q} = A_{p+1,q+1}$ for $1 \leq p, q \leq n-1$, we have

$$\prod_{q=1}^{n-1} B_{\ell_q,q} = \prod_{k=2}^{n} A_{j_k,k}.$$

Hence,

$$\det A = A_{1,1} \sum_{\tau \in \operatorname{perm}(n-1)} (\operatorname{sign} \tau) \prod_{q=1}^{n-1} B_{\ell_q,q} = A_{1,1} \det B.$$

∎

Exercise 9C3

Suppose $T \in \mathcal{L}(V)$ is nilpotent. Prove that $\det(I+T) = 1$.

Solution. By 8.18, V has a basis with respect to which $\mathcal{M}(T)$ is strictly upper triangular. By 3.35, $\mathcal{M}(I+T) = I + \mathcal{M}(T)$ is upper triangular with all diagonal entries equal to 1. By 9.53 and 9.48, $\det(I+T) = 1$. ∎

Exercise 9C4

Suppose V is an inner product space and $S \in \mathcal{L}(V)$. Prove that S is unitary if and only if $|\det S| = \|S\| = 1$.

Solution. Let $s_1 \geq \cdots \geq s_n$ denote the singular values of S. By 7.88(a) and 9.60, $\|S\| = s_1$ and $|\det S| = s_1 \cdots s_n$.
$\boxed{\Rightarrow}$ Suppose S is unitary. Then S is an isometry, so 7.69 gives $s_1 = \cdots = s_n = 1$, and hence $|\det S| = s_1 \cdots s_n = 1$ and $\|S\| = s_1 = 1$.

$\boxed{\Leftarrow}$ Suppose $|\det S| = \|S\| = 1$. Then $s_1 = \|S\| = 1$ and $s_1 \cdots s_n = |\det S| = 1$. Because $0 \le s_k \le s_1 = 1$ for all k, we must have $s_k = 1$ for all k. Since $\det S \neq 0$, 9.50 yields that S is invertible. Therefore, S is an invertible isometry by 7.69, and hence unitary. ■

Exercise 9C5

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. Proceed by induction on m. The base case $m = 1$ is trivial, since $\det A = \det A_1$. Suppose $m > 1$ and the result holds for block upper-triangular matrices with $m - 1$ diagonal blocks. Write A as

$$A = \begin{pmatrix} A_1 & * \\ 0 & B \end{pmatrix},$$

where B is block upper triangular with diagonal blocks $A_2, \ldots, A_m$. By the induction hypothesis, it suffices to show that $\det A = (\det A_1)(\det B)$. Let A_1 be n-by-n, B be p-by-p, and let $e_1, \ldots, e_n$ denote the standard basis of $\mathbf{F}^n$. For $v_1, \ldots, v_n \in \mathbf{F}^n$, let $\alpha(v_1, \ldots, v_n)$ be the determinant of A with the first n columns replaced by $(v_1, 0), \ldots, (v_n, 0)$, where 0 denotes the zero vector in $\mathbf{F}^p$. Since the map $v \mapsto (v, 0)$ is linear from $\mathbf{F}^n$ to $\mathbf{F}^{n+p}$, 9.45 implies that α is an alternating n-linear form on $\mathbf{F}^n$. By 9.36 applied to the standard basis and 9.46, we have

$$\alpha(v_1, \ldots, v_n) = \alpha(e_1, \ldots, e_n) \det(v_1, \ldots, v_n).$$

Observe that

$$\alpha(e_1, \ldots, e_n) = \det \begin{pmatrix} I & * \\ 0 & B \end{pmatrix} = \det B,$$

where Exercise 9C2 is applied n times to expand the determinant along the first n columns. By taking $v_1, \ldots, v_n$ to be the columns of A_1, we get $\det A = (\det A_1)(\det B)$, as desired. ■

Exercise 9C6

Suppose $A = \begin{pmatrix} v_1 & \cdots & v_n \end{pmatrix}$ is an n-by-n matrix, with v_k denoting the kth column of A. Show that if $(m_1, \ldots, m_n) \in \operatorname{perm} n$, then

$$\det \begin{pmatrix} v_{m_1} & \cdots & v_{m_n} \end{pmatrix} = (\operatorname{sign}(m_1, \ldots, m_n)) \det A.$$

Solution. By 9.45, $\alpha(u_1, \ldots, u_n) = \det \begin{pmatrix} u_1 & \cdots & u_n \end{pmatrix}$ is an alternating n-linear form on $\mathbf{F}^n$. Applying 9.35 to α and the list $v_1, \ldots, v_n$, we get

$$\alpha(v_{m_1}, \ldots, v_{m_n}) = \operatorname{sign}(m_1, \ldots, m_n) \alpha(v_1, \ldots, v_n),$$

where $\alpha(v_1, \ldots, v_n) = \det A$, which completes the proof. ■

Exercise 9C7

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. Let $n = \dim V$. By 9.65, $p(0) = (-1)^n \det T$, which is nonzero by 9.50. Note that for $z \neq 0$,

$$zI - T^{-1} = (-zI)T^{-1}\left(\frac{1}{z}I - T\right).$$

By 9.49(a), 9.53, 9.48, and 9.50, we get

$$\begin{aligned} q(z) &= \det\left((-zI)T^{-1}\left(\frac{1}{z}I - T\right)\right) \\ &= \det(-zI)\det(T^{-1})\det\left(\frac{1}{z}I - T\right) \\ &= (-z)^n \det(T^{-1}) p\left(\frac{1}{z}\right) \\ &= \frac{z^n}{(-1)^n \det T} p\left(\frac{1}{z}\right) \\ &= \frac{z^n}{p(0)} p\left(\frac{1}{z}\right). \end{aligned}$$

■

Exercise 9C8

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. Let $n = \dim V$ and let p be the characteristic polynomial of T. By 9.65, p is monic with $\deg p = n$ and $p(0) = (-1)^n \det T$. By 9.51, p has no real zeros, so 4.16 yields that p has no real linear factors, and

$$p(z) = \prod_{j=1}^{M} (z^2 + b_j z + c_j),$$

where $b_j, c_j \in \mathbf{R}$ and $b_j^2 < 4c_j$ for each j. Comparing degrees gives $n = 2M$, and since $b_j^2 < 4c_j$, we have $c_j > 0$ for each j, so

$$\det T = p(0) = \prod_{j=1}^{M} c_j > 0.$$

■

Exercise 9C9

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. Let $n = \dim V$ and let p be the characteristic polynomial of T. By 9.65, p is monic with $\deg p = n$ and $p(0) = (-1)^n \det T$. By 4.16, we may write

$$p(z) = \prod_{j=1}^{m} (z - \lambda_j) \prod_{k=1}^{M} (z^2 + b_k z + c_k),$$

where $\lambda_j \in \mathbf{R}$, $b_k, c_k \in \mathbf{R}$, and $c_k > b_k^2/4 \geq 0$ for each k. Since $\deg p$ is even and the second product has degree $2M$, it follows that m must also be even. Evaluating p at 0 gives

$$p(0) = \prod_{j=1}^{m}(-\lambda_j)\prod_{k=1}^{M} c_k < 0.$$

By 9.51, each λ_j is an eigenvalue of T. Suppose T has at most one eigenvalue. Then either $m = 0$, in which case $p(0) = \prod_{k=1}^{M} c_k \geq 0$, or $m \geq 2$, in which case all λ_j are equal to some λ, and $p(0) = \prod_{j=1}^{m}(-\lambda)\prod_{k=1}^{M} c_k \geq 0$. In either case, we get a contradiction with $p(0) < 0$, so T must have at least two distinct eigenvalues. ■

Exercise 9C10

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.

Axler Note. *This result was previously proved without using determinants or the characteristic polynomial—see 5.34.*

Solution. Let $n = \dim V$ and let p be the characteristic polynomial of T. By 9.65, p is monic with $\deg p = n$. By 4.16, we may write

$$p(z) = \prod_{j=1}^{m}(z - \lambda_j)\prod_{k=1}^{M}(z^2 + b_k z + c_k),$$

where $\lambda_j \in \mathbf{R}$, $b_k, c_k \in \mathbf{R}$, and $b_k^2 < 4c_k$ for each k. Comparing degrees, we have $n = m + 2M$, which forces $m \geq 1$ since n is odd. Then $p(\lambda_1) = 0$, so λ_1 is an eigenvalue of T by 9.51. ■

Exercise 9C11

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.

Axler Note. *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. Suppose V is a real vector space with $\dim V = 2$ and basis v_1, v_2. By the linear map lemma, define $T \in \mathcal{L}(V)$ by $Tv_1 = -v_1$ and $Tv_2 = -2v_2$. Then $\mathcal{M}(T)$ with respect to the basis v_1, v_2 is diagonal with diagonal entries -1 and -2, so $\det T = 2 > 0$ by 9.53 and 9.48. Moreover, for any $\lambda, \mu \in \mathbf{R}$, we have

$$(T + I)(\lambda v_1 + \mu v_2) = \lambda(-v_1 + v_1) + \mu(-2v_2 + v_2) = -\mu v_2,$$

so $E(-1, T) = \operatorname{null}(T + I) = \operatorname{span} v_1$.

Suppose T has a square root S, so $S^2 = T$. Then $ST = S^3 = TS$, so $E(-1, T)$ is invariant under S by 5.75, and $Sv_1 = \nu v_1$ for some $\nu \in \mathbf{R}$. However, $-v_1 = Tv_1 = S^2 v_1 = S(\nu v_1) = \nu S v_1 = \nu^2 v_1$ gives $v_1 = 0$ or $\nu^2 = -1$. Since v_1 is a basis vector and $\nu \in \mathbf{R}$, this is impossible. Therefore, T cannot have a square root. ■

Exercise 9C12

Suppose $S, T \in \mathcal{L}(V)$ and S is invertible. Define $p\colon \mathbf{F} \to \mathbf{F}$ 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. Note that $zS - T = S(zI - S^{-1}T)$, so by 9.49(a), $p(z) = (\det S)q(z)$, where $q(z)$ is the characteristic polynomial of $S^{-1}T$. By 9.65, q is monic with $\deg q = \dim V$. By 9.50, $\det S \neq 0$, so p is a polynomial of degree $\dim V$ and the coefficient of $z^{\dim V}$ in p is $\det S$. ■

Exercise 9C13

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. By 9.62, the characteristic polynomial of T is

$$\det(zI - T) = \prod_{k=1}^{n}(z - \lambda_k).$$

In general, the coefficient of z^{n-k} in the characteristic polynomial of T is $(-1)^k$ times the sum of all products of k distinct eigenvalues of T.

(a) With $r = 2$, the coefficient of z^{n-2} is

$$(-1)^2 \sum_{1 \le j < k \le n} \lambda_j \lambda_k = \sum_{1 \le i < j \le n} \lambda_i \lambda_j.$$

(b) With $r = n - 1$, the coefficient of z is

$$(-1)^{n-1} \sum_{1 \le j_1 < \cdots < j_{n-1} \le n} \lambda_{j_1} \cdots \lambda_{j_{n-1}}.$$

■

Exercise 9C14

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 7.39, $(\sqrt{T})^2 = T$, so $\det T = (\det \sqrt{T})^2$ by 9.49(a). Since $\sqrt{T}$ is positive, $\det \sqrt{T} \ge 0$ by 9.59, so taking square roots gives $\det \sqrt{T} = \sqrt{\det T}$. ■

Exercise 9C15

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, $T = S\sqrt{T^*T}$ for some unitary operator S, where T^*T is positive by 7.64(a). By 9.58 and 9.59, $|\det S| = 1$ and $\det \sqrt{T^*T} \ge 0$. Therefore, 9.49(a) gives

$$|\det T| = |\det S||\det \sqrt{T^*T}| = \det \sqrt{T^*T},$$

and Exercise 9C14 gives $\det \sqrt{T^*T} = \sqrt{\det(T^*T)}$. ■

Exercise 9C16

Suppose $T \in \mathcal{L}(V)$. Define $g\colon \mathbf{F} \to \mathbf{F}$ by $g(x) = \det(I + xT)$. Show that $g'(0) = \operatorname{tr} T$.

Axler Hint. *Look for a clean solution to this exercise, without using the explicit but complicated formula for the determinant of a matrix.*

Solution. Let $n = \dim V$. By 9.65 and 8.56 applied to $-T$,

$$\det(zI + T) = z^n + (\operatorname{tr} T)z^{n-1} + \sum_{k=0}^{n-2} \lambda_k z^k$$

for $\lambda_0, \dots, \lambda_{n-2} \in \mathbf{F}$. For nonzero x, we have $I + xT = x(\frac{1}{x}I + T)$. By 9.53, 3.38, and 9.57(c) applied to each column of $x\mathcal{M}(\frac{1}{x}I + T)$, we have $\det(x(\frac{1}{x}I + T)) = x^n \det(\frac{1}{x}I + T)$. Using the substitution $z = \frac{1}{x}$ and the fact that $g(0) = \det I = 1$ by 9.53 and 9.48, we have for all $x \in \mathbf{F}$ that

$$g(x) = 1 + (\operatorname{tr} T)x + \sum_{k=2}^{n} \lambda_{n-k} x^k.$$

Since each exponent in the sum is at least 2, evaluating $(x^k)'$ at $x = 0$ gives 0 for each $k \geq 2$, so $g'(0) = \operatorname{tr} T$. ∎

Exercise 9C17

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 be the ellipsoid in question. Let Ω be the unit ball in $\mathbf{R}^3$ with volume$(\Omega) = \frac{4}{3}\pi$. Define $T \in \mathcal{L}(\mathbf{R}^3)$ by $T(x, y, z) = (ax, by, cz)$. We claim that $T(\Omega) = E$.

$\boxed{\subseteq}$ Suppose $(x, y, z) \in \Omega$. Then $x^2 + y^2 + z^2 < 1$, so

$$\frac{(ax)^2}{a^2} + \frac{(by)^2}{b^2} + \frac{(cz)^2}{c^2} = x^2 + y^2 + z^2 < 1.$$

Hence, $(ax, by, cz) \in E$, so $T(\Omega) \subseteq E$.

$\boxed{\supseteq}$ Suppose $(p, q, r) \in E$. Then

$$\frac{p^2}{a^2} + \frac{q^2}{b^2} + \frac{r^2}{c^2} = \left(\frac{p}{a}\right)^2 + \left(\frac{q}{b}\right)^2 + \left(\frac{r}{c}\right)^2 < 1,$$

so $(\frac{p}{a}, \frac{q}{b}, \frac{r}{c}) \in \Omega$, and thus $T(\frac{p}{a}, \frac{q}{b}, \frac{r}{c}) = (p, q, r) \in T(\Omega)$, whence $E \subseteq T(\Omega)$.

Therefore, $T(\Omega) = E$. With respect to the standard basis of $\mathbf{R}^3$, the matrix of T is diagonal with diagonal entries a, b, and c, so $\det T = abc$ by 9.53 and 9.48. By 9.61, volume $E = |\det T|(\text{volume } \Omega) = \frac{4}{3}abc\pi$. ∎

Exercise 9C18

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. Since A is invertible, $Ax = 0$ only when $x = 0$, so the columns $v_1, \dots, v_n$ of A are linearly independent by 3.50. By QR factorization, we may write $A = QR$, where Q is unitary and R is upper triangular with positive diagonal entries.

Let $e_1, \dots, e_n$ be the standard basis of $\mathbf{F}^n$ and let $T \in \mathcal{L}(\mathbf{F}^n)$ be such that $\mathcal{M}(T, (e_1, \dots, e_n)) = Q$. By 7.57(d), $Q^*Q = I$, where $(Q^*Q)_{j,k} = \langle Q_{\cdot,k}, Q_{\cdot,j} \rangle$, so the columns of Q are orthonormal, and hence form an orthonormal basis of $\mathbf{F}^n$ by 6.28. By 3.75, $Te_k = Q_{\cdot,k}$ for each $k = 1, \dots, n$, so T is unitary by 7.53(d). By 9.43 and 9.58, $|\det Q| = |\det T| = 1$. By 9.49(b), 9.48, and $R_{k,k} > 0$ for each k, we have

$$|\det A| = |\det Q||\det R| = \prod_{k=1}^{n} R_{k,k}.$$

Since $v_k = QR_{\cdot,k}$ by 3.48, we have $\|v_k\| = \|R_{\cdot,k}\|$ by 7.57(c). Because R is upper triangular,

$$\|R_{\cdot,k}\|^2 = R_{k,k}^2 + \sum_{j<k} |R_{j,k}|^2.$$

Since $0 < R_{k,k} \le \|R_{\cdot,k}\|$ for each k, Hadamard's inequality for A can be read as

$$|\det A| = \prod_{k=1}^{n} R_{k,k} \le \prod_{k=1}^{n} \|R_{\cdot,k}\| = \prod_{k=1}^{n} \|v_k\|.$$

The products are equal if and only if $\|R_{\cdot,k}\| = R_{k,k}$ for each $k = 1, \dots, n$, i.e., $R_{j,k} = 0$ for all k and $j < k$ if and only if R is diagonal.

$\boxed{\Rightarrow}$ Suppose equality holds in Hadamard's inequality for A. Then R is diagonal, so $v_k = R_{k,k}Q_{\cdot,k}$ for each $k = 1, \dots, n$. Since the columns of Q are orthonormal, $\langle v_j, v_k \rangle = 0$ for all $j \ne k$, i.e., each column of A is orthogonal to the other columns.

$\boxed{\Leftarrow}$ Suppose the columns $v_1, \dots, v_n$ of A are mutually orthogonal, and let $e_1, \dots, e_n$ be the standard orthonormal basis of $\mathbf{F}^n$. Let $T \in \mathcal{L}(\mathbf{F}^n)$ be such that $\mathcal{M}(T, (e_1, \dots, e_n)) = A$. By 7.9 and 3.43, $\mathcal{M}(T^*T) = A^*A$. By 9.43, 9.53, and 9.60, $|\det A| = |\det T| = \sqrt{\det(T^*T)} = \sqrt{\det(A^*A)}$. By 3.46, the (j,k)th entry of A^*A is $\langle v_k, v_j \rangle$. By orthogonality of the columns, the off-diagonal entries of A^*A are zero, and the diagonal entries are $\langle v_k, v_k \rangle = \|v_k\|^2$. By 9.48,

$$\det(A^*A) = \prod_{k=1}^{n} \|v_k\|^2.$$

Taking square roots yields the desired result. ∎

Exercise 9C19

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 $\det T \le \prod_{k=1}^{n} \langle Te_k, e_k \rangle$.

(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. Let $A = \mathcal{M}(\sqrt{T}, (e_1, \dots, e_n))$. Since $\sqrt{T}$ is positive, it is self-adjoint. By 3.75, $A_{\cdot,k} = \mathcal{M}(\sqrt{T}e_k)$ for each $k = 1, \dots, n$. By 6.30(c) and self-adjointness,

$$\langle A_{\cdot,k}, A_{\cdot,j} \rangle = \langle \sqrt{T}e_k, \sqrt{T}e_j \rangle = \langle Te_k, e_j \rangle,$$

and therefore, $\|A_{\cdot,k}\|^2 = \langle Te_k, e_k \rangle$ for each $k = 1, \dots, n$. By 9.53 and Exercise 9C14, $\det A = \det \sqrt{T} = \sqrt{\det T} \ge 0$, so $\det T = |\det A|^2$.

(a) By Hadamard's inequality,

$$\det T = |\det A|^2 \le \prod_{k=1}^{n} \|A_{\cdot,k}\|^2 = \prod_{k=1}^{n} \langle Te_k, e_k \rangle.$$

(b) $\boxed{\Rightarrow}$ Suppose the inequality in (a) is an equality. Since T is invertible, $\det T \neq 0$, so $\det A = \sqrt{\det T} \neq 0$, and thus $\sqrt{T}$ is invertible by 9.50, whence A is also invertible by 3.86. Taking nonnegative square roots, equality holds in Hadamard's inequality for A. Thus, $\langle Te_k, e_j \rangle = 0$ for all $j \neq k$ by Exercise 9C18. By 6.30(a), $Te_k = \langle Te_k, e_k \rangle e_k$ for each $k = 1, \dots, n$, so e_k is an eigenvector of T.

$\boxed{\Leftarrow}$ Suppose e_k is an eigenvector of T for each $k = 1, \dots, n$. Then $Te_k = \lambda_k e_k$ for some $\lambda_k \in \mathbf{F}$ and each $k = 1, \dots, n$. It follows that $\mathcal{M}(T, (e_1, \dots, e_n))$ is diagonal with diagonal entries $\lambda_k = \langle Te_k, e_k \rangle$ for each $k = 1, \dots, n$. By 9.53 and 9.48,

$$\det T = \prod_{k=1}^{n} \lambda_k = \prod_{k=1}^{n} \langle Te_k, e_k \rangle. \qquad \blacksquare$$

Exercise 9C20

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

$$|\det A| \leq c^n n^{n/2}.$$

Axler Note. *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. Note that $c \geq |A_{1,1}| \geq 0$. For each $k = 1, \dots, n$, the column v_k of A satisfies

$$\|v_k\|^2 = \sum_{j=1}^{n} |A_{j,k}|^2 \leq c^2 n.$$

By Hadamard's inequality,

$$|\det A| \leq \prod_{k=1}^{n} \|v_k\| \leq \prod_{k=1}^{n} \sqrt{c^2 n} = c^n n^{n/2}. \qquad \blacksquare$$

Exercise 9C21

Suppose n is a positive integer and $\delta \colon \mathbf{C}^{n,n} \to \mathbf{C}$ 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 $A \in \mathbf{C}^{n,n}$.

Axler Note. *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 8D10, which shows that the trace is uniquely determined by its algebraic properties.*

Solution. Let $e_1, \dots, e_n$ be the standard basis of $\mathbf{C}^n$. Since I is diagonal, $\delta(I) = 1$. For every invertible matrix $P \in \mathbf{C}^{n,n}$, we have $\delta(P)\delta(P^{-1}) = \delta(PP^{-1}) = \delta(I) = 1$, so $\delta(P^{-1}XP) = \delta(X)$ for all $X \in \mathbf{C}^{n,n}$. Note that $\det I = 1$ by 9.48, so by 9.49(b), $\det(P^{-1}XP) = \det(P^{-1})(\det X)(\det P) = \det X$ for all $X \in \mathbf{C}^{n,n}$ and invertible P. Therefore, $\delta(X) = \det X$ whenever $X = P^{-1}DP$ for some invertible P and diagonal matrix D. We split by whether X is invertible or not.

- Suppose X is invertible. By 5.47 and 3.84, $X = P^{-1}YP$ for some invertible P and upper-triangular matrix Y, where Y has diagonal entries $\lambda_1, \dots, \lambda_n$. Since $\det Y = \det X \neq 0$ by 9.48 and 9.50, $\lambda_k \neq 0$ for all k. Let D be diagonal

with diagonal entries λ_k/k for each $k = 1, \ldots, n$. Then $Y' = D^{-1}Y$ is upper triangular with entries $1, \ldots, n$ on the diagonal. By 5.41, the operator $T \in \mathcal{L}(\mathbf{C}^n)$ such that $\mathcal{M}(T, (e_1, \ldots, e_n)) = Y'$ has eigenvalues $1, \ldots, n$. By 5.58, T is diagonalizable, so 3.84 gives $Y' = Q^{-1}D'Q$ for invertible Q and diagonal D'. By the beginning calculation, $\det Y' = \det D'$, so $\det Y = (\det D)(\det Y') = (\det D)(\det D')$ by 9.49(b). Therefore,

$$\delta(X) = \delta(P^{-1}YP) = \delta(Y) = \delta(D)\delta(Y') = (\det D)(\det D') = \det Y = \det X.$$

- Suppose X is not invertible. Let $S \in \mathcal{L}(\mathbf{C}^n)$ be such that $\mathcal{M}(S, (e_1, \ldots, e_n)) = X$. By 3.86, S is not invertible, so it is not injective by 3.65. Then there exists nonzero $v \in \mathbf{C}^n$ such that $Sv = 0$, whence $Xv = \mathcal{M}(Sv) = 0$ by 3.76. By 2.32, extend v to a basis $v, v_2, \ldots, v_n$ of $\mathbf{C}^n$, and let Q be the matrix whose columns are $v, v_2, \ldots, v_n$. Since $Q = \mathcal{M}(I, (v, v_2, \ldots, v_n), (e_1, e_2, \ldots, e_n))$, 3.82 gives that Q is invertible. Let $M = Q^{-1}XQ$. Then M has the first column equal to 0, and thus $M = ME$, where E is diagonal with diagonal entries $0, 1, \ldots, 1$. By 9.53 and 9.50, we have $\det X = \det S = 0$. Therefore,

$$\delta(X) = \delta(M) = \delta(M)\delta(E) = \delta(M) \cdot 0 = 0 = \det X.$$

■

9D Tensor Products

Exercise 9D1

Suppose $v \in V$ and $w \in W$. Prove that $v \otimes w = 0$ if and only if $v = 0$ or $w = 0$.

Solution.

$\boxed{\Rightarrow}$ We proceed by contrapositive. Suppose $v \neq 0$ and $w \neq 0$ and let $n = \dim V$. Then the list v is linearly independent, so it extends to a basis $v, v_2, \ldots, v_n$ of V by 2.32. Let $\varphi \in V'$ be the first element of the dual basis of $v, v_2, \ldots, v_n$, so $\varphi(v) = 1$. A similar argument applies to w to obtain $\psi \in W'$ such that $\psi(w) = 1$. Then

$$(v \otimes w)(\varphi, \psi) = \varphi(v)\psi(w) = 1 \neq 0,$$

so $v \otimes w \neq 0$.

$\boxed{\Leftarrow}$ Suppose $v = 0$ or $w = 0$. Then for all $\varphi \in V'$ and $\psi \in W'$, we have

$$(v \otimes w)(\varphi, \psi) = \varphi(v)\psi(w) = 0,$$

and so $v \otimes w = 0$. ■

Exercise 9D2

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 e_1, e_2, e_3 be the standard basis of $\mathbf{R}^3$. Define the vectors

$$\begin{aligned} v_1 &= e_1, & v_2 &= e_2, & v_3 &= e_1 + 2e_2, \\ w_1 &= e_3, & w_2 &= 2e_3, & w_3 &= -e_3. \end{aligned}$$

Let $x = e_1 \otimes e_3$ and $y = e_2 \otimes e_3$. By 9.73, $v_1 \otimes w_1 = x$, $v_2 \otimes w_2 = 2y$, and $v_3 \otimes w_3 = -x - 2y$, which sum to 0. By 9.74(a), the list x, y is linearly independent. Note that $\lambda x + \mu(-x - 2y) = (\lambda - \mu)x - 2\mu y$, and $\lambda(2y) + \mu(-x - 2y) = -\mu x + (2\lambda - 2\mu)y$ for $\lambda, \mu \in \mathbf{R}$. Since x and y are linearly independent, these expressions can only be zero if $\lambda = \mu = 0$. Therefore, no two of $x, 2y, -x - 2y$ are scalar multiples of each other. ■

Exercise 9D3 ✔

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. Let $n = \dim W$ and let $u_1, \ldots, u_n$ be a basis of W. Write $w_k = \sum_{j=1}^{n} \lambda_{j,k} u_j$ for $\lambda_{j,k} \in \mathbf{F}$. By 9.73, we have

$$0 = \sum_{k=1}^{m} v_k \otimes w_k = \sum_{k=1}^{m} v_k \otimes \sum_{j=1}^{n} \lambda_{j,k} u_j = \sum_{k=1}^{m} \sum_{j=1}^{n} \lambda_{j,k} (v_k \otimes u_j).$$

By 9.74(a), the list $\{v_k \otimes u_j\}_{k=1,\dots,m;j=1,\dots,n}$ is linearly independent. Therefore, $\lambda_{j,k} = 0$ for all j and k, and hence $w_k = 0$ for all k. ■

Exercise 9D4 ✔

Suppose $\dim V > 1$ and $\dim W > 1$. Prove that

$$\{v \otimes w : (v, w) \in V \times W\}$$

is not a subspace of $V \otimes W$.

Axler Note. *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 $n = \dim V$ and $m = \dim W$. Fix bases $v_1, \dots, v_n$ of V and $w_1, \dots, w_m$ of W. Let $u = v_1 \otimes w_1 + v_2 \otimes w_2$. We claim that $u \neq v \otimes w$ for all $(v, w) \in V \times W$. To see this, write $v = \sum_{k=1}^{n} \lambda_k v_k$. By 9.73,

$$v \otimes w - u = \sum_{k=1}^{n} v_k \otimes x_k, \quad x_k = \begin{cases} \lambda_k w - w_k, & k = 1, 2 \\ \lambda_k w, & k > 2. \end{cases}$$

By 2.22, at least one of w_1 or w_2 is not in $\operatorname{span}(w)$. If $w_k \notin \operatorname{span}(w)$, then $x_k \neq 0$, so $v \otimes w - u \neq 0$ by Exercise 9D3. Therefore, the set contains $v_1 \otimes w_1$ and $v_2 \otimes w_2$, but not their sum u, so it is not closed under addition and hence not a subspace of $V \otimes W$. ■

Exercise 9D5

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.

$\boxed{\subseteq}$ Note that the kth column of $v \otimes w$ is $w_k v$. By 2.6, the span of the columns is a subspace of $\operatorname{span}(v)$, so the rank of $v \otimes w$ is at most $\dim \operatorname{span}(v)$ by 2.37. Since $\dim \operatorname{span}(v) \leq 1$ by 2.30, the rank of $v \otimes w$ is at most one.

$\boxed{\supseteq}$ Suppose A is an m-by-n matrix of rank at most one. If $A = 0$, then $A = 0 \otimes 0$. Otherwise, $\operatorname{rank} A = 1$, so by 3.56, $A = BC$ for $B \in \mathbf{F}^{m,1}$ and $C \in \mathbf{F}^{1,n}$. Let $x = (B_{1,1}, \dots, B_{m,1})$ and $y = (C_{1,1}, \dots, C_{1,n})$. By 3.46, $A_{j,k} = B_{j,1} C_{1,k}$ for all j and k, so $A = x \otimes y$ by 9.76. ■

Exercise 9D6

Suppose m and n are positive integers. Give a description, analogous to Exercise 9D5, of the set of m-by-n matrices (with entries in $\mathbf{F}$) that have rank at most two.

Solution. Identify each element of $\mathbf{F}^m \otimes \mathbf{F}^n$ with its coefficient matrix with respect to the basis of 9.74(b). Note that this identification is linear and extends 9.76. We claim that the set

$$\{v_1 \otimes w_1 + v_2 \otimes w_2 : v_1, v_2 \in \mathbf{F}^m, w_1, w_2 \in \mathbf{F}^n\}$$

is the set of m-by-n matrices that have rank at most two.

$\boxed{\subseteq}$ Let $w_{p,q}$ denote the qth entry of w_p for $p = 1, 2$ and $q = 1, \ldots, n$, and let $u = v_1 \otimes w_1 + v_2 \otimes w_2$. Note that the kth column of u is $w_{1,k}v_1 + w_{2,k}v_2$. By 2.6, the span of the columns is a subspace of $\operatorname{span}(v_1, v_2)$, so the rank of u is at most $\dim \operatorname{span}(v_1, v_2)$ by 2.37. Since $\dim \operatorname{span}(v_1, v_2) \leq 2$ by 2.30, the rank of u is at most two.
$\boxed{\supseteq}$ Suppose A is an m-by-n matrix of rank at most two. If A has rank at most one, then Exercise 9D5 and Exercise 9D1 imply that $A = v \otimes w + 0 \otimes 0$ for some $v \in \mathbf{F}^m$ and $w \in \mathbf{F}^n$. If A has rank two, then by 3.56, $A = BC$ for some $B \in \mathbf{F}^{m,2}$ and $C \in \mathbf{F}^{2,n}$. Let x_ℓ be the ℓth column of B and y_ℓ be the ℓth row of C for $\ell = 1, 2$. By 3.46, $A_{j,k} = B_{j,1}C_{1,k} + B_{j,2}C_{2,k}$ for all j and k, so $A = x_1 \otimes y_1 + x_2 \otimes y_2$. ■

Exercise 9D7 ✔

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. Let $n = \dim V$ and $m = \dim W$. Fix bases $e_1, \ldots, e_n$ of V and $f_1, \ldots, f_m$ of W. Let $u = e_1 \otimes f_1 + e_2 \otimes f_2 + e_3 \otimes f_3$. As in Exercise 9D4, we claim that $u \neq v_1 \otimes w_1 + v_2 \otimes w_2$ for any $v_1, v_2 \in V$ and $w_1, w_2 \in W$. Indeed, write $v_j = \sum_{k=1}^n \lambda_{j,k} e_k$. By 9.73,

$$v_1 \otimes w_1 + v_2 \otimes w_2 - u = \sum_{k=1}^n e_k \otimes x_k, \quad x_k = \begin{cases} \lambda_{1,k}w_1 + \lambda_{2,k}w_2 - f_k & \text{for } k \leq 3, \\ \lambda_{1,k}w_1 + \lambda_{2,k}w_2 & \text{for } k > 3. \end{cases}$$

By 2.22, not all of f_1, f_2, f_3 lie in $\operatorname{span}(w_1, w_2)$. Choose $k \leq 3$ such that $f_k \notin \operatorname{span}(w_1, w_2)$. Since $x_k + f_k \in \operatorname{span}(w_1, w_2)$, if $x_k = 0$, then $f_k \in \operatorname{span}(w_1, w_2)$, which is a contradiction, so $x_k \neq 0$. Hence, $v_1 \otimes w_1 + v_2 \otimes w_2 - u \neq 0$ by Exercise 9D3, and $u \notin \{v_1 \otimes w_1 + v_2 \otimes w_2 : v_1, v_2 \in V \text{ and } w_1, w_2 \in W\}$. ■

Exercise 9D8

Suppose $v_1, \ldots, v_m \in V$ and $w_1, \ldots, w_m \in W$ 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. By 9.79(a), 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 $w \in W$. By linearity and 3.10,

$$\sum_{k=1}^m \Gamma(v_k, w_k) = \sum_{k=1}^m \hat{\Gamma}(v_k \otimes w_k) = \hat{\Gamma}\left(\sum_{k=1}^m v_k \otimes w_k\right) = \hat{\Gamma}(0) = 0.$$ ■

Exercise 9D9

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 $w \in W$.
Axler Note. *In an abuse of notation, the operator on $V \otimes W$ given by this exercise is often called $S \otimes T$.*

Solution. Note that the map $\Gamma : V \times W \to V \otimes W$ given by $\Gamma(v, w) = Sv \otimes Tw$ is bilinear, since each slot is linear by linearity of S and T and 9.73. By 9.79(a) with $U = V \otimes W$, there exists a unique linear map $\hat{\Gamma} : V \otimes W \to V \otimes W$ such that $\hat{\Gamma}(v \otimes w) = \Gamma(v, w) = Sv \otimes Tw$ for all $v \in V$ and $w \in W$. ■

Exercise 9D10

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 $(S \otimes T)^{-1} = S^{-1} \otimes T^{-1}$, where we are using the notation from the comment after Exercise 9D9.

Solution.
$\boxed{\Rightarrow}$ Suppose $S \otimes T$ is invertible, so it is injective. Fix $w \in W$ such that $w \neq 0$. If $Sv = 0$ for some $v \in V$, then Exercise 9D9 and Exercise 9D1 give that $(S \otimes T)(v \otimes w) = 0 \otimes Tw = 0$. Therefore, $v \otimes w = 0$ by 3.15. By Exercise 9D1, $v = 0$, so $\operatorname{null} S = \{0\}$, so S is injective by 3.15. A similar argument shows that T is injective, and thus S and T are invertible by 3.65.
$\boxed{\Leftarrow}$ Suppose S and T are invertible. Then for all $v \in V$ and $w \in W$,

$$(S \otimes T)(S^{-1} \otimes T^{-1})(v \otimes w) = (S \otimes T)(S^{-1}v \otimes T^{-1}w) = SS^{-1}v \otimes TT^{-1}w = v \otimes w.$$

Note that both operators $I_{V\otimes W}$ and $(S\otimes T)(S^{-1}\otimes T^{-1})$ send $v\otimes w$ to itself for all $v \in V$ and $w \in W$, so $(S\otimes T)(S^{-1}\otimes T^{-1}) = I_{V\otimes W}$ by the uniqueness in Exercise 9D9, with I_V and I_W in place of S and T. By 3.68, $(S^{-1} \otimes T^{-1})(S \otimes T) = I_{V\otimes W}$ as well, so $S \otimes T$ is invertible with inverse $S^{-1} \otimes T^{-1}$. ∎

Exercise 9D11

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 9D9.

Solution. For all $v_1, v_2 \in V$ and $w_1, w_2 \in W$, we have

$$\begin{aligned}\langle (S \otimes T)(v_1 \otimes w_1), v_2 \otimes w_2 \rangle &= \langle Sv_1 \otimes Tw_1, v_2 \otimes w_2 \rangle \\ &= \langle Sv_1, v_2 \rangle \langle Tw_1, w_2 \rangle \\ &= \langle v_1, S^*v_2 \rangle \langle w_1, T^*w_2 \rangle \\ &= \langle v_1 \otimes w_1, S^*v_2 \otimes T^*w_2 \rangle \\ &= \langle v_1 \otimes w_1, (S^* \otimes T^*)(v_2 \otimes w_2) \rangle,\end{aligned}$$

where the first equality is by Exercise 9D9 applied to S and T; the second and fourth equalities are by 9.80; and the last equality is by Exercise 9D9 applied to S^* and T^*. By 9.74(b), tensors of the form $v \otimes w$ for $v \in V$ and $w \in W$ span $V \otimes W$. By the linearity of $S \otimes T$ and $S^* \otimes T^*$ and by 6.6(a), (d), and (e), the equality $\langle (S \otimes T)x, y \rangle = \langle x, (S^* \otimes T^*)y \rangle$ extends to all $x, y \in V \otimes W$. Therefore, by the uniqueness of the adjoint, $(S \otimes T)^* = S^* \otimes T^*$. ∎

Exercise 9D12

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 $V_1 \times \cdots \times V_m$.
Axler Note. *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. By 6.35, there exists an orthonormal basis $e_1^k, \dots, e_{n_k}^k$ of V_k for each $k = 1, \dots, m$. For each $J = (j_1, \dots, j_m) \in \{1, \dots, n_1\} \times \cdots \times \{1, \dots, n_m\}$, define $E_J = e_{j_1}^1 \otimes \cdots \otimes e_{j_m}^m$. By 9.90, $\mathcal{E} = \{E_J\}_J$ is a basis of $V_1 \otimes \cdots \otimes V_m$. By the lemma proved in Exercise 7B24, there exists an inner product on $V_1 \otimes \cdots \otimes V_m$ such that $\mathcal{E}$ is an orthonormal basis. Then for any $w \in V_1 \otimes \cdots \otimes V_m$, $\langle w, E_J \rangle$ is the coefficient of E_J in the expansion of w in the basis $\mathcal{E}$ by 6.30(a) and 2.28. For any $(v_1, \dots, v_m), (u_1, \dots, u_m) \in V_1 \times \cdots \times V_m$, fix $x = v_1 \otimes \cdots \otimes v_m$ and $y = u_1 \otimes \cdots \otimes u_m$. Define

$$\lambda_J = \prod_{k=1}^m \langle v_k, e_{j_k}^k \rangle \quad \text{and} \quad \mu_J = \prod_{k=1}^m \langle u_k, e_{j_k}^k \rangle.$$

Since each $\varphi_k \in V_k'$ in 9.88 is linear, the map $(w_1, \dots, w_m) \mapsto w_1 \otimes \cdots \otimes w_m$ is multilinear, so expanding each v_k in x by 6.30(a) gives $x = \sum_J \lambda_J E_J$ and likewise $y = \sum_J \mu_J E_J$. Therefore, by the observation above, 4.4, and 6.30(c),

$$\langle x, y \rangle = \sum_J \lambda_J \overline{\mu_J} = \prod_{k=1}^m \sum_{j=1}^{n_k} \langle v_k, e_j^k \rangle \overline{\langle u_k, e_j^k \rangle} = \prod_{k=1}^m \langle v_k, u_k \rangle,$$

where the last equality follows from 6.30(c) applied in each V_k.

For any inner product that satisfies the condition, applying the condition to the tensors E_J and $E_{J'}$ fixes $\langle E_J, E_{J'} \rangle$ for all J and J'. By 9.90, 6.6(a), (d), and (e), the inner product is determined by its values on pairs $(E_J, E_{J'})$ for all J and J', so the inner product is uniquely determined. ∎

Exercise 9D13

Suppose that $V_1, \dots, V_m$ 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 9D12. Suppose $e_1^k, \dots, e_{n_k}^k$ is an orthonormal basis of V_k for each $k = 1, \dots, m$. Show that the list

$$\{e_{j_1}^1 \otimes \cdots \otimes e_{j_m}^m\}_{j_1 = 1, \dots, n_1; \dots; j_m = 1, \dots, n_m}$$

is an orthonormal basis of $V_1 \otimes \cdots \otimes V_m$.

Solution. By 9.90, the list in the statement is a basis of $V_1 \otimes \cdots \otimes V_m$. By Exercise 9D12,

$$\langle e_{j_1}^1 \otimes \cdots \otimes e_{j_m}^m, e_{\ell_1}^1 \otimes \cdots \otimes e_{\ell_m}^m \rangle = \prod_{k=1}^m \langle e_{j_k}^k, e_{\ell_k}^k \rangle = \prod_{k=1}^m \delta_{j_k, \ell_k} = \delta_{(j_1, \dots, j_m), (\ell_1, \dots, \ell_m)},$$

where $j_k, \ell_k \in \{1, \dots, n_k\}$ for each $k = 1, \dots, m$. Therefore, the list is orthonormal, and thus an orthonormal basis. ∎

A Appendix: Alternate Proofs of Exercises

A1 Vector Spaces

No alternate proofs yet!

A2 Finite-Dimensional Vector Spaces

Exercise 2A20 ✔

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 $\mathcal{P}_m(\mathbf{F})$.

blanketism. Since $p_k(2) = 0$ for each $k \in \{0, \ldots, m\}$, we have

$$p_k(x) = (x-2)q_k(x)$$

for some polynomial $q_k(x) \in \mathcal{P}_{m-1}(\mathbf{F})$. Then we may write

$$p_0(x), p_1(x), \ldots, p_m(x) = (x-2)q_0(x), (x-2)q_1(x), \ldots, (x-2)q_m(x).$$

Since $\mathcal{P}_{m-1}(\mathbf{F})$ can be spanned by the m polynomials $1, x, \ldots, x^{m-1}$, the list of $m+1$ polynomials $q_0, q_1, \ldots, q_m$ must be linearly dependent. Thus, there exist scalars $\alpha_0, \ldots, \alpha_m$, not all zero, such that

$$\alpha_0 q_0(x) + \cdots + \alpha_m q_m(x) = 0.$$

Then we may write

$$\alpha_0 p_0(x) + \cdots + \alpha_m p_m(x) = (x-2)(\alpha_0 q_0(x) + \cdots + \alpha_m q_m(x)) = 0.$$

Since not all the scalars $\alpha_0, \ldots, \alpha_m$ are zero, this shows that $p_0, p_1, \ldots, p_m$ is linearly dependent in $\mathcal{P}_m(\mathbf{F})$. ∎

Exercise 2C8 ✔

Suppose $v_1, \ldots, v_m$ is linearly independent in V and $w \in V$. Prove that

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

Civil Service Pigeon. Let $S = \operatorname{span}(v_1 + w, \ldots, v_m + w)$. Since each $v_i = (v_i + w) - w$, we have

$$S + \operatorname{span}(w) = \operatorname{span}(v_1, \ldots, v_m, w).$$

Since $v_1, \ldots, v_m$ are linearly independent, we have

$$m \leq \dim\big(S + \operatorname{span}(w)\big) \leq \dim S + \dim \operatorname{span}(w) \leq \dim S + 1 \implies \dim S \geq m - 1.$$ ∎

A3 Linear Maps

Exercise 3B29 ✔

Suppose $p \in \mathcal{P}(\mathbf{R})$. Prove that there exists a polynomial $q \in \mathcal{P}(\mathbf{R})$ such that $5q'' + 3q' = p$.

Coolempire93. Let $p(x) = a_0 + a_1x + \cdots + a_nx^n$ and suppose there exists a polynomial $q(x) = b_0 + b_1x + \cdots + b_mx^m$ satisfying $5q'' + 3q' = p$. If $\deg q = m \geq 2$, we can say q' has degree $m-1$ and q'' has degree $m-2$, so $5q'' + 3q'$ has degree $m-1$. To satisfy the given equation, then, we must have $m - 1 = n = \deg p$, and we may write $q(x) = b_0 + b_1x + \cdots + b_{n+1}x^{n+1}$. Before we deal with the general case, we will quickly consider when $n = 0$ (which would force $m < 2$):

If $p(x) = a_0$, then let $q(x) = b_0 + b_1x$. Because $5q'' + 3q' = p$, we have $5q'' + 3q' = 5(0) + 3b_1 = a_0$, so we set $b_1 = \frac{a_0}{3}$, and we have a polynomial q such that $5q'' + 3q' = p$ for any value of b_0. As we will see below, the constant term of q never depends on p.

For the remainder of this proof, we take $m = n + 1$ and suppose $n \geq 1$. We have

$$\begin{aligned} q'(x) &= b_1 + 2b_2x + \cdots + (n+1)b_{n+1}x^n \\ q''(x) &= 2b_2 + 6b_3x + \cdots + n(n+1)b_{n+1}x^{n-1} \end{aligned}$$

Equating coefficients of $5q'' + 3q'$ with those of p, we obtain the system

$$\begin{cases} a_0 = 3b_1 + 10b_2 \\ a_1 = 6b_2 + 30b_3 \\ \vdots \\ a_{n-1} = 3nb_n + 5n(n+1)b_{n+1} \\ a_n = 3(n+1)b_{n+1} \end{cases}.$$

That is, $a_i = 3(i+1)b_{i+1} + 5(i+1)(i+2)b_{i+2}$ for $0 \leq i < n$ and $a_n = 3(n+1)b_{n+1}$. Finally, we solve for the coefficients of q by back-substitution. Noting that

$$b_j = \frac{a_{j-1} - 5j(j+1)b_{j+1}}{3j} = \frac{a_{j-1}}{3j} - \frac{5}{3}(j+1)b_{j+1},$$

we have

$$\begin{cases} b_{n+1} = \frac{a_n}{3(n+1)} \\ b_n = \frac{a_{n-1}}{3n} - \frac{5}{3}(n+1)b_{n+1} = \frac{a_{n-1}}{3n} - \frac{5}{9}a_n \\ b_{n-1} = \frac{a_{n-2}}{3(n-1)} - \frac{5}{3}nb_n = \frac{a_{n-2}}{3(n-1)} - \frac{5}{9}a_{n-1} + \frac{25}{27}na_n \\ \vdots \\ b_1 = \frac{a_0 - 10b_2}{3} \end{cases}$$

which has a valid solution† since no division by 0 occurs. Finally, we note that b_0 is a free variable, so there is an infinite family of polynomials $q \in \mathcal{P}(\mathbf{R})$ with the given coefficients such that $5q'' + 3q' = p$. ∎

† Solving this recurrence is outside the scope of this course, but, specifically, the coefficients of q are

$$\begin{aligned} b_j &= \frac{a_{j-1}}{3j} - \frac{5}{9}a_j + \frac{25}{27}(j+1)a_{j+1} - \frac{125}{81}(j+1)(j+2)a_{j+2} + \cdots + (-1)^{n-j+1}\frac{5^{n-j+1}}{3^{n-j+2}}(j+1)(j+2)\cdots na_n \\ &= \frac{1}{3}\left(\frac{a_{j-1}}{j} + \sum_{i=j}^{n}\left(-\frac{5}{3}\right)^{i-j+1}\frac{i!}{j!}a_i\right). \end{aligned}$$

Exercise 3D19 ✔

Suppose V is 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.

The manual solution provides a constructive proof utilizing sums of the form $v_1 + v_k$ to eliminate the k-th column, along with $v_2 + v_1$ to eliminate the 1st column. Here, we present different constructions. Each proof assumes the declaration of the matrix to be A and that $\dim V \geq 2$.

This method shows that we can swap two *arbitrary* basis vectors to force the same value of all off-diagonal entries as well as the same value of all diagonal entries. Then, we can negate one basis vector to force all off-diagonal entries to be 0 via the linear independence of the basis vectors.

Coolempire93. Let i, j be arbitrary with $1 \leq i, j \leq n$ and $i \neq j$. Consider the basis $w_1, \ldots, w_n$ of V defined by $w_i = v_j$, $w_j = v_i$ and $w_k = v_k$ otherwise. Then

$$\begin{aligned} T(w_j - v_i) &= Tw_j - Tv_i \\ T(0) &= A_{1,j}w_1 + A_{2,j}w_2 + \cdots + A_{n,j}w_n - (A_{1,i}v_1 + A_{2,i}v_2 + \cdots + A_{n,i}v_n) \\ 0 &= (A_{j,j} - A_{i,i})v_i + (A_{i,j} - A_{j,i})v_j + \sum_{\substack{k \neq i \\ k \neq j}} (A_{k,j} - A_{k,i})v_k. \end{aligned}$$

By linear independence of $v_1, \ldots, v_n$, we have $A_{i,i} = A_{j,j}$, $A_{i,j} = A_{j,i}$ and $A_{k,j} = A_{k,i}$ for all $k \neq i, j$. Because i and j were arbitrary, we have $A_{1,1} = A_{2,2} = \cdots = A_{n,n} = \lambda$, as well as A is symmetric, and all off-diagonal entries match across rows, so we can say $A_{i,j} = \gamma$ for all $i \neq j$. To finish, consider the basis $u_1, \ldots, u_n$ defined by $u_1 = -v_1$ and $u_k = v_k$ otherwise. Then

$$\begin{aligned} T(v_1 + u_1) &= Tv_1 + Tu_1 \\ T(0) &= \lambda v_1 + \gamma(v_2 + \cdots + v_n) + (\lambda u_1 + \gamma(u_2 + \cdots + u_n)) \\ 0 &= 2\gamma(v_2 + \cdots + v_n). \end{aligned}$$

Again by linear independence of $v_1, \ldots, v_n$, we have $2\gamma = 0$ and thus $\gamma = 0$. Therefore, $A = \lambda I$ and T is a scalar multiple of the identity. ∎

This method shows that we can negate one *arbitrary* basis vector to force all off-diagonal entries to be 0 via the linear independence of the basis vectors. Then, we can swap v_1 with any other basis vector to force all diagonal entries to be the same.

Coolempire93. Let j be arbitrary and consider the basis $w_1, \ldots, w_n$ of V defined by $w_j = -v_j$ and $w_i = v_i$ otherwise. Then

$$\begin{aligned} T(w_j + v_j) &= Tw_j + Tv_j \\ T(0) &= A_{1,j}w_1 + A_{2,j}w_2 + \cdots + A_{n,j}w_n + A_{1,j}v_1 + A_{2,j}v_2 + \cdots + A_{n,j}v_n \\ 0 &= 2\sum_{k \neq j} A_{k,j}v_k. \end{aligned}$$

By linear independence of $v_1, \ldots, v_n$, we have $A_{k,j} = 0$ for all $k \neq j$. Since j was arbitrary, we have that all off-diagonal entries of A are 0.

We turn to one more basis to prove that all diagonal entries of A must be the same. Once again, let j be arbitrary, and

this time consider the basis $u_1, \ldots, u_n$ defined by $u_1 = v_j$, $u_j = v_1$ and $u_i = v_i$ otherwise. Then

$$\begin{aligned} T(v_1 - u_j) &= Tv_1 - Tu_j \\ T(0) &= A_{1,1}v_1 - A_{j,j}u_j \\ 0 &= (A_{1,1} - A_{j,j})v_1. \end{aligned}$$

As v_1 belongs to a basis for V, it is nonzero, and thus as j was arbitrary, $A_{1,1} = A_{2,2} = \cdots = A_{n,n} = \lambda$ for some scalar λ, and we have $A = \lambda I$, i.e., T is a multiple of the identity operator. ■

Exercise 3D20 ✔

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 $x \in \mathbf{R}$.

Coolempire93. We go by the same technique as in Appendix 3B29.

Let $q(x) = a_0 + a_1x + \cdots + a_nx^n$ and suppose that there exists a polynomial $p(x) = b_0 + b_1x + \cdots + b_mx^m$ satisfying the constraint $q(x) = (x^2+x)p''(x)+2xp'(x)+p(3)$. If q is constant, then let $p = q = a_0$. We have $(x^2+x)p''(x)+2xp'(x)+p(3) = (x^2 + x)0 + 2x(0) + a_0 = a_0 = q(x)$. If q is linear, say $q(x) = a_0 + a_1x$, then let $p(x) = b_0 + b_1x$. We want

$$q(x) = (x^2 + x)p''(x) + 2xp'(x) + p(3) = (x^2 + x)0 + 2x(b_1) + b_0 + 3b_1 = 2b_1x + 3b_1 + b_0,$$

so letting $b_1 = \frac{a_1}{2}$ and $b_0 = a_0 - \frac{3a_1}{2}$ yields a polynomial p satisfying the given equation. Otherwise, consider when $n \geq 2$. If p has degree $m \geq 2$, then p' has degree $m - 1$ and p'' has degree $m - 2$, so $(x^2 + x)p'' + 2xp'$ has degree m. To satisfy the constraint, we must have $m = n \geq 2$, so we may write $p(x) = b_0 + b_1x + \cdots + b_nx^n$. Evaluating, we obtain

$$\begin{aligned} p'(x) = b_1 + 2b_2x + \cdots + nb_nx^{n-1} &\implies 2xp'(x) = 2b_1x + 4b_2x^2 + \cdots + 2nb_nx^n \\ p''(x) = 2b_2 + 6b_3x + \cdots + n(n-1)b_nx^{n-2} &\implies \begin{cases} x^2p''(x) = 2b_2x^2 + 6b_3x^3 + \cdots + n(n-1)b_nx^n \\ xp''(x) = 2b_2x + 6b_3x^2 + \cdots + n(n-1)b_nx^{n-1} \end{cases}. \end{aligned}$$

Equating coefficients of $(x^2 + x)p'' + 2xp' + p(3)$ with those of q, we have that

$$\begin{cases} a_0 = p(3) = b_0 + 3b_1 + 9b_2 + \cdots + 3^nb_n \\ a_i = i(i+1)(b_i + b_{i+1}) \qquad \text{for } 1 \leq i \leq n-1 \\ a_n = n(n-1)b_n + 2nb_n = n(n+1)b_n \end{cases},$$

with the formula for a_i coming from the fact that the coefficient of x^i in $(x^2 + x)p''(x) + 2xp'(x)$ is

$$(i(i-1)b_i + i(i+1)b_{i+1}) + 2ib_i = (i^2 - i + 2i)b_i + i(i+1)b_{i+1} = i(i+1)(b_i + b_{i+1}).$$

As before, we solve for the coefficients of p by back-substitution. Noting that for $1 \leq i \leq n$, we have

$$b_i = \frac{a_i}{i(i+1)} - b_{i+1},$$

we obtain the system

$$\begin{cases} b_n = \frac{a_n}{n(n+1)} \\ b_{n-1} = \frac{a_{n-1}}{n(n-1)} - b_n = \frac{a_{n-1}}{n(n-1)} - \frac{a_n}{n(n+1)} \\ b_{n-2} = \frac{a_{n-2}}{(n-1)(n-2)} - b_{n-1} = \frac{a_{n-2}}{(n-1)(n-2)} - \frac{a_{n-1}}{n(n-1)} + \frac{a_n}{n(n+1)} \\ \vdots \\ b_1 = \frac{a_1}{2} - b_2 \\ b_0 = a_0 - 3b_1 - 9b_2 - \cdots - 3^n b_n \end{cases}$$

which has a valid solution† since no division by 0 occurs. ∎

† As before, solving this recurrence is outside the scope of this course, but, specifically, the coefficients of p are

$$b_i = \frac{a_i}{i(i+1)} - \frac{a_{i+1}}{(i+1)(i+2)} + \cdots + (-1)^{n-i}\frac{a_n}{n(n+1)} = \sum_{j=i}^{n}(-1)^{j-i}\frac{a_j}{j(j+1)}$$

for $1 \le i \le n$ and $b_0 = a_0 - \sum_{j=1}^{n} 3^j b_j$.

Exercise 3E11 ✔

Suppose $U = \{(x_1, x_2, \ldots) \in \mathbf{F}^\infty : x_k \neq 0 \text{ for only finitely many } k\}$.

(a) Show that U is a subspace of $\mathbf{F}^\infty$.

(b) Prove that $\mathbf{F}^\infty / U$ is infinite-dimensional.

Civil Service Pigeon. This solution assumes that the reader is familiar with partitioning $\mathbb{N}$ into infinitely many disjoint infinite sets.

(a) Since $(0, 0, \ldots) \in U$, U is nonempty. Suppose $x, y \in U$ and $c \in \mathbb{F}$. Let S_x and S_y be the sets of indices where x and y have nonzero terms, respectively. By definition, S_x and S_y are finite. $cx + y$ can only have nonzero terms at indices in $S_x \cup S_y$. Since the union of two finite sets is finite, $cx + y \in U$, which makes U a subspace of $\mathbb{F}^\infty$.

(b) Partition $\mathbb{Z}^+$ into a countably infinite family of disjoint infinite subsets $A_1, A_2, \ldots$. For each n, define $v_n \in \mathbb{F}^\infty$ such that its j-th term is 1 if $j \in A_n$ and 0 otherwise.

Suppose that $\sum_{i=1}^{m} c_i(v_i + U) = 0 + U$ for some positive integer m and scalars $c_1, \ldots, c_m \in \mathbb{F}$. Then, $w = \sum_{i=1}^{m} c_i v_i \in U$. By construction, for $i \in \{1, \ldots, m\}$, the j-th term of w is c_i for all $j \in A_i$. Since $w \in U$, it only has a finite number of nonzero terms. But each A_i is infinite, so this forces $c_i = 0$ for all $i \in \{1, \ldots, m\}$. Hence, $v_1 + U, v_2 + U, \ldots$ are linearly independent in $\mathbb{F}^\infty / U$.

Because $\mathbb{F}^\infty / U$ contains an infinite sequence of linearly independent vectors, it is infinite-dimensional. ∎

Exercise 3F11 ✔

Suppose $v_1, \ldots, v_n$ is a basis of V and $\varphi_1, \ldots, \varphi_n$ is the corresponding dual basis of V'. Suppose $\psi \in V'$. Prove that

$$\psi = \psi(v_1)\varphi_1 + \cdots + \psi(v_n)\varphi_n.$$

Civil Service Pigeon. Because $\varphi_1, \ldots, \varphi_n$ is the dual basis, $\varphi_i(v_j) = \delta_{i,j}$. For each basis vector v_j,

$$\left(\sum_{i=1}^{n} \psi(v_i)\varphi_i\right)(v_j) = \sum_{i=1}^{n} \psi(v_i)\varphi_i(v_j) = \psi(v_j).$$

Since the two linear functionals agree on a basis of V, they are equal on all of V. ∎

A4 Polynomials

Exercise 4-9 ✔

Prove that every polynomial of odd degree with real coefficients has a real zero.

This solution comes from a calculus perspective, using the intermediate value theorem.

blanketism. Let $p(x)$ be a polynomial of odd degree with real coefficients. As $x \to \infty$, $p(x) \to \infty$ or $p(x) \to -\infty$ depending on the leading coefficient. Similarly, as $x \to -\infty$, $p(x) \to -\infty$ or $p(x) \to \infty$. By the intermediate value theorem, since $p(x)$ is continuous and takes on both positive and negative values, there must exist some $x \in \mathbf{R}$ such that $p(x) = 0$. Therefore, p has a real zero. ∎

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.

This solution examines the roots of the polynomial p and how they contribute to the polynomial q.

blanketism. If $p = 0$, then $q = 0$ is a polynomial with real coefficients. Otherwise, let $n = \deg p$. Then we may write

$$p(z) = \alpha(z - \lambda_1)(z - \lambda_2)\cdots(z - \lambda_n),$$

where $\alpha \in \mathbf{C}$ and $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbf{C}$ are the roots of p. Then

$$q(z) = p(z)\overline{p(\overline{z})} = \alpha(z - \lambda_1)(z - \lambda_2)\cdots(z - \lambda_n) \cdot \overline{\alpha}(z - \overline{\lambda_1})(z - \overline{\lambda_2})\cdots(z - \overline{\lambda_n}).$$

Note that the product contains the term $\alpha\overline{\alpha} = |\alpha|^2 \in \mathbf{R}$, as well as the pairs

$$(z - \lambda_i)(z - \overline{\lambda_i}) = z^2 - 2\,\mathrm{Re}(\lambda_i)z + |\lambda_i|^2.$$

Each such pair contributes a quadratic polynomial with real coefficients. Since the product of polynomials with real coefficients also has real coefficients, it follows that $q(z)$ is a polynomial with real coefficients. ∎

A5 Eigenvalues and Eigenvectors

Exercise 5A27 ✔

Suppose that V is finite-dimensional and $k \in \{1, \dots, \dim V - 1\}$. 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.

This solution is more elegant than the one provided in the main text, as it avoids the explicit construction of subspaces and instead relies on the intersection of all k-dimensional invariant subspaces containing a given vector.

Civil Service Pigeon. Let $v \in V$ be nonzero. For any $u \notin \operatorname{span}(v)$, we can extend v, u to a basis and construct a k-dimensional subspace that contains v but not u since $k \in \{1, \dots, \dim V - 1\}$. So, the intersection of all k-dimensional subspaces containing v is precisely $\operatorname{span}(v)$.

The intersection of invariant subspaces is invariant, so $\operatorname{span}(v)$ is invariant. This means that every nonzero vector is an eigenvector of T, meaning that T is a scalar multiple of the identity by Exercise 5A26. ∎

Exercise 5A42 ✔

Define $T \in \mathcal{L}(\mathbf{F}^n)$ by $T(x_1, x_2, x_3, \dots, x_n) = (x_1, 2x_2, 3x_3, \dots, nx_n)$.

(a) Find all eigenvalues and eigenvectors of T.

(b) Find all subspaces of $\mathbf{F}^n$ that are invariant under T.

The solution is meant to generalize to any diagonal operator, showing that the invariant subspaces are precisely the spans of subsets of the standard basis vectors. The original solution is meant to show the student the mechanism by which invariant subspaces of a diagonal operator can be determined using the standard basis vectors.

Civil Service Pigeon. Let $e_1, \dots, e_n$ be the standard basis of $\mathbb{F}^n$.

(a) For each $k \in \{1, \dots, n\}$, we have $Te_k = ke_k$. Hence, the eigenvalues of T are 1, 2, …, n and the eigenvectors corresponding to k are the nonzero scalar multiples of e_k. Note that this list of eigenvalues (and eigenvectors) is exhaustive because $\dim(\mathbb{F}^n) = n$.

(b) Let $U \subseteq \mathbb{F}^n$ be an invariant subspace under T. Take any $u = \sum_{i=1}^n \alpha_i e_i \in U$ and suppose $\alpha_k \neq 0$ for some k. By Lagrange interpolation, there exists a polynomial p such that $p(i) = \delta_{i,k}$ for all $i \in \{1, \dots, n\}$. Since U is invariant under T, it is also invariant under $p(T)$, and so

$$p(T)u = \sum_{i=1}^n \alpha_i p(T)e_i = \sum_{i=1}^n \alpha_i p(i)e_i = \alpha_k e_k \in U \implies e_k \in U.$$

Repeating this argument for every index with a nonzero coefficient in some element of U shows that U contains each of the basis vectors $e_1, \dots, e_n$. Hence, $U = \operatorname{span}\{e_i : i \in S\}$ for some $S \subseteq \{1, \dots, n\}$.

Conversely, any such span is invariant because $Te_i = ie_i$ belongs to that span. There are 2^n such subspaces, and they exhaust all invariant subspaces of T. ∎

Exercise 5A43 ✔

Suppose that V is finite-dimensional, $\dim V > 1$, and $T \in \mathcal{L}(V)$. Prove that $\{p(T) : p \in \mathcal{P}(\mathbf{F})\} \neq \mathcal{L}(V)$.

This solution uses the commutativity of operators in $\{p(T) : p \in \mathcal{P}(\mathbf{F})\}$ to show that not all operators in $\mathcal{L}(V)$ can be expressed as a polynomial in T.

blanketism. Let $v_1, \ldots, v_n$ be a basis for V, with $n = \dim V$. Consider the operators $T_1, T_2 \in \mathcal{L}(V)$ defined by

$$T_1 v_1 = v_2, \quad T_1 v_i = 0 \text{ for } i = 2, \ldots, n,$$

and

$$T_2 v_1 = v_1, \quad T_2 v_i = 0 \text{ for } i = 2, \ldots, n.$$

Then $T_2 T_1 v_1 = T_2(T_1 v_1) = T_2 v_2 = 0$, and $T_1 T_2 v_1 = T_1(T_2 v_1) = T_1 v_1 = v_2$. Hence, $T_2 T_1 \neq T_1 T_2$, so T_1 and T_2 do not commute. Let $S = \{p(T) : p \in \mathcal{P}(\mathbf{F})\}$. Let $p, q \in \mathcal{P}$ be defined as

$$p(z) = a_0 + a_1 z + \cdots + a_m z^m, \quad q(z) = b_0 + b_1 z + \cdots + b_k z^k.$$

Since $p(T)q(T) = q(T)p(T)$, it follows that all operators in S commute. However, if $T_1 = r(T)$ and $T_2 = s(T)$ for some polynomials $r, s \in \mathcal{P}$, then T_1 and T_2 would commute, which is a contradiction. Hence, T_1 and T_2 cannot both be in S, so $S \neq \mathcal{L}(V)$. ∎

Exercise 5B29 ✔

Show that every operator on a finite-dimensional vector space of dimension at least two has an invariant subspace of dimension two.

Axler Note. Exercise 5C6 *will give an improvement of this result when* $\mathbf{F} = \mathbf{C}$.

This solution uses upper-triangular matrices, which have not appeared in 5B yet, but they provide a convenient way to identify invariant subspaces of dimension two.

Civil Service Pigeon. If $\mathbb{F} = \mathbb{C}$, the result is immediate: every operator on a finite-dimensional complex vector space has an upper-triangular matrix with respect to some basis $v_1, \ldots, v_n$. Then $\operatorname{span}(v_1, v_2)$ is a 2-dimensional invariant subspace. Hence, we restrict to $\mathbb{F} = \mathbb{R}$ from now on.

Assume $\dim V \geq 2$ and let $T \in \mathcal{L}(V)$ have minimal polynomial p. Over $\mathbb{R}$, p factors into linear and irreducible quadratic factors.

Case 1: p has an irreducible quadratic factor. Write $p(z) = (z^2 + az + b)m(z)$. Pick $v \in V$ such that $x = m(T)v \neq 0$; then $(T^2 + aT + bI)x = 0$. Set $U = \operatorname{span}(x, Tx)$. Clearly, T maps U to itself since $T(Tx) = -bx - aTx \in U$. If x and Tx were linearly dependent, x would be an eigenvector and its eigenvalue would be a real root of $z^2 + az + b$, contradicting irreducibility. Hence $\dim U = 2$.

Case 2: p splits completely into linear factors. Then T is upper-triangularizable over $\mathbb{R}$. Let $v_1, \ldots, v_n$ be a basis with respect to which the matrix of T is upper-triangular. Then, $Tv_1 \in \operatorname{span}(v_1)$ and $Tv_2 \in \operatorname{span}(v_1, v_2)$. Therefore $\operatorname{span}(v_1, v_2)$ is a 2-dimensional invariant subspace.

These two cases cover all possibilities over $\mathbb{R}$, and so combining with trivial complex case finishes the proof. ∎

Exercise 5C3 ✔

Suppose $T \in \mathcal{L}(V)$ is invertible and $v_1, \ldots, v_n$ is a basis of V with respect to which the matrix of T is upper triangular, with $\lambda_1, \ldots, \lambda_n$ on the diagonal. Show that the matrix of T^{-1} is also upper triangular with respect to the basis $v_1, \ldots, v_n$, with

$$\frac{1}{\lambda_1}, \ldots, \frac{1}{\lambda_n}$$

on the diagonal.

This solution attempts for a direct computation of the entries of T^{-1}.

blanketism. Let A be the matrix of T with respect to the basis $v_1, \ldots, v_n$. Since T is invertible, A is also invertible. Let $A = (a_{ij})$ be the upper-triangular matrix with $\lambda_1, \ldots, \lambda_n$ on the diagonal, and let $B = (b_{ij})$ be the matrix of T^{-1} with respect to the same basis. Denote b_j as the j-th column of B. Then we have

$$Ab_j = e_j,$$

where e_j is the j-th column of the identity matrix. In the nth row, we have $\lambda_n b_{nj} = \delta_{nj}$, hence $b_{nj} = 0$ for $j < n$ and $b_{nn} = \lambda_n^{-1}$. Now, consider the $(n-1)$-th row. We have

$$a_{n-1,n-1}b_{n-1,j} + a_{n-1,n}b_{n,j} = \delta_{n-1,j}.$$

For $j < n-1$, we have $b_{n,j} = 0$ and $\delta_{n-1,j} = 0$, so $b_{n-1,j} = 0$. For $j = n-1$, we have $b_{n,n-1} = 0$ and $\delta_{n-1,n-1} = 1$, so $b_{n-1,n-1} = a_{n-1,n-1}^{-1} = \lambda_{n-1}^{-1}$. Assume inductively that for some $k \in \{1, \ldots, n-1\}$, we have $b_{ij} = 0$ for all $i > j$ and $b_{ii} = \lambda_i^{-1}$ for all $i \geq k+1$. Now, consider the k-th row. We have

$$a_{kk}b_{kj} + a_{k,k+1}b_{k+1,j} + \cdots + a_{kn}b_{nj} = \delta_{kj}.$$

For $j < k$, we have $b_{ij} = 0$ for all $i > j \geq k$, so the left-hand side reduces to $a_{kk}b_{kj} = 0$, hence $b_{kj} = 0$. For $j = k$, we have $b_{ij} = 0$ for all $i > k$, so the left-hand side reduces to $a_{kk}b_{kk} = 1$, hence $b_{kk} = a_{kk}^{-1} = \lambda_k^{-1}$. By induction, we conclude that B is upper-triangular with $\lambda_1^{-1}, \ldots, \lambda_n^{-1}$ on the diagonal. ∎

A6 Inner Product Spaces

Exercise 6A23 ✔

Suppose $v_1, \dots, v_m \in V$ are such that $\|v_k\| \leq 1$ for each $k = 1, \dots, m$. Show that there exist $a_1, \dots, a_m \in \{1, -1\}$ such that

$$\|a_1 v_1 + \cdots + a_m v_m\| \leq \sqrt{m}.$$

Manual Remark. *Note that* $\min\{A, B\}^2 \leq \frac{A^2+B^2}{2}$ *so long as* $A, B \geq 0$. *This follows from the fact that either* $A \leq B$ *or* $B \leq A$ *forces* $\min\{A, B\}^2 \leq A^2$ *and* $\min\{A, B\}^2 \leq B^2$.

Coolempire93. Note that if the average value M of the squared norm $\|a_1 v_1 + \cdots + a_m v_m\|^2$ over all possible assignments of $a_1, \dots, a_m \in \{1, -1\}$ is less than or equal to m, then there must exist some assignment $a_1, \dots, a_m \in \{1, -1\}$ such that $\|a_1 v_1 + \cdots + a_m v_m\|^2 \leq m$. Taking the square root of both sides of this inequality would yield our desired inequality. By Exercise 6A13, this value is less than or equal to the average value M of the squared norm $\|a_1 v_1 + \cdots + a_m v_m\|^2$, so we may use M to bound our desired average.

Consider that there are 2^m possible assignments of $a_1, \dots, a_m$, so

$$\begin{aligned}
M &= \frac{1}{2^m} \sum_{a_1,\dots,a_m \in \{1,-1\}} \|a_1 v_1 + \cdots, a_m v_m\|^2 \\
&= \frac{1}{2^m} \sum_{a_1,\dots,a_m \in \{1,-1\}} \left\| \sum_{k=1}^{m} a_k v_k \right\|^2 \\
&= \frac{1}{2^m} \sum_{a_1,\dots,a_m \in \{1,-1\}} \left\langle \sum_{k=1}^{m} a_k v_k, \sum_{k=1}^{m} a_k v_k \right\rangle \\
&= \frac{1}{2^m} \sum_{a_1,\dots,a_m \in \{1,-1\}} \left\langle \sum_{i=1}^{m} a_i v_i, \sum_{j=1}^{m} a_j v_j \right\rangle \\
&= \frac{1}{2^m} \sum_{a_1,\dots,a_m \in \{1,-1\}} \sum_{i=1}^{m} \sum_{j=1}^{m} \langle a_i v_i, a_j v_j \rangle \\
&= \frac{1}{2^m} \sum_{a_1,\dots,a_m \in \{1,-1\}} \sum_{i=1}^{m} \sum_{j=1}^{m} a_i a_j \langle v_i, v_j \rangle,
\end{aligned}$$

with the last steps being by additivity in both slots and homogeneity ($\overline{a_j} = a_j$ since a_j is real). Consider $a_i a_j \langle v_i, v_j \rangle$ when $i \neq j$. Because we are summing over all possible assignments of $a_1, \dots, a_m$, and there are exactly as many assignments with $a_i a_j \langle v_i, v_j \rangle = \langle v_i, v_j \rangle$ as there are assignments with $a_i a_j \langle v_i, v_j \rangle = -\langle v_i, v_j \rangle$, these cancel out. We may simplify to only consider terms where $i = j$:

$$\begin{aligned}
M &= \frac{1}{2^m} \sum_{a_1,\dots,a_m \in \{1,-1\}} \sum_{i=1}^{m} \sum_{j=1}^{m} a_i a_j \langle v_i, v_j \rangle \\
&= \frac{1}{2^m} \sum_{a_1,\dots,a_m \in \{1,-1\}} \sum_{k=1}^{m} a_k a_k \langle v_k, v_k \rangle \\
&= \frac{1}{2^m} \sum_{a_1,\dots,a_m \in \{1,-1\}} \sum_{k=1}^{m} a_k^2 \|v_k\|^2
\end{aligned}$$

But $1^2 = 1$ and $(-1)^2 = 1$, so $a_k^2 = 1$:

$$\begin{aligned} M &= \frac{1}{2^m} \sum_{a_1,\ldots,a_m \in \{1,-1\}} \sum_{k=1}^{m} \|v_k\|^2 \\ &= \sum_{k=1}^{m} \|v_k\|^2 \cdot \frac{1}{2^m} \sum_{a_1,\ldots,a_m \in \{1,-1\}} 1 \\ &= \sum_{k=1}^{m} \|v_k\|^2 \cdot \frac{1}{2^m} 2^m \\ &= \sum_{k=1}^{m} \|v_k\|^2 \\ &\le m \end{aligned}$$

because there are 2^m assignments of $a_1, \ldots, a_m \in \{1, -1\}$, and our original assumption had $\|v_k\| \le 1$.

Therefore, since the average of $\|a_1 v_1 + \cdots + a_m v_m\|^2$ over all possible assignments of $a_1, \ldots, a_m \in \{1, -1\}$ is less than or equal to m, there is an assignment $a_1, \ldots, a_m \in \{1, -1\}$ such that $\|a_1 v_1 + \cdots + a_m v_m\|^2 \le m$. Taking the square root of both sides yields that for the given assignment, $\|a_1 v_1 + \cdots + a_m v_m\| \le \sqrt{m}$. ■

Exercise 6C16 ✔

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.

A more analysis-based solution to part (a) is as follows.

blanketism. Let $g \in U^\perp$, so $\langle f, g \rangle = 0$ for all $f \in U$. For any $\varepsilon > 0$, define $f_\varepsilon \in C[-1, 1]$ by $f_\varepsilon(x) = g(x)$ for $|x| \ge \varepsilon$, $f_\varepsilon(0) = 0$, with f_ε linear on $[-\varepsilon, 0]$ and on $[0, \varepsilon]$ so that f_ε is continuous at $x = 0$ and $x = \pm\varepsilon$. Then $f_\varepsilon \in U$, so $\langle f_\varepsilon, g \rangle = 0$. Let $M = \max_{[-1,1]} |g|$. Since $f_\varepsilon = g$ outside $[-\varepsilon, \varepsilon]$,

$$|\langle g, g \rangle| = |\langle f_\varepsilon - g, g \rangle| = \left| \int_{-\varepsilon}^{\varepsilon} (f_\varepsilon - g) g \right| \le \int_{-\varepsilon}^{\varepsilon} |f_\varepsilon - g|\,|g| \le 2M \cdot M \cdot 2\varepsilon = 4\varepsilon M^2.$$

Since $\varepsilon > 0$ is arbitrary, $\langle g, g \rangle = 0$, so $g = 0$. Hence, $U^\perp = \{0\}$. ■

A7 Operators on Inner Product Spaces

Exercise 7B14 ✔

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 $\lambda_1, \ldots, \lambda_m$ denote the distinct eigenvalues of T.

Manual Remark. *This exercise adds another characterization to the real spectral theorem, where T is self-adjoint if and only if the eigenvectors corresponding to distinct eigenvalues are orthogonal and V decomposes as the direct sum of its eigenspaces. In particular, this is the orthogonal analogue of Theorem 5.55.*

This solution rewrites the backwards direction of the proof, appealing to an orthogonal projection argument rather than the real spectral theorem.

Civil Service Pigeon. Let P_i be the orthogonal projection onto $E(\lambda_i, T)$. By hypothesis, eigenvectors corresponding to distinct eigenvalues are orthogonal, so the eigenspaces are mutually orthogonal. Since

$$V = \bigoplus_{i=1}^{m} E(\lambda_i, T),$$

these eigenspaces orthogonally decompose V, yielding $\sum_{i=1}^{m} P_i = I$. For any $v \in V$,

$$Tv = T\left(\sum_{i=1}^{m} P_i v\right) = \sum_{i=1}^{m} \lambda_i P_i v \implies T = \sum_{i=1}^{m} \lambda_i P_i.$$

Since each P_i an orthogonal projection onto mutually orthogonal subspaces, $P_i^* = P_i$ and $P_i P_j = 0$ for $i \neq j$. ∎

Exercise 7B15 ✔

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 $\lambda_1, \ldots, \lambda_m$ denote the distinct eigenvalues of T.

Manual Remark. *This has the same remark as Exercise 7B14 but in the context of the complex spectral theorem. It also uses normal operators instead of self-adjoint operators.*

Civil Service Pigeon. Since $T^* = \sum_{i=1}^{m} \overline{\lambda_i} P_i$,

$$TT^* = \left(\sum_{i=1}^{m} \lambda_i P_i\right)\left(\sum_{j=1}^{m} \overline{\lambda_j} P_j\right) = \sum_{i=1}^{m} |\lambda_i|^2 P_i.$$

By symmetry, T^*T yields the same sum, so $TT^* = T^*T$ and T is normal. ∎

Exercise 7C2 ✔

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.

The construction of B in the following solution comes from some intuition and T's structure. In particular, the choice of B is such that B^*B reproduces the tridiagonal structure of T, which makes it straightforward to verify that T is positive.

Civil Service Pigeon. Let $B \in \mathcal{L}\left(\mathbb{F}^4, \mathbb{F}^5\right)$ be defined by

$$B(x_1, x_2, x_3, x_4) = (x_1, x_1 - x_2, x_2 - x_3, x_3 - x_4, x_4).$$

It is easy to verify that $\mathcal{M}(B^*B) = \mathcal{M}(T)$; hence T is self-adjoint.

For any $x \in \mathbb{F}^4$,

$$\langle Tx, x\rangle = \langle B^*Bx, x\rangle = \langle Bx, Bx\rangle = \|Bx\|^2 \geq 0,$$

which implies T is a positive operator.

If $Tx = 0$, then

$$0 = \langle Tx, x\rangle = \|Bx\|^2 \implies Bx = 0 \implies x = 0.$$

Since $\operatorname{null} T = \{0\}$, T is injective and therefore invertible. ∎

Exercise 7C8 ✔

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.

This solution uses some future machinery, namely the pseudoinverse of an operator and its properties which is expounded in 7E. While the manual solution requires an explicit witness in the form of an orthonormal basis and appealing to Exercise 7B25, the below approach only leverages the structural properties of the pseudoinverse and the decomposition of the vector space into the range and null space of T.

Civil Service Pigeon. Suppose T is a positive operator. Since T is self-adjoint, which means its pseudoinverse is also self-adjoint since

$$\left(T^\dagger\right)^* = (T^*)^\dagger = T^\dagger.$$

Also, $V = \operatorname{range} T \oplus \operatorname{null} T$. For any $v \in V$, we can write $v = u + w$ where $u \in \operatorname{range} T$ and $w \in \operatorname{null} T$. By the definition of the pseudoinverse,

$$T^\dagger w = 0, \qquad T^\dagger u \in (\operatorname{null} T)^\perp = \operatorname{range} T.$$

Hence $T^\dagger u$ and w are orthogonal. So,

$$\left\langle T^\dagger v, v\right\rangle = \left\langle T^\dagger(u + w), u + w\right\rangle = \left\langle T^\dagger u, u + w\right\rangle = \left\langle T^\dagger u, u\right\rangle.$$

Setting $x = T^\dagger u$ yields $Tx = T\left(T^\dagger u\right) = u$. Thus,

$$\left\langle T^\dagger u, u\right\rangle = \langle x, Tx\rangle = \langle Tx, x\rangle \geq 0$$

since T is positive. Since $\left\langle T^\dagger v, v\right\rangle \geq 0$ for all $v \in V$ and $T^\dagger$ is self-adjoint, $T^\dagger$ is a positive operator. ∎

Exercise 7C12 ✔

Suppose $T \in \mathcal{L}(V)$ is a positive operator. Prove that T^k is a positive operator for every positive integer k.

This solution goes through a parity argument in k. Moreover, it again does not require an explicit orthonormal basis for the vector space.

Civil Service Pigeon. Since T is a positive operator, it is self-adjoint. Thus

$$\left(T^k\right)^* = (T^*)^k = T^k \implies T^k \text{ is self adjoint for all } k \in \mathbb{Z}^+.$$

Let $v \in V$.

- If $k = 2m$ for some integer $m \geq 1$, then

$$\left\langle T^k v, v \right\rangle = \left\langle T^{2m} v, v \right\rangle = \langle T^m v, T^m v \rangle = \|T^m v\|^2 \geq 0.$$

- If $k = 2m + 1$ for some integer $m \geq 0$, then

$$\left\langle T^k v, v \right\rangle = \left\langle T^{2m+1} v, v \right\rangle = \langle T\left(T^m v\right), T^m v \rangle.$$

Since T is positive, $\langle Tu, u \rangle \geq 0$ for all $u \in V$. Taking $u = T^m v$ yields $\langle T\left(T^m v\right), T^m v \rangle \geq 0$.
Therefore, T^k is a positive operator. ■

Exercise 7C19 ✔

Show that the identity operator on $\mathbf{F}^2$ has infinitely many self-adjoint square roots.

An alternate argument below uses no specific matrix representation of the operator.

Civil Service Pigeon. For any subspace U of $\mathbb{F}^2$, let P_U denote the orthogonal projection onto U. Define $T = 2P_U - I$. Since orthogonal projections are self-adjoint, T is also self-adjoint. Since $P_U^2 = P_U$,

$$T^2 = (2P_U - I)^2 = 4P_U^2 - 4P_U + I = 4P_U - 4P_U + I = I,$$

meaning T is a square root of the identity. Because $\mathbb{F}^2$ contains infinitely many one-dimensional subspaces, there are infinitely many distinct orthogonal projections on $\mathbb{F}^2$. Since $P_U = \frac{1}{2}(T + I)$, distinct orthogonal projections yield distinct operators T. Therefore, the identity operator on $\mathbb{F}^2$ has infinitely many self-adjoint square roots. ■

Exercise 7D10 ✔

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.

An alternate argument using the spectral theorem is given for part (b), where the eigenvalues $\lambda + i\sqrt{1 - \lambda^2}$ are made explicit on the unit circle.

Civil Service Pigeon. Since T is self-adjoint, there exists an orthonormal basis of eigenvectors of T. For any such eigenvector with eigenvalue λ, the condition $\|Tv\| \leq \|v\|$ implies $\lambda \in [-1, 1]$. The operator $T + i\sqrt{1 - T^2}$ acts on this eigenvector by multiplication with $\lambda + i\sqrt{1 - \lambda^2}$, which has modulus 1. Because there exists an orthonormal basis of eigenvectors whose corresponding eigenvalues all have modulus 1, the operator is unitary. ■

Exercise 7F2 ✔

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

$$\|T\| = \max\{|\lambda| : \lambda \text{ is an eigenvalue of } T\}.$$

Civil Service Pigeon. If $V = \{0\}$, the statement holds trivially.

Otherwise, since T is self-adjoint or complex normal, the spectral theorem guarantees an orthonormal basis $e_1, \ldots, e_n$ of V consisting of eigenvectors of T with corresponding eigenvalues $\lambda_1, \ldots, \lambda_n$. For any vector $v = \sum_{i=1}^{n} c_i e_i \in V$,

$$\|Tv\|^2 = \sum_{i=1}^{n} |\lambda_i|^2 |c_i|^2 \leq \left(\max_i |\lambda_i|\right)^2 \sum_{i=1}^{n} |c_i|^2 = \left(\max_i |\lambda_i|\right)^2 \|v\|^2,$$

which implies $\|T\| \leq \max_i |\lambda_i|$.

Conversely, letting j be an index where the maximum is attained, $\|Te_j\| = |\lambda_j| \|e_j\|$, so

$$\|T\| \geq \frac{\|Te_j\|}{\|e_j\|} = |\lambda_j| = \max_i |\lambda_i|.$$

Therefore,

$$\|T\| = \max\{|\lambda| : \lambda \text{ is an eigenvalue of } T\}.$$

■

A8 Generalized Eigenvectors and Nilpotent Operators

Exercise 8A20 ✔

Suppose V is an inner product space and $T \in \mathcal{L}(V)$ is normal and nilpotent. Prove that $T = 0$.

Civil Service Pigeon. If $T \neq 0$, let $k \geq 2$ be the smallest integer such that $T^k = 0$. Then there exists $v \in V$ with $T^{k-1}v \neq 0$. Then $T(T^{k-1}v) = T^k v = 0$, meaning $T^{k-1}v \in \operatorname{null} T$. Since T is normal, $\operatorname{null} T = \operatorname{null} T^*$, so $T^*(T^{k-1}v) = 0$. Thus

$$\|T^{k-1}v\|^2 = \langle T^{k-1}v, T^{k-1}v\rangle = \langle T^{k-2}v, T^*T^{k-1}v\rangle = 0 \implies T^{k-1}v = 0,$$

contradicting the choice of k. Therefore, $T = 0$. ∎

Exercise 8B20 ✔

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 $p_1 \cdots p_m$.

Civil Service Pigeon. For each eigenvalue $\lambda \in \mathbb{C}$, since $V = \bigoplus_{k=1}^{m} V_k$ with each V_k invariant under T, we have $G(\lambda, T) = \bigoplus_{k=1}^{m} G\left(\lambda, T|_{V_k}\right)$. This implies

$$\dim G(\lambda, T) = \sum_{k=1}^{m} \dim G\left(\lambda, T|_{V_k}\right).$$

The characteristic polynomial of T is

$$\prod_{\lambda}(z-\lambda)^{\dim G(\lambda,T)} = \prod_{\lambda}(z-\lambda)^{\sum_{k=1}^{m} \dim G\left(\lambda,T|_{V_k}\right)} = \prod_{k=1}^{m}\prod_{\lambda}(z-\lambda)^{\dim G\left(\lambda,T|_{V_k}\right)} = p_1 \cdots p_m.$$

∎

A9 Multilinear Algebra and Determinants

Exercise 9A4 ✔

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 $u, v \in V$.

Axler Note. *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.*

This proof uses the Riesz representation theorem to each functional $u \mapsto \beta(u, v)$ separately and describes, in detail, where the assumption $\mathbf{F} = \mathbf{R}$ is used.

Civil Service Pigeon. For each $v \in V$, the map $u \mapsto \beta(u, v)$ is a linear functional on V. By the Riesz representation theorem, there exists a unique vector $Tv \in V$ such that $\beta(u, v) = \langle u, Tv \rangle$ for all $u \in V$. This defines a map $T : V \to V$.

For any $v, w \in V$ and $\lambda, \mu \in \mathbb{R}$ and $u \in V$, we have

$$\langle u, T(\lambda v + \mu w) \rangle = \beta(u, \lambda v + \mu w) = \lambda \beta(u, v) + \mu \beta(u, w) = \langle u, \lambda Tv + \mu Tw \rangle .$$

Since this holds for all $u \in V$, it follows that $T(\lambda v + \mu w) = \lambda Tv + \mu Tw$, so $T \in \mathcal{L}(V)$.

To prove uniqueness, if $S \in \mathcal{L}(V)$ also satisfies $\beta(u, v) = \langle u, Sv \rangle$ for all $u, v \in V$, then $\langle u, Tv - Sv \rangle = 0$ for all $u \in V$, which implies $Tv = Sv$ for all $v \in V$ by choosing $u = Tv - Sv$. Hence, $T = S$. ∎

Exercise 9B1 ✔

Suppose m is a positive integer. Show that $\dim V^{(m)} = (\dim V)^m$.

Civil Service Pigeon. Proceed by induction on m. For $m = 1$, $V^{(1)} = V'$ and $\dim V^{(1)} = \dim V$. Assume the claim holds for $m - 1$, so $\dim V^{(m-1)} = (\dim V)^{m-1}$. There is a natural vector space isomorphism between $V^{(m)}$ and $\mathcal{L}\left(V, V^{(m-1)}\right)$ given by mapping each $\beta \in V^{(m)}$ to the linear map $v \mapsto \beta(v, \dots)$. Consequently,

$$\dim V^{(m)} = (\dim V)(\dim V^{(m-1)}) = (\dim V)(\dim V)^{m-1} = (\dim V)^m,$$

completing the induction. ∎

Exercise 9D3 ✔

Suppose that $v_1, \dots, v_m$ is a linearly independent list in V. Suppose also that $w_1, \dots, w_m$ is a list in W such that

$$v_1 \otimes w_1 + \cdots + v_m \otimes w_m = 0.$$

Prove that $w_1 = \cdots = w_m = 0$.

The proof below uses at dual basis functionals rather than relying on expanding in a basis.

Civil Service Pigeon. Since $v_1, \dots, v_m$ is linearly independent, there exist linear functionals $\phi_1, \dots, \phi_m \in V'$ such that $\phi_j(v_k) = \delta_{j,k}$. For any $j \in \{1, \dots, m\}$ and any arbitrary $\tau \in W'$,

$$0 = \left(\sum_{k=1}^{m} v_k \otimes w_k \right)(\phi_j, \tau) = \sum_{k=1}^{m} \phi_j(v_k)\tau(w_k) = \tau(w_j).$$

Since $\tau(w_j) = 0$ for all $\tau \in W'$, $w_j = 0$ for each j. ∎

Exercise 9D4 ✔

Suppose $\dim V > 1$ and $\dim W > 1$. Prove that

$$\{v \otimes w : (v, w) \in V \times W\}$$

is not a subspace of $V \otimes W$.

Axler Note. *This exercise implies that if* $\dim V > 1$ *and* $\dim W > 1$, *then*

$$\{v \otimes w : (v, w) \in V \times W\} \neq V \otimes W.$$

Similar to Appendix 9D3.

Civil Service Pigeon. Since $\dim V > 1$ and $\dim W > 1$, choose linearly independent $v_1, v_2 \in V$ and linearly independent $w_1, w_2 \in W$. For the sake of contradiction, suppose that $v_1 \otimes w_1 + v_2 \otimes w_2 = v \otimes w$ for some $v \in V$ and $w \in W$. Choose linear functionals $\phi, \varphi \in V'$ such that $\phi(v_1) = 1, \phi(v_2) = 0$ and $\varphi(v_1) = 0, \varphi(v_2) = 1$.

Applying the map $V \otimes W \to W$ given by $x \otimes y \mapsto f(x)y$ for each functional yields

$$w_1 = \phi(v)w, \qquad w_2 = \varphi(v)w.$$

Hence, w_1 and w_2 are both scalar multiples of w, contradicting linear independence. Thus, the set is not a subspace of $V \otimes W$. ∎

Exercise 9D7 ✔

Suppose $\dim V > 2$ and $\dim W > 2$. Prove that

$$\{v_1 \otimes w_1 + v_2 \otimes w_2 : v_1, v_2 \in V \text{ and } w_1, w_2 \in W\} \neq V \otimes W.$$

This proof gives tensors as linear maps from V' to W, then compares ranks.

Civil Service Pigeon. Since $\dim V > 2$ and $\dim W > 2$, choose linearly independent vectors $e_1, e_2, e_3 \in V$ and $f_1, f_2, f_3 \in W$. Under the natural isomorphism between $V \otimes W$ and $\mathcal{L}(V', W)$, $v \otimes w$ corresponds to a linear map of rank at most 1, implying any element of the form $v_1 \otimes w_1 + v_2 \otimes w_2$ corresponds to a map of rank at most 2. However, $u = e_1 \otimes f_1 + e_2 \otimes f_2 + e_3 \otimes f_3 \in V \otimes W$ corresponds to a linear map of rank 3. Therefore, u cannot be written as the sum of two pure tensors, proving the set is not all of $V \otimes W$. ∎