

# *Invariant Subspaces of Matrices with Applications*

I. GOHBERG

*School of Mathematical Sciences  
Tel Aviv University  
Tel Aviv, Israel*

P. LANCASTER

*Department of Mathematics and Statistics  
The University of Calgary  
Calgary, Alberta, Canada*

L. RODMAN

*School of Mathematical Sciences  
Tel Aviv University  
Tel Aviv, Israel*

*and*

*Department of Mathematics  
Arizona State University  
Tempe, Arizona*

A Wiley-Interscience Publication

JOHN WILEY & SONS

New York • Chichester • ~~Brisbane~~ • Toronto • Singapore

***Invariant Subspaces  
of Matrices  
with Applications***

CANADIAN MATHEMATICAL SOCIETY SERIES OF  
MONOGRAPHS AND ADVANCED TEXTS

*Monographies et Études de la Société Mathématique  
du Canada*

EDITORIAL BOARD

Frederick V. Atkinson, Bernhard Banaschewski, Colin W. Clark,  
Erwin O. Kreyszig (Chairman) and John B. Walsh



*Frank H. Clarke Optimization and Nonsmooth Analysis*

*Erwin Klein and Anthony C. Thompson Theory of Correspondences:  
Including Applications to Mathematical Economics*

*I. Gohberg, P. Lancaster, and L. Rodman Invariant Subspaces of  
Matrices with Applications*

*Jonathan M. Borwein and Peter Borwein Pi and the AGM—A Study  
in Analytic Number Theory and Computational Complexity*

# *Invariant Subspaces of Matrices with Applications*

I. GOHBERG

*School of Mathematical Sciences  
Tel Aviv University  
Tel Aviv, Israel*

P. LANCASTER

*Department of Mathematics and Statistics  
The University of Calgary  
Calgary, Alberta, Canada*

L. RODMAN

*School of Mathematical Sciences  
Tel Aviv University  
Tel Aviv, Israel  
and  
Department of Mathematics  
Arizona State University  
Tempe, Arizona*

A Wiley-Interscience Publication

JOHN WILEY & SONS

New York • Chichester • Brisbane • Toronto • Singapore

*To our wives  
Bella, Edna, and Ella*

Copyright © 1986 by John Wiley & Sons, Inc.

All rights reserved. Published simultaneously in Canada.

Reproduction or translation of any part of this work beyond that permitted by Section 107 or 108 of the 1976 United States Copyright Act without the permission of the copyright owner is unlawful. Requests for permission or further information should be addressed to the Permissions Department, John Wiley & Sons, Inc.

***Library of Congress Cataloging in Publication Data:***

Gohberg, I. (Israel), 1928-

Invariant subspaces of matrices with applications.

(Canadian Mathematical Society series of monographs and advanced texts)

Bibliography; p.

1. Invariant subspaces. 2. Matrices. I. Lancaster, Peter, 1929- . II. Rodman, L. III. Title.

IV. Series.

QA322.G649 1986 515.7'246 86-11091

ISBN 0-471-84260-5

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

## *Preface*

---

This is a book in advanced linear algebra in which invariant subspaces of matrices are the central notion and the main tool. To the authors' knowledge it is the first book written with such a theme. It contains a reasonably comprehensive treatment of geometrical, algebraic, topological and analytical properties of invariant subspaces. As well, an important part of the work consists of applications to matrix polynomials, rational matrix functions, linear systems, and matrix quadratic equations.

Parts of the book are written like a textbook and are easily accessible for undergraduate students. Gradually, the exposition changes to approach the style, and admit the content, of a monograph. Here, recent achievements and some unsolved problems are presented. A large portion of the content of the book has not appeared before in books. The fundamental character of the mathematics, its accessibility, and its importance in applications should make this a widely useful work for experts and students in mathematics, science, and engineering.

This is the third book written jointly by the authors. The first book is *Matrix Polynomials*, published by Academic Press in 1982, and the second is *Matrices and Indefinite Scalar Products*, published by Birkhäuser Verlag in 1983. These three books are connected and, to some extent, one led to another. Material that could not be included in one of the books became the starting point for the next. Moreover, invariant subspaces play an important role in the first two books and indicated to us the need for a systematic treatment of this subject.

The authors are pleased to acknowledge continuing support, throughout the development of this work, from the Natural Sciences and Engineering Research Council of Canada and from the Nathan and Lily Silver Chair in Mathematical Analysis and Operator Theory of Tel Aviv University. Constant support has also been provided by the staffs of the School of Mathematical Sciences, Tel Aviv University, and the Department of Mathematics and Statistics of the University of Calgary. We are especially grateful to Jacqueline Gorsky in Israel and Pat Dalgetty in Canada for skillful and speedy work on our typescripts. In addition, the second named author is pleased to acknowledge a Killam Resident Fellowship awarded for

the Winter Term of 1985 to assist in the completion of the book, and the third named author gratefully acknowledges support from the Basic Research Fund administered by the Israel Academy for Science and Humanities, and also from the U.S. National Science Foundation.

I. GOHBERG  
P. LANCASTER  
L. RODMAN

Calgary, Alberta, Canada  
July 1986

## Contents

<i>Introduction</i>	1
 <i>Part One Fundamental Properties of Invariant Subspaces and Applications</i>	 3
<i>Chapter One Invariant Subspaces: Definition, Examples, and First Properties</i>	5
1.1 Definition and Examples	5
1.2 Eigenvalues and Eigenvectors	10
1.3 Jordan Chains	12
1.4 Invariant Subspaces and Basic Operations on Linear Transformations	16
1.5 Invariant Subspaces and Projectors	20
1.6 Angular Transformations and Matrix Quadratic Equations	25
1.7 Transformations in Factor Spaces	28
1.8 The Lattice of Invariant Subspaces	31
1.9 Triangular Matrices and Complete Chains of Invariant Subspaces	37
1.10 Exercises	40
 <i>Chapter Two Jordan Form and Invariant Subspaces</i>	 45
2.1 Root Subspaces	45
2.2 The Jordan Form and Partial Multiplicities	52
2.3 Proof of the Jordan Form	58
2.4 Spectral Subspaces	60
2.5 Irreducible Invariant Subspaces and Unicellular Transformations	65
2.6 Generators of Invariant Subspaces	69
2.7 Maximal Invariant Subspace in a Given Subspace	72
2.8 Minimal Invariant Subspace over a Given Subspace	78
2.9 Marked Invariant Subspaces	83



2.10	Functions of Transformations	85
2.11	Partial Multiplicities and Invariant Subspaces of Functions of Transformations	92
2.12	Exercises	95
<b>Chapter Three Coinvariant and Semiinvariant Subspaces</b>		<b>105</b>
3.1	Coinvariant Subspaces	105
3.2	Reducing Subspaces	109
3.3	Semiinvariant Subspaces	112
3.4	Special Classes of Transformations	116
3.5	Exercises	119
<b>Chapter Four Jordan Form for Extensions and Completions</b>		<b>121</b>
4.1	Extensions from an Invariant Subspace	121
4.2	Completions from a Pair of Invariant and Coinvariant Subspaces	128
4.3	The Sigal Inequalities	133
4.4	Special Case of Completions	136
4.5	Exercises	142
<b>Chapter Five Applications to Matrix Polynomials</b>		<b>144</b>
5.1	Linearizations, Standard Triples, and Representations of Monic Matrix Polynomials	144
5.2	Multiplication of Monic Matrix Polynomials and Partial Multiplicities of a Product	153
5.3	Divisibility of Monic Matrix Polynomials	156
5.4	Proof of Theorem 5.3.2	161
5.5	Example	167
5.6	Factorization into Several Factors and Chains of Invariant Subspaces	171
5.7	Differential Equations	175
5.8	Difference Equations	180
5.9	Exercises	183
<b>Chapter Six Invariant Subspaces for Transformations Between Different Spaces</b>		<b>189</b>
6.1	$[A \ B]$ -Invariant Subspaces	189
6.2	Block Similarity	192
6.3	Analysis of the Brunovsky Canonical Form	197
6.4	Description of $[A \ B]$ -Invariant Subspaces	200
6.5	The Spectral Assignment Problem	203
6.6	Some Dual Concepts	207
6.7	Exercises	209

<b>Chapter Seven Rational Matrix Functions</b>		<b>212</b>
7.1	Realizations of Rational Matrix Functions	212
7.2	Partial Multiplicities and Multiplication	218
7.3	Minimal Factorization of Rational Matrix Functions	225
7.4	Example	230
7.5	Minimal Factorizations into Several Factors and Chains of Invariant Subspaces	234
7.6	Linear Fractional Transformations	238
7.7	Linear Fractional Decompositions and Invariant Subspaces of Nonsquare Matrices	244
7.8	Linear Fractional Decompositions: Further Deductions	251
7.9	Exercises	255
<b>Chapter Eight Linear Systems</b>		<b>262</b>
8.1	Reductions, Dilations, and Transfer Functions	262
8.2	Minimal Linear Systems: Controllability and Observability	265
8.3	Cascade Connections of Linear Systems	270
8.4	The Disturbance Decoupling Problem	274
8.5	The Output Stabilization Problem	279
8.6	Exercises	285
Notes to Part 1.		290
<b>Part Two Algebraic Properties of Invariant Subspaces</b>		<b>293</b>
<b>Chapter Nine Commuting Matrices and Hyperinvariant Subspaces</b>		<b>295</b>
9.1	Commuting Matrices	295
9.2	Common Invariant Subspaces for Commuting Matrices	301
9.3	Common Invariant Subspaces for Matrices with Rank 1 Commutators	303
9.4	Hyperinvariant Subspaces	305
9.5	Proof of Theorem 9.4.2	307
9.6	Further Properties of Hyperinvariant Subspaces	311
9.7	Exercises	313
<b>Chapter Ten Description of Invariant Subspaces and Linear Transformations with the Same Invariant Subspaces</b>		<b>316</b>
10.1	Description of Irreducible Subspaces	316
10.2	Transformations Having the Same Set of Invariant Subspaces	323

- 10.3 Proof of Theorem 10.2.1 328  
 10.4 Exercises 338

**Chapter Eleven Algebras of Matrices and Invariant Subspaces 339**

- 11.1 Finite-Dimensional Algebras 339  
 11.2 Chains of Invariant Subspaces 340  
 11.3 Proof of Theorem 11.2.1 343  
 11.4 Reflexive Lattices 346  
 11.5 Reductive and Self-Adjoint Algebras 350  
 11.6 Exercises 355

**Chapter Twelve Real Linear Transformations 359**

- 12.1 Definition, Examples, and First Properties of Invariant Subspaces 359  
 12.2 Root Subspaces and the Real Jordan Form 363  
 12.3 Complexification and Proof of the Real Jordan Form 366  
 12.4 Commuting Matrices 371  
 12.5 Hyperinvariant Subspaces 374  
 12.6 Real Transformations with the Same Invariant Subspaces 378  
 12.7 Exercises 380  
 Notes to Part 2. 384

**Part Three Topological Properties of Invariant Subspaces and Stability 385**

**Chapter Thirteen The Metric Space of Subspaces 387**

- 13.1 The Gap Between Subspaces 387  
 13.2 The Minimal Angle and the Spherical Gap 392  
 13.3 Minimal Opening and Angular Linear Transformations 396  
 13.4 The Metric Space of Subspaces 400  
 13.5 Kernels and Images of Linear Transformations 406  
 13.6 Continuous Families of Subspaces 408  
 13.7 Applications to Generalized Inverses 411  
 13.8 Subspaces of Normed Spaces 415  
 13.9 Exercises 420

**Chapter Fourteen The Metric Space of Invariant Subspaces 423**

- 14.1 Connected Components: The Case of One Eigenvalue 423  
 14.2 Connected Components: The General Case 426  
 14.3 Isolated Invariant Subspaces 428  
 14.4 Reducing Invariant Subspaces 432  
 14.5 Coinvariant and Semiinvariant Subspaces 437  
 14.6 The Real Case 439  
 14.7 Exercises 443

**Chapter Fifteen Continuity and Stability of Invariant Subspaces 444**

- 15.1 Sequences of Invariant Subspaces 444  
 15.2 Stable Invariant Subspaces: The Main Result 447  
 15.3 Proof of Theorem 15.2.1 in the General Case 451  
 15.4 Perturbed Stable Invariant Subspaces 455  
 15.5 Lipschitz Stable Invariant Subspaces 459  
 15.6 Stability of Lattices of Invariant Subspaces 463  
 15.7 Stability in Metric of the Lattice of Invariant Subspaces 464  
 15.8 Stability of  $[A \ B]$ -Invariant Subspaces 468  
 15.9 Stable Invariant Subspaces for Real Transformations 470  
 15.10 Partial Multiplicities of Close Linear Transformations 475  
 15.11 Exercises 479

**Chapter Sixteen Perturbations of Lattices of Invariant Subspaces with Restrictions on the Jordan Structure 482**

- 16.1 Preservation of Jordan Structure and Isomorphism of Lattices 482  
 16.2 Properties of Linear Isomorphisms of Lattices: The Case of Similar Transformations 486  
 16.3 Distance Between Invariant Subspaces for Transformations with the Same Jordan Structure 492  
 16.4 Transformations with the Same Derogatory Jordan Structure 497  
 16.5 Proofs of Theorems 16.4.1 and 16.4.4 500  
 16.6 Distance between Invariant Subspaces for Transformations with Different Jordan Structures 507  
 16.7 Conjectures 510  
 16.8 Exercises 513



**Chapter Seventeen Applications**

514

- 17.1 Stable Factorizations of Matrix Polynomials: Preliminaries 514
- 17.2 Stable Factorizations of Matrix Polynomials: Main Results 520
- 17.3 Lipschitz Stable Factorizations of Monic Matrix Polynomials 525
- 17.4 Stable Minimal Factorizations of Rational Matrix Functions: The Main Result 528
- 17.5 Proof of the Auxiliary Lemmas 532
- 17.6 Stable Minimal Factorizations of Rational Matrix Functions: Further Deductions 537
- 17.7 Stability of Linear Fractional Decompositions of Rational Matrix Functions 540
- 17.8 Isolated Solutions of Matrix Quadratic Equations 545
- 17.9 Stability of Solutions of Matrix Quadratic Equations 551
- 17.10 The Real Case 553
- 17.11 Exercises 557
- Notes to Part 3. 561

**Part Four Analytic Properties of Invariant Subspaces 563****Chapter Eighteen Analytic Families of Subspaces**

565

- 18.1 Definition and Examples 565
- 18.2 Kernel and Image of Analytic Families of Transformations 569
- 18.3 Global Properties of Analytic Families of Subspaces 575
- 18.4 Proof of Theorem 18.3.1 (Compact Sets) 578
- 18.5 Proof of Theorem 18.3.1 (General Case) 584
- 18.6 Direct Complements for Analytic Families of Subspaces 590
- 18.7 Analytic Families of Invariant Subspaces 594
- 18.8 Analytic Dependence of the Set of Invariant Subspaces and Fixed Jordan Structure 596
- 18.9 Analytic Dependence on a Real Variable 599
- 18.10 Exercises 601

**Chapter Nineteen Jordan Form of Analytic Matrix Functions**

604

- 19.1 Local Behaviour of Eigenvalues and Eigenvectors 604
- 19.2 Global Behaviour of Eigenvalues and Eigenvectors 607

- 19.3 Proof of Theorem 19.2.3 613
- 19.4 Analytic Extendability of Invariant Subspaces 616
- 19.5 Analytic Matrix Functions of a Real Variable 620
- 19.6 Exercises 622

**Chapter Twenty Applications**

624

- 20.1 Factorization of Monic Matrix Polynomials 624
- 20.2 Rational Matrix Functions Depending Analytically on a Parameter 627
- 20.3 Minimal Factorizations of Rational Matrix Functions 634
- 20.4 Matrix Quadratic Equations 639
- 20.5 Exercises 642
- Notes to Part 4. 645

**Appendix. Equivalence of Matrix Polynomials**

646

- A.1 The Smith Form: Existence 646
- A.2 The Smith Form: Uniqueness 651
- A.3 Invariant Polynomials, Elementary Divisors, and Partial Multiplicities 654
- A.4 Equivalence of Linear Matrix Polynomials 659
- A.5 Strict Equivalence of Linear Matrix Polynomials: Regular Case 662
- A.6 The Reduction Theorem for Singular Polynomials 666
- A.7 Minimal Indices and Strict Equivalence of Linear Matrix Polynomials (General Case) 672
- A.8 Notes to the Appendix 678

**List of Notations and Conventions**

679

**References**

683

**Author Index**

687

**Subject Index**

689

***Invariant Subspaces  
of Matrices  
with Applications***

## *Introduction*

Invariant subspaces are a central notion of linear algebra. However, in existing texts and expositions the notion is not easily or systematically followed. Perhaps because the whole structure is very rich, the treatment becomes fragmented as other related ideas and notions intervene. In particular, the notion of an invariant subspace as an entity is often lost in the discussion of eigenvalues, eigenvectors, generalized eigenvectors, and so on. The importance of invariant subspaces becomes clearer in the context of operator theory on spaces of infinite dimension. Here, it can be argued that the structure is poorer and this is one of the few available tools for the study of many classes of operators. Probably for this reason, the first books on invariant subspaces appeared in the framework of infinite-dimensional spaces. It seems to the authors that now there is a case for developing a treatment of linear algebra in which the central role of invariant subspace is systematically followed up.

The need for such a treatment has become more apparent in recent years because of developments in different fields of application and especially in linear systems theory, where concepts such as controllability, feedback, factorization, and realization of matrix functions are commonplace. In the treatment of such problems new concepts and theories have been developed that form complete new chapters in the body of linear algebra. As examples of new concepts of linear algebra developed to meet the needs of systems theory, we should mention invariant subspaces for nonsquare matrices and similarity of such matrices.

In this book the reader will find a treatment of certain aspects of linear algebra that meets the two objectives: to develop systematically the central role of invariant subspaces in the analysis of linear transformations and to include relevant recent developments of linear algebra stimulated by linear systems theory. The latter are not dealt with separately, but are integrated into the text in a way that is natural in the development of the mathematical structure.

The first part of the book, taken alone or together with selections from the other parts, can be used as a text for undergraduate courses in mathematics, having only a first course in linear algebra as prerequisite. At the same time, the book will be of interest to graduate students in science and engineering. We trust that experts will also find the exposition and new results interesting. The authors anticipate that the book will also serve as a valuable reference work for mathematicians, scientists, and engineers. A set of exercises is included in each chapter. In general, they are designed to provide illustrations and training rather than extensions of the theory.

The first part of the book is devoted mainly to geometric properties of invariant subspaces and their applications in three fields. The fields in question are matrix polynomials, rational matrix functions, and linear systems theory. They are each presented in self-contained form, and—rather than being exhaustive—the focus is on those problems in which invariant subspaces of square and nonsquare matrices play a central role. These problems include factorization and linear fractional decompositions for matrix functions; problems of realization for rational matrix functions; and the problem of describing connections, or cascades, of linear systems, pole assignment, output stabilization, and disturbance decoupling.

The second part is of a more algebraic character in which other properties of invariant subspaces are analyzed. It contains an analysis of the extent to which the invariant subspaces determine the parent matrix, invariant subspaces common to commuting matrices, and lattices of subspaces for a single matrix and for algebras of matrices.

The numerical computation of invariant subspaces is a difficult task as, in general, it makes sense to compute only those invariant subspaces that change very little after small changes in the transformation. Thus it is important to have appropriate notions of “stable” invariant subspaces. Such an analysis of the stability of invariant subspaces and their generalizations is the main subject of Part 3. This analysis leads to applications in some of the problem areas mentioned above.

The subject of Part 4 is analytic families of invariant subspaces and has many useful applications. Here, the analysis is influenced by the theory of complex vector bundles, although we do not make use of this theory. The study of the connections between local and global problems is one of the main problems studied in this part. Within reasonable bounds, Part 4 relies only on the theory developed in this book. The material presented here appears for the first time in a book on linear algebra and is thereby made accessible to a wider audience.

## *Part One*

# *Fundamental Properties of Invariant Subspaces and Applications*

Part 1 of this work comprises almost half of the entire book. It includes what can be described as a self-contained course in linear algebra with emphasis on invariant subspaces, together with substantial developments of applications to the theory of polynomial and rational matrix-valued functions, and to systems theory. These applications demand extensions of the standard material in linear algebra that are included in our treatment in a natural way. They also serve to breathe new life into an otherwise familiar body of knowledge. Thus there is a considerable amount of material here (including all of Chapters 3, 4, and 6) that cannot be found in other books on linear algebra.

Almost all of the material in this part can be understood by readers who have completed a beginning course in linear algebra, although there are places where basic ideas of calculus and complex analysis are required.

## Chapter One

# *Invariant Subspaces: Definition, Examples, and First Properties*

This chapter is mainly introductory. It contains the simplest properties of invariant subspaces of a linear transformation. Some basic tools (projectors, factor spaces, angular transformations, triangular forms) for the study of invariant subspaces are developed. We also study the behaviour of invariant subspaces of a transformation when the operations of similarity and taking adjoints are applied to the transformation. The lattice of invariant subspaces of a linear transformation—a notion that will be important in the sequel—is introduced. The presentation of the material here is elementary and does not even require use of the Jordan form.

### 1.1 DEFINITION AND EXAMPLES

Let  $A: \mathbb{C}^n \rightarrow \mathbb{C}^n$  be a linear transformation. A subspace  $\mathcal{M} \subset \mathbb{C}^n$  is called *invariant* for the transformation  $A$ , or *A invariant*, if  $Ax \in \mathcal{M}$  for every vector  $x \in \mathcal{M}$ . In other words,  $\mathcal{M}$  is invariant for  $A$  means that the image of  $\mathcal{M}$  under  $A$  is contained in  $\mathcal{M}$ ;  $A\mathcal{M} \subset \mathcal{M}$ . Trivial examples of invariant subspaces are  $\{0\}$  and  $\mathbb{C}^n$ . Less trivial examples are the subspaces

$$\text{Ker } A = \{x \in \mathbb{C}^n \mid Ax = 0\}$$

and

$$\text{Im } A = \{Ax \mid x \in \mathbb{C}^n\}$$

Indeed, as  $Ax = 0 \in \text{Ker } A$  for every  $x \in \text{Ker } A$ , the subspace  $\text{Ker } A$  is  $A$  invariant. Also, for every  $x \in \mathbb{C}^n$ , the vector  $Ax$  belongs to  $\text{Im } A$ ; in particular,  $A(\text{Im } A) \subset \text{Im } A$ , and  $\text{Im } A$  is  $A$  invariant.

More generally, the subspaces

$$\text{Ker } A^m = \{x \in \mathbb{C}^n \mid A^m x = 0\}, \quad m = 1, 2, \dots$$

and

$$\text{Im } A^m = \{A^m x \mid x \in \mathbb{C}^n\}, \quad m = 1, 2, \dots$$

are  $A$  invariant. To verify this, let  $x \in \text{Ker } A^m$ , so  $A^m x = 0$ . Then  $A^m(Ax) = A(A^m x) = 0$ , that is,  $Ax \in \text{Ker } A^m$ . This means that  $\text{Ker } A^m$  is  $A$  invariant. Further, let  $x \in \text{Im } A^m$ , so  $x = A^m y$  for some  $y \in \mathbb{C}^n$ . Then  $Ax = A(A^m y) = A^{m+1} y$ , which implies that  $Ax \in \text{Im } A^m$ . So  $\text{Im } A^m$  is  $A$  invariant as well.

When convenient, we shall often assume implicitly that a linear transformation from  $\mathbb{C}^m$  into  $\mathbb{C}^n$  is given by an  $n \times m$  matrix with respect to the standard orthonormal bases  $e_1 = \langle 1, 0, \dots, 0 \rangle$ ,  $e_2 = \langle 0, 1, 0, \dots, 0 \rangle$ ,  $e_n = \langle 0, 0, \dots, 0, 1 \rangle$  in  $\mathbb{C}^n$ ,  $e_1, \dots, e_m$  in  $\mathbb{C}^m$ .

The following three examples of transformations and their invariant subspaces are basic and are often used in the sequel.

EXAMPLE 1.1.1. Let

$$A = \begin{bmatrix} \lambda_0 & 1 & \cdots & 0 \\ 0 & \lambda_0 & & \vdots \\ \vdots & & & 1 \\ 0 & \cdots & 0 & \lambda_0 \end{bmatrix}, \quad \lambda_0 \in \mathbb{C}$$

(the  $n \times n$  Jordan block with  $\lambda_0$  on the main diagonal). Every nonzero  $A$ -invariant subspace is of the form  $\text{Span}\{e_1, \dots, e_k\}$ , where  $e_i$  is the vector  $\langle 0, \dots, 0, 1, 0, \dots, 0 \rangle$  with 1 in the  $i$ th place. Indeed, let  $\mathcal{M}$  be a nonzero  $A$ -invariant subspace, and let

$$x = \sum_{i=1}^n \alpha_i e_i, \quad \alpha_i \in \mathbb{C}$$

be a vector from  $\mathcal{M}$  for which the index  $k = \max\{m \mid 1 \leq m \leq n, \alpha_m \neq 0\}$  is maximal. Then clearly

$$\mathcal{M} \subset \text{Span}\{e_1, \dots, e_k\}$$

On the other hand, the vector  $x = \sum_{i=1}^k \alpha_i e_i$ ,  $\alpha_k \neq 0$  belongs to  $\mathcal{M}$ . Hence, since  $\mathcal{M}$  is  $A$  invariant, the vectors

$$\begin{aligned} x_1 &= Ax - \lambda_0 x = \sum_{i=2}^k \alpha_i e_{i-1} \\ x_2 &= Ax_1 - \lambda_0 x_1 = \sum_{i=3}^k \alpha_i e_{i-2} \\ &\vdots \\ x_{k-1} &= Ax_{k-2} - \lambda_0 x_{k-2} = \alpha_k e_1 \end{aligned}$$

also belong to  $\mathcal{M}$ . Hence the vectors

$$\begin{aligned} e_1 &= \frac{1}{\alpha_k} x_{k-1} \\ e_2 &= \frac{1}{\alpha_k} (x_{k-2} - \alpha_{k-1} e_1) \\ &\vdots \\ e_k &= \frac{1}{\alpha_k} \left( x - \sum_{i=1}^{k-1} \alpha_i e_i \right) \end{aligned}$$

belong to  $\mathcal{M}$  as well. So

$$\text{Span}\{e_1, \dots, e_k\} \subset \mathcal{M}$$

and the equality

$$\text{Span}\{e_1, \dots, e_k\} = \mathcal{M}$$

follows. As for every  $y = \sum_{i=1}^k \beta_i e_i \in \text{Span}\{e_1, \dots, e_k\}$  we have

$$Ay = \lambda_0 y + \sum_{i=2}^k \beta_i e_{i-1} \in \text{Span}\{e_1, \dots, e_k\}$$

The subspace  $\text{Span}\{e_1, \dots, e_k\}$  is indeed  $A$  invariant. The total number of  $A$ -invariant subspaces (including  $\{0\}$  and  $\mathbb{C}^n$ ) is thus  $n+1$ .

In this example we have

$$\text{Ker } A = \begin{cases} \{0\} & \text{if } \lambda_0 \neq 0 \\ \text{Span}\{e_1\} & \text{if } \lambda_0 = 0 \end{cases}$$

and

$$\text{Im } A = \begin{cases} \mathbb{C}^n & \text{if } \lambda_0 \neq 0 \\ \text{Span}\{e_1, \dots, e_{n-1}\} & \text{if } \lambda_0 = 0 \end{cases}$$

As expected, these subspaces are  $A$  invariant.  $\square$

EXAMPLE 1.1.2. Let  $A = \lambda_0 I$ , where  $I$  is the  $n \times n$  identity matrix. Clearly, every subspace in  $\mathbb{C}^n$  is  $A$  invariant. Here the number of  $A$ -invariant subspaces is infinite (if  $n > 1$ ).

Note that the set  $\text{Inv}(A)$  of all  $A$ -invariant subspaces is uncountably infinite. Indeed, for linearly independent vectors  $x, y \in \mathbb{C}^n$  the one-dimensional subspaces  $\text{Span}\{x + \alpha y\}$ ,  $\alpha \in \mathbb{R}$  are all different and belong to  $\text{Inv}(A)$ . So they form an uncountable set of  $A$ -invariant subspaces.

Conversely, if every one-dimensional subspace of  $\mathbb{C}^n$  is  $A$  invariant for a linear transformation  $A$ , then  $A = \lambda_0 I$  for some  $\lambda_0$ . Indeed, for every  $x \neq 0$  the subspace  $\text{Span}\{x\}$  is  $A$  invariant, so  $Ax = \lambda(x)x$ , where  $\lambda(x)$  is a complex number that may, *a priori*, depend on  $x$ . Now if  $\lambda(x_1) \neq \lambda(x_2)$  for linearly independent vectors  $x_1$  and  $x_2$ , then  $\text{Span}\{x_1 + x_2\}$  is not  $A$  invariant, because

$$A(x_1 + x_2) = \lambda(x_1)x_1 + \lambda(x_2)x_2 \notin \text{Span}\{x_1 + x_2\}$$

Hence we must have  $\lambda_0 = \lambda(x)$  is independent of  $x \neq 0$ , so actually  $A = \lambda_0 I$ .  $\square$

Later (see Proposition 2.5.4) we shall see that the set of all  $A$ -invariant subspaces of an  $n \times n$  complex matrix  $A$  is never countably infinite; it is either finite or uncountably infinite.

EXAMPLE 1.1.3. Let

$$A = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ & \lambda_2 & & \vdots \\ & & \ddots & \vdots \\ 0 & \cdots & & \lambda_n \end{bmatrix} \quad (n \geq 2)$$

where the complex numbers  $\lambda_1, \dots, \lambda_n$  are distinct. For any indices  $1 \leq i_1 < \dots < i_k \leq n$  the subspace  $\text{Span}\{e_{i_1}, \dots, e_{i_k}\}$  is  $A$  invariant. Indeed, for

$$x = \sum_{j=1}^k \alpha_j e_{i_j} \in \text{Span}\{e_{i_1}, \dots, e_{i_k}\}$$

we have

$$Ax = \sum_{j=1}^k \alpha_j \lambda_{i_j} e_{i_j} \in \text{Span}\{e_{i_1}, \dots, e_{i_k}\}$$

It turns out that these are all the invariant subspaces for  $A$ . The proof of this fact for a general  $n$  is given later in a more general framework. So the total number of  $A$ -invariant subspaces is

$$\sum_{k=0}^n \binom{n}{k} = 2^n$$

Here we shall check only that the  $2 \times 2$  matrix

$$A = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}, \quad \lambda_1 \neq \lambda_2$$

has exactly two nontrivial invariant subspaces,  $\text{Span}\{e_1\}$  and  $\text{Span}\{e_2\}$ . Indeed, let  $\mathcal{M}$  be any one-dimensional  $A$ -invariant subspace

$$\mathcal{M} = \text{Span}\{x\}, \quad x = \alpha_1 e_1 + \alpha_2 e_2 \neq 0$$

Then  $Ax = \alpha_1 \lambda_1 e_1 + \alpha_2 \lambda_2 e_2$  should belong to  $\mathcal{M}$  and thus is a scalar multiple of  $x$ :

$$\alpha_1 \lambda_1 e_1 + \alpha_2 \lambda_2 e_2 = \beta \alpha_1 e_1 + \beta \alpha_2 e_2$$

for some  $\beta \in \mathbb{C}$ . Comparing coefficients, we see that we obtain a contradiction  $\lambda_1 = \lambda_2$  unless  $\alpha_1 = 0$  or  $\alpha_2 = 0$ . In the former case  $\mathcal{M} = \text{Span}\{e_2\}$  and in the latter case  $\mathcal{M} = \text{Span}\{e_1\}$ .

In this example we have  $\text{Ker } A = \text{Span}\{e_{i_0}\}$  (when  $\det A = 0$ ), where  $i_0$  is the index for which  $\lambda_{i_0} = 0$  (as we have assumed that the  $\lambda_i$  are distinct and  $\det A = 0$ , there is exactly one such index), and  $\text{Im } A = \text{Span}\{e_i \mid i \neq i_0\}$ .  $\square$

The following observation is often useful in proving that a given subspace is  $A$  invariant: A subspace  $\mathcal{M} = \text{Span}\{x_1, \dots, x_k\}$  is  $A$  invariant if and only if  $Ax_i \in \mathcal{M}$  for  $i = 1, \dots, k$ . The proof of this fact is an easy exercise.

For a given transformation  $A: \mathbb{C}^n \rightarrow \mathbb{C}^n$  and a given vector  $x \in \mathbb{C}^n$ , consider the subspace

$$\mathcal{M} = \text{Span}\{x, Ax, A^2x, \dots\}$$

We now appeal to the Cayley-Hamilton theorem, which states that  $\sum_{j=0}^n \alpha_j A^j = 0$ , where the complex numbers  $\alpha_0, \dots, \alpha_n$  are the coefficients of the characteristic polynomial  $\det(\lambda I - A)$  of  $A$ :

$$\det(\lambda I - A) = \sum_{j=0}^n \alpha_j \lambda^j$$

(By writing  $A$  as an  $n \times n$  matrix in some basis in  $\mathbb{C}^n$ , we easily see from the definition of the determinant that  $\det(\lambda I - A)$  is a polynomial of degree  $n$  with  $\alpha_n = 1$ .) Hence  $A^k x$  with  $k \geq n$  is a linear combination of  $x, Ax, \dots, A^{n-1}x$ , so actually

$$\mathcal{M} = \text{Span}\{x, Ax, A^2x, \dots, A^{n-1}x\}$$

The preceding observation shows immediately that  $\mathcal{M}$  is  $A$  invariant. Any



$A$ -invariant subspace  $\mathcal{L}$  that contains  $x$  also contains all the vectors  $Ax, A^2x, \dots$ , and hence contains  $\mathcal{M}$ . It follows that  $\mathcal{M}$  is the smallest  $A$ -invariant subspace that contains the vector  $x$ .

We conclude this section with another useful fact regarding invariant subspaces. Namely, a subspace  $\mathcal{M} \subset \mathbb{C}^n$  is  $A$  invariant for a transformation  $A: \mathbb{C}^n \rightarrow \mathbb{C}^n$  if and only if it is  $(\alpha A + \beta I)$  invariant, where  $\alpha, \beta$  are arbitrary complex numbers such that  $\alpha \neq 0$ . Indeed, assume that  $\mathcal{M}$  is  $A$  invariant. Then for every  $x \in \mathcal{M}$  we see that the vector

$$(\alpha A + \beta I)x = \alpha Ax + \beta x$$

belongs to  $\mathcal{M}$ . So  $\mathcal{M}$  is  $(\alpha A + \beta I)$  invariant. As

$$A = \frac{1}{\alpha}(\alpha A + \beta I) - \frac{\beta}{\alpha}I$$

the same reasoning shows that any  $(\alpha A + \beta I)$  invariant subspace is also  $A$  invariant.

## 1.2 EIGENVALUES AND EIGENVECTORS

The most primitive nontrivial invariant subspaces are those with dimension equal to one. For a transformation  $A: \mathbb{C}^n \rightarrow \mathbb{C}^n$  and some nonzero  $x \in \mathbb{C}^n$ , therefore, we consider an  $A$ -invariant subspace of the form  $\mathcal{M} = \text{Span}\{x\}$ . In this case there must be a  $\lambda_0 \in \mathbb{C}$  such that  $Ax = \lambda_0 x$ . Since we then have  $A(\alpha x) = \alpha(Ax) = \lambda_0(\alpha x)$  for any  $\alpha \in \mathbb{C}$ , the number  $\lambda_0$  does not depend on the choice of the nonzero vector in  $\mathcal{M}$ . We call  $\lambda_0$  an *eigenvalue* of  $A$ , and, when  $Ax = \lambda_0 x$  with  $0 \neq x \in \mathbb{C}^n$ , we call  $x$  an *eigenvector* of  $A$  (corresponding to the eigenvalue  $\lambda_0$ ). Observe that, since  $(\lambda_0 I - A)x = 0$ , the eigenvalues of  $A$  can also be characterized as the set of complex zeros of the *characteristic polynomial* of  $A$ ;  $\varphi_A(\lambda) \stackrel{\text{def}}{=} \det(\lambda I - A)$ .

The set of all eigenvalues of  $A$  is called the *spectrum* of  $A$  and is denoted by  $\sigma(A)$ . We have seen that any one-dimensional  $A$ -invariant subspace is spanned by some eigenvector. Conversely, if  $x_0$  is an eigenvector of  $A$  corresponding to some eigenvalue  $\lambda_0$ , then  $\text{Span}\{x_0\}$  is  $A$  invariant. (In other words,  $A$  is the operator of multiplication by  $\lambda_0$  when restricted to  $\text{Span}\{x_0\}$ .)

Let us have a closer look at the eigenvalues. As the characteristic polynomial  $\varphi_A(\lambda) = \det(\lambda I - A)$  is a polynomial of degree  $n$ , by the fundamental theorem of algebra,  $\varphi_A(\lambda)$  has  $n$  (in general, complex) zeros when counted with multiplicities. These zeros are exactly the eigenvalues of  $A$ . Since the characteristic polynomial and eigenvalues are independent of the choice of basis producing the matrix representation, they are properties of the underlying transformation. So a transformation  $A: \mathbb{C}^n \rightarrow \mathbb{C}^n$  has exactly

$n$  eigenvalues when counted with multiplicities, and, in any event, the number of *distinct* eigenvalues of  $A$  does not exceed  $n$ . Note that this is a property of transformations over the field of complex numbers (or, more generally, over an algebraically closed field). As we shall see later, a transformation from  $\mathbb{R}^n$  into  $\mathbb{R}^n$  does not always have (real) eigenvalues. Since at least one eigenvector corresponds to any eigenvalue  $\lambda_0$  of  $A$  it follows that every linear transformation  $A: \mathbb{C}^n \rightarrow \mathbb{C}^n$  has at least one one-dimensional invariant subspace. Example 1.1.1 shows that in certain cases a linear transformation has exactly one one-dimensional invariant subspace.

We pass now to the description of two-dimensional  $A$ -invariant subspaces in terms of eigenvalues and eigenvectors. So assume that  $\mathcal{M}$  is a two-dimensional  $A$ -invariant subspace. Then, in a natural way,  $A$  determines a transformation from  $\mathcal{M}$  into  $\mathcal{M}$ . We have seen above that for every transformation in a (complex) finite-dimensional vector space (which can be identified with  $\mathbb{C}^m$  for some  $m$ ) there is an eigenvalue and a corresponding eigenvector. So there exists an  $x_0 \in \mathcal{M} \setminus \{0\}$  and a complex number  $\lambda_0$  such that  $Ax_0 = \lambda_0 x_0$ . Now let  $x_1$  be a vector in  $\mathcal{M}$  for which  $\{x_0, x_1\}$  is a linearly independent set; in other words,  $\mathcal{M} = \text{Span}\{x_0, x_1\}$ . Since  $\mathcal{M}$  is  $A$  invariant it follows that

$$Ax_1 = \mu_0 x_0 + \mu_1 x_1$$

for some complex numbers  $\mu_0$  and  $\mu_1$ . If  $\mu_0 = 0$ , then  $x_1$  is an eigenvector of  $A$  corresponding to the eigenvalue  $\mu_1$ . If  $\mu_0 \neq 0$  and  $\mu_1 \neq \lambda_0$ , then the vector  $y = -\mu_0 x_0 + (\lambda_0 - \mu_1)x_1$  is an eigenvector of  $A$  corresponding to  $\mu_1$  for which  $\{x_0, y\}$  is a linearly independent set. Indeed

$$\begin{aligned} Ay &= -\mu_0 Ax_0 + (\lambda_0 - \mu_1)Ax_1 = -\mu_0 \lambda_0 x_0 + (\lambda_0 - \mu_1)(\mu_0 x_0 + \mu_1 x_1) \\ &= (\lambda_0 - \mu_1)\mu_1 x_1 - \mu_1 \mu_0 x_0 = \mu_1 y \end{aligned}$$

Finally, if  $\mu_0 \neq 0$  and  $\mu_1 = \lambda_0$ , then  $x_0$  is the only eigenvector (up to multiplication by a nonzero complex number) of  $A$  in  $\mathcal{M}$ . To check this, assume that  $\alpha_0 x_0 + \alpha_1 x_1$ ,  $\alpha_1 \neq 0$ , is an eigenvector of  $A$  corresponding to an eigenvalue  $\nu_0$ . Then

$$A(\alpha_0 x_0 + \alpha_1 x_1) = \nu_0 \alpha_0 x_0 + \nu_0 \alpha_1 x_1 \quad (1.2.1)$$

But the left-hand side of this equality is

$$\alpha_0 Ax_0 + \alpha_1 Ax_1 = \alpha_0 \lambda_0 x_0 + \alpha_1 (\mu_0 x_0 + \lambda_0 x_1)$$

and comparing this with equality (2.1), we obtain

$$\lambda_0 \alpha_1 = \nu_0 \alpha_1, \quad \alpha_0 \lambda_0 + \alpha_1 \mu_0 = \nu_0 \alpha_0$$

which (with  $\alpha_1 \neq 0$ ) implies  $\lambda_0 = \nu_0$  and  $\alpha_1 \mu_0 = 0$ , a contradiction with the assumption  $\mu_0 \neq 0$ . However, note that the vectors  $z = (1/\mu_0)x_1$  and  $x_0$  form a linearly independent set and  $z$  has the property that  $Az - \lambda_0 z = x_0$ . Such a vector  $z$  will be called a *generalized eigenvector* of  $A$  corresponding to the eigenvector  $x_0$ .

In conclusion, the two-dimensional invariant subspace  $\mathcal{M}$  is spanned by two eigenvectors if and only if either  $\mu_0 = 0$  or  $\mu_0 \neq 0$  and  $\mu_1 \neq \lambda_0$ . If  $\mu_0 \neq 0$  and  $\mu_1 = \lambda_0$ , then  $\mathcal{M}$  is spanned by an eigenvector and a corresponding generalized eigenvector.

A study of invariant subspaces of dimension greater than 2 along these lines becomes tedious. Nevertheless, it can be done and leads to the well-known Jordan normal form of a matrix (or transformation) (see Chapter 2).

Using eigenvectors, one can generally produce numerous invariant subspaces, as demonstrated by the following proposition.

### Proposition 1.2.1

Let  $\lambda_1, \dots, \lambda_k$  be eigenvalues of  $A$  (not necessarily distinct), and let  $x_i$  be an eigenvector of  $A$  corresponding to  $\lambda_i$ ,  $i = 1, \dots, k$ . Then  $\text{Span}\{x_1, \dots, x_k\}$  is an  $A$ -invariant subspace.

*Proof.* For any  $x = \sum_{i=1}^k \alpha_i x_i \in \text{Span}\{x_1, \dots, x_k\}$ , where  $\alpha_i \in \mathbb{C}$ , we have

$$Ax = \sum_{i=1}^k \alpha_i Ax_i = \sum_{i=1}^k \alpha_i \lambda_i x_i$$

so indeed  $\text{Span}\{x_1, \dots, x_k\}$  is  $A$  invariant.  $\square$

For some transformations all invariant subspaces are spanned by eigenvectors as in Proposition 1.2.1, and for some transformations not all invariant subspaces are of this form. Indeed, in Example 1.1.1 only one of the  $n$  nonzero invariant subspaces is spanned by eigenvectors. On the other hand, in Example 1.1.2 every nonzero vector is an eigenvector corresponding to  $\lambda_0$ , so obviously every  $A$ -invariant subspace is spanned by eigenvectors.

## 1.3 JORDAN CHAINS

We have seen in the description of two-dimensional invariant subspaces that eigenvectors alone are not always sufficient for description of all invariant subspaces. This fact necessitates consideration of generalized eigenvectors as well. Let us make a general definition that will include this notion. Let  $\lambda_0$  be an eigenvalue of a linear transformation  $A: \mathbb{C}^n \rightarrow \mathbb{C}^n$ . A chain of vectors

$x_0, x_1, \dots, x_k$  is called a *Jordan chain* of  $A$  corresponding to  $\lambda_0$  if  $x_0 \neq 0$  and the following relations hold:

$$\begin{aligned} Ax_0 &= \lambda_0 x_0 \\ Ax_1 - \lambda_0 x_1 &= x_0 \\ Ax_2 - \lambda_0 x_2 &= x_1 \\ &\vdots \\ Ax_k - \lambda_0 x_k &= x_{k-1} \end{aligned} \quad (1.3.1)$$

The first equation (together with  $x_0 \neq 0$ ) means that  $x_0$  is an eigenvector of  $A$  corresponding to  $\lambda_0$ . The vectors  $x_1, \dots, x_k$  are called *generalized eigenvectors* of  $A$  corresponding to the eigenvalue  $\lambda_0$  and the eigenvector  $x_0$ .

For example, let

$$A = \begin{bmatrix} \lambda_0 & 1 & \cdots & 0 \\ 0 & \lambda_0 & & \vdots \\ \vdots & & \ddots & 1 \\ 0 & \cdots & 0 & \lambda_0 \end{bmatrix}, \quad \lambda_0 \in \mathbb{C}$$

as in Example 1.1.1. Then  $e_1$  is an eigenvector of  $A$  corresponding to  $\lambda_0$ , and  $e_1, e_2, \dots, e_n$  is a Jordan chain. This Jordan chain is by no means unique; for instance,  $e_1, e_2 + \alpha e_1, \dots, e_n + \alpha e_{n-1}$  is again a Jordan chain of  $A$ , where  $\alpha \in \mathbb{C}$  is any number.

In Example 1.1.3 the matrix  $A$  does not have generalized eigenvectors at all; that is, every Jordan chain consists of an eigenvector only. Indeed, we have  $A = \text{diag}[\lambda_1, \lambda_2, \dots, \lambda_n]$ , where  $\lambda_1, \dots, \lambda_n$  are distinct complex numbers; therefore

$$\det(\lambda I - A) = (\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_n)$$

So  $\lambda_1, \dots, \lambda_n$  are exactly the eigenvalues of  $A$ . It is easily seen that any eigenvector of  $A$  corresponding to  $\lambda_{i_0}$  is of the form  $\alpha e_{i_0}$  with a nonzero scalar  $\alpha$ . Assuming that there is a Jordan chain  $\alpha e_{i_0}, x$  of  $A$  corresponding to  $\lambda_{i_0}$ , equations (1.3.1) imply

$$Ax - \lambda_{i_0} x = \alpha e_{i_0} \quad (1.3.2)$$

Write  $x = \sum_{i=1}^n \beta_i e_i$ , then  $Ax = \sum_{i=1}^n \lambda_i \beta_i e_i$ , and equality (1.3.2) gives

$$\sum_{i=1}^n (\lambda_i - \lambda_{i_0}) \beta_i e_i = \alpha e_{i_0} \quad (1.3.3)$$