

# *Linear Algebra Done Right*

preliminary version of fourth edition

7 September 2023

© 2024 Sheldon Axler

## Sheldon Axler

This document contains the front matter and Chapter 5 of the future fourth edition of *Linear Algebra Done Right*. Suggestions for improvements are most welcome. Please send them to [linear@axler.net](mailto:linear@axler.net).

The fourth edition of *Linear Algebra Done Right* will be an Open Access book, which means that the electronic version will be legally free to the world. The print version will be published by Springer and will be reasonably priced. Both the electronic and the print versions will become available in late November 2023.

Because this is not the final version of Chapter 5, please do not post this document elsewhere on the web, although it is fine to post a link to it (<https://linear.axler.net/>).

$$F_n = \frac{1}{\sqrt{5}} \left[ \left( \frac{1 + \sqrt{5}}{2} \right)^n - \left( \frac{1 - \sqrt{5}}{2} \right)^n \right]$$

# About the Author

Sheldon Axler received his undergraduate degree from Princeton University, followed by a PhD in mathematics from the University of California at Berkeley.

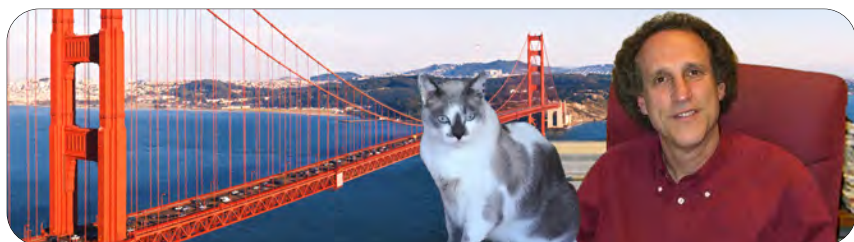
As a postdoctoral Moore Instructor at MIT, Axler received a university-wide teaching award. He was then an assistant professor, associate professor, and professor at Michigan State University, where he received the first J. Sutherland Frame Teaching Award and the Distinguished Faculty Award.

Axler received the Lester R. Ford Award for expository writing from the Mathematical Association of America in 1996, for a paper that eventually expanded into this book. In addition to publishing numerous research papers, he is the author of six mathematics textbooks, ranging from freshman to graduate level. Previous editions of this book have been adopted as a textbook at over 375 universities and colleges and have been translated into three languages.

Axler has served as Editor-in-Chief of the *Mathematical Intelligencer* and Associate Editor of the *American Mathematical Monthly*. He has been a member of the Council of the American Mathematical Society and of the Board of Trustees of the Mathematical Sciences Research Institute. He has also served on the editorial board of Springer's series Undergraduate Texts in Mathematics, Graduate Texts in Mathematics, Universitext, and Springer Monographs in Mathematics.

Axler is a Fellow of the American Mathematical Society and has been a recipient of numerous grants from the National Science Foundation.

Axler joined San Francisco State University as chair of the Mathematics Department in 1997. He served as dean of the College of Science & Engineering from 2002 to 2015, when he returned to a regular faculty appointment as a professor in the Mathematics Department.



Carrie Heeler, Bishnu Sarangi

*The author and his cat Moon.*

*Cover equation:* Formula for the  $n^{\text{th}}$  Fibonacci number. Exercise 21 in Section 5D uses linear algebra to derive this formula.

# Contents

*About the Author*   v

*Preface for Students*   xii

*Preface for Instructors*   xiii

*Acknowledgments*   xvii

Chapter 1

*Vector Spaces*   1

1A  $\mathbf{R}^n$  and  $\mathbf{C}^n$    2

Complex Numbers   2

Lists   5

$\mathbf{F}^n$    6

Digression on Fields   10

Exercises 1A   10

1B Definition of Vector Space   12

Exercises 1B   16

1C Subspaces   18

Sums of Subspaces   19

Direct Sums   21

Exercises 1C   24

Chapter 2

*Finite-Dimensional Vector Spaces*   27

2A Span and Linear Independence   28

Linear Combinations and Span   28

Linear Independence   31

Exercises 2A   37

2B Bases	39
Exercises 2B	42
2C Dimension	44
Exercises 2C	48

### Chapter 3

## **Linear Maps** 51

3A Vector Space of Linear Maps	52
Definition and Examples of Linear Maps	52
Algebraic Operations on $\mathcal{L}(V, W)$	55
Exercises 3A	57
3B Null Spaces and Ranges	59
Null Space and Injectivity	59
Range and Surjectivity	61
Fundamental Theorem of Linear Maps	62
Exercises 3B	66
3C Matrices	69
Representing a Linear Map by a Matrix	69
Addition and Scalar Multiplication of Matrices	71
Matrix Multiplication	72
Column–Row Factorization and Rank of a Matrix	77
Exercises 3C	79
3D Invertibility and Isomorphisms	82
Invertible Linear Maps	82
Isomorphic Vector Spaces	86
Linear Maps Thought of as Matrix Multiplication	88
Change of Basis	90
Exercises 3D	93
3E Products and Quotients of Vector Spaces	96
Products of Vector Spaces	96
Quotient Spaces	98
Exercises 3E	103
3F Duality	105
Dual Space and Dual Map	105
Null Space and Range of Dual of Linear Map	109

Matrix of Dual of Linear Map	113
Exercises 3F	115

## Chapter 4

***Polynomials*** 119

Zeros of Polynomials	122
Division Algorithm for Polynomials	123
Factorization of Polynomials over $\mathbf{C}$	124
Factorization of Polynomials over $\mathbf{R}$	127
Exercises 4	129

## Chapter 5

***Eigenvalues and Eigenvectors*** 132

5A Invariant Subspaces	133
Eigenvalues	133
Polynomials Applied to Operators	137
Exercises 5A	139
5B The Minimal Polynomial	143
Existence of Eigenvalues on Complex Vector Spaces	143
Eigenvalues and the Minimal Polynomial	144
Eigenvalues on Odd-Dimensional Real Vector Spaces	149
Exercises 5B	150
5C Upper-Triangular Matrices	154
Exercises 5C	160
5D Diagonalizable Operators	163
Diagonal Matrices	163
Conditions for Diagonalizability	165
Gershgorin Disk Theorem	170
Exercises 5D	172
5E Commuting Operators	175
Exercises 5E	179

## Chapter 6

***Inner Product Spaces*** 181

6A Inner Products and Norms	182
Inner Products	182

Norms 186

Exercises 6A 191

## 6B Orthonormal Bases 197

Orthonormal Lists and the Gram–Schmidt Procedure 197

Linear Functionals on Inner Product Spaces 204

Exercises 6B 207

## 6C Orthogonal Complements and Minimization Problems 211

Orthogonal Complements 211

Minimization Problems 217

Pseudoinverse 220

Exercises 6C 224

## Chapter 7

### *Operators on Inner Product Spaces* 227

#### 7A Self-Adjoint and Normal Operators 228

Adjoints 228

Self-Adjoint Operators 233

Normal Operators 235

Exercises 7A 239

#### 7B Spectral Theorem 243

Real Spectral Theorem 243

Complex Spectral Theorem 246

Exercises 7B 247

#### 7C Positive Operators 251

Exercises 7C 255

#### 7D Isometries, Unitary Operators, and Matrix Factorization 258

Isometries 258

Unitary Operators 260

QR Factorization 263

Cholesky Factorization 266

Exercises 7D 268

#### 7E Singular Value Decomposition 270

Singular Values 270

SVD for Linear Maps and for Matrices 273

Exercises 7E 278

7F	Consequences of Singular Value Decomposition	280
	Norms of Linear Maps	280
	Approximation by Linear Maps with Lower-Dimensional Range	283
	Polar Decomposition	285
	Operators Applied to Ellipsoids and Parallelepipeds	287
	Volume via Singular Values	291
	Properties of an Operator as Determined by Its Eigenvalues	293
	Exercises 7F	294

## Chapter 8

### *Operators on Complex Vector Spaces* 297

8A	Generalized Eigenvectors and Nilpotent Operators	298
	Null Spaces of Powers of an Operator	298
	Generalized Eigenvectors	300
	Nilpotent Operators	303
	Exercises 8A	306
8B	Generalized Eigenspace Decomposition	308
	Generalized Eigenspaces	308
	Multiplicity of an Eigenvalue	310
	Block Diagonal Matrices	314
	Exercises 8B	316
8C	Consequences of Generalized Eigenspace Decomposition	319
	Square Roots of Operators	319
	Jordan Form	321
	Exercises 8C	324
8D	Trace: A Connection Between Matrices and Operators	326
	Exercises 8D	330

## Chapter 9

### *Multilinear Algebra and Determinants* 332

9A	Bilinear Forms and Quadratic Forms	333
	Bilinear Forms	333
	Symmetric Bilinear Forms	337
	Quadratic Forms	341
	Exercises 9A	344

9B	Alternating Multilinear Forms	346
	Multilinear Forms	346
	Alternating Multilinear Forms and Permutations	348
	Exercises 9B	352
9C	Determinants	354
	Defining the Determinant	354
	Properties of Determinants	357
	Exercises 9C	367
9D	Tensor Products	370
	Tensor Product of Two Vector Spaces	370
	Tensor Product of Inner Product Spaces	376
	Tensor Product of Multiple Vector Spaces	378
	Exercises 9D	380
	<i>Photo Credits</i>	383
	<i>Symbol Index</i>	384
	<i>Index</i>	385
	<i>Colophon: Notes on Typesetting</i>	390



# *Preface for Students*

You are probably about to begin your second exposure to linear algebra. Unlike your first brush with the subject, which probably emphasized Euclidean spaces and matrices, this encounter will focus on abstract vector spaces and linear maps. These terms will be defined later, so don't worry if you do not know what they mean. This book starts from the beginning of the subject, assuming no knowledge of linear algebra. The key point is that you are about to immerse yourself in serious mathematics, with an emphasis on attaining a deep understanding of the definitions, theorems, and proofs.

You cannot read mathematics the way you read a novel. If you zip through a page in less than an hour, you are probably going too fast. When you encounter the phrase “as you should verify”, you should indeed do the verification, which will usually require some writing on your part. When steps are left out, you need to supply the missing pieces. You should ponder and internalize each definition. For each theorem, you should seek examples to show why each hypothesis is necessary. Discussions with other students should help.

As a visual aid, definitions are in yellow boxes and theorems are in blue boxes (in color versions of the book). Each theorem has an informal descriptive name.

Please check the website below for additional information about the book, including a link to videos that are freely available to accompany the book.

Your suggestions, comments, and corrections are most welcome.

Best wishes for success and enjoyment in learning linear algebra!

Sheldon Axler

San Francisco State University

website: <https://linear.axler.net>

e-mail: [linear@axler.net](mailto:linear@axler.net)

# *Preface for Instructors*

You are about to teach a course that will probably give students their second exposure to linear algebra. During their first brush with the subject, your students probably worked with Euclidean spaces and matrices. In contrast, this course will emphasize abstract vector spaces and linear maps.

The title of this book deserves an explanation. Most linear algebra textbooks use determinants to prove that every linear operator on a finite-dimensional complex vector space has an eigenvalue. Determinants are difficult, nonintuitive, and often defined without motivation. To prove the theorem about existence of eigenvalues on complex vector spaces, most books must define determinants, prove that a linear operator is not invertible if and only if its determinant equals 0, and then define the characteristic polynomial. This tortuous (torturous?) path gives students little feeling for why eigenvalues exist.

In contrast, the simple determinant-free proofs presented here (for example, see 5.19) offer more insight. Once determinants have been moved to the end of the book, a new route opens to the main goal of linear algebra—understanding the structure of linear operators.

This book starts at the beginning of the subject, with no prerequisites other than the usual demand for suitable mathematical maturity. A few examples and exercises involve calculus concepts such as continuity, differentiation, and integration. You can easily skip those examples and exercises if your students have not had calculus. If your students have had calculus, then those examples and exercises can enrich their experience by showing connections between different parts of mathematics.

Even if your students have already seen some of the material in the first few chapters, they may be unaccustomed to working exercises of the type presented here, most of which require an understanding of proofs.

Here is a chapter-by-chapter summary of the highlights of the book:

- Chapter 1: Vector spaces are defined in this chapter, and their basic properties are developed.
- Chapter 2: Linear independence, span, basis, and dimension are defined in this chapter, which presents the basic theory of finite-dimensional vector spaces.
- Chapter 3: This chapter introduces linear maps. The key result here is the fundamental theorem of linear maps: if  $T$  is a linear map on  $V$ , then  $\dim V = \dim \text{null } T + \dim \text{range } T$ . Quotient spaces and duality are topics in this chapter at a higher level of abstraction than most of the book; these topics can be skipped (except that duality is needed for tensor products in Section 9D).

- Chapter 4: The part of the theory of polynomials that will be needed to understand linear operators is presented in this chapter. This chapter contains no linear algebra. It can be covered quickly, especially if your students are already familiar with these results.
- Chapter 5: The idea of studying a linear operator by restricting it to small subspaces leads to eigenvectors in the early part of this chapter. The highlight of this chapter is a simple proof that on complex vector spaces, eigenvalues always exist. This result is then used to show that each linear operator on a complex vector space has an upper-triangular matrix with respect to some basis. The minimal polynomial plays an important role here and later in the book. For example, this chapter gives a characterization of the diagonalizable operators in terms of the minimal polynomial. Section 5E can be skipped if you want to save some time.
- Chapter 6: Inner product spaces are defined in this chapter, and their basic properties are developed along with tools such as orthonormal bases and the Gram–Schmidt procedure. This chapter also shows how orthogonal projections can be used to solve certain minimization problems. The pseudoinverse is then introduced as a useful tool when the inverse does not exist. The material on the pseudoinverse can be skipped if you want to save some time.
- Chapter 7: The spectral theorem, which characterizes the linear operators for which there exists an orthonormal basis consisting of eigenvectors, is one of the highlights of this book. The work in earlier chapters pays off here with especially simple proofs. This chapter also deals with positive operators, isometries, unitary operators, matrix factorizations, and especially the singular value decomposition, which leads to the polar decomposition and norms of linear maps.
- Chapter 8: This chapter shows that for each operator on a complex vector space, there is a basis of the vector space consisting of generalized eigenvectors of the operator. Then the generalized eigenspace decomposition describes a linear operator on a complex vector space. The multiplicity of an eigenvalue is defined as the dimension of the corresponding generalized eigenspace. These tools are used to prove that every invertible linear operator on a complex vector space has a square root. Then the chapter gives a proof that every linear operator on a complex vector space can be put into Jordan form. The chapter concludes with an investigation of the trace of operators.
- Chapter 9: This chapter begins by looking at bilinear forms and showing that the vector space of bilinear forms is the direct sum of the subspaces of symmetric bilinear forms and alternating bilinear forms. Then quadratic forms are diagonalized. Moving to multilinear forms, the chapter shows that the subspace of alternating  $n$ -linear forms on an  $n$ -dimensional vector space has dimension one. This result leads to a clean basis-free definition of the determinant of an operator. For complex vector spaces, the determinant turns out to equal the product of the eigenvalues, with each eigenvalue included in the product as many times as its multiplicity. The chapter concludes with an introduction to tensor products.

This book usually develops linear algebra simultaneously for real and complex vector spaces by letting  $\mathbf{F}$  denote either the real or the complex numbers. If you and your students prefer to think of  $\mathbf{F}$  as an arbitrary field, then see the comments at the end of Section 1A. I prefer avoiding arbitrary fields at this level because they introduce extra abstraction without leading to any new linear algebra. Also, students are more comfortable thinking of polynomials as functions instead of the more formal objects needed for polynomials with coefficients in finite fields. Finally, even if the beginning part of the theory were developed with arbitrary fields, inner product spaces would push consideration back to just real and complex vector spaces.

You probably cannot cover everything in this book in one semester. Going through all the material in the first seven or eight chapters during a one-semester course may require a rapid pace. If you must reach Chapter 9, then consider skipping the material on quotient spaces in Section 3E, skipping Section 3F on duality (unless you intend to cover tensor products in Section 9D), covering Chapter 4 on polynomials in a half hour, skipping Section 5E on commuting operators, and skipping the subsection in Section 6C on the pseudoinverse.

A goal more important than teaching any particular theorem is to develop in students the ability to understand and manipulate the objects of linear algebra. Mathematics can be learned only by doing. Fortunately, linear algebra has many good homework exercises. When teaching this course, during each class I usually assign as homework several of the exercises, due the next class. Going over the homework might take up significant time in a typical class.

Some of the exercises are intended to lead curious students into important topics beyond what might usually be included in a basic second course in linear algebra.

### The author's top ten

Listed below are the author's ten favorite results in the book, in order of their appearance in the book. Students who leave your course with a good understanding of these crucial results will have an excellent foundation in linear algebra.

- any two bases of a vector space have the same length (2.34)
- fundamental theorem of linear maps (3.21)
- existence of eigenvalues if  $\mathbf{F} = \mathbf{C}$  (5.19)
- upper-triangular form always exists if  $\mathbf{F} = \mathbf{C}$  (5.47)
- Cauchy–Schwarz inequality (6.14)
- Gram–Schmidt procedure (6.32)
- spectral theorem (7.29 and 7.31)
- singular value decomposition (7.70)
- generalized eigenspace decomposition theorem when  $\mathbf{F} = \mathbf{C}$  (8.22)
- dimension of alternating  $n$ -linear forms on  $V$  is 1 if  $\dim V = n$  (9.37)

## Major improvements and additions for the fourth edition

- Over 250 new exercises and over 70 new examples.
- Increasing use of the minimal polynomial to provide cleaner proofs of multiple results, including necessary and sufficient conditions for an operator to have an upper-triangular matrix with respect to some basis (see Section 5C), necessary and sufficient conditions for diagonalizability (see Section 5D), and the real spectral theorem (see Section 7B).
- New section on commuting operators (see Section 5E).
- New subsection on pseudoinverse (see Section 6C).
- New subsections on QR factorization/Cholesky factorization (see Section 7D).
- Singular value decomposition now done for linear maps from an inner product space to another (possibly different) inner product space, rather than only dealing with linear operators from an inner product space to itself (see Section 7E).
- Polar decomposition now proved from singular value decomposition, rather than in the opposite order; this has led to cleaner proofs of both the singular value decomposition (see Section 7E) and the polar decomposition (see Section 7F).
- New subsection on norms of linear maps on finite-dimensional inner product spaces, using the singular value decomposition to avoid even mentioning supremum in the definition of the norm of a linear map (see Section 7F).
- New subsection on approximation by linear maps with lower-dimensional range (see Section 7F).
- New elementary proof of the important result that if  $T$  is an operator on a finite-dimensional complex vector space  $V$ , then there exists a basis of  $V$  consisting of generalized eigenvectors of  $T$  (see 8.9).
- New Chapter 9 on multilinear algebra, including bilinear forms, quadratic forms, multilinear forms, and tensor products. Determinants now are defined using a basis-free approach via alternating multilinear forms.
- New formatting to improve the student-friendly appearance of the book. For example, the definition and result boxes now have rounded corners instead of right-angle corners, for a gentler look. The main font size has been reduced from 11 point to 10.5 point.

Please check the website below for additional links and information about the book. Your suggestions, comments, and corrections are most welcome.

Best wishes for teaching a successful linear algebra class!

Sheldon Axler  
San Francisco State University

website: <https://linear.axler.net>

e-mail: [linear@axler.net](mailto:linear@axler.net)

*Contact the author, or Springer if the author is not available, for permission for translations or other commercial reuse of the contents of this book.*

# Acknowledgments

I owe a huge intellectual debt to all the mathematicians who created linear algebra over the past two centuries. The results in this book belong to the common heritage of mathematics. A special case of a theorem may first have been proved long ago, then sharpened and improved by many mathematicians in different time periods. Bestowing proper credit on all contributors would be a difficult task that I have not undertaken. In no case should the reader assume that any result presented here represents my original contribution.

Many people helped make this a better book. The three previous editions of this book were used as a textbook at over 375 universities and colleges around the world. I received thousands of suggestions and comments from faculty and students who used the book. Many of those suggestions led to improvements in this edition. The manuscript for this fourth edition was class tested at 30 universities. I am extremely grateful for the useful feedback that I received from faculty and students during this class testing.

The long list of people who should be thanked for their suggestions would fill up many pages. Lists are boring to read. Thus to represent all contributors to this edition, I will mention only Noel Hinton, a graduate student at Australian National University, who sent me more suggestions and corrections for this fourth edition than anyone else. To everyone who contributed suggestions, let me say how truly grateful I am to all of you. Many many thanks!

I thank Springer for providing me with help when I needed it and for allowing me the freedom to make the final decisions about the content and appearance of this book. Special thanks to the two terrific mathematics editors at Springer who worked with me on this project—Loretta Bartolini during the first half of my work on the fourth edition, and Elizabeth Loew during the second half of my work on the fourth edition. I am deeply indebted to David Kramer, who did a magnificent job of copyediting and prevented me from making many mistakes.

Extra special thanks to my fantastic partner Carrie Heeter. Her understanding and encouragement enabled me to work intensely on this new edition. Our wonderful cat Moon, whose picture appears on the *About the Author* page, provided sweet breaks throughout the writing process. Moon died suddenly due to a blood clot as this book was being finished. We are grateful for five precious years with him.

Sheldon Axler

# *Eigenvalues and Eigenvectors*

Linear maps from one vector space to another vector space were the objects of study in Chapter 3. Now we begin our investigation of operators, which are linear maps from a vector space to itself. Their study constitutes the most important part of linear algebra.

To learn about an operator, we might try restricting it to a smaller subspace. Asking for that restriction to be an operator will lead us to the notion of invariant subspaces. Each one-dimensional invariant subspace arises from a vector that the operator maps into a scalar multiple of the vector. This path will lead us to eigenvectors and eigenvalues.

We will then prove one of the most important results in linear algebra: every operator on a finite-dimensional nonzero complex vector space has an eigenvalue. This result will allow us to show that for each operator on a finite-dimensional complex vector space, there is a basis of the vector space with respect to which the matrix of the operator has at least almost half its entries equal to 0.

*standing assumptions for this chapter*

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



Hans-Peter Postel CC BY

*Statue of Leonardo of Pisa (1170–1250, approximate dates), also known as Fibonacci.*

*Exercise 21 in Section 5D shows how linear algebra can be used to find the explicit formula for the Fibonacci sequence shown on the front cover.*

## 5A Invariant Subspaces

### Eigenvalues

#### 5.1 definition: operator

A linear map from a vector space to itself is called an *operator*.

Suppose  $T \in \mathcal{L}(V)$ . If  $m \geq 2$  and

$$V = V_1 \oplus \cdots \oplus V_m,$$

Recall that we defined the notation  $\mathcal{L}(V)$  to mean  $\mathcal{L}(V, V)$ .

where each  $V_k$  is a nonzero subspace of  $V$ , then to understand the behavior of  $T$  we only need to understand the behavior of each  $T|_{V_k}$ ; here  $T|_{V_k}$  denotes the restriction of  $T$  to the smaller domain  $V_k$ . Dealing with  $T|_{V_k}$  should be easier than dealing with  $T$  because  $V_k$  is a smaller vector space than  $V$ .

However, if we intend to apply tools useful in the study of operators (such as taking powers), then we have a problem:  $T|_{V_k}$  may not map  $V_k$  into itself; in other words,  $T|_{V_k}$  may not be an operator on  $V_k$ . Thus we are led to consider only decompositions of  $V$  of the form above in which  $T$  maps each  $V_k$  into itself. Hence we now give a name to subspaces of  $V$  that get mapped into themselves by  $T$ .

#### 5.2 definition: invariant subspace

Suppose  $T \in \mathcal{L}(V)$ . A subspace  $U$  of  $V$  is called *invariant* under  $T$  if  $Tu \in U$  for every  $u \in U$ .

Thus  $U$  is invariant under  $T$  if  $T|_U$  is an operator on  $U$ .

#### 5.3 example: subspace invariant under differentiation operator

Suppose that  $T \in \mathcal{L}(\mathcal{P}(\mathbf{R}))$  is defined by  $Tp = p'$ . Then  $\mathcal{P}_4(\mathbf{R})$ , which is a subspace of  $\mathcal{P}(\mathbf{R})$ , is invariant under  $T$  because if  $p \in \mathcal{P}(\mathbf{R})$  has degree at most 4, then  $p'$  also has degree at most 4.

#### 5.4 example: four invariant subspaces, not necessarily all different

If  $T \in \mathcal{L}(V)$ , then the following subspaces of  $V$  are all invariant under  $T$ .

$\{0\}$  The subspace  $\{0\}$  is invariant under  $T$  because if  $u \in \{0\}$ , then  $u = 0$  and hence  $Tu = 0 \in \{0\}$ .

$V$  The subspace  $V$  is invariant under  $T$  because if  $u \in V$ , then  $Tu \in V$ .

$\text{null } T$  The subspace  $\text{null } T$  is invariant under  $T$  because if  $u \in \text{null } T$ , then  $Tu = 0$ , and hence  $Tu \in \text{null } T$ .

$\text{range } T$  The subspace  $\text{range } T$  is invariant under  $T$  because if  $u \in \text{range } T$ , then  $Tu \in \text{range } T$ .



Must an operator  $T \in \mathcal{L}(V)$  have any invariant subspaces other than  $\{0\}$  and  $V$ ? Later we will see that this question has an affirmative answer if  $V$  is finite-dimensional and  $\dim V > 1$  (for  $\mathbf{F} = \mathbf{C}$ ) or  $\dim V > 2$  (for  $\mathbf{F} = \mathbf{R}$ ); see 5.19 and Exercise 29 in Section 5B.

The previous example noted that  $\text{null } T$  and  $\text{range } T$  are invariant under  $T$ . However, these subspaces do not necessarily provide easy answers to the question above about the existence of invariant subspaces other than  $\{0\}$  and  $V$ , because  $\text{null } T$  may equal  $\{0\}$  and  $\text{range } T$  may equal  $V$  (this happens when  $T$  is invertible).

We will return later to a deeper study of invariant subspaces. Now we turn to an investigation of the simplest possible nontrivial invariant subspaces—invariant subspaces of dimension one.

Take any  $v \in V$  with  $v \neq 0$  and let  $U$  equal the set of all scalar multiples of  $v$ :

$$U = \{\lambda v : \lambda \in \mathbf{F}\} = \text{span}(v).$$

Then  $U$  is a one-dimensional subspace of  $V$  (and every one-dimensional subspace of  $V$  is of this form for an appropriate choice of  $v$ ). If  $U$  is invariant under an operator  $T \in \mathcal{L}(V)$ , then  $Tv \in U$ , and hence there is a scalar  $\lambda \in \mathbf{F}$  such that

$$Tv = \lambda v.$$

Conversely, if  $Tv = \lambda v$  for some  $\lambda \in \mathbf{F}$ , then  $\text{span}(v)$  is a one-dimensional subspace of  $V$  invariant under  $T$ .

The equation  $Tv = \lambda v$ , which we have just seen is intimately connected with one-dimensional invariant subspaces, is important enough that the scalars  $\lambda$  and vectors  $v$  satisfying it are given special names.

### 5.5 definition: *eigenvalue*

Suppose  $T \in \mathcal{L}(V)$ . A number  $\lambda \in \mathbf{F}$  is called an *eigenvalue* of  $T$  if there exists  $v \in V$  such that  $v \neq 0$  and  $Tv = \lambda v$ .

In the definition above, we require that  $v \neq 0$  because every scalar  $\lambda \in \mathbf{F}$  satisfies  $T0 = \lambda 0$ .

The comments above show that  $V$  has a one-dimensional subspace invariant under  $T$  if and only if  $T$  has an eigenvalue.

*The word **eigenvalue** is half-German, half-English. The German prefix **eigen** means “own” in the sense of characterizing an intrinsic property.*

### 5.6 example: *eigenvalue*

Define an operator  $T \in \mathcal{L}(\mathbf{F}^3)$  by

$$T(x, y, z) = (7x + 3z, 3x + 6y + 9z, -6y)$$

for  $(x, y, z) \in \mathbf{F}^3$ . Then  $T(3, 1, -1) = (18, 6, -6) = 6(3, 1, -1)$ . Thus 6 is an eigenvalue of  $T$ .

The equivalences in the next result, along with many deep results in linear algebra, are valid only in the context of finite-dimensional vector spaces.

### 5.7 equivalent conditions to be an eigenvalue

Suppose  $V$  is finite-dimensional,  $T \in \mathcal{L}(V)$ , and  $\lambda \in \mathbf{F}$ . Then the following are equivalent.

- (a)  $\lambda$  is an eigenvalue of  $T$ .
- (b)  $T - \lambda I$  is not injective.
- (c)  $T - \lambda I$  is not surjective.
- (d)  $T - \lambda I$  is not invertible.

*Reminder:  $I \in \mathcal{L}(V)$  is the identity operator. Thus  $Iv = v$  for all  $v \in V$ .*

**Proof** Conditions (a) and (b) are equivalent because the equation  $Tv = \lambda v$  is equivalent to the equation  $(T - \lambda I)v = 0$ . Conditions (b), (c), and (d) are equivalent by 3.65. ■

### 5.8 definition: eigenvector

Suppose  $T \in \mathcal{L}(V)$  and  $\lambda \in \mathbf{F}$  is an eigenvalue of  $T$ . A vector  $v \in V$  is called an *eigenvector* of  $T$  corresponding to  $\lambda$  if  $v \neq 0$  and  $Tv = \lambda v$ .

In other words, a nonzero vector  $v \in V$  is an eigenvector of an operator  $T \in \mathcal{L}(V)$  if and only if  $Tv$  is a scalar multiple of  $v$ . Because  $Tv = \lambda v$  if and only if  $(T - \lambda I)v = 0$ , a vector  $v \in V$  with  $v \neq 0$  is an eigenvector of  $T$  corresponding to  $\lambda$  if and only if  $v \in \text{null}(T - \lambda I)$ .

### 5.9 example: eigenvalues and eigenvectors

Suppose  $T \in \mathcal{L}(\mathbf{F}^2)$  is defined by  $T(w, z) = (-z, w)$ .

- (a) First consider the case  $\mathbf{F} = \mathbf{R}$ . Then  $T$  is a counterclockwise rotation by  $90^\circ$  about the origin in  $\mathbf{R}^2$ . An operator has an eigenvalue if and only if there exists a nonzero vector in its domain that gets sent by the operator to a scalar multiple of itself. A  $90^\circ$  counterclockwise rotation of a nonzero vector in  $\mathbf{R}^2$  cannot equal a scalar multiple of itself. Conclusion: if  $\mathbf{F} = \mathbf{R}$ , then  $T$  has no eigenvalues (and thus has no eigenvectors).
- (b) Now consider the case  $\mathbf{F} = \mathbf{C}$ . To find eigenvalues of  $T$ , we must find the scalars  $\lambda$  such that  $T(w, z) = \lambda(w, z)$  has some solution other than  $w = z = 0$ . The equation  $T(w, z) = \lambda(w, z)$  is equivalent to the simultaneous equations

$$5.10 \quad -z = \lambda w, \quad w = \lambda z.$$

Substituting the value for  $w$  given by the second equation into the first equation gives

$$-z = \lambda^2 z.$$

Now  $z$  cannot equal 0 [otherwise 5.10 implies that  $w = 0$ ; we are looking for solutions to 5.10 such that  $(w, z)$  is not the 0 vector], so the equation above leads to the equation

$$-1 = \lambda^2.$$

The solutions to this equation are  $\lambda = i$  and  $\lambda = -i$ .

You can verify that  $i$  and  $-i$  are eigenvalues of  $T$ . Indeed, the eigenvectors corresponding to the eigenvalue  $i$  are the vectors of the form  $(w, -wi)$ , with  $w \in \mathbb{C}$  and  $w \neq 0$ . Furthermore, the eigenvectors corresponding to the eigenvalue  $-i$  are the vectors of the form  $(w, wi)$ , with  $w \in \mathbb{C}$  and  $w \neq 0$ .

In the next proof, we again use the equivalence

$$Tv = \lambda v \iff (T - \lambda I)v = 0.$$

### 5.11 linearly independent eigenvectors

Suppose  $T \in \mathcal{L}(V)$ . Then every list of eigenvectors of  $T$  corresponding to distinct eigenvalues of  $T$  is linearly independent.

**Proof** Suppose the desired result is false. Then there exists a smallest positive integer  $m$  such that there exists a linearly dependent list  $v_1, \dots, v_m$  of eigenvectors of  $T$  corresponding to distinct eigenvalues  $\lambda_1, \dots, \lambda_m$  of  $T$  (note that  $m \geq 2$  because an eigenvector is, by definition, nonzero). Thus there exist  $a_1, \dots, a_m \in \mathbb{F}$ , none of which are 0 (because of the minimality of  $m$ ), such that

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

Apply  $T - \lambda_m I$  to both sides of the equation above, getting

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

Because the eigenvalues  $\lambda_1, \dots, \lambda_m$  are distinct, none of the coefficients above equal 0. Thus  $v_1, \dots, v_{m-1}$  is a linearly dependent list of  $m - 1$  eigenvectors of  $T$  corresponding to distinct eigenvalues, contradicting the minimality of  $m$ . This contradiction completes the proof. ■

The result above leads to a short proof of the result below, which puts an upper bound on the number of distinct eigenvalues that an operator can have.

### 5.12 operator cannot have more eigenvalues than dimension of vector space

Suppose  $V$  is finite-dimensional. Then each operator on  $V$  has at most  $\dim V$  distinct eigenvalues.

**Proof** Let  $T \in \mathcal{L}(V)$ . Suppose  $\lambda_1, \dots, \lambda_m$  are distinct eigenvalues of  $T$ . Let  $v_1, \dots, v_m$  be corresponding eigenvectors. Then 5.11 implies that the list  $v_1, \dots, v_m$  is linearly independent. Thus  $m \leq \dim V$  (see 2.22), as desired. ■

## Polynomials Applied to Operators

The main reason that a richer theory exists for operators (which map a vector space into itself) than for more general linear maps is that operators can be raised to powers. In this subsection we define that notion and the concept of applying a polynomial to an operator. This concept will be the key tool that we use in the next section when we prove that every operator on a nonzero finite-dimensional complex vector space has an eigenvalue.

If  $T$  is an operator, then  $TT$  makes sense (see 3.7) and is also an operator on the same vector space as  $T$ . We usually write  $T^2$  instead of  $TT$ . More generally, we have the following definition of  $T^m$ .

### 5.13 notation: $T^m$

Suppose  $T \in \mathcal{L}(V)$  and  $m$  is a positive integer.

- $T^m \in \mathcal{L}(V)$  is defined by  $T^m = \underbrace{T \cdots T}_{m \text{ times}}$ .
- $T^0$  is defined to be the identity operator  $I$  on  $V$ .
- If  $T$  is invertible with inverse  $T^{-1}$ , then  $T^{-m} \in \mathcal{L}(V)$  is defined by

$$T^{-m} = (T^{-1})^m.$$

You should verify that if  $T$  is an operator, then

$$T^m T^n = T^{m+n} \quad \text{and} \quad (T^m)^n = T^{mn},$$

where  $m$  and  $n$  are arbitrary integers if  $T$  is invertible and are nonnegative integers if  $T$  is not invertible.

Having defined powers of an operator, we can now define what it means to apply a polynomial to an operator.

### 5.14 notation: $p(T)$

Suppose  $T \in \mathcal{L}(V)$  and  $p \in \mathcal{P}(\mathbf{F})$  is a polynomial given by

$$p(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_m z^m$$

for all  $z \in \mathbf{F}$ . Then  $p(T)$  is the operator on  $V$  defined by

$$p(T) = a_0 I + a_1 T + a_2 T^2 + \cdots + a_m T^m.$$

This is a new use of the symbol  $p$  because we are applying  $p$  to operators, not just elements of  $\mathbf{F}$ . The idea here is that to evaluate  $p(T)$ , we simply replace  $z$  with  $T$  in the expression defining  $p$ . Note that the constant term  $a_0$  in  $p(z)$  becomes the operator  $a_0 I$  (which is a reasonable choice because  $a_0 = a_0 z^0$  and thus we should replace  $a_0$  with  $a_0 T^0$ , which equals  $a_0 I$ ).

## 5.15 example: a polynomial applied to the differentiation operator

Suppose  $D \in \mathcal{L}(\mathcal{P}(\mathbf{R}))$  is the differentiation operator defined by  $Dq = q'$  and  $p$  is the polynomial defined by  $p(x) = 7 - 3x + 5x^2$ . Then  $p(D) = 7I - 3D + 5D^2$ . Thus

$$(p(D))q = 7q - 3q' + 5q''$$

for every  $q \in \mathcal{P}(\mathbf{R})$ .

If we fix an operator  $T \in \mathcal{L}(V)$ , then the function from  $\mathcal{P}(\mathbf{F})$  to  $\mathcal{L}(V)$  given by  $p \mapsto p(T)$  is linear, as you should verify.

## 5.16 definition: product of polynomials

If  $p, q \in \mathcal{P}(\mathbf{F})$ , then  $pq \in \mathcal{P}(\mathbf{F})$  is the polynomial defined by

$$(pq)(z) = p(z)q(z)$$

for all  $z \in \mathbf{F}$ .

The order does not matter in taking products of polynomials of a single operator, as shown by (b) in the next result.

## 5.17 multiplicative properties

Suppose  $p, q \in \mathcal{P}(\mathbf{F})$  and  $T \in \mathcal{L}(V)$ . Then

- (a)  $(pq)(T) = p(T)q(T)$ ;
- (b)  $p(T)q(T) = q(T)p(T)$ .

*Informal proof: When a product of polynomials is expanded using the distributive property, it does not matter whether the symbol is  $z$  or  $T$ .*

## Proof

- (a) Suppose  $p(z) = \sum_{j=0}^m a_j z^j$  and  $q(z) = \sum_{k=0}^n b_k z^k$  for all  $z \in \mathbf{F}$ . Then

$$(pq)(z) = \sum_{j=0}^m \sum_{k=0}^n a_j b_k z^{j+k}.$$

Thus

$$\begin{aligned} (pq)(T) &= \sum_{j=0}^m \sum_{k=0}^n a_j b_k T^{j+k} \\ &= \left( \sum_{j=0}^m a_j T^j \right) \left( \sum_{k=0}^n b_k T^k \right) \\ &= p(T)q(T). \end{aligned}$$

- (b) Using (a) twice, we have  $p(T)q(T) = (pq)(T) = (qp)(T) = q(T)p(T)$ . ■

We observed earlier that if  $T \in \mathcal{L}(V)$ , then the subspaces  $\text{null } T$  and  $\text{range } T$  are invariant under  $T$  (see 5.4). Now we show that the null space and the range of every polynomial of  $T$  are also invariant under  $T$ .

**5.18** *null space and range of  $p(T)$  are invariant under  $T$*

Suppose  $T \in \mathcal{L}(V)$  and  $p \in \mathcal{P}(\mathbf{F})$ . Then  $\text{null } p(T)$  and  $\text{range } p(T)$  are invariant under  $T$ .

**Proof** Suppose  $u \in \text{null } p(T)$ . Then  $p(T)u = 0$ . Thus

$$(p(T))(Tu) = T(p(T)u) = T(0) = 0.$$

Hence  $Tu \in \text{null } p(T)$ . Thus  $\text{null } p(T)$  is invariant under  $T$ , as desired.

Suppose  $u \in \text{range } p(T)$ . Then there exists  $v \in V$  such that  $u = p(T)v$ . Thus

$$Tu = T(p(T)v) = p(T)(Tv).$$

Hence  $Tu \in \text{range } p(T)$ . Thus  $\text{range } p(T)$  is invariant under  $T$ , as desired. ■

## Exercises 5A

- Suppose  $T \in \mathcal{L}(V)$  and  $U$  is a subspace of  $V$ .
  - Prove that if  $U \subseteq \text{null } T$ , then  $U$  is invariant under  $T$ .
  - Prove that if  $\text{range } T \subseteq U$ , then  $U$  is invariant under  $T$ .
- Suppose that  $T \in \mathcal{L}(V)$  and  $V_1, \dots, V_m$  are subspaces of  $V$  invariant under  $T$ . Prove that  $V_1 + \dots + V_m$  is invariant under  $T$ .
- Suppose  $T \in \mathcal{L}(V)$ . Prove that the intersection of every collection of subspaces of  $V$  invariant under  $T$  is invariant under  $T$ .
- 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  $U = V$ .
- Suppose  $T \in \mathcal{L}(\mathbf{R}^2)$  is defined by  $T(x, y) = (-3y, x)$ . Find the eigenvalues of  $T$ .
- Define  $T \in \mathcal{L}(\mathbf{F}^2)$  by  $T(w, z) = (z, w)$ . Find all eigenvalues and eigenvectors of  $T$ .
- 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  $T$ .
- Suppose  $P \in \mathcal{L}(V)$  is such that  $P^2 = P$ . Prove that if  $\lambda$  is an eigenvalue of  $P$ , then  $\lambda = 0$  or  $\lambda = 1$ .

- 9 Define  $T: \mathcal{P}(\mathbf{R}) \rightarrow \mathcal{P}(\mathbf{R})$  by  $Tp = p'$ . Find all eigenvalues and eigenvectors of  $T$ .
- 10 Define  $T \in \mathcal{L}(\mathcal{P}_4(\mathbf{R}))$  by  $(Tp)(x) = xp'(x)$  for all  $x \in \mathbf{R}$ . Find all eigenvalues and eigenvectors of  $T$ .
- 11 Suppose  $V$  is finite-dimensional,  $T \in \mathcal{L}(V)$ , and  $\alpha \in \mathbf{F}$ . Prove that there exists  $\delta > 0$  such that  $T - \lambda I$  is invertible for all  $\lambda \in \mathbf{F}$  such that  $0 < |\alpha - \lambda| < \delta$ .
- 12 Suppose  $V = U \oplus W$ , where  $U$  and  $W$  are nonzero subspaces of  $V$ . Define  $P \in \mathcal{L}(V)$  by  $P(u + w) = u$  for each  $u \in U$  and each  $w \in W$ . Find all eigenvalues and eigenvectors of  $P$ .
- 13 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.
  - What is the relationship between the eigenvectors of  $T$  and the eigenvectors of  $S^{-1}TS$ ?
- 14 Give an example of an operator on  $\mathbf{R}^4$  that has no (real) eigenvalues.
- 15 Suppose  $V$  is finite-dimensional,  $T \in \mathcal{L}(V)$ , and  $\lambda \in \mathbf{F}$ . Show that  $\lambda$  is an eigenvalue of  $T$  if and only if  $\lambda$  is an eigenvalue of the dual operator  $T' \in \mathcal{L}(V')$ .
- 16 Suppose  $v_1, \dots, v_n$  is a basis of  $V$  and  $T \in \mathcal{L}(V)$ . Prove that if  $\lambda$  is an eigenvalue of  $T$ , then

$$|\lambda| \leq n \max\{|\mathcal{M}(T)_{j,k}| : 1 \leq j, k \leq n\},$$

where  $\mathcal{M}(T)_{j,k}$  denotes the entry in row  $j$ , column  $k$  of the matrix of  $T$  with respect to the basis  $v_1, \dots, v_n$ .

See Exercise 19 in Section 6A for a different bound on  $|\lambda|$ .

- 17 Suppose  $\mathbf{F} = \mathbf{R}$ ,  $T \in \mathcal{L}(V)$ , and  $\lambda \in \mathbf{R}$ . Prove that  $\lambda$  is an eigenvalue of  $T$  if and only if  $\lambda$  is an eigenvalue of the complexification  $T_{\mathbf{C}}$ .
- See Exercise 33 in Section 3B for the definition of  $T_{\mathbf{C}}$ .
- 18 Suppose  $\mathbf{F} = \mathbf{R}$ ,  $T \in \mathcal{L}(V)$ , and  $\lambda \in \mathbf{C}$ . Prove that  $\lambda$  is an eigenvalue of the complexification  $T_{\mathbf{C}}$  if and only if  $\bar{\lambda}$  is an eigenvalue of  $T_{\mathbf{C}}$ .
- 19 Show that the forward shift operator  $T \in \mathcal{L}(\mathbf{F}^{\infty})$  defined by

$$T(z_1, z_2, \dots) = (0, z_1, z_2, \dots)$$

has no eigenvalues.

- 20 Define the backward shift operator  $S \in \mathcal{L}(\mathbf{F}^{\infty})$  by

$$S(z_1, z_2, z_3, \dots) = (z_2, z_3, \dots).$$

- Show that every element of  $\mathbf{F}$  is an eigenvalue of  $S$ .
- Find all eigenvectors of  $S$ .

- 21** Suppose  $T \in \mathcal{L}(V)$  is invertible.
- Suppose  $\lambda \in \mathbf{F}$  with  $\lambda \neq 0$ . Prove that  $\lambda$  is an eigenvalue of  $T$  if and only if  $\frac{1}{\lambda}$  is an eigenvalue of  $T^{-1}$ .
  - Prove that  $T$  and  $T^{-1}$  have the same eigenvectors.
- 22** Suppose  $T \in \mathcal{L}(V)$  and there exist nonzero vectors  $u$  and  $w$  in  $V$  such that
- $$Tu = 3w \quad \text{and} \quad Tw = 3u.$$
- Prove that 3 or  $-3$  is an eigenvalue of  $T$ .
- 23** Suppose  $V$  is finite-dimensional and  $S, T \in \mathcal{L}(V)$ . Prove that  $ST$  and  $TS$  have the same eigenvalues.
- 24** Suppose  $A$  is an  $n$ -by- $n$  matrix with entries in  $\mathbf{F}$ . Define  $T \in \mathcal{L}(\mathbf{F}^n)$  by  $Tx = Ax$ , where elements of  $\mathbf{F}^n$  are thought of as  $n$ -by-1 column vectors.
- Suppose the sum of the entries in each row of  $A$  equals 1. Prove that 1 is an eigenvalue of  $T$ .
  - Suppose the sum of the entries in each column of  $A$  equals 1. Prove that 1 is an eigenvalue of  $T$ .
- 25** Suppose  $T \in \mathcal{L}(V)$  and  $u, w$  are eigenvectors of  $T$  such that  $u + w$  is also an eigenvector of  $T$ . Prove that  $u$  and  $w$  are eigenvectors of  $T$  corresponding to the same eigenvalue.
- 26** Suppose  $T \in \mathcal{L}(V)$  is such that every nonzero vector in  $V$  is an eigenvector of  $T$ . Prove that  $T$  is a scalar multiple of the identity operator.
- 27** 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.
- 28** Suppose  $V$  is finite-dimensional and  $T \in \mathcal{L}(V)$ . Prove that  $T$  has at most  $1 + \dim \text{range } T$  distinct eigenvalues.
- 29** Suppose  $T \in \mathcal{L}(\mathbf{R}^3)$  and  $-4, 5$ , and  $\sqrt{7}$  are eigenvalues of  $T$ . Prove that there exists  $x \in \mathbf{R}^3$  such that  $Tx - 9x = (-4, 5, \sqrt{7})$ .
- 30** Suppose  $T \in \mathcal{L}(V)$  and  $(T - 2I)(T - 3I)(T - 4I) = 0$ . Suppose  $\lambda$  is an eigenvalue of  $T$ . Prove that  $\lambda = 2$  or  $\lambda = 3$  or  $\lambda = 4$ .
- 31** Give an example of  $T \in \mathcal{L}(\mathbf{R}^2)$  such that  $T^4 = -I$ .
- 32** Suppose  $T \in \mathcal{L}(V)$  has no eigenvalues and  $T^4 = I$ . Prove that  $T^2 = -I$ .
- 33** Suppose  $T \in \mathcal{L}(V)$  and  $m$  is a positive integer.
- Prove that  $T$  is injective if and only if  $T^m$  is injective.
  - Prove that  $T$  is surjective if and only if  $T^m$  is surjective.



**34** 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.

**35** Suppose that  $\lambda_1, \dots, \lambda_n$  is a list of distinct real numbers. Prove that the list  $e^{\lambda_1 x}, \dots, e^{\lambda_n x}$  is linearly independent in the vector space of real-valued functions on  $\mathbf{R}$ .

*Hint: Let  $V = \text{span}(e^{\lambda_1 x}, \dots, e^{\lambda_n x})$ , and define an operator  $D \in \mathcal{L}(V)$  by  $Df = f'$ . Find eigenvalues and eigenvectors of  $D$ .*

**36** Suppose that  $\lambda_1, \dots, \lambda_n$  is a list of distinct positive numbers. Prove that the list  $\cos(\lambda_1 x), \dots, \cos(\lambda_n x)$  is linearly independent in the vector space of real-valued functions on  $\mathbf{R}$ .

**37** Suppose  $V$  is finite-dimensional and  $T \in \mathcal{L}(V)$ . Define  $\mathcal{A} \in \mathcal{L}(\mathcal{L}(V))$  by

$$\mathcal{A}(S) = TS$$

for each  $S \in \mathcal{L}(V)$ . Prove that the set of eigenvalues of  $T$  equals the set of eigenvalues of  $\mathcal{A}$ .

**38** Suppose  $V$  is finite-dimensional,  $T \in \mathcal{L}(V)$ , and  $U$  is a subspace of  $V$  invariant under  $T$ . The *quotient operator*  $T/U \in \mathcal{L}(V/U)$  is defined by

$$(T/U)(v + U) = Tv + U$$

for each  $v \in V$ .

(a) Show that the definition of  $T/U$  makes sense (which requires using the condition that  $U$  is invariant under  $T$ ) and show that  $T/U$  is an operator on  $V/U$ .

(b) Show that each eigenvalue of  $T/U$  is an eigenvalue of  $T$ .

**39** Suppose  $V$  is finite-dimensional and  $T \in \mathcal{L}(V)$ . Prove that  $T$  has an eigenvalue if and only if there exists a subspace of  $V$  of dimension  $\dim V - 1$  that is invariant under  $T$ .

**40** Suppose  $S, T \in \mathcal{L}(V)$  and  $S$  is invertible. Suppose  $p \in \mathcal{P}(\mathbf{F})$  is a polynomial. Prove that

$$p(STS^{-1}) = Sp(T)S^{-1}.$$

**41** Suppose  $T \in \mathcal{L}(V)$  and  $U$  is a subspace of  $V$  invariant under  $T$ . Prove that  $U$  is invariant under  $p(T)$  for every polynomial  $p \in \mathcal{P}(\mathbf{F})$ .

**42** 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  $T$ .

**43** Suppose that  $V$  is finite-dimensional,  $\dim V > 1$ , and  $T \in \mathcal{L}(V)$ . Prove that  $\{p(T) : p \in \mathcal{P}(\mathbf{F})\} \neq \mathcal{L}(V)$ .

## 5B The Minimal Polynomial

### Existence of Eigenvalues on Complex Vector Spaces

Now we come to one of the central results about operators on finite-dimensional complex vector spaces.

#### 5.19 existence of eigenvalues

Every operator on a finite-dimensional nonzero complex vector space has an eigenvalue.

**Proof** Suppose  $V$  is a finite-dimensional complex vector space of dimension  $n > 0$  and  $T \in \mathcal{L}(V)$ . Choose  $v \in V$  with  $v \neq 0$ . Then

$$v, Tv, T^2v, \dots, T^n v$$

is not linearly independent, because  $V$  has dimension  $n$  and this list has length  $n + 1$ . Hence some linear combination (with not all the coefficients equal to 0) of the vectors above equals 0. Thus there exists a nonconstant polynomial  $p$  of smallest degree such that

$$p(T)v = 0.$$

By the first version of the fundamental theorem of algebra (see 4.12), there exists  $\lambda \in \mathbb{C}$  such that  $p(\lambda) = 0$ . Hence there exists a polynomial  $q \in \mathcal{P}(\mathbb{C})$  such that

$$p(z) = (z - \lambda)q(z)$$

for every  $z \in \mathbb{C}$  (see 4.6). This implies (using 5.17) that

$$0 = p(T)v = (T - \lambda I)(q(T)v).$$

Because  $q$  has smaller degree than  $p$ , we know that  $q(T)v \neq 0$ . Thus the equation above implies that  $\lambda$  is an eigenvalue of  $T$  with eigenvector  $q(T)v$ . ■

The proof above makes crucial use of the fundamental theorem of algebra. The comment following Exercise 16 helps explain why the fundamental theorem of algebra is so tightly connected to the result above.

The hypothesis in the result above that  $\mathbf{F} = \mathbb{C}$  cannot be replaced with the hypothesis that  $\mathbf{F} = \mathbb{R}$ , as shown by Example 5.9. The next example shows that the finite-dimensional hypothesis in the result above also cannot be deleted.

#### 5.20 example: an operator on a complex vector space with no eigenvalues

Define  $T \in \mathcal{L}(\mathcal{P}(\mathbb{C}))$  by  $(Tp)(z) = zp(z)$ . If  $p \in \mathcal{P}(\mathbb{C})$  is a nonzero polynomial, then the degree of  $Tp$  is one more than the degree of  $p$ , and thus  $Tp$  cannot equal a scalar multiple of  $p$ . Hence  $T$  has no eigenvalues.

Because  $\mathcal{P}(\mathbb{C})$  is infinite-dimensional, this example does not contradict the result above.

## Eigenvalues and the Minimal Polynomial

In this subsection we introduce an important polynomial associated with each operator. We begin with the following definition.

### 5.21 definition: *monic polynomial*

A *monic polynomial* is a polynomial whose highest-degree coefficient equals 1.

For example, the polynomial  $2 + 9z^2 + z^7$  is a monic polynomial of degree 7.

### 5.22 existence, uniqueness, and degree of minimal polynomial

Suppose  $V$  is finite-dimensional and  $T \in \mathcal{L}(V)$ . Then there is a unique monic polynomial  $p \in \mathcal{P}(\mathbf{F})$  of smallest degree such that  $p(T) = 0$ . Furthermore,  $\deg p \leq \dim V$ .

**Proof** If  $\dim V = 0$ , then  $I$  is the zero operator on  $V$  and thus we take  $p$  to be the constant polynomial 1.

Now use induction on  $\dim V$ . Thus assume that  $\dim V > 0$  and that the desired result is true for all operators on all complex vector spaces of smaller dimension. Let  $v \in V$  be such that  $v \neq 0$ . The list  $v, Tv, \dots, T^{\dim V}v$  has length  $1 + \dim V$  and thus is linearly dependent. By the linear dependence lemma (2.19), there is a smallest positive integer  $m \leq \dim V$  such that  $T^m v$  is a linear combination of  $v, Tv, \dots, T^{m-1}v$ . Thus there exist scalars  $c_0, c_1, c_2, \dots, c_{m-1} \in \mathbf{F}$  such that

$$5.23 \quad c_0 v + c_1 Tv + \dots + c_{m-1} T^{m-1}v + T^m v = 0.$$

Define a monic polynomial  $q \in \mathcal{P}_m(\mathbf{F})$  by

$$q(z) = c_0 + c_1 z + \dots + c_{m-1} z^{m-1} + z^m.$$

Then 5.23 implies that  $q(T)v = 0$ .

If  $k$  is a nonnegative integer, then

$$q(T)(T^k v) = T^k(q(T)v) = T^k(0) = 0.$$

The linear dependence lemma (2.19) shows that  $v, Tv, \dots, T^{m-1}v$  is linearly independent. Thus the equation above implies that  $\dim \text{null } q(T) \geq m$ . Hence

$$\dim \text{range } q(T) = \dim V - \dim \text{null } q(T) \leq \dim V - m.$$

Because  $\text{range } q(T)$  is invariant under  $T$  (by 5.18), we can apply our induction hypothesis to the operator  $T|_{\text{range } q(T)}$  on the vector space  $\text{range } q(T)$ . Thus there is a monic polynomial  $s \in \mathcal{P}(\mathbf{F})$  with

$$\deg s \leq \dim V - m \quad \text{and} \quad s(T|_{\text{range } q(T)}) = 0.$$

Hence for all  $v \in V$  we have

$$(sq)(T)(v) = s(T)(q(T)v) = 0$$

because  $q(T)v \in \text{range } q(T)$  and  $s(T)|_{\text{range } q(T)} = s(T|_{\text{range } q(T)}) = 0$ . Thus  $sq$  is a monic polynomial such that  $\deg sq \leq \dim V$  and  $(sq)(T) = 0$ .

The paragraph above shows that there is a monic polynomial of degree at most  $\dim V$  that when applied to  $T$  gives the 0 operator. Thus there is a monic polynomial of smallest degree with this property, completing the existence part of this result.

Let  $p \in \mathcal{P}(\mathbf{F})$  be a monic polynomial of smallest degree such that  $p(T) = 0$ . To prove the uniqueness part of the result, suppose  $r \in \mathcal{P}(\mathbf{F})$  is a monic polynomial of the same degree as  $p$  and  $r(T) = 0$ . Then  $(p - r)(T) = 0$  and also  $\deg(p - r) < \deg p$ . If  $p - r$  were not equal to 0, then we could divide  $p - r$  by the coefficient of the highest-order term in  $p - r$  to get a monic polynomial (of smaller degree than  $p$ ) that when applied to  $T$  gives the 0 operator. Thus  $p - r = 0$ , as desired. ■

The previous result justifies the following definition.

5.24 definition: *minimal polynomial*

Suppose  $V$  is finite-dimensional and  $T \in \mathcal{L}(V)$ . Then the *minimal polynomial* of  $T$  is the unique monic polynomial  $p \in \mathcal{P}(\mathbf{F})$  of smallest degree such that  $p(T) = 0$ .

To compute the minimal polynomial of an operator  $T \in \mathcal{L}(V)$ , we need to find the smallest positive integer  $m$  such that the equation

$$c_0 I + c_1 T + \cdots + c_{m-1} T^{m-1} = -T^m$$

has a solution  $c_0, c_1, \dots, c_{m-1} \in \mathbf{F}$ . If we pick a basis of  $V$  and replace  $T$  in the equation above with the matrix of  $T$ , then the equation above can be thought of as a system of  $(\dim V)^2$  linear equations in the  $m$  unknowns  $c_0, c_1, \dots, c_{m-1} \in \mathbf{F}$ . Gaussian elimination or another fast method of solving systems of linear equations can tell us whether a solution exists, testing successive values  $m = 1, 2, \dots$  until a solution exists. By 5.22, a solution exists for some smallest positive integer  $m \leq \dim V$ . The minimal polynomial of  $T$  is then  $c_0 + c_1 z + \cdots + c_{m-1} z^{m-1} + z^m$ .

Even faster (usually), pick  $v \in V$  and consider the equation

$$5.25 \quad c_0 v + c_1 T v + \cdots + c_{\dim V - 1} T^{\dim V - 1} v = -T^{\dim V} v.$$

Use a basis of  $V$  to convert the equation above to a system of  $\dim V$  linear equations in  $\dim V$  unknowns  $c_0, c_1, \dots, c_{\dim V - 1}$ . If this system of equations has a unique solution  $c_0, c_1, \dots, c_{\dim V - 1}$  (as happens most of the time), then the scalars  $c_0, c_1, \dots, c_{\dim V - 1}, 1$  are the coefficients of the minimal polynomial of  $T$  (because 5.22 states that the degree of the minimal polynomial is at most  $\dim V$ ).

Consider operators on  $\mathbf{R}^4$  (thought of as 4-by-4 matrices with respect to the standard basis), and take  $v = (1, 0, 0, 0)$

*These estimates are based on testing millions of random matrices.*

in the paragraph above. The faster method described above works on over 99.8% of the 4-by-4 matrices with integer entries in the interval  $[-10, 10]$  and on over 99.999% of the 4-by-4 matrices with integer entries in  $[-100, 100]$ .

The next example illustrates the faster procedure discussed above.

5.26 example: *minimal polynomial of an operator on  $\mathbf{F}^5$*

Suppose  $T \in \mathcal{L}(\mathbf{F}^5)$  and

$$\mathcal{M}(T) = \begin{pmatrix} 0 & 0 & 0 & 0 & -3 \\ 1 & 0 & 0 & 0 & 6 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

with respect to the standard basis  $e_1, e_2, e_3, e_4, e_5$ . Taking  $v = e_1$  for 5.25, we have

$$\begin{aligned} Te_1 &= e_2, & T^4e_1 &= T(T^3e_1) = Te_4 = e_5, \\ T^2e_1 &= T(Te_1) = Te_2 = e_3, & T^5e_1 &= T(T^4e_1) = Te_5 = -3e_1 + 6e_2. \\ T^3e_1 &= T(T^2e_1) = Te_3 = e_4, \end{aligned}$$

Thus  $3e_1 - 6Te_1 = -T^5e_1$ . The list  $e_1, Te_1, T^2e_1, T^3e_1, T^4e_1$ , which equals the list  $e_1, e_2, e_3, e_4, e_5$ , is linearly independent, so no other linear combination of this list equals  $-T^5e_1$ . Hence the minimal polynomial of  $T$  is  $3 - 6z + z^5$ .

Recall that by definition, eigenvalues of operators on  $V$  and zeros of polynomials in  $\mathcal{P}(\mathbf{F})$  must be elements of  $\mathbf{F}$ . In particular, if  $\mathbf{F} = \mathbf{R}$ , then eigenvalues and zeros must be real numbers.

5.27 *eigenvalues are the zeros of the minimal polynomial*

Suppose  $V$  is finite-dimensional and  $T \in \mathcal{L}(V)$ .

- The zeros of the minimal polynomial of  $T$  are the eigenvalues of  $T$ .
- If  $V$  is a complex vector space, then the minimal polynomial of  $T$  has the form

$$(z - \lambda_1) \cdots (z - \lambda_m),$$

where  $\lambda_1, \dots, \lambda_m$  is a list of all eigenvalues of  $T$ , possibly with repetitions.

**Proof** Let  $p$  be the minimal polynomial of  $T$ .

- First suppose  $\lambda \in \mathbf{F}$  is a zero of  $p$ . Then  $p$  can be written in the form

$$p(z) = (z - \lambda)q(z),$$

where  $q$  is a monic polynomial with coefficients in  $\mathbf{F}$  (see 4.6). Because  $p(T) = 0$ , we have

$$0 = (T - \lambda I)(q(T)v)$$

for all  $v \in V$ . Because the degree of  $q$  is less than the degree of the minimal polynomial  $p$ , there exists at least one vector  $v \in V$  such that  $q(T)v \neq 0$ . The equation above thus implies that  $\lambda$  is an eigenvalue of  $T$ , as desired.

To prove that every eigenvalue of  $T$  is a zero of  $p$ , now suppose  $\lambda \in \mathbf{F}$  is an eigenvalue of  $T$ . Thus there exists  $v \in V$  with  $v \neq 0$  such that  $Tv = \lambda v$ . Repeated applications of  $T$  to both sides of this equation show that  $T^k v = \lambda^k v$  for every nonnegative integer  $k$ . Thus

$$p(T)v = p(\lambda)v.$$

Because  $p$  is the minimal polynomial of  $T$ , we have  $p(T)v = 0$ . Hence the equation above implies that  $p(\lambda) = 0$ . Thus  $\lambda$  is a zero of  $p$ , as desired.

- (b) To get the desired result, use (a) and the second version of the fundamental theorem of algebra (see 4.13). ■

A nonzero polynomial has at most as many distinct zeros as its degree (see 4.8). Thus (a) of the previous result, along with the result that the minimal polynomial of an operator on  $V$  has degree at most  $\dim V$ , gives an alternative proof of 5.12, which states that an operator on  $V$  has at most  $\dim V$  distinct eigenvalues.

Every monic polynomial is the minimal polynomial of some operator, as shown by Exercise 16, which generalizes Example 5.26. Thus 5.27(a) shows that finding exact expressions for the eigenvalues of an operator is equivalent to the problem of finding exact expressions for the zeros of a polynomial (and thus is not possible for some operators).

5.28 example: *An operator whose eigenvalues cannot be found exactly*

Let  $T \in \mathcal{L}(\mathbf{C}^5)$  be the operator defined by

$$T(z_1, z_2, z_3, z_4, z_5) = (-3z_5, z_1 + 6z_5, z_2, z_3, z_4).$$

The matrix of  $T$  with respect to the standard basis of  $\mathbf{C}^5$  is the 5-by-5 matrix in Example 5.26. As we showed in that example, the minimal polynomial of  $T$  is the polynomial

$$3 - 6z + z^5.$$

No zero of the polynomial above can be expressed using rational numbers, roots of rational numbers, and the usual rules of arithmetic (a proof of this would take us considerably beyond linear algebra). Because the zeros of the polynomial above are the eigenvalues of  $T$  [by 5.27(a)], we cannot find an exact expression for any eigenvalue of  $T$  in any familiar form.

Numeric techniques, which we will not discuss here, show that the zeros of the polynomial above, and thus the eigenvalues of  $T$ , are approximately the following five complex numbers:

$$-1.67, \quad 0.51, \quad 1.40, \quad -0.12 + 1.59i, \quad -0.12 - 1.59i.$$

Note that the two nonreal zeros of this polynomial are complex conjugates of each other, as we expect for a polynomial with real coefficients (see 4.14).

The next result completely characterizes the polynomials that when applied to an operator give the 0 operator.

5.29  $q(T) = 0 \iff q$  is a polynomial multiple of the minimal polynomial

Suppose  $V$  is finite-dimensional,  $T \in \mathcal{L}(V)$ , and  $q \in \mathcal{P}(\mathbf{F})$ . Then  $q(T) = 0$  if and only if  $q$  is a polynomial multiple of the minimal polynomial of  $T$ .

**Proof** Let  $p$  denote the minimal polynomial of  $T$ .

First suppose  $q(T) = 0$ . By the division algorithm for polynomials (4.9), there exist polynomials  $s, r \in \mathcal{P}(\mathbf{F})$  such that

$$5.30 \quad q = ps + r$$

and  $\deg r < \deg p$ . We have

$$0 = q(T) = p(T)s(T) + r(T) = r(T).$$

The equation above implies that  $r = 0$  (otherwise, dividing  $r$  by its highest-degree coefficient would produce a monic polynomial that when applied to  $T$  gives 0; this polynomial would have a smaller degree than the minimal polynomial, which would be a contradiction). Thus 5.30 becomes the equation  $q = ps$ . Hence  $q$  is a polynomial multiple of  $p$ , as desired.

To prove the other direction, now suppose  $q$  is a polynomial multiple of  $p$ . Thus there exists a polynomial  $s \in \mathcal{P}(\mathbf{F})$  such that  $q = ps$ . We have

$$q(T) = p(T)s(T) = 0s(T) = 0,$$

as desired. ■

The next result is a nice consequence of the result above.

5.31 *minimal polynomial of a restriction operator*

Suppose  $V$  is finite-dimensional,  $T \in \mathcal{L}(V)$ , and  $U$  is a subspace of  $V$  that is invariant under  $T$ . Then the minimal polynomial of  $T$  is a polynomial multiple of the minimal polynomial of  $T|_U$ .

**Proof** Suppose  $p$  is the minimal polynomial of  $T$ . Thus  $p(T)v = 0$  for all  $v \in V$ . In particular,

$$p(T)u = 0 \text{ for all } u \in U.$$

Thus  $p(T|_U) = 0$ . Now 5.29, applied to the operator  $T|_U$  in place of  $T$ , implies that  $p$  is a polynomial multiple of the minimal polynomial of  $T|_U$ . ■

See Exercise 25 for a result about quotient operators that is analogous to the result above.

The next result shows that the constant term of the minimal polynomial of an operator determines whether the operator is invertible.

5.32  $T$  not invertible  $\iff$  constant term of minimal polynomial of  $T$  is 0

Suppose  $V$  is finite-dimensional and  $T \in \mathcal{L}(V)$ . Then  $T$  is not invertible if and only if the constant term of the minimal polynomial of  $T$  is 0.

**Proof** Suppose  $T \in \mathcal{L}(V)$  and  $p$  is the minimal polynomial of  $T$ . Then

$$\begin{aligned} T \text{ is not invertible} &\iff 0 \text{ is an eigenvalue of } T \\ &\iff 0 \text{ is a zero of } p \\ &\iff \text{the constant term of } p \text{ is } 0, \end{aligned}$$

where the first equivalence holds by 5.7, the second equivalence holds by 5.27(a), and the last equivalence holds because the constant term of  $p$  equals  $p(0)$ . ■

### Eigenvalues on Odd-Dimensional Real Vector Spaces

The next result will be the key tool that we use to show that every operator on an odd-dimensional real vector space has an eigenvalue.

5.33 even-dimensional null space

Suppose  $\mathbf{F} = \mathbf{R}$  and  $V$  is finite-dimensional. Suppose also that  $T \in \mathcal{L}(V)$  and  $b, c \in \mathbf{R}$  with  $b^2 < 4c$ . Then  $\dim \text{null}(T^2 + bT + cI)$  is an even number.

**Proof** Recall that  $\text{null}(T^2 + bT + cI)$  is invariant under  $T$  (by 5.18). By replacing  $V$  with  $\text{null}(T^2 + bT + cI)$  and replacing  $T$  with  $T$  restricted to  $\text{null}(T^2 + bT + cI)$ , we can assume that  $T^2 + bT + cI = 0$ ; we now need to prove that  $\dim V$  is even.

Suppose  $\lambda \in \mathbf{R}$  and  $v \in V$  are such that  $Tv = \lambda v$ . Then

$$0 = (T^2 + bT + cI)v = (\lambda^2 + b\lambda + c)v = \left( \left( \lambda + \frac{b}{2} \right)^2 + c - \frac{b^2}{4} \right)v.$$

The term in large parentheses above is a positive number. Thus the equation above implies that  $v = 0$ . Hence we have shown that  $T$  has no eigenvectors.

Let  $U$  be a subspace of  $V$  that is invariant under  $T$  and has the largest dimension among all subspaces of  $V$  that are invariant under  $T$  and have even dimension. If  $U = V$ , then we are done; otherwise assume there exists  $w \in V$  such that  $w \notin U$ .

Let  $W = \text{span}(w, Tw)$ . Then  $W$  is invariant under  $T$  because  $T(Tw) = -bTw - cw$ . Furthermore,  $\dim W = 2$  because otherwise  $w$  would be an eigenvector of  $T$ . Now

$$\dim(U + W) = \dim U + \dim W - \dim(U \cap W) = \dim U + 2,$$

where  $U \cap W = \{0\}$  because otherwise  $U \cap W$  would be a one-dimensional subspace of  $V$  that is invariant under  $T$  (impossible because  $T$  has no eigenvectors).

Because  $U + W$  is invariant under  $T$ , the equation above shows that there exists a subspace of  $V$  invariant under  $T$  of even dimension larger than  $\dim U$ . Thus the assumption that  $U \neq V$  was incorrect. Hence  $V$  has even dimension. ■



The next result states that on odd-dimensional vector spaces, every operator has an eigenvalue. We already know this result for finite-dimensional complex vector spaces (without the odd hypothesis). Thus in the proof below, we will assume that  $\mathbf{F} = \mathbf{R}$ .

### 5.34 operators on odd-dimensional vector spaces have eigenvalues

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

**Proof** Suppose  $\mathbf{F} = \mathbf{R}$  and  $V$  is finite-dimensional. Let  $n = \dim V$ , and suppose  $n$  is an odd number. Let  $T \in \mathcal{L}(V)$ . We will use induction on  $n$  in steps of size two to show that  $T$  has an eigenvalue. To get started, note that the desired result holds if  $\dim V = 1$  because then every nonzero vector in  $V$  is an eigenvector of  $T$ .

Now suppose that  $n \geq 3$  and the desired result holds for all operators on all odd-dimensional vector spaces of dimension less than  $n$ . Let  $p$  denote the minimal polynomial of  $T$ . If  $p$  is a polynomial multiple of  $x - \lambda$  for some  $\lambda \in \mathbf{R}$ , then  $\lambda$  is an eigenvalue of  $T$  [by 5.27(a)] and we are done. Thus we can assume that there exist  $b, c \in \mathbf{R}$  such that  $b^2 < 4c$  and  $p$  is a polynomial multiple of  $x^2 + bx + c$  (see 4.16).

There exists a monic polynomial  $q \in \mathcal{P}(\mathbf{R})$  such that  $p(x) = q(x)(x^2 + bx + c)$  for all  $x \in \mathbf{R}$ . Now

$$0 = p(T) = (q(T))(T^2 + bT + cI),$$

which means that  $q(T)$  equals 0 on  $\text{range}(T^2 + bT + cI)$ . Because  $\deg q < \deg p$  and  $p$  is the minimal polynomial of  $T$ , this implies that  $\text{range}(T^2 + bT + cI) \neq V$ .

The fundamental theorem of linear maps (3.21) tells us that

$$\dim V = \dim \text{null}(T^2 + bT + cI) + \dim \text{range}(T^2 + bT + cI).$$

Because  $\dim V$  is odd (by hypothesis) and  $\dim \text{null}(T^2 + bT + cI)$  is even (by 5.33), the equation above shows that  $\dim \text{range}(T^2 + bT + cI)$  is odd.

Hence  $\text{range}(T^2 + bT + cI)$  is a subspace of  $V$  that is invariant under  $T$  (by 5.18) and that has odd dimension less than  $\dim V$ . Our induction hypothesis now implies that  $T$  restricted to  $\text{range}(T^2 + bT + cI)$  has an eigenvalue, which means that  $T$  has an eigenvalue. ■

See Exercise 23 in Section 8B and Exercise 10 in Section 9C for alternative proofs of the result above.

## Exercises 5B

- 1 Suppose  $T \in \mathcal{L}(V)$ . Prove that 9 is an eigenvalue of  $T^2$  if and only if 3 or  $-3$  is an eigenvalue of  $T$ .
- 2 Suppose  $V$  is a complex vector space and  $T \in \mathcal{L}(V)$  has no eigenvalues. Prove that every subspace of  $V$  invariant under  $T$  is either  $\{0\}$  or infinite-dimensional.

- 3 Suppose  $n$  is a positive integer and  $T \in \mathcal{L}(\mathbf{F}^n)$  is defined by

$$T(x_1, \dots, x_n) = (x_1 + \dots + x_n, \dots, x_1 + \dots + x_n).$$

- (a) Find all eigenvalues and eigenvectors of  $T$ .  
 (b) Find the minimal polynomial of  $T$ .

*The matrix of  $T$  with respect to the standard basis of  $\mathbf{F}^n$  consists of all 1's.*

- 4 Suppose  $\mathbf{F} = \mathbf{C}$ ,  $T \in \mathcal{L}(V)$ ,  $p \in \mathcal{P}(\mathbf{C})$ , and  $\alpha \in \mathbf{C}$ . Prove that  $\alpha$  is an eigenvalue of  $p(T)$  if and only if  $\alpha = p(\lambda)$  for some eigenvalue  $\lambda$  of  $T$ .
- 5 Give an example of an operator on  $\mathbf{R}^2$  that shows the result in Exercise 4 does not hold if  $\mathbf{C}$  is replaced with  $\mathbf{R}$ .

- 6 Suppose  $T \in \mathcal{L}(\mathbf{F}^2)$  is defined by  $T(w, z) = (-z, w)$ . Find the minimal polynomial of  $T$ .

- 7 (a) Give an example of  $S, T \in \mathcal{L}(\mathbf{F}^2)$  such that the minimal polynomial of  $ST$  does not equal the minimal polynomial of  $TS$ .  
 (b) Suppose  $V$  is finite-dimensional and  $S, T \in \mathcal{L}(V)$ . Prove that if at least one of  $S, T$  is invertible, then the minimal polynomial of  $ST$  equals the minimal polynomial of  $TS$ .

*Hint: Show that if  $S$  is invertible and  $p \in \mathcal{P}(\mathbf{F})$ , then  $p(TS) = S^{-1}p(ST)S$ .*

- 8 Suppose  $T \in \mathcal{L}(\mathbf{R}^2)$  is the operator of counterclockwise rotation by  $1^\circ$ . Find the minimal polynomial of  $T$ .

*Because  $\dim \mathbf{R}^2 = 2$ , the degree of the minimal polynomial of  $T$  is at most 2. Thus the minimal polynomial of  $T$  is not the tempting polynomial  $x^{180} + 1$ , even though  $T^{180} = -I$ .*

- 9 Suppose  $T \in \mathcal{L}(V)$  is such that with respect to some basis of  $V$ , all entries of the matrix of  $T$  are rational numbers. Explain why all coefficients of the minimal polynomial of  $T$  are rational numbers.
- 10 Suppose  $V$  is finite-dimensional,  $T \in \mathcal{L}(V)$ , and  $v \in V$ . Prove that

$$\text{span}(v, Tv, \dots, T^m v) = \text{span}(v, Tv, \dots, T^{\dim V - 1} v)$$

for all integers  $m \geq \dim V - 1$ .

- 11 Suppose  $V$  is a two-dimensional vector space,  $T \in \mathcal{L}(V)$ , and the matrix of  $T$  with respect to some basis of  $V$  is  $\begin{pmatrix} a & c \\ b & d \end{pmatrix}$ .

- (a) Show that  $T^2 - (a + d)T + (ad - bc)I = 0$ .  
 (b) Show that the minimal polynomial of  $T$  equals

$$\begin{cases} z - a & \text{if } b = c = 0 \text{ and } a = d, \\ z^2 - (a + d)z + (ad - bc) & \text{otherwise.} \end{cases}$$

- 12 Define  $T \in \mathcal{L}(\mathbf{F}^n)$  by  $T(x_1, x_2, x_3, \dots, x_n) = (x_1, 2x_2, 3x_3, \dots, nx_n)$ . Find the minimal polynomial of  $T$ .
- 13 Suppose  $T \in \mathcal{L}(V)$  and  $p \in \mathcal{P}(\mathbf{F})$ . Prove that there exists a unique  $r \in \mathcal{P}(\mathbf{F})$  such that  $p(T) = r(T)$  and  $\deg r$  is less than the degree of the minimal polynomial of  $T$ .
- 14 Suppose  $V$  is finite-dimensional and  $T \in \mathcal{L}(V)$  has minimal polynomial  $4 + 5z - 6z^2 - 7z^3 + 2z^4 + z^5$ . Find the minimal polynomial of  $T^{-1}$ .
- 15 Suppose  $V$  is a finite-dimensional complex vector space with  $\dim V > 0$  and  $T \in \mathcal{L}(V)$ . Define  $f: \mathbf{C} \rightarrow \mathbf{R}$  by

$$f(\lambda) = \dim \text{range}(T - \lambda I).$$

Prove that  $f$  is not a continuous function.

- 16 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

$$a_0 + a_1 z + \dots + a_{n-1} z^{n-1} + z^n.$$

*The matrix above is called the **companion matrix** of the polynomial above. This exercise shows that every monic polynomial is the minimal polynomial of some operator. Hence a formula or an algorithm that could produce exact eigenvalues for each operator on each  $\mathbf{F}^n$  could then produce exact zeros for each polynomial [by 5.27(a)]. Thus there is no such formula or algorithm. However, efficient numeric methods exist for obtaining very good approximations for the eigenvalues of an operator.*

- 17 Suppose  $V$  is finite-dimensional,  $T \in \mathcal{L}(V)$ , and  $p$  is the minimal polynomial of  $T$ . Suppose  $\lambda \in \mathbf{F}$ . Show that the minimal polynomial of  $T - \lambda I$  is the polynomial  $q$  defined by  $q(z) = p(z + \lambda)$ .
- 18 Suppose  $V$  is finite-dimensional,  $T \in \mathcal{L}(V)$ , and  $p$  is the minimal polynomial of  $T$ . Suppose  $\lambda \in \mathbf{F} \setminus \{0\}$ . Show that the minimal polynomial of  $\lambda T$  is the polynomial  $q$  defined by  $q(z) = \lambda^{\deg p} p\left(\frac{z}{\lambda}\right)$ .

- 19 Suppose  $V$  is finite-dimensional and  $T \in \mathcal{L}(V)$ . Let  $\mathcal{E}$  be the subspace of  $\mathcal{L}(V)$  defined by

$$\mathcal{E} = \{q(T) : q \in \mathcal{P}(\mathbf{F})\}.$$

Prove that  $\dim \mathcal{E}$  equals the degree of the minimal polynomial of  $T$ .

- 20 Suppose  $T \in \mathcal{L}(\mathbf{F}^4)$  is such that the eigenvalues of  $T$  are 3, 5, 8. Prove that  $(T - 3I)^2(T - 5I)^2(T - 8I)^2 = 0$ .

- 21 Suppose  $V$  is finite-dimensional and  $T \in \mathcal{L}(V)$ . Prove that the minimal polynomial of  $T$  has degree at most  $1 + \dim \text{range } T$ .

*If  $\dim \text{range } T < \dim V - 1$ , then this exercise gives a better upper bound than 5.22 for the degree of the minimal polynomial of  $T$ .*

- 22 Suppose  $V$  is finite-dimensional and  $T \in \mathcal{L}(V)$ . Prove that  $T$  is invertible if and only if  $I \in \text{span}(T, T^2, \dots, T^{\dim V})$ .

- 23 Suppose  $V$  is finite-dimensional and  $T \in \mathcal{L}(V)$ . Let  $n = \dim V$ . Prove that if  $v \in V$ , then  $\text{span}(v, Tv, \dots, T^{n-1}v)$  is invariant under  $T$ .

- 24 Suppose  $V$  is a finite-dimensional complex vector space. Suppose  $T \in \mathcal{L}(V)$  is such that 5 and 6 are eigenvalues of  $T$  and that  $T$  has no other eigenvalues. Prove that  $(T - 5I)^{\dim V - 1}(T - 6I)^{\dim V - 1} = 0$ .

- 25 Suppose  $V$  is finite-dimensional,  $T \in \mathcal{L}(V)$ , and  $U$  is a subspace of  $V$  that is invariant under  $T$ .

(a) Prove that the minimal polynomial of  $T$  is a polynomial multiple of the minimal polynomial of the quotient operator  $T/U$ .

(b) Prove that

$$(\text{minimal polynomial of } T|_U) \times (\text{minimal polynomial of } T/U)$$

is a polynomial multiple of the minimal polynomial of  $T$ .

*The quotient operator  $T/U$  was defined in Exercise 38 in Section 5A.*

- 26 Suppose  $V$  is finite-dimensional,  $T \in \mathcal{L}(V)$ , and  $U$  is a subspace of  $V$  that is invariant under  $T$ . Prove that the set of eigenvalues of  $T$  equals the union of the set of eigenvalues of  $T|_U$  and the set of eigenvalues of  $T/U$ .

- 27 Suppose  $\mathbf{F} = \mathbf{R}$ ,  $V$  is finite-dimensional, and  $T \in \mathcal{L}(V)$ . Prove that the minimal polynomial of  $T_{\mathbf{C}}$  equals the minimal polynomial of  $T$ .

*The complexification  $T_{\mathbf{C}}$  was defined in Exercise 33 of Section 3B.*

- 28 Suppose  $V$  is finite-dimensional and  $T \in \mathcal{L}(V)$ . Prove that the minimal polynomial of  $T' \in \mathcal{L}(V')$  equals the minimal polynomial of  $T$ .

*The dual map  $T'$  was defined in Section 3F.*

- 29 Show that every operator on a finite-dimensional vector space of dimension at least two has an invariant subspace of dimension two.

*Exercise 6 in Section 5C will give an improvement of this result when  $\mathbf{F} = \mathbf{C}$ .*

## 5C Upper-Triangular Matrices

In Chapter 3 we defined the matrix of a linear map from a finite-dimensional vector space to another finite-dimensional vector space. That matrix depends on a choice of basis of each of the two vector spaces. Now that we are studying operators, which map a vector space to itself, the emphasis is on using only one basis.

### 5.35 definition: matrix of an operator, $\mathcal{M}(T)$

Suppose  $T \in \mathcal{L}(V)$ . The *matrix of  $T$*  with respect to a basis  $v_1, \dots, v_n$  of  $V$  is the  $n$ -by- $n$  matrix

$$\mathcal{M}(T) = \begin{pmatrix} A_{1,1} & \cdots & A_{1,n} \\ \vdots & & \vdots \\ A_{n,1} & \cdots & A_{n,n} \end{pmatrix}$$

whose entries  $A_{j,k}$  are defined by

$$Tv_k = A_{1,k}v_1 + \cdots + A_{n,k}v_n.$$

The notation  $\mathcal{M}(T, (v_1, \dots, v_n))$  is used if the basis is not clear from the context.

Operators have square matrices (meaning that the number of rows equals the number of columns), rather than the more general rectangular matrices that we considered earlier for linear maps.

If  $T$  is an operator on  $\mathbf{F}^n$  and no basis is specified, assume that the basis in question is the standard one (where the  $k^{\text{th}}$  basis vector is 1 in the  $k^{\text{th}}$  slot and 0 in all other slots). You can then think of the  $k^{\text{th}}$  column of  $\mathcal{M}(T)$  as  $T$  applied to the  $k^{\text{th}}$  basis vector, where we identify  $n$ -by-1 column vectors with elements of  $\mathbf{F}^n$ .

*The  $k^{\text{th}}$  column of the matrix  $\mathcal{M}(T)$  is formed from the coefficients used to write  $Tv_k$  as a linear combination of the basis  $v_1, \dots, v_n$ .*

### 5.36 example: matrix of an operator with respect to standard basis

Define  $T \in \mathcal{L}(\mathbf{F}^3)$  by  $T(x, y, z) = (2x + y, 5y + 3z, 8z)$ . Then the matrix of  $T$  with respect to the standard basis of  $\mathbf{F}^3$  is

$$\mathcal{M}(T) = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 5 & 3 \\ 0 & 0 & 8 \end{pmatrix},$$

as you should verify.

A central goal of linear algebra is to show that given an operator  $T$  on a finite-dimensional vector space  $V$ , there exists a basis of  $V$  with respect to which  $T$  has a reasonably simple matrix. To make this vague formulation a bit more precise, we might try to choose a basis of  $V$  such that  $\mathcal{M}(T)$  has many 0's.

If  $V$  is a finite-dimensional complex vector space, then we already know enough to show that there is a basis of  $V$  with respect to which the matrix of  $T$  has 0's everywhere in the first column, except possibly the first entry. In other words, there is a basis of  $V$  with respect to which the matrix of  $T$  looks like

$$\begin{pmatrix} \lambda & & \\ 0 & * & \\ \vdots & & \\ 0 & & \end{pmatrix};$$

here  $*$  denotes the entries in all columns other than the first column. To prove this, let  $\lambda$  be an eigenvalue of  $T$  (one exists by 5.19) and let  $v$  be a corresponding eigenvector. Extend  $v$  to a basis of  $V$ . Then the matrix of  $T$  with respect to this basis has the form above. Soon we will see that we can choose a basis of  $V$  with respect to which the matrix of  $T$  has even more 0's.

5.37 definition: *diagonal of a matrix*

The *diagonal* of a square matrix consists of the entries on the line from the upper left corner to the bottom right corner.

For example, the diagonal of the matrix

$$\mathcal{M}(T) = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 5 & 3 \\ 0 & 0 & 8 \end{pmatrix}$$

from Example 5.36 consists of the entries 2, 5, 8, which are shown in red in the matrix above.

5.38 definition: *upper-triangular matrix*

A square matrix is called *upper triangular* if all entries below the diagonal are 0.

For example, the 3-by-3 matrix above is upper triangular.

Typically we represent an upper-triangular matrix in the form

$$\begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix};$$

the 0 in the matrix above indicates that all entries below the diagonal in this  $n$ -by- $n$  matrix equal 0. Upper-triangular matrices can be considered reasonably simple—if  $n$  is large, then at least almost half the entries in an  $n$ -by- $n$  upper-triangular matrix are 0.

We often use  $*$  to denote matrix entries that we do not know or that are irrelevant to the questions being discussed.

The next result provides a useful connection between upper-triangular matrices and invariant subspaces.

### 5.39 conditions for upper-triangular matrix

Suppose  $T \in \mathcal{L}(V)$  and  $v_1, \dots, v_n$  is a basis of  $V$ . Then the following are equivalent.

- (a) The matrix of  $T$  with respect to  $v_1, \dots, v_n$  is upper triangular.
- (b)  $\text{span}(v_1, \dots, v_k)$  is invariant under  $T$  for each  $k = 1, \dots, n$ .
- (c)  $Tv_k \in \text{span}(v_1, \dots, v_k)$  for each  $k = 1, \dots, n$ .

**Proof** First suppose (a) holds. To prove that (b) holds, suppose  $k \in \{1, \dots, n\}$ . If  $j \in \{1, \dots, n\}$ , then

$$Tv_j \in \text{span}(v_1, \dots, v_j)$$

because the matrix of  $T$  with respect to  $v_1, \dots, v_n$  is upper triangular. Because  $\text{span}(v_1, \dots, v_j) \subseteq \text{span}(v_1, \dots, v_k)$  if  $j \leq k$ , we see that

$$Tv_j \in \text{span}(v_1, \dots, v_k)$$

for each  $j \in \{1, \dots, k\}$ . Thus  $\text{span}(v_1, \dots, v_k)$  is invariant under  $T$ , completing the proof that (a) implies (b).

Now suppose (b) holds, so  $\text{span}(v_1, \dots, v_k)$  is invariant under  $T$  for each  $k = 1, \dots, n$ . In particular,  $Tv_k \in \text{span}(v_1, \dots, v_k)$  for each  $k = 1, \dots, n$ . Thus (b) implies (c).

Now suppose (c) holds, so  $Tv_k \in \text{span}(v_1, \dots, v_k)$  for each  $k = 1, \dots, n$ . This means that when writing each  $Tv_k$  as a linear combination of the basis vectors  $v_1, \dots, v_n$ , we need to use only the vectors  $v_1, \dots, v_k$ . Hence all entries under the diagonal of  $\mathcal{M}(T)$  are 0. Thus  $\mathcal{M}(T)$  is an upper-triangular matrix, completing the proof that (c) implies (a).

We have shown that (a)  $\implies$  (b)  $\implies$  (c)  $\implies$  (a), which shows that (a), (b), and (c) are equivalent. ■

The next result tells us that if  $T \in \mathcal{L}(V)$  and with respect to some basis of  $V$  we have

$$\mathcal{M}(T) = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix},$$

then  $T$  satisfies a simple equation depending on  $\lambda_1, \dots, \lambda_n$ .

### 5.40 equation satisfied by operator with upper-triangular matrix

Suppose  $T \in \mathcal{L}(V)$  and  $V$  has a basis with respect to which  $T$  has an upper-triangular matrix with diagonal entries  $\lambda_1, \dots, \lambda_n$ . Then

$$(T - \lambda_1 I) \cdots (T - \lambda_n I) = 0.$$

**Proof** Let  $v_1, \dots, v_n$  denote a basis of  $V$  with respect to which  $T$  has an upper-triangular matrix with diagonal entries  $\lambda_1, \dots, \lambda_n$ . Then  $Tv_1 = \lambda_1 v_1$ , which means that  $(T - \lambda_1 I)v_1 = 0$ , which implies that  $(T - \lambda_1 I) \cdots (T - \lambda_m I)v_1 = 0$  for  $m = 1, \dots, n$  (using the commutativity of each  $T - \lambda_j I$  with each  $T - \lambda_k I$ ).

Note that  $(T - \lambda_2 I)v_2 \in \text{span}(v_1)$ . Thus  $(T - \lambda_1 I)(T - \lambda_2 I)v_2 = 0$  (by the previous paragraph), which implies that  $(T - \lambda_1 I) \cdots (T - \lambda_m I)v_1 = 0$  for  $m = 2, \dots, n$  (using the commutativity of each  $T - \lambda_j I$  with each  $T - \lambda_k I$ ).

Note that  $(T - \lambda_3 I)v_3 \in \text{span}(v_1, v_2)$ . Thus by the previous paragraph,  $(T - \lambda_1 I)(T - \lambda_2 I)(T - \lambda_3 I)v_3 = 0$ , which implies that  $(T - \lambda_1 I) \cdots (T - \lambda_m I)v_1 = 0$  for  $m = 3, \dots, n$  (using the commutativity of each  $T - \lambda_j I$  with each  $T - \lambda_k I$ ).

Continuing this pattern, we see that  $(T - \lambda_1 I) \cdots (T - \lambda_n I)v_k = 0$  for each  $k = 1, \dots, n$ . Thus  $(T - \lambda_1 I) \cdots (T - \lambda_n I)$  is the 0 operator because it is 0 on each vector in a basis of  $V$ . ■

Unfortunately no method exists for exactly computing the eigenvalues of an operator from its matrix. However, if we are fortunate enough to find a basis with respect to which the matrix of the operator is upper triangular, then the problem of computing the eigenvalues becomes trivial, as the next result shows.

#### 5.41 *determination of eigenvalues from upper-triangular matrix*

Suppose  $T \in \mathcal{L}(V)$  has an upper-triangular matrix with respect to some basis of  $V$ . Then the eigenvalues of  $T$  are precisely the entries on the diagonal of that upper-triangular matrix.

**Proof** Suppose  $v_1, \dots, v_n$  is a basis of  $V$  with respect to which  $T$  has an upper-triangular matrix

$$\mathcal{M}(T) = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

Because  $Tv_1 = \lambda_1 v_1$ , we see that  $\lambda_1$  is an eigenvalue of  $T$ .

Suppose  $k \in \{2, \dots, n\}$ . Then  $(T - \lambda_k I)v_k \in \text{span}(v_1, \dots, v_{k-1})$ . Thus  $T - \lambda_k I$  maps  $\text{span}(v_1, \dots, v_k)$  into  $\text{span}(v_1, \dots, v_{k-1})$ . Because

$$\dim \text{span}(v_1, \dots, v_k) = k \quad \text{and} \quad \dim \text{span}(v_1, \dots, v_{k-1}) = k - 1,$$

this implies that  $T - \lambda_k I$  restricted to  $\text{span}(v_1, \dots, v_k)$  is not injective (by 3.22). Thus there exists  $v \in \text{span}(v_1, \dots, v_k)$  such that  $v \neq 0$  and  $(T - \lambda_k I)v = 0$ . Thus  $\lambda_k$  is an eigenvalue of  $T$ . Hence we have shown that every entry on the diagonal of  $\mathcal{M}(T)$  is an eigenvalue of  $T$ .

To prove  $T$  has no other eigenvalues, let  $q$  be the polynomial defined by  $q(z) = (z - \lambda_1) \cdots (z - \lambda_n)$ . Then  $q(T) = 0$  (by 5.40). Hence  $q$  is a polynomial multiple of the minimal polynomial of  $T$  (by 5.29). Thus every zero of the minimal polynomial of  $T$  is a zero of  $q$ . Because the zeros of the minimal polynomial of  $T$  are the eigenvalues of  $T$  (by 5.27), this implies that every eigenvalue of  $T$  is a zero of  $q$ . Hence the eigenvalues of  $T$  are all contained in the list  $\lambda_1, \dots, \lambda_n$ . ■



5.42 example: *eigenvalues via an upper-triangular matrix*

Define  $T \in \mathcal{L}(\mathbf{F}^3)$  by  $T(x, y, z) = (2x + y, 5y + 3z, 8z)$ . The matrix of  $T$  with respect to the standard basis is

$$\mathcal{M}(T) = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 5 & 3 \\ 0 & 0 & 8 \end{pmatrix}.$$

Now 5.41 implies that the eigenvalues of  $T$  are 2, 5, and 8.

The next example illustrates 5.44: an operator has an upper-triangular matrix with respect to some basis if and only if the minimal polynomial of the operator is the product of polynomials of degree 1.

5.43 example: *whether  $T$  has an upper-triangular matrix can depend on  $\mathbf{F}$* 

Define  $T \in \mathcal{L}(\mathbf{F}^4)$  by

$$T(z_1, z_2, z_3, z_4) = (-z_2, z_1, 2z_1 + 3z_3, z_3 + 3z_4).$$

Thus with respect to the standard basis of  $\mathbf{F}^4$ , the matrix of  $T$  is

$$\begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 2 & 0 & 3 & 0 \\ 0 & 0 & 1 & 3 \end{pmatrix}.$$

You can ask a computer to verify that the minimal polynomial of  $T$  is the polynomial  $p$  defined by

$$p(z) = 9 - 6z + 10z^2 - 6z^3 + z^4.$$

First consider the case  $\mathbf{F} = \mathbf{R}$ . Then the polynomial  $p$  factors as

$$p(z) = (z^2 + 1)(z - 3)(z - 3),$$

with no further factorization of  $z^2 + 1$  as the product of two polynomials of degree 1 with real coefficients. Thus 5.44 states that there does not exist a basis of  $\mathbf{R}^4$  with respect to which  $T$  has an upper-triangular matrix.

Now consider the case  $\mathbf{F} = \mathbf{C}$ . Then the polynomial  $p$  factors as

$$p(z) = (z - i)(z + i)(z - 3)(z - 3),$$

where all factors above have the form  $z - \lambda_k$ . Thus 5.44 states that there is a basis of  $\mathbf{C}^4$  with respect to which  $T$  has an upper-triangular matrix. Indeed, you can verify that with respect to the basis  $(4 - 3i, -3 - 4i, -3 + i, 1)$ ,  $(4 + 3i, -3 + 4i, -3 - i, 1)$ ,  $(0, 0, 0, 1)$ ,  $(0, 0, 1, 0)$  of  $\mathbf{C}^4$ , the operator  $T$  has the upper-triangular matrix

$$\begin{pmatrix} i & 0 & 0 & 0 \\ 0 & -i & 0 & 0 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 3 \end{pmatrix}.$$

## 5.44 necessary and sufficient condition to have an upper-triangular matrix

Suppose  $V$  is finite-dimensional and  $T \in \mathcal{L}(V)$ . Then  $T$  has an upper-triangular matrix with respect to some basis of  $V$  if and only if the minimal polynomial of  $T$  equals  $(z - \lambda_1) \cdots (z - \lambda_m)$  for some  $\lambda_1, \dots, \lambda_m \in \mathbf{F}$ .

**Proof** First suppose  $T$  has an upper-triangular matrix with respect to some basis of  $V$ . Let  $\alpha_1, \dots, \alpha_n$  denote the diagonal entries of that matrix. Define a polynomial  $q \in \mathcal{P}(\mathbf{F})$  by

$$q(z) = (z - \alpha_1) \cdots (z - \alpha_n).$$

Then  $q(T) = 0$ , by 5.40. Hence  $q$  is a polynomial multiple of the minimal polynomial of  $T$ , by 5.29. Thus the minimal polynomial of  $T$  equals  $(z - \lambda_1) \cdots (z - \lambda_m)$  for some  $\lambda_1, \dots, \lambda_m \in \mathbf{F}$  with  $\{\lambda_1, \dots, \lambda_m\} \subseteq \{\alpha_1, \dots, \alpha_n\}$ .

To prove the implication in the other direction, now suppose the minimal polynomial of  $T$  equals  $(z - \lambda_1) \cdots (z - \lambda_m)$  for some  $\lambda_1, \dots, \lambda_m \in \mathbf{F}$ . We will use induction on  $m$ . To get started, if  $m = 1$  then  $z - \lambda_1$  is the minimal polynomial of  $T$ , which implies that  $T = \lambda_1 I$ , which implies that the matrix of  $T$  (with respect to any basis of  $V$ ) is upper triangular.

Now suppose  $m > 1$  and the desired result holds for all smaller positive integers. Let

$$U = \text{range}(T - \lambda_m I).$$

Then  $U$  is invariant under  $T$  [this is a special case of 5.18 with  $p(z) = z - \lambda_m$ ]. Thus  $T|_U$  is an operator on  $U$ .

If  $u \in U$ , then  $u = (T - \lambda_m I)v$  for some  $v \in V$  and

$$(T - \lambda_1 I) \cdots (T - \lambda_{m-1} I)u = (T - \lambda_1 I) \cdots (T - \lambda_m I)v = 0.$$

Hence  $(z - \lambda_1) \cdots (z - \lambda_{m-1})$  is a polynomial multiple of the minimal polynomial of  $T|_U$ , by 5.29. Thus the minimal polynomial of  $T|_U$  is the product of at most  $m - 1$  terms of the form  $z - \lambda_k$ .

By our induction hypothesis, there is a basis  $u_1, \dots, u_M$  of  $U$  with respect to which  $T|_U$  has an upper-triangular matrix. Thus for each  $k \in \{1, \dots, M\}$ , we have (using 5.39)

$$5.45 \quad Tu_k = (T|_U)(u_k) \in \text{span}(u_1, \dots, u_k).$$

Extend  $u_1, \dots, u_M$  to a basis  $u_1, \dots, u_M, v_1, \dots, v_N$  of  $V$ . For each  $k \in \{1, \dots, N\}$ , we have

$$Tv_k = (T - \lambda_m I)v_k + \lambda_m v_k.$$

The definition of  $U$  shows that  $(T - \lambda_m I)v_k \in U = \text{span}(u_1, \dots, u_M)$ . Thus the equation above shows that

$$5.46 \quad Tv_k \in \text{span}(u_1, \dots, u_M, v_1, \dots, v_k).$$

From 5.45 and 5.46, we conclude (using 5.39) that  $T$  has an upper-triangular matrix with respect to the basis  $u_1, \dots, u_M, v_1, \dots, v_N$  of  $V$ , as desired. ■

The set of numbers  $\{\lambda_1, \dots, \lambda_m\}$  from the previous result equals the set of eigenvalues of  $T$  (because the set of zeros of the minimal polynomial of  $T$  equals the set of eigenvalues of  $T$ , by 5.27), although the list  $\lambda_1, \dots, \lambda_m$  in the previous result may contain repetitions.

In Chapter 8 we will improve even the wonderful result below; see 8.37 and 8.46.

5.47 if  $\mathbf{F} = \mathbf{C}$ , then every operator on  $V$  has an upper-triangular matrix

Suppose  $V$  is a finite-dimensional complex vector space and  $T \in \mathcal{L}(V)$ . Then  $T$  has an upper-triangular matrix with respect to some basis of  $V$ .

**Proof** The desired result follows immediately from 5.44 and the second version of the fundamental theorem of algebra (see 4.13). ■

For an extension of the result above to two operators  $S$  and  $T$  such that

$$ST = TS,$$

see 5.80. Also, for an extension to more than two operators, see Exercise 9(b) in Section 5E.

**Caution:** If an operator  $T \in \mathcal{L}(V)$  has a upper-triangular matrix with respect to some basis  $v_1, \dots, v_n$  of  $V$ , then the eigenvalues of  $T$  are exactly the entries on the diagonal of  $\mathcal{M}(T)$ , as shown by 5.41, and furthermore  $v_1$  is an eigenvector of  $T$ . However,  $v_2, \dots, v_n$  need not be eigenvectors of  $T$ . Indeed, a basis vector  $v_k$  is an eigenvector of  $T$  if and only if all entries in the  $k^{\text{th}}$  column of the matrix of  $T$  are 0, except possibly the  $k^{\text{th}}$  entry.

You may recall from a previous course that every matrix of numbers can be changed to a matrix in what is called row echelon form. If one begins with a square matrix, the matrix in row echelon form will be an upper-triangular matrix. Do not confuse this upper-triangular matrix with the upper-triangular matrix of an operator with respect to some basis whose existence is proclaimed by 5.47 (if  $\mathbf{F} = \mathbf{C}$ )—there is no connection between these upper-triangular matrices.

*The row echelon form of the matrix of an operator does not give us a list of the eigenvalues of the operator. In contrast, an upper-triangular matrix with respect to some basis gives us a list of all the eigenvalues of the operator. However, there is no method for computing exactly such an upper-triangular matrix, even though 5.47 guarantees its existence if  $\mathbf{F} = \mathbf{C}$ .*

## Exercises 5C

- 1 Prove or give a counterexample: If  $T \in \mathcal{L}(V)$  and  $T^2$  has an upper-triangular matrix with respect to some basis of  $V$ , then  $T$  has an upper-triangular matrix with respect to some basis of  $V$ .

- 2 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  $B$ .
- Show that  $A + B$  is an upper-triangular matrix with  $\alpha_1 + \beta_1, \dots, \alpha_n + \beta_n$  on the diagonal.
  - Show that  $AB$  is an upper-triangular matrix with  $\alpha_1\beta_1, \dots, \alpha_n\beta_n$  on the diagonal.

*The results in this exercise are used in the proof of 5.81.*

- 3 Suppose  $T \in \mathcal{L}(V)$  is invertible and  $v_1, \dots, v_n$  is a basis of  $V$  with respect to which the matrix of  $T$  is upper triangular, with  $\lambda_1, \dots, \lambda_n$  on the diagonal. Show that the matrix of  $T^{-1}$  is also upper triangular with respect to the basis  $v_1, \dots, v_n$ , with

$$\frac{1}{\lambda_1}, \dots, \frac{1}{\lambda_n}$$

on the diagonal.

- 4 Give an example of an operator whose matrix with respect to some basis contains only 0's on the diagonal, but the operator is invertible.

*This exercise and the exercise below show that 5.41 fails without the hypothesis that an upper-triangular matrix is under consideration.*

- 5 Give an example of an operator whose matrix with respect to some basis contains only nonzero numbers on the diagonal, but the operator is not invertible.

- 6 Suppose  $F = \mathbf{C}$ ,  $V$  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  $T$ .

- 7 Suppose  $V$  is finite-dimensional,  $T \in \mathcal{L}(V)$ , and  $v \in V$ .

- Prove that there exists a unique monic polynomial  $p_v$  of smallest degree such that  $p_v(T)v = 0$ .
- Prove that the minimal polynomial of  $T$  is a polynomial multiple of  $p_v$ .

- 8 Suppose  $V$  is finite-dimensional,  $T \in \mathcal{L}(V)$ , and there exists a nonzero vector  $v \in V$  such that  $T^2v + 2Tv = -2v$ .

- Prove that if  $F = \mathbf{R}$ , then there does not exist a basis of  $V$  with respect to which  $T$  has an upper-triangular matrix.
- Prove that if  $F = \mathbf{C}$  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  $A$ .

- 9 Suppose  $B$  is a square matrix with complex entries. Prove that there exists an invertible square matrix  $A$  with complex entries such that  $A^{-1}BA$  is an upper-triangular matrix.

- 10** Suppose  $T \in \mathcal{L}(V)$  and  $v_1, \dots, v_n$  is a basis of  $V$ . Show that the following are equivalent.
- (a) The matrix of  $T$  with respect to  $v_1, \dots, v_n$  is lower triangular.
  - (b)  $\text{span}\langle v_k, \dots, v_n \rangle$  is invariant under  $T$  for each  $k = 1, \dots, n$ .
  - (c)  $Tv_k \in \text{span}\langle v_k, \dots, v_n \rangle$  for each  $k = 1, \dots, n$ .

A square matrix is called **lower triangular** if all entries above the diagonal are 0.

- 11** Suppose  $F = \mathbb{C}$  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.
- 12** Suppose  $V$  is finite-dimensional,  $T \in \mathcal{L}(V)$  has an upper-triangular matrix with respect to some basis of  $V$ , and  $U$  is a subspace of  $V$  that is invariant under  $T$ .
- (a) Prove that  $T|_U$  has an upper-triangular matrix with respect to some basis of  $U$ .
  - (b) Prove that the quotient operator  $T/U$  has an upper-triangular matrix with respect to some basis of  $V/U$ .

The quotient operator  $T/U$  was defined in Exercise 38 in Section 5A.

- 13** Suppose  $V$  is finite-dimensional and  $T \in \mathcal{L}(V)$ . Suppose there exists a subspace  $U$  of  $V$  that is invariant under  $T$  such that  $T|_U$  has an upper-triangular matrix with respect to some basis of  $U$  and also  $T/U$  has an upper-triangular matrix with respect to some basis of  $V/U$ . Prove that  $T$  has an upper-triangular matrix with respect to some basis of  $V$ .
- 14** Suppose  $V$  is finite-dimensional and  $T \in \mathcal{L}(V)$ . Prove that  $T$  has an upper-triangular matrix with respect to some basis of  $V$  if and only if the dual operator  $T'$  has an upper-triangular matrix with respect to some basis of the dual space  $V'$ .

## 5D Diagonalizable Operators

### Diagonal Matrices

5.48 definition: *diagonal matrix*

A *diagonal matrix* is a square matrix that is 0 everywhere except possibly on the diagonal.

5.49 example: *diagonal matrix*

$$\begin{pmatrix} 8 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 5 \end{pmatrix}$$

is a diagonal matrix.

If an operator has a diagonal matrix with respect to some basis, then the entries on the diagonal are precisely the eigenvalues of the operator; this follows from 5.41 (or find an easier direct proof for diagonal matrices).

*Every diagonal matrix is upper triangular. Diagonal matrices typically have many more 0's than most upper-triangular matrices of the same size.*

5.50 definition: *diagonalizable*

An operator on  $V$  is called *diagonalizable* if the operator has a diagonal matrix with respect to some basis of  $V$ .

5.51 example: *diagonalization may require a different basis*

Define  $T \in \mathcal{L}(\mathbf{R}^2)$  by

$$T(x, y) = (41x + 7y, -20x + 74y).$$

The matrix of  $T$  with respect to the standard basis of  $\mathbf{R}^2$  is

$$\begin{pmatrix} 41 & 7 \\ -20 & 74 \end{pmatrix},$$

which is not a diagonal matrix. However,  $T$  is diagonalizable. Specifically, the matrix of  $T$  with respect to the basis  $(1, 4), (7, 5)$  is

$$\begin{pmatrix} 69 & 0 \\ 0 & 46 \end{pmatrix}$$

because  $T(1, 4) = (69, 276) = 69(1, 4)$  and  $T(7, 5) = (322, 230) = 46(7, 5)$ .

For  $\lambda \in \mathbf{F}$ , we will find it convenient to have a name and a notation for the set of vectors that an operator  $T$  maps to  $\lambda$  times the vector.

5.52 definition: *eigenspace*,  $E(\lambda, T)$

Suppose  $T \in \mathcal{L}(V)$  and  $\lambda \in \mathbf{F}$ . The *eigenspace* of  $T$  corresponding to  $\lambda$  is the subspace  $E(\lambda, T)$  of  $V$  defined by

$$E(\lambda, T) = \text{null}(T - \lambda I) = \{v \in V : Tv = \lambda v\}.$$

Hence  $E(\lambda, T)$  is the set of all eigenvectors of  $T$  corresponding to  $\lambda$ , along with the 0 vector.

For  $T \in \mathcal{L}(V)$  and  $\lambda \in \mathbf{F}$ , the set  $E(\lambda, T)$  is a subspace of  $V$  because the null space of each linear map on  $V$  is a subspace of  $V$ . The definitions imply that  $\lambda$  is an eigenvalue of  $T$  if and only if  $E(\lambda, T) \neq \{0\}$ .

5.53 example: *eigenspaces of an operator*

Suppose the matrix of an operator  $T \in \mathcal{L}(V)$  with respect to a basis  $v_1, v_2, v_3$  of  $V$  is the matrix in Example 5.49. Then

$$E(8, T) = \text{span}(v_1), \quad E(5, T) = \text{span}(v_2, v_3).$$

If  $\lambda$  is an eigenvalue of an operator  $T \in \mathcal{L}(V)$ , then  $T$  restricted to  $E(\lambda, T)$  is just the operator of multiplication by  $\lambda$ .

5.54 *sum of eigenspaces is a direct sum*

Suppose  $T \in \mathcal{L}(V)$  and  $\lambda_1, \dots, \lambda_m$  are distinct eigenvalues of  $T$ . Then

$$E(\lambda_1, T) + \dots + E(\lambda_m, T)$$

is a direct sum. Furthermore, if  $V$  is finite-dimensional, then

$$\dim E(\lambda_1, T) + \dots + \dim E(\lambda_m, T) \leq \dim V.$$

**Proof** To show that  $E(\lambda_1, T) + \dots + E(\lambda_m, T)$  is a direct sum, suppose

$$v_1 + \dots + v_m = 0,$$

where each  $v_k$  is in  $E(\lambda_k, T)$ . Because eigenvectors corresponding to distinct eigenvalues are linearly independent (by 5.11), this implies that each  $v_k$  equals 0. Thus  $E(\lambda_1, T) + \dots + E(\lambda_m, T)$  is a direct sum (by 1.45), as desired.

Now suppose  $V$  is finite-dimensional. Then

$$\begin{aligned} \dim E(\lambda_1, T) + \dots + \dim E(\lambda_m, T) &= \dim(E(\lambda_1, T) \oplus \dots \oplus E(\lambda_m, T)) \\ &\leq \dim V, \end{aligned}$$

where the first line follows from 3.94 and the second line follows from 2.37. ■

## Conditions for Diagonalizability

The following characterizations of diagonalizable operators will be useful.

### 5.55 conditions equivalent to diagonalizability

Suppose  $V$  is finite-dimensional and  $T \in \mathcal{L}(V)$ . Let  $\lambda_1, \dots, \lambda_m$  denote the distinct eigenvalues of  $T$ . Then the following are equivalent.

- (a)  $T$  is diagonalizable.
- (b)  $V$  has a basis consisting of eigenvectors of  $T$ .
- (c)  $V = E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)$ .
- (d)  $\dim V = \dim E(\lambda_1, T) + \cdots + \dim E(\lambda_m, T)$ .

**Proof** An operator  $T \in \mathcal{L}(V)$  has a diagonal matrix

$$\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

with respect to a basis  $v_1, \dots, v_n$  of  $V$  if and only if  $Tv_k = \lambda_k v_k$  for each  $k$ . Thus (a) and (b) are equivalent.

Suppose (b) holds; thus  $V$  has a basis consisting of eigenvectors of  $T$ . Hence every vector in  $V$  is a linear combination of eigenvectors of  $T$ , which implies that

$$V = E(\lambda_1, T) + \cdots + E(\lambda_m, T).$$

Now 5.54 shows that (c) holds, proving that (b) implies (c).

That (c) implies (d) follows immediately from 3.94.

Finally, suppose (d) holds; thus

$$5.56 \quad \dim V = \dim E(\lambda_1, T) + \cdots + \dim E(\lambda_m, T).$$

Choose a basis of each  $E(\lambda_k, T)$ ; put all these bases together to form a list  $v_1, \dots, v_n$  of eigenvectors of  $T$ , where  $n = \dim V$  (by 5.56). To show that this list is linearly independent, suppose

$$a_1 v_1 + \cdots + a_n v_n = 0,$$

where  $a_1, \dots, a_n \in \mathbb{F}$ . For each  $k = 1, \dots, m$ , let  $u_k$  denote the sum of all the terms  $a_j v_j$  such that  $v_j \in E(\lambda_k, T)$ . Thus each  $u_k$  is in  $E(\lambda_k, T)$ , and

$$u_1 + \cdots + u_m = 0.$$

Because eigenvectors corresponding to distinct eigenvalues are linearly independent (see 5.11), this implies that each  $u_k$  equals 0. Because each  $u_k$  is a sum of terms  $a_j v_j$ , where the  $v_j$ 's were chosen to be a basis of  $E(\lambda_k, T)$ , this implies that all  $a_j$ 's equal 0. Thus  $v_1, \dots, v_n$  is linearly independent and hence is a basis of  $V$  (by 2.38). Thus (d) implies (b), completing the proof. ■

For additional conditions equivalent to diagonalizability, see 5.62, Exercises 5 and 15 in this section, Exercise 24 in Section 7B, and Exercise 15 in Section 8A.



As we know, every operator on a finite-dimensional complex vector space has an eigenvalue. However, not every operator on a finite-dimensional complex vector space has enough eigenvectors to be diagonalizable, as shown by the next example.

**5.57** example: *an operator that is not diagonalizable*

Define an operator  $T \in \mathcal{L}(\mathbf{F}^3)$  by  $T(a, b, c) = (b, c, 0)$ . The matrix of  $T$  with respect to the standard basis of  $\mathbf{F}^3$  is

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix},$$

which is an upper-triangular matrix but is not a diagonal matrix.

As you should verify, 0 is the only eigenvalue of  $T$  and furthermore

$$E(0, T) = \{(a, 0, 0) \in \mathbf{F}^3 : a \in \mathbf{F}\}.$$

Hence conditions (b), (c), and (d) of 5.55 fail (of course, because these conditions are equivalent, it is sufficient to check that only one of them fails). Thus condition (a) of 5.55 also fails. Hence  $T$  is not diagonalizable, regardless of whether  $\mathbf{F} = \mathbf{R}$  or  $\mathbf{F} = \mathbf{C}$ .

The next result shows that if an operator has as many distinct eigenvalues as the dimension of its domain, then the operator is diagonalizable.

**5.58** *enough eigenvalues implies diagonalizability*

Suppose  $V$  is finite-dimensional and  $T \in \mathcal{L}(V)$  has  $\dim V$  distinct eigenvalues. Then  $T$  is diagonalizable.

**Proof** Suppose  $T$  has distinct eigenvalues  $\lambda_1, \dots, \lambda_{\dim V}$ . For each  $k$ , let  $v_k \in V$  be an eigenvector corresponding to the eigenvalue  $\lambda_k$ . Because eigenvectors corresponding to distinct eigenvalues are linearly independent (see 5.11),  $v_1, \dots, v_{\dim V}$  is linearly independent.

A linearly independent list of  $\dim V$  vectors in  $V$  is a basis of  $V$  (see 2.38); thus  $v_1, \dots, v_{\dim V}$  is a basis of  $V$ . With respect to this basis consisting of eigenvectors,  $T$  has a diagonal matrix. ■

In later chapters we will find additional conditions that imply that certain operators are diagonalizable. For example, see the real spectral theorem (7.29) and the complex spectral theorem (7.31).

The result above gives a sufficient condition for an operator to be diagonalizable. However, this condition is not necessary. For example, the operator  $T$  on  $\mathbf{F}^3$  defined by  $T(x, y, z) = (6x, 6y, 7z)$  has only two eigenvalues (6 and 7) and  $\dim \mathbf{F}^3 = 3$ , but  $T$  is diagonalizable (by the standard basis of  $\mathbf{F}^3$ ).

The next example illustrates the importance of diagonalization, which can be used to compute high powers of an operator, taking advantage of the equation  $T^k v = \lambda^k v$  if  $v$  is an eigenvector of  $T$  with eigenvalue  $\lambda$ .

*For a spectacular application of these techniques, see Exercise 21, which shows how to use diagonalization to find an exact formula for the  $n^{\text{th}}$  term of the Fibonacci sequence.*

5.59 example: using diagonalization to compute  $T^{100}$

Define  $T \in \mathcal{L}(\mathbf{F}^3)$  by  $T(x, y, z) = (2x + y, 5y + 3z, 8z)$ . With respect to the standard basis, the matrix of  $T$  is

$$\begin{pmatrix} 2 & 1 & 0 \\ 0 & 5 & 3 \\ 0 & 0 & 8 \end{pmatrix}.$$

The matrix above is an upper-triangular matrix but it is not a diagonal matrix. By 5.41, the eigenvalues of  $T$  are 2, 5, and 8. Because  $T$  is an operator on a vector space of dimension three and  $T$  has three distinct eigenvalues, 5.58 assures us that there exists a basis of  $\mathbf{F}^3$  with respect to which  $T$  has a diagonal matrix.

To find this basis, we only have to find an eigenvector for each eigenvalue. In other words, we have to find a nonzero solution to the equation

$$T(x, y, z) = \lambda(x, y, z)$$

for  $\lambda = 2$ , then for  $\lambda = 5$ , and then for  $\lambda = 8$ . Solving these simple equations shows that for  $\lambda = 2$  we have an eigenvector  $(1, 0, 0)$ , for  $\lambda = 5$  we have an eigenvector  $(1, 3, 0)$ , and for  $\lambda = 8$  we have an eigenvector  $(1, 6, 6)$ .

Thus  $(1, 0, 0)$ ,  $(1, 3, 0)$ ,  $(1, 6, 6)$  is a basis of  $\mathbf{F}^3$  consisting of eigenvectors of  $T$ , and with respect to this basis the matrix of  $T$  is the diagonal matrix

$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 8 \end{pmatrix}.$$

To compute  $T^{100}(0, 0, 1)$ , for example, write  $(0, 0, 1)$  as a linear combination of our basis of eigenvectors:

$$(0, 0, 1) = \frac{1}{6}(1, 0, 0) - \frac{1}{3}(1, 3, 0) + \frac{1}{6}(1, 6, 6).$$

Now apply  $T^{100}$  to both sides of the equation above, getting

$$\begin{aligned} T^{100}(0, 0, 1) &= \frac{1}{6}(T^{100}(1, 0, 0)) - \frac{1}{3}(T^{100}(1, 3, 0)) + \frac{1}{6}(T^{100}(1, 6, 6)) \\ &= \frac{1}{6}(2^{100}(1, 0, 0) - 2 \cdot 5^{100}(1, 3, 0) + 8^{100}(1, 6, 6)) \\ &= \frac{1}{6}(2^{100} - 2 \cdot 5^{100} + 8^{100}, 6 \cdot 8^{100} - 6 \cdot 5^{100}, 6 \cdot 8^{100}). \end{aligned}$$

We saw earlier that an operator  $T$  on a finite-dimensional vector space  $V$  has an upper-triangular matrix with respect to some basis of  $V$  if and only if the minimal polynomial of  $T$  equals  $(z - \lambda_1) \cdots (z - \lambda_m)$  for some  $\lambda_1, \dots, \lambda_m \in \mathbf{F}$  (see 5.44). As we previously noted (see 5.47), this condition is always satisfied if  $\mathbf{F} = \mathbf{C}$ .

Our next result 5.62 states that an operator  $T \in \mathcal{L}(V)$  has a diagonal matrix with respect to some basis of  $V$  if and only if the minimal polynomial of  $T$  equals  $(z - \lambda_1) \cdots (z - \lambda_m)$  for some **distinct**  $\lambda_1, \dots, \lambda_m \in \mathbf{F}$ . Before formally stating this result, we give two examples of using it.

5.60 example: *diagonalizable, but with no known exact eigenvalues*

Define  $T \in \mathcal{L}(\mathbf{C}^5)$  by

$$T(z_1, z_2, z_3, z_4, z_5) = (-3z_5, z_1 + 6z_5, z_2, z_3, z_4).$$

The matrix of  $T$  is shown in Example 5.26, where we showed that the minimal polynomial of  $T$  is  $3 - 6z + z^5$ .

As mentioned in Example 5.28, no exact expression is known for any of the zeros of this polynomial, but numeric techniques show that the zeros of this polynomial are approximately  $-1.67$ ,  $0.51$ ,  $1.40$ ,  $-0.12 + 1.59i$ ,  $-0.12 - 1.59i$ .

The software that produces these approximations is accurate to more than three digits. Thus these approximations are good enough to show that the five numbers above are distinct. The minimal polynomial of  $T$  equals the fifth degree monic polynomial with these zeros. Now 5.62 shows that  $T$  is diagonalizable.

5.61 example: *showing that an operator is not diagonalizable*

Define  $T \in \mathcal{L}(\mathbf{F}^3)$  by

$$T(z_1, z_2, z_3) = (6z_1 + 3z_2 + 4z_3, 6z_2 + 2z_3, 7z_3).$$

The matrix of  $T$  with respect to the standard basis of  $\mathbf{F}^3$  is

$$\begin{pmatrix} 6 & 3 & 4 \\ 0 & 6 & 2 \\ 0 & 0 & 7 \end{pmatrix}.$$

The matrix above is an upper-triangular matrix but is not a diagonal matrix. Might  $T$  have a diagonal matrix with respect to some other basis of  $\mathbf{F}^3$ ?

To answer this question, we will find the minimal polynomial of  $T$ . First note that the eigenvalues of  $T$  are the diagonal entries of the matrix above (by 5.41). Thus the zeros of the minimal polynomial of  $T$  are 6, 7 [by 5.27(a)]. The diagonal of the matrix above tells us that  $(T - 6I)^2(T - 7I) = 0$  (by 5.40). The minimal polynomial of  $T$  has degree at most 3 (by 5.22). Putting all this together, we see that the minimal polynomial of  $T$  is either  $(z - 6)(z - 7)$  or  $(z - 6)^2(z - 7)$ .

A simple computation shows that  $(T - 6I)(T - 7I) \neq 0$ . Thus the minimal polynomial of  $T$  is  $(z - 6)^2(z - 7)$ .

Now 5.62 shows that  $T$  is not diagonalizable.

5.62 *necessary and sufficient condition for diagonalizability*

Suppose  $V$  is finite-dimensional and  $T \in \mathcal{L}(V)$ . Then  $T$  is diagonalizable if and only if the minimal polynomial of  $T$  equals  $(z - \lambda_1) \cdots (z - \lambda_m)$  for some list of distinct numbers  $\lambda_1, \dots, \lambda_m \in \mathbf{F}$ .

**Proof** First suppose  $T$  is diagonalizable. Thus there is a basis  $v_1, \dots, v_n$  of  $V$  consisting of eigenvectors of  $T$ . Let  $\lambda_1, \dots, \lambda_m$  be the distinct eigenvalues of  $T$ . Then for each  $v_j$ , there exists  $\lambda_k$  with  $(T - \lambda_k I)v_j = 0$ . Thus

$$(T - \lambda_1 I) \cdots (T - \lambda_m I)v_j = 0,$$

which implies that the minimal polynomial of  $T$  equals  $(z - \lambda_1) \cdots (z - \lambda_m)$ .

To prove the implication in the other direction, now suppose the minimal polynomial of  $T$  equals  $(z - \lambda_1) \cdots (z - \lambda_m)$  for some list of distinct numbers  $\lambda_1, \dots, \lambda_m \in \mathbf{F}$ . Thus

$$5.63 \quad (T - \lambda_1 I) \cdots (T - \lambda_m I) = 0.$$

We will prove that  $T$  is diagonalizable by induction on  $m$ . To get started, suppose  $m = 1$ . Then  $T - \lambda_1 I = 0$ , which means that  $T$  is a scalar multiple of the identity operator, which implies that  $T$  is diagonalizable.

Now suppose that  $m > 1$  and the desired result holds for all smaller values of  $m$ . The subspace  $\text{range}(T - \lambda_m I)$  is invariant under  $T$  [this is a special case of 5.18 with  $p(z) = z - \lambda_m$ ]. Thus  $T$  restricted to  $\text{range}(T - \lambda_m I)$  is an operator on  $\text{range}(T - \lambda_m I)$ .

If  $u \in \text{range}(T - \lambda_m I)$ , then  $u = (T - \lambda_m I)v$  for some  $v \in V$ , and 5.63 implies

$$5.64 \quad (T - \lambda_1 I) \cdots (T - \lambda_{m-1} I)u = (T - \lambda_1 I) \cdots (T - \lambda_m I)v = 0.$$

Hence  $(z - \lambda_1) \cdots (z - \lambda_{m-1})$  is a polynomial multiple of the minimal polynomial of  $T$  restricted to  $\text{range}(T - \lambda_m I)$  [by 5.29]. Thus by our induction hypothesis, there is a basis of  $\text{range}(T - \lambda_m I)$  consisting of eigenvectors of  $T$ .

Suppose that  $u \in \text{range}(T - \lambda_m I) \cap \text{null}(T - \lambda_m I)$ . Then  $Tu = \lambda_m u$ . Now 5.64 implies that

$$\begin{aligned} 0 &= (T - \lambda_1 I) \cdots (T - \lambda_{m-1} I)u \\ &= (\lambda_m - \lambda_1) \cdots (\lambda_m - \lambda_{m-1})u. \end{aligned}$$

Because  $\lambda_1, \dots, \lambda_m$  are distinct, the equation above implies that  $u = 0$ . Hence  $\text{range}(T - \lambda_m I) \cap \text{null}(T - \lambda_m I) = \{0\}$ .

Thus  $\text{range}(T - \lambda_m I) + \text{null}(T - \lambda_m I)$  is a direct sum (by 1.46) whose dimension is  $\dim V$  (by 3.94 and 3.21). Hence  $\text{range}(T - \lambda_m I) \oplus \text{null}(T - \lambda_m I) = V$ . Every vector in  $\text{null}(T - \lambda_m I)$  is an eigenvector of  $T$  with eigenvalue  $\lambda_m$ . Earlier in this proof we saw that there is a basis of  $\text{range}(T - \lambda_m I)$  consisting of eigenvectors of  $T$ . Adjoining to that basis a basis of  $\text{null}(T - \lambda_m I)$  gives a basis of  $V$  consisting of eigenvectors of  $T$ . The matrix of  $T$  with respect to this basis is a diagonal matrix, as desired. ■

No formula exists for the zeros of polynomials of degree 5 or greater. However, the previous result can be used to determine whether an operator on a complex vector space is diagonalizable without even finding approximations of the zeros of the minimal polynomial—see Exercise 15.

The next result will be a key tool when we prove a result about the simultaneous diagonalization of two operators; see 5.76. Note how the use of the characterization of diagonalizable operators in terms of the minimal polynomial (see 5.62) leads to a short proof of the next result.

#### 5.65 restriction of diagonalizable operator to invariant subspace

Suppose  $T \in \mathcal{L}(V)$  is diagonalizable and  $U$  is a subspace of  $V$  that is invariant under  $T$ . Then  $T|_U$  is a diagonalizable operator on  $U$ .

**Proof** Because the operator  $T$  is diagonalizable, the minimal polynomial of  $T$  equals  $(z - \lambda_1) \cdots (z - \lambda_m)$  for some list of distinct numbers  $\lambda_1, \dots, \lambda_m \in \mathbf{F}$  (by 5.62). The minimal polynomial of  $T$  is a polynomial multiple of the minimal polynomial of  $T|_U$  (by 5.31). Hence the minimal polynomial of  $T|_U$  has the form required by 5.62, which shows that  $T|_U$  is diagonalizable. ■

## Gershgorin Disk Theorem

#### 5.66 definition: Gershgorin disks

Suppose  $T \in \mathcal{L}(V)$  and  $v_1, \dots, v_n$  is a basis of  $V$ . Let  $A$  denote the matrix of  $T$  with respect to this basis. A *Gershgorin disk* of  $T$  with respect to the basis  $v_1, \dots, v_n$  is a set of the form

$$\left\{ z \in \mathbf{F} : |z - A_{j,j}| \leq \sum_{\substack{k=1 \\ k \neq j}}^n |A_{j,k}| \right\},$$

where  $j \in \{1, \dots, n\}$ .

Because there are  $n$  choices for  $j$  in the definition above,  $T$  has  $n$  Gershgorin disks. If  $\mathbf{F} = \mathbf{C}$ , then for each  $j \in \{1, \dots, n\}$ , the corresponding Gershgorin disk is a closed disk in  $\mathbf{C}$  centered at  $A_{j,j}$ , which is the  $j^{\text{th}}$  entry on the diagonal of  $A$ . The radius of this closed disk is the sum of the absolute values of the entries in row  $j$  of  $A$ , excluding the diagonal entry. If  $\mathbf{F} = \mathbf{R}$ , then the Gershgorin disks are closed intervals in  $\mathbf{R}$ .

In the special case that the square matrix  $A$  above is a diagonal matrix, each Gershgorin disk consists of a single point that is a diagonal entry of  $A$  (and each eigenvalue of  $T$  is one of those points, as required by the next result). One consequence of our next result is that if the nondiagonal entries of  $A$  are small, then each eigenvalue of  $T$  is near a diagonal entry of  $A$ .

## 5.67 Gershgorin disk theorem

Suppose  $T \in \mathcal{L}(V)$  and  $v_1, \dots, v_n$  is a basis of  $V$ . Then each eigenvalue of  $T$  is contained in some Gershgorin disk of  $T$  with respect to the basis  $v_1, \dots, v_n$ .

**Proof** Suppose  $\lambda \in \mathbf{F}$  is an eigenvalue of  $T$ . Let  $w \in V$  be a corresponding eigenvector. There exist  $c_1, \dots, c_n \in \mathbf{F}$  such that

$$5.68 \quad w = c_1 v_1 + \dots + c_n v_n.$$

Let  $A$  denote the matrix of  $T$  with respect to the basis  $v_1, \dots, v_n$ . Applying  $T$  to both sides of the equation above gives

$$\begin{aligned} 5.69 \quad \lambda w &= \sum_{k=1}^n c_k T v_k \\ &= \sum_{k=1}^n c_k \sum_{j=1}^n A_{j,k} v_j \\ 5.70 \quad &= \sum_{j=1}^n \left( \sum_{k=1}^n A_{j,k} c_k \right) v_j. \end{aligned}$$

Let  $j \in \{1, \dots, n\}$  be such that

$$|c_j| = \max\{|c_1|, \dots, |c_n|\}.$$

Using 5.68, we see that the coefficient of  $v_j$  on the left side of 5.69 equals  $\lambda c_j$ , which must equal the coefficient of  $v_j$  on the right side of 5.70. In other words,

$$\lambda c_j = \sum_{k=1}^n A_{j,k} c_k.$$

Subtract  $A_{j,j} c_j$  from each side of the equation above and then divide both sides by  $c_j$  to get

$$\begin{aligned} |\lambda - A_{j,j}| &= \left| \sum_{\substack{k=1 \\ k \neq j}}^n A_{j,k} \frac{c_k}{c_j} \right| \\ &\leq \sum_{\substack{k=1 \\ k \neq j}}^n |A_{j,k}|. \end{aligned}$$

Thus  $\lambda$  is in the  $j^{\text{th}}$  Gershgorin disk with respect to the basis  $v_1, \dots, v_n$ . ■

Exercise 22 gives a nice application of the Gershgorin disk theorem.

Exercise 23 states that the radius of each Gershgorin disk could be changed to the sum of the absolute values of corresponding column entries (instead of row entries), excluding the diagonal entry, and the theorem above would still hold.

*The Gershgorin disk theorem is named for Semyon Aronovich Gershgorin, who published this result in 1931.*

Exercises 5D

---

- 1 Suppose  $V$  is a finite-dimensional complex vector space and  $T \in \mathcal{L}(V)$ .
  - (a) Prove that if  $T^4 = I$ , then  $T$  is diagonalizable.
  - (b) Prove that if  $T^4 = T$ , then  $T$  is diagonalizable.
  - (c) Give an example of an operator  $T \in \mathcal{L}(\mathbb{C}^2)$  such that  $T^4 = T^2$  and  $T$  is not diagonalizable.
- 2 Suppose  $T \in \mathcal{L}(V)$  has a diagonal matrix  $A$  with respect to some basis of  $V$ . Prove that if  $\lambda \in \mathbb{F}$ , then  $\lambda$  appears on the diagonal of  $A$  precisely  $\dim E(\lambda, T)$  times.
- 3 Suppose  $V$  is finite-dimensional and  $T \in \mathcal{L}(V)$ . Prove that if the operator  $T$  is diagonalizable, then  $V = \text{null } T \oplus \text{range } T$ .
- 4 Suppose  $V$  is finite-dimensional and  $T \in \mathcal{L}(V)$ . Prove that the following are equivalent.
  - (a)  $V = \text{null } T \oplus \text{range } T$ .
  - (b)  $V = \text{null } T + \text{range } T$ .
  - (c)  $\text{null } T \cap \text{range } T = \{0\}$ .
- 5 Suppose  $V$  is a finite-dimensional complex vector space and  $T \in \mathcal{L}(V)$ . Prove that  $T$  is diagonalizable if and only if

$$V = \text{null}(T - \lambda I) \oplus \text{range}(T - \lambda I)$$

for every  $\lambda \in \mathbb{C}$ .

- 6 Suppose  $T \in \mathcal{L}(\mathbb{F}^5)$  and  $\dim E(8, T) = 4$ . Prove that  $T - 2I$  or  $T - 6I$  is invertible.
- 7 Suppose  $T \in \mathcal{L}(V)$  is invertible. Prove that
 
$$E(\lambda, T) = E\left(\frac{1}{\lambda}, T^{-1}\right)$$
 for every  $\lambda \in \mathbb{F}$  with  $\lambda \neq 0$ .
- 8 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
 
$$\dim E(\lambda_1, T) + \dots + \dim E(\lambda_m, T) \leq \dim \text{range } T.$$
- 9 Suppose  $R, T \in \mathcal{L}(\mathbb{F}^3)$  each have 2, 6, 7 as eigenvalues. Prove that there exists an invertible operator  $S \in \mathcal{L}(\mathbb{F}^3)$  such that  $R = S^{-1}TS$ .
- 10 Find  $R, T \in \mathcal{L}(\mathbb{F}^4)$  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  $S \in \mathcal{L}(\mathbb{F}^4)$  such that  $R = S^{-1}TS$ .

- 11 Find  $T \in \mathcal{L}(\mathbb{C}^3)$  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  $\mathbb{C}^3$ .
- 12 Suppose  $T \in \mathcal{L}(\mathbb{C}^3)$  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  $\mathbb{C}^3$ . Prove that there exists  $(z_1, z_2, z_3) \in \mathbb{C}^3$  such that

$$T(z_1, z_2, z_3) = (6 + 8z_1, 7 + 8z_2, 13 + 8z_3).$$

- 13 Suppose  $A$  is a diagonal matrix with distinct entries on the diagonal and  $B$  is a matrix of the same size as  $A$ . Show that  $AB = BA$  if and only if  $B$  is a diagonal matrix.
- 14 (a) Give an example of a finite-dimensional complex vector space and an operator  $T$  on that vector space such that  $T^2$  is diagonalizable but  $T$  is not diagonalizable.
- (b) Suppose  $F = \mathbb{C}$ ,  $k$  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.
- 15 Suppose  $V$  is a finite-dimensional complex vector space,  $T \in \mathcal{L}(V)$ , and  $p$  is the minimal polynomial of  $T$ . Prove that the following are equivalent.
- (a)  $T$  is diagonalizable.
- (b) There does not exist  $\lambda \in \mathbb{C}$  such that  $p$  is a polynomial multiple of  $(z - \lambda)^2$ .
- (c)  $p$  and its derivative  $p'$  have no zeros in common.
- (d) The greatest common divisor of  $p$  and  $p'$  is the constant polynomial 1.

*The **greatest common divisor** of  $p$  and  $p'$  is the monic polynomial  $q$  of largest degree such that  $p$  and  $p'$  are both polynomial multiples of  $q$ . The Euclidean algorithm for polynomials (look it up) can quickly determine the greatest common divisor of two polynomials, without requiring any information about the zeros of the polynomials. Thus the equivalence of (a) and (d) above shows that we can determine whether  $T$  is diagonalizable without knowing anything about the zeros of  $p$ .*

- 16 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  $U = U_1 \oplus \dots \oplus U_m$ .
- 17 Suppose  $V$  is finite-dimensional. Prove that  $\mathcal{L}(V)$  has a basis consisting of diagonalizable operators.
- 18 Suppose that  $T \in \mathcal{L}(V)$  is diagonalizable and  $U$  is a subspace of  $V$  that is invariant under  $T$ . Prove that the quotient operator  $T/U$  is a diagonalizable operator on  $V/U$ .

*The quotient operator  $T/U$  was defined in Exercise 38 in Section 5A.*



- 19 Prove or give a counterexample: If  $T \in \mathcal{L}(V)$  and there exists a subspace  $U$  of  $V$  that is invariant under  $T$  such that  $T|_U$  and  $T/U$  are both diagonalizable, then  $T$  is diagonalizable.

*See Exercise 13 in Section 5C for an analogous statement about upper-triangular matrices.*

- 20 Suppose  $V$  is finite-dimensional and  $T \in \mathcal{L}(V)$ . Prove that  $T$  is diagonalizable if and only if the dual operator  $T'$  is diagonalizable.
- 21 The Fibonacci sequence  $F_0, F_1, F_2, \dots$  is defined by

$$F_0 = 0, F_1 = 1, \text{ and } F_n = F_{n-2} + F_{n-1} \text{ for } n \geq 2.$$

Define  $T \in \mathcal{L}(\mathbf{R}^2)$  by  $T(x, y) = (y, x + y)$ .

- (a) Show that  $T^n(0, 1) = (F_n, F_{n+1})$  for each nonnegative integer  $n$ .  
 (b) Find the eigenvalues of  $T$ .  
 (c) Find a basis of  $\mathbf{R}^2$  consisting of eigenvectors of  $T$ .  
 (d) Use the solution to (c) to compute  $T^n(0, 1)$ . Conclude that

$$F_n = \frac{1}{\sqrt{5}} \left[ \left( \frac{1 + \sqrt{5}}{2} \right)^n - \left( \frac{1 - \sqrt{5}}{2} \right)^n \right]$$

for each nonnegative integer  $n$ .

- (e) Use (d) to conclude that if  $n$  is a nonnegative integer, then the Fibonacci number  $F_n$  is the integer that is closest to

$$\frac{1}{\sqrt{5}} \left( \frac{1 + \sqrt{5}}{2} \right)^n.$$

*Each  $F_n$  is a nonnegative integer, even though the right side of the formula in (d) does not look like an integer. The number*

$$\frac{1 + \sqrt{5}}{2}$$

*is called the **golden ratio**.*

- 22 Suppose  $T \in \mathcal{L}(V)$  and  $A$  is an  $n$ -by- $n$  matrix that is the matrix of  $T$  with respect to some basis of  $V$ . Prove that if

$$|A_{j,j}| > \sum_{\substack{k=1 \\ k \neq j}}^n |A_{j,k}|$$

for each  $j \in \{1, \dots, n\}$ , then  $T$  is invertible.

*This exercise states that if the diagonal entries of the matrix of  $T$  are large compared to the nondiagonal entries, then  $T$  is invertible.*

- 23 Suppose the definition of the Gershgorin disks is changed so that the radius of the  $k^{\text{th}}$  disk is the sum of the absolute values of the entries in column (instead of row)  $k$  of  $A$ , excluding the diagonal entry. Show that the Gershgorin disk theorem (5.67) still holds with this changed definition.

## 5E Commuting Operators

### 5.71 definition: *commute*

- Two operators  $S$  and  $T$  on the same vector space *commute* if  $ST = TS$ .
- Two square matrices  $A$  and  $B$  of the same size *commute* if  $AB = BA$ .

For example, if  $T$  is an operator and  $p, q \in \mathcal{P}(\mathbf{F})$ , then  $p(T)$  and  $q(T)$  commute [see 5.17(b)].

As another example, if  $I$  is the identity operator on  $V$ , then  $I$  commutes with every operator on  $V$ .

### 5.72 example: *partial differentiation operators commute*

Suppose  $m$  is a nonnegative integer. Let  $\mathcal{P}_m(\mathbf{R}^2)$  denote the real vector space of polynomials (with real coefficients) in two real variables and of degree at most  $m$ , with the usual operations of addition and scalar multiplication of real-valued functions. Thus the elements of  $\mathcal{P}_m(\mathbf{R}^2)$  are functions  $p$  on  $\mathbf{R}^2$  of the form

$$5.73 \quad p = \sum_{j+k \leq m} a_{j,k} x^j y^k,$$

where the indices  $j$  and  $k$  take on all nonnegative integer values such that  $j+k \leq m$ , each  $a_{j,k}$  is in  $\mathbf{R}$ , and  $x^j y^k$  denotes the function on  $\mathbf{R}^2$  defined by  $(x, y) \mapsto x^j y^k$ .

Define operators  $D_x, D_y \in \mathcal{L}(\mathcal{P}_m(\mathbf{R}^2))$  by

$$D_x p = \frac{\partial p}{\partial x} = \sum_{j+k \leq m} j a_{j,k} x^{j-1} y^k \quad \text{and} \quad D_y p = \frac{\partial p}{\partial y} = \sum_{j+k \leq m} k a_{j,k} x^j y^{k-1},$$

where  $p$  is as in 5.73. The operators  $D_x$  and  $D_y$  are called partial differentiation operators because each of these operators differentiates with respect to one of the variables while pretending that the other variable is a constant.

The operators  $D_x$  and  $D_y$  commute because if  $p$  is as in 5.73, then

$$(D_x D_y) p = \sum_{j+k \leq m} j k a_{j,k} x^{j-1} y^{k-1} = (D_y D_x) p.$$

The equation  $D_x D_y = D_y D_x$  on  $\mathcal{P}_m(\mathbf{R}^2)$  illustrates a more general result that the order of partial differentiation does not matter for nice functions.

Commuting matrices are unusual. For example, there are 214,358,881 pairs of 2-by-2 matrices all of whose entries are integers in the interval  $[-5, 5]$ . Only about 0.3% of these pairs of matrices commute.

All 214,358,881 (which equals  $11^8$ ) pairs of the 2-by-2 matrices under consideration were checked by a computer to discover that only 674,609 of these pairs of matrices commute.

The next result shows that two operators commute if and only if their matrices (with respect to the same basis) commute.

**5.74** *commuting operators correspond to commuting matrices*

Suppose  $S, T \in \mathcal{L}(V)$  and  $v_1, \dots, v_n$  is a basis of  $V$ . Then  $S$  and  $T$  commute if and only if  $\mathcal{M}(S, (v_1, \dots, v_n))$  and  $\mathcal{M}(T, (v_1, \dots, v_n))$  commute.

**Proof** We have

$$\begin{aligned} S \text{ and } T \text{ commute} &\iff ST = TS \\ &\iff \mathcal{M}(ST) = \mathcal{M}(TS) \\ &\iff \mathcal{M}(S)\mathcal{M}(T) = \mathcal{M}(T)\mathcal{M}(S) \\ &\iff \mathcal{M}(S) \text{ and } \mathcal{M}(T) \text{ commute,} \end{aligned}$$

as desired. ■

The next result shows that if two operators commute, then every eigenspace for one operator is invariant under the other operator. This result, which we will use several times, is one of the main reasons why a pair of commuting operators behaves better than a pair of operators that does not commute.

**5.75** *eigenspace is invariant under commuting operator*

Suppose  $S, T \in \mathcal{L}(V)$  commute and  $\lambda \in \mathbf{F}$ . Then  $E(\lambda, S)$  is invariant under  $T$ .

**Proof** Suppose  $v \in E(\lambda, S)$ . Then

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

The equation above shows that  $Tv \in E(\lambda, S)$ . Thus  $E(\lambda, S)$  is invariant under  $T$ . ■

Suppose we have two operators, each of which is diagonalizable. If we want to do computations involving both operators (for example, involving their sum), then we want the two operators to be diagonalizable by the same basis, which according to the next result is possible when the two operators commute.

**5.76** *simultaneous diagonalizability  $\iff$  commutativity*

Two diagonalizable operators on the same vector space have diagonal matrices with respect to the same basis if and only if the two operators commute.

**Proof** First suppose  $S, T \in \mathcal{L}(V)$  have diagonal matrices with respect to the same basis. The product of two diagonal matrices of the same size is the diagonal matrix obtained by multiplying the corresponding elements of the two diagonals. Thus any two diagonal matrices of the same size commute. Thus  $S$  and  $T$  commute, by 5.74.

To prove the implication in the other direction, now suppose that  $S, T \in \mathcal{L}(V)$  are diagonalizable operators that commute. Let  $\lambda_1, \dots, \lambda_m$  denote the distinct eigenvalues of  $S$ . Because  $S$  is diagonalizable, 5.55(c) shows that

$$5.77 \quad V = E(\lambda_1, S) \oplus \cdots \oplus E(\lambda_m, S).$$

For each  $k = 1, \dots, m$ , the subspace  $E(\lambda_k, S)$  is invariant under  $T$  (by 5.75). Because  $T$  is diagonalizable, 5.65 implies that  $T|_{E(\lambda_k, S)}$  is diagonalizable for each  $k$ . Hence for each  $k = 1, \dots, m$ , there is a basis of  $E(\lambda_k, S)$  consisting of eigenvectors of  $T$ . Putting these bases together gives a basis of  $V$  (because of 5.77), with each vector in this basis being an eigenvector of both  $S$  and  $T$ . Thus  $S$  and  $T$  both have diagonal matrices with respect to this basis, as desired. ■

See Exercise 2 for an extension of the result above to more than two operators.

Suppose  $V$  is a finite-dimensional nonzero complex vector space. Then every operator on  $V$  has an eigenvector (see 5.19). The next result shows that if two operators on  $V$  commute, then there is a vector in  $V$  that is an eigenvector for both operators (but the two commuting operators might not have a common eigenvalue). For an extension of the next result to more than two operators, see Exercise 9(a).

#### 5.78 common eigenvector for commuting operators

Every pair of commuting operators on a finite-dimensional nonzero complex vector space has a common eigenvector.

**Proof** Suppose  $V$  is a finite-dimensional nonzero complex vector space and  $S, T \in \mathcal{L}(V)$  commute. Let  $\lambda$  be an eigenvalue of  $S$  (5.19 tells us that  $S$  does indeed have an eigenvalue). Thus  $E(\lambda, S) \neq \{0\}$ . Also,  $E(\lambda, S)$  is invariant under  $T$  (by 5.75).

Thus  $T|_{E(\lambda, S)}$  has an eigenvector (again using 5.19), which is an eigenvector for both  $S$  and  $T$ , completing the proof. ■

#### 5.79 example: common eigenvector for partial differentiation operators

Let  $\mathcal{P}_m(\mathbf{R}^2)$  be as in Example 5.72 and let  $D_x, D_y \in \mathcal{L}(\mathcal{P}_m(\mathbf{R}^2))$  be the commuting partial differentiation operators in that example. As you can verify, 0 is the only eigenvalue of each of these operators. Also

$$E(0, D_x) = \left\{ \sum_{k=0}^m a_k y^k : a_0, \dots, a_m \in \mathbf{R} \right\},$$

$$E(0, D_y) = \left\{ \sum_{j=0}^m c_j x^j : c_0, \dots, c_m \in \mathbf{R} \right\}.$$

The intersection of these two eigenspaces is the set of common eigenvectors of the two operators. Because  $E(0, D_x) \cap E(0, D_y)$  is the set of constant functions, we see that  $D_x$  and  $D_y$  indeed have a common eigenvector, as promised by 5.78.

The next result extends 5.47 (the existence of a basis that gives an upper-triangular matrix) to two commuting operators.

**5.80** *commuting operators are simultaneously upper triangularizable*

Suppose  $V$  is a finite-dimensional complex vector space and  $S, T$  are commuting operators on  $V$ . Then there is a basis of  $V$  with respect to which both  $S$  and  $T$  have upper-triangular matrices.

**Proof** Let  $n = \dim V$ . We will use induction on  $n$ . The desired result holds if  $n = 1$  because all 1-by-1 matrices are upper triangular. Now suppose  $n > 1$  and the desired result holds for all complex vector spaces whose dimension is  $n - 1$ .

Let  $v_1$  be any common eigenvector of  $S$  and  $T$  (using 5.78). Hence  $sv_1 \in \text{span}(v_1)$  and  $tv_1 \in \text{span}(v_1)$ . Let  $W$  be a subspace of  $V$  such that

$$V = \text{span}(v_1) \oplus W;$$

see 2.33 for the existence of  $W$ . Define a linear map  $P: V \rightarrow W$  by

$$P(av_1 + w) = w$$

for each  $a \in \mathbb{C}$  and each  $w \in W$ . Define  $\hat{S}, \hat{T} \in \mathcal{L}(W)$  by

$$\hat{S}w = P(Sw) \quad \text{and} \quad \hat{T}w = P(Tw)$$

for each  $w \in W$ . To apply our induction hypothesis to  $\hat{S}$  and  $\hat{T}$ , we must first show that these two operators on  $W$  commute. To do this, suppose  $w \in W$ . Then there exists  $a \in \mathbb{C}$  such that

$$(\hat{S}\hat{T})w = \hat{S}(P(Tw)) = \hat{S}(Tw - av_1) = P(S(Tw - av_1)) = P((ST)w),$$

where the last equality holds because  $v_1$  is an eigenvector of  $S$  and  $Pv_1 = 0$ . Similarly,

$$(\hat{T}\hat{S})w = P((TS)w).$$

Because the operators  $S$  and  $T$  commute, the last two displayed equations show that  $(\hat{S}\hat{T})w = (\hat{T}\hat{S})w$ . Hence  $\hat{S}$  and  $\hat{T}$  commute.

Thus we can use our induction hypothesis to state that there exists a basis  $v_2, \dots, v_n$  of  $W$  such that  $\hat{S}$  and  $\hat{T}$  both have upper-triangular matrices with respect to this basis. The list  $v_1, \dots, v_n$  is a basis of  $V$ .

If  $k \in \{2, \dots, n\}$ , then there exist  $a_k, b_k \in \mathbb{C}$  such that

$$Sv_k = a_kv_1 + \hat{S}v_k \quad \text{and} \quad Tv_k = b_kv_1 + \hat{T}v_k.$$

Because  $\hat{S}$  and  $\hat{T}$  have upper-triangular matrices with respect to  $v_2, \dots, v_n$ , we know that  $\hat{S}v_k \in \text{span}(v_2, \dots, v_k)$  and  $\hat{T}v_k \in \text{span}(v_2, \dots, v_k)$ . Hence the equations above imply that

$$Sv_k \in \text{span}(v_1, \dots, v_k) \quad \text{and} \quad Tv_k \in \text{span}(v_1, \dots, v_k).$$

Thus  $S$  and  $T$  have upper-triangular matrices with respect to  $v_1, \dots, v_n$ , as desired. ■

Exercise 9(b) extends the result above to more than two operators.

In general, it is not possible to determine the eigenvalues of the sum or product of two operators from the eigenvalues of the two operators. However, the next result shows that something nice happens when the two operators commute.

### 5.81 eigenvalues of sum and product of commuting operators

Suppose  $V$  is a finite-dimensional complex vector space and  $S, T$  are commuting operators on  $V$ . Then

- every eigenvalue of  $S + T$  is an eigenvalue of  $S$  plus an eigenvalue of  $T$ ,
- every eigenvalue of  $ST$  is an eigenvalue of  $S$  times an eigenvalue of  $T$ .

**Proof** There is a basis of  $V$  with respect to which both  $S$  and  $T$  have upper-triangular matrices (by 5.80). With respect to that basis,

$$\mathcal{M}(S + T) = \mathcal{M}(S) + \mathcal{M}(T) \quad \text{and} \quad \mathcal{M}(ST) = \mathcal{M}(S)\mathcal{M}(T),$$

as stated in 3.35 and 3.43.

The definition of matrix addition shows that each entry on the diagonal of  $\mathcal{M}(S + T)$  equals the sum of the corresponding entries on the diagonals of  $\mathcal{M}(S)$  and  $\mathcal{M}(T)$ . Similarly, because  $\mathcal{M}(S)$  and  $\mathcal{M}(T)$  are upper-triangular matrices, the definition of matrix multiplication shows that each entry on the diagonal of  $\mathcal{M}(ST)$  equals the product of the corresponding entries on the diagonals of  $\mathcal{M}(S)$  and  $\mathcal{M}(T)$ . Furthermore,  $\mathcal{M}(S + T)$  and  $\mathcal{M}(ST)$  are upper-triangular matrices (see Exercise 2 in Section 5B).

Every entry on the diagonal of  $\mathcal{M}(S)$  is an eigenvalue of  $S$ , and every entry on the diagonal of  $\mathcal{M}(T)$  is an eigenvalue of  $T$  (by 5.41). Every eigenvalue of  $S + T$  is on the diagonal of  $\mathcal{M}(S + T)$ , and every eigenvalue of  $ST$  is on the diagonal of  $\mathcal{M}(ST)$  (these assertions follow from 5.41). Putting all this together, we conclude that every eigenvalue of  $S + T$  is an eigenvalue of  $S$  plus an eigenvalue of  $T$ , and every eigenvalue of  $ST$  is an eigenvalue of  $S$  times an eigenvalue of  $T$ . ■

## Exercises 5E

- 1 Give an example of two commuting operators  $S, T$  on  $\mathbf{F}^4$  such that there is a subspace of  $\mathbf{F}^4$  that is invariant under  $S$  but not under  $T$  and there is a subspace of  $\mathbf{F}^4$  that is invariant under  $T$  but not under  $S$ .
- 2 Suppose  $\mathcal{E}$  is a subset of  $\mathcal{L}(V)$  and every element of  $\mathcal{E}$  is diagonalizable. Prove that there exists a basis of  $V$  with respect to which every element of  $\mathcal{E}$  has a diagonal matrix if and only if every pair of elements of  $\mathcal{E}$  commutes.

*This exercise extends 5.76, which considers the case in which  $\mathcal{E}$  contains only two elements. For this exercise,  $\mathcal{E}$  may contain any number of elements, and  $\mathcal{E}$  may even be an infinite set.*

3 Suppose  $S, T \in \mathcal{L}(V)$  are such that  $ST = TS$ . Suppose  $p \in \mathcal{P}(\mathbb{F})$ .

- (a) Prove that  $\text{null } p(S)$  is invariant under  $T$ .
- (b) Prove that  $\text{range } p(S)$  is invariant under  $T$ .

See 5.18 for the special case  $S = T$ .

4 Prove or give a counterexample: If  $A$  is a diagonal matrix and  $B$  is an upper-triangular matrix of the same size as  $A$ , then  $A$  and  $B$  commute.

5 Prove that a pair of operators on a finite-dimensional vector space commute if and only if their dual operators commute.

See 3.118 for the definition of the dual of an operator.

6 Suppose  $V$  is a finite-dimensional complex vector space and  $S, T \in \mathcal{L}(V)$  commute. Prove that there exist  $\alpha, \lambda \in \mathbb{C}$  such that

$$\text{range}(S - \alpha I) + \text{range}(T - \lambda I) \neq V.$$

7 Suppose  $V$  is a complex vector space,  $S \in \mathcal{L}(V)$  is diagonalizable, and  $T \in \mathcal{L}(V)$  commutes with  $S$ . Prove that there is a basis of  $V$  such that  $S$  has a diagonal matrix with respect to this basis and  $T$  has an upper-triangular matrix with respect to this basis.

8 Suppose  $m = 3$  in Example 5.72 and  $D_x, D_y$  are the commuting partial differentiation operators on  $\mathcal{P}_3(\mathbb{R}^2)$  from that example. Find a basis of  $\mathcal{P}_3(\mathbb{R}^2)$  with respect to which  $D_x$  and  $D_y$  each have an upper-triangular matrix.

9 Suppose  $V$  is a finite-dimensional nonzero complex vector space. Suppose that  $\mathcal{E} \subseteq \mathcal{L}(V)$  is such that  $S$  and  $T$  commute for all  $S, T \in \mathcal{E}$ .

- (a) Prove that there is a vector in  $V$  that is an eigenvector for every element of  $\mathcal{E}$ .
- (b) Prove that there is a basis of  $V$  with respect to which every element of  $\mathcal{E}$  has an upper-triangular matrix.

*This exercise extends 5.78 and 5.80, which consider the case in which  $\mathcal{E}$  contains only two elements. For this exercise,  $\mathcal{E}$  may contain any number of elements, and  $\mathcal{E}$  may even be an infinite set.*

10 Give an example of two commuting operators  $S, T$  on a finite-dimensional real vector space such that  $S + T$  has a eigenvalue that does not equal an eigenvalue of  $S$  plus an eigenvalue of  $T$  and  $ST$  has a eigenvalue that does not equal an eigenvalue of  $S$  times an eigenvalue of  $T$ .

*This exercise shows that 5.81 does not hold on real vector spaces.*