Matrix Polynomials

I. Gohberg

P. Lancaster

L. Rodman

ACADEMIC PRESS, INC.
111 Fifth Avenue, New York, New York 10003

United Kingdom Edition published by
ACADEMIC PRESS, INC. (LONDON) LTD.
24/28 Oval Road, London NW1 7DX

Library of Congress Cataloging in Publication Data

Gohberg, I. TS. (Izrail' TSudikovich), Date.
Matrix polynomials.

(Computer science and applied mathematics series)
Bibliography: p.
Includes index.
1. Matrices. 2. Polynomials. I. Lancaster, Peter, mathematician. II. Rodman, L. III. Title. IV. Series.
QA188.G64 512.9'434 81-22792
ISBN 0-12-287160-X AACR2

PRINTED IN THE UNITED STATES OF AMERICA

82 83 84 85 9 8 7 6 5 4 3 2 1

Preface

This book provides a comprehensive treatment of the theory of matrix polynomials. By a matrix polynomial (sometimes known as a λ-matrix) is understood a polynomial of a complex variable with matrix coefficients. Basic matrix theory (including the Jordan form, etc.) may be viewed as a theory of matrix polynomials $I\lambda - A$ of first degree. The theory developed here is a natural extension to polynomials of higher degrees, and forms an important new part of linear algebra for which the main concepts and results have been arrived at during the past five years. The material has important applications in differential equations, boundary value problems, the Wiener–Hopf technique, system theory, analysis of vibrations, network theory, filtering of multiple time series, numerical analysis, and other areas. The mathematical tools employed are accessible even for undergraduate students who have studied matrix theory and complex analysis. Consequently, the book will be useful to a wide audience of engineers, scientists, mathematicians, and students working in the fields mentioned, and it is to this audience that the work is addressed.

Collaboration among the authors on problems concerning matrix polynomials started in early 1976. We came to the subject with quite different backgrounds in operator theory and in applied mathematics, but had in common a desire to understand matrix polynomials better from the point of view of spectral theory. After bringing together our points of view, expertise, and tools, the solution to some problems for monic polynomials could be seen already by the summer of 1976. Then the theory evolved rapidly to include deeper analysis, more general (not monic) problems on the one hand, and more highly structured (self-adjoint) problems on the other. This work, enjoyable and exciting, was initially carried out at Tel-Aviv University, Israel, and the University of Calgary, Canada.

Very soon after active collaboration began, colleagues in Amsterdam, The Netherlands, and Haifa, Israel, were attracted to the subject and began to make substantial contributions. We have in mind H. Bart and M. A. Kaashoek of the Free University, Amsterdam, and L. Lerer of the Technion, Haifa. It is a pleasure to acknowledge their active participation in the development of the work we present and to express our gratitude for thought-provoking discussions. This three-way international traffic of ideas and personalities has been a fruitful and gratifying experience for the authors.

The past four years have shown that, indeed, a theory has evolved with its own structure and applications. The need to present a connected treatment of this material provided the motivation for writing this monograph. However, the material that we present could not be described as closed, or complete. There is related material in the literature which we have not included, and there are still many open questions to be answered.

Many colleagues have given us the benefit of discussion, criticism, or access to unpublished papers. It is a pleasure to express our appreciation for such assistance from E. Bohl, K. Clancey, N. Cohen, M. Cowen, P. Dewilde, R. G. Douglas, H. Dym, C. Foias, P. Fuhrmann, S. Goldberg, B. Gramsch, J. W. Helton, T. Kailath, R. E. Kalman, B. Lawruk, D. C. Lay, J. D. Pincus, A. Ran, B. Rowley, P. Van Dooren, F. Van Schagen, J. Willems, and H. K. Wimmer.

The authors acknowledge financial support for some or all of them from the Natural Sciences and Engineering Research Council of Canada, and the National Science Foundation of the United States. We are also very grateful to our home departments at Tel-Aviv University, the Weizmann Institute, and the University of Calgary for understanding and support. In particular, the second author is grateful for the award of a Killam Resident Fellowship at the University of Calgary. At different times the authors have made extended visits to the Free University, Amsterdam, the State University of New York at Stony Brook, and the University of Münster. These have been important in the development of our work, and the support of these institutions is cordially acknowledged.

Several members of the secretarial staff of the Department of Mathematics and Statistics of the University of Calgary have worked diligently and skillfully on the preparation of drafts and the final typescript. The authors much appreciate their efforts, especially those of Liisa Torrence, whose contributions far exceeded the call of duty.

Contents

Chapter 3 Multiplication and Divisibility

Chapter 4 Spectral Divisors and Canonical Factorization

Chapter 5 Perturbation and Stability of Divisors

Chapter 6 Extension Problems

Part II NONMONIC MATRIX POLYNOMIALS

Chapter 7 Spectral Properties and Representations

Introduction

This is probably the first book to contain a comprehensive theory of matrix polynomials. Although the importance of matrix polynomials is quite clear, books on linear algebra and matrix theory generally present only a modest treatment, if any. The important treatise of Gantmacher [22], for example, gives the subject some emphasis, but mainly as a device for developing the Jordan structure of a square matrix. The authors are aware of only two earlier works devoted primarily to matrix polynomials, both of which are strongly motivated by the theory of vibrating systems: one by Frazer, Duncan, and Collar in 1938 [19], and the other by one of the present authors in 1966 [52b].

By a matrix polynomial, sometimes known as a λ-matrix, we understand a matrix-valued function of a complex variable of the form $L(\lambda) = \sum_{i=0}^{l} A_i \lambda^i$, where $A_0, A_1, \ldots, A_l$ are $n \times n$ matrices of complex numbers. For the time being, we suppose that $A_l = I$, the identity matrix, in which case $L(\lambda)$ is said to be *monic*. Generally, the student first meets with matrix polynomials when studying systems of ordinary differential equations (of order $l > 1$) with constant coefficients, i.e., a system of the form

$$\sum_{i=0}^{l} A_i \left(\frac{d}{dt}\right)^i u(t) = 0.$$

Looking for solutions of the form $u(t) = x_0 e^{\lambda_0 t}$, with x_0, λ_0 independent of t, immediately leads to the eigenvalue–eigenvector problem for a matrix polynomial: $L(\lambda_0)x_0 = 0$.

More generally, the function

$$u(t) = \left\{\frac{t^k}{k!} x_0 + \cdots + \frac{t}{1!} x_{k-1} + x_k\right\} e^{\lambda_0 t}$$

is a solution of the differential equation if and only if the set of vectors $x_0, x_1, \ldots, x_k$, with $x_0 \neq 0$, satisfies the relations

$$\sum_{p=0}^{j} \frac{1}{p!} L^{(p)}(\lambda_0) x_{j-p} = 0, \qquad j = 0, 1, \ldots, k.$$

Such a set of vectors $x_0, x_1, \ldots, x_k$ is called a *Jordan chain* of length $k + 1$ associated with eigenvalue λ_0 and eigenvector x_0.

It is this information on eigenvalues with associated multiplicities and Jordan chains which we refer to as the spectral data for the matrix polynomial, and is first to be organized in a concise and systematic way. The spectral theory we are to develop must include as a special case the classical theory for polynomials of first degree (when we may write $L(\lambda) = I\lambda - A$). Another familiar special case which will be included is that of scalar polynomials, when the A_i are simply complex numbers and, consequently, the analysis of a single (scalar) constant coefficient ordinary differential equation of order l.

Now, what we understand by spectral theory must contain a complete and explicit description of the polynomial itself in terms of the spectral data. When $L(\lambda) = I\lambda - A$, this is obtained when a Jordan form J for A is known together with a transforming matrix X for which $A = XJX^{-1}$, for we then have $L(\lambda) = X(I\lambda - J)X^{-1}$. Furthermore, X can be interpreted explicitly in terms of the eigenvector structure of A (or of $I\lambda - A$ in our terminology). The full generalization of this to matrix polynomials $L(\lambda)$ of degree l is presented here and is, surprisingly, of very recent origin.

The generalization referred to includes a Jordan matrix J of size ln which contains all the information about the eigenvalues of $L(\lambda)$ and their multiplicities. In addition, we organize complete information about Jordan chains in a single $n \times ln$ matrix X. This is done by associating with a typical $k \times k$ Jordan block of J with eigenvalue λ_0 k columns of X in the corresponding (consecutive) positions which consist of the vectors in an associated Jordan chain of length k. When $l = 1$, J and X reduce precisely to the classical case mentioned above. The representation of the coefficients of $L(\lambda)$ is then obtained in terms of this pair of matrices, which we call a *Jordan pair* for $L(\lambda)$. In fact, if we define the $ln \times ln$ matrix

$$Q = \begin{bmatrix} X \\ XJ \\ \vdots \\ XJ^{l-1} \end{bmatrix},$$

then Q is necessarily nonsingular and the coefficient matrices of $L(\lambda)$ are given by

$$[A_0 \quad A_1 \quad \cdots \quad A_{l-1}] = -XJ^l Q^{-1}.$$

Such a representation for the polynomial coefficients gives a solution of the inverse problem: to determine a matrix polynomial in terms of its spectral data. It also suggests further problems. Given only partial spectral data, when can it be extended to complete spectral data for some monic matrix polynomial? Problems of this kind are soluble by the methods we develop.

There is, of course, a close connection between the systems of eigenvectors for $L(\lambda)$ (the *right* eigenvectors) and those of the transposed polynomial $L^T(\lambda) = \sum_{i=0}^{l} A_i^T \lambda^i$ (the *left* eigenvectors). In fact, complete information about the left Jordan chains can be obtained from the $ln \times n$ matrix Y defined in terms of a Jordan pair by

$$\begin{bmatrix} X \\ XJ \\ \vdots \\ XJ^{l-1} \end{bmatrix} Y = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ I \end{bmatrix}.$$

The three matrices X, J, Y are then described as a Jordan triple.

A representation theorem can now be formulated for the *inverse* of a monic matrix polynomial. We have

$$L^{-1}(\lambda) = X(I\lambda - J)^{-1}Y.$$

Results of this kind admit compact closed form solutions of the corresponding differential equations with either initial or two-point boundary conditions.

It is important to note that the role of a Jordan pair X, J can, for many aspects of the theory, be played by any pair, say V, T related to X, J by similarity as follows:

$$V = XS^{-1}, \qquad T = SJS^{-1}.$$

Any such pair V, T is called a *standard pair* for $L(\lambda)$. In particular, if C_1 is the companion matrix for $L(\lambda)$ defined by

$$C_1 = \begin{bmatrix} 0 & I & 0 & \cdots & 0 \\ \vdots & & I & \ddots & \vdots \\ 0 & 0 & & \cdots & I \\ -A_0 & -A_1 & & \cdots & -A_{l-1} \end{bmatrix},$$

then J is also a Jordan form for C_1 and there is an invertible S for which $C_1 = SJS^{-1}$. But more can be said. If U is the $n \times ln$ matrix $[I \quad 0 \quad \cdots \quad 0]$, and Q is the invertible matrix introduced above, then

$$U = XQ^{-1}, \qquad C_1 = QJQ^{-1}.$$

Thus U, C_1 is a standard pair. This particular standard pair admits the formulation of several important results in terms of either spectral data or the coefficients of L. Thus, the theory is not limited by the difficulties associated with the calculation of Jordan normal forms, for example.

A standard triple V, T, W is obtained by adjoining to a standard pair V, T the (unique) $ln \times n$ matrix W for which

$$\begin{bmatrix} V \\ VT \\ \vdots \\ VT^{l-1} \end{bmatrix} W = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ I \end{bmatrix}.$$

The next main feature of the theory is the application of the spectral analysis to factorization problems. In terms of the motivation via differential equations, such results can be seen as replacing an lth-order system by a composition of two lower-order systems, with natural advantages in the analysis of solutions. However, factorizations arise naturally in several other situations. For example, the notion of decoupling in systems theory requires such a factorization, as also the design of filters for multiple time series, and the study of Toeplitz matrices and Wiener–Hopf equations.

An essential part of the present theory of factorization involves a geometrical characterization of right divisors. To introduce this, consider any standard pair V, T of $L(\lambda)$. Let $\mathscr{S}$ be an invariant subspace of T, let V_0, T_0 denote the restrictions of V and T to $\mathscr{S}$, and define a linear transformation $Q_k \colon \mathscr{S} \to \mathbb{C}^{nk}$ $(1 \le k \le l-1)$ by

$$Q_k = \begin{bmatrix} V_0 \\ V_0 T_0 \\ \vdots \\ V_0 T_0^{k-1} \end{bmatrix}.$$

Then $\mathscr{S}$ is said to be a *supporting subspace* (with respect to T) if Q_k is invertible. It is found that to each such supporting subspace corresponds a monic right divisor of degree k and, conversely, each monic right divisor has an associated supporting subspace. Furthermore, when $\mathscr{S}$ is a supporting subspace with respect to T, the pair V_0, T_0 is a standard pair for the associated right divisor.

This characterization allows one to write down explicit formulas for a right divisor and the corresponding quotient. Thus, if the coefficients of the monic right divisor of degree k are $B_0, B_1, \ldots, B_{k-1}$, then

$$[B_0 \quad B_1 \quad \cdots \quad B_{k-1}] = -V_0 T_0^k Q_k^{-1}.$$

To obtain the coefficients of the corresponding monic quotient of degree $l - k$, we introduce the third component W of the standard tiple V, T, W and any projector P which acts *along* the supporting subspace $\mathscr{S}$, i.e., $\operatorname{Ker} P = \mathscr{S}$. Then define a linear transformation $R_{l-k}\colon \mathbb{C}^{n(l-k)} \to \operatorname{Im} P$ by

$$R_{l-k} = [PW \quad PTPW \quad \cdots \quad PT^{l-k-1}PW].$$

This operator is found to be invertible and, if $C_0, C_1, \ldots, C_{l-k-1}$ are the coefficients of the quotient, then

$$\begin{bmatrix} C_0 \\ C_1 \\ \vdots \\ C_{l-k-1} \end{bmatrix} = R_{l-k}^{-1}.$$

Several applications concern the existence of divisors which have spectrum localized in the complex plane in some way (in a half-plane, or the unit circle, for example). For such cases the geometrical approach admits the transformation of the problem to the construction of invariant subspaces with associated properties. This interpretation of the problem gives new insights into the theory of factorization and admits a comprehensive treatment of questions concerning existence, perturbations, stability, and explicit representations, for example. In addition, extensions of the theory to the study of *systems* of matrix polynomials provide a useful geometrical approach to least common multiples and greatest common divisors.

The spectral theory developed here is also extended to include matrix polynomials for which the leading coefficient is not necessarily the identity matrix, more exactly, the case in which $\det L(\lambda) \not\equiv 0$. In this case it is necessary to include the possibility of a point of spectrum at infinity, and this needs careful treatment. It is easily seen that if $\det L(\lambda_0) \neq 0$, and we define $\mu = \lambda + \lambda_0$, then $M(\mu) = \mu^l L^{-1}(\lambda_0) L(\mu^{-1})$ is monic. Such a transformation is basic, but it is still necessary to analyze the effect of this transformation on the original problem.

Several important problems give rise to *self-adjoint* polynomials, i.e., in which all coefficients $A_0, A_1, \ldots, A_l$ are hermitian matrices. The prototype problem of this kind is that of damped vibrations, when $l = 2$ and A_2 is positive definite. The more familiar special cases which a theory for the self-adjoint case must include are $L(\lambda) = I\lambda - A$ with $A^* = A$ and the case of a scalar polynomial with *real* coefficients. The general theory developed leads to the introduction of a new invariant for a self-adjoint polynomial called the *sign characteristic*.

For self-adjoint matrix polynomials it is natural to seek the corresponding symmetry in a Jordan triple. What is going to be the simplest relationship between X and Y in a Jordan triple X, J, Y? It turns out that knowledge of the sign characteristic is vital in providing an answer to this question. Also, when a factorization involves a real eigenvalue common to both divisor and quotient, then the sign characteristic plays an important part.

The account given to this point serves to clarify the intent of the theory presented and to describe, in broad terms, the class of problems to which it is applied. This is the content of the spectral theory of matrix polynomials, as understood by the authors. Now we would like to mention some related developments which have influenced the authors' thinking.

In the mathematical literature many papers have appeared in the last three decades on operator polynomials and on more general operator-valued functions. Much of this work is concerned with operators acting on infinite-dimensional spaces and it is, perhaps, surprising that the complete theory for the finite-dimensional case has been overlooked. It should also be mentioned, however, that a number of important results discussed in this book are valid, and were first discovered, in the more general case.

On reflection, the authors find that four particular developments in operator theory have provided them with ideas and stimulus. The first is the early work of Keldysh [49a], which motivated, and made attractive, the study of spectral theory for operator polynomials; see also [32b]. Then the work of Krein and Langer [51] led to an appreciation of the importance of monic divisors for spectral theory and the role of the methods of indefinite scalar product spaces. Third, the theory of characteristic functions developed by Brodskii and Livsic [9] gave a clue for the characterization of right divisors by supporting subspaces. Finally, the work of Marcus and Mereutsa [62a] was very helpful in the attack on problems concerning greatest common divisors and least common multiples. In each case, the underlying vector spaces are infinite dimensional, so that the emphasis and objectives may be rather different from those which appear in the theory developed here.

In parallel with the operator theoretic approach, a large body of work on systems theory was evolving, mainly in the engineering literature. The authors learned about this more recently and although it has had its effect on us, the theory presented here is largely independent of systems theory. Even so, several formulas and notions have striking similarities and we know the connection to be very strong. The most obvious manifestation of this is the observation that, from the systems theory point of view, we study here systems for which the transfer function is the inverse of a matrix polynomial. However, this is the tip of the iceberg, and a more complete account of these connections would take us too far afield.

Another influential topic for the authors is the study of Toeplitz matrices and Wiener–Hopf equations. Although we do not pursue the connections in this book, they are strong. In particular, the spectral theory approach admits the calculation of the "partial indices"; a concept of central importance in that theory.

Finally, a short description of the contents of the book is presented. There are 19 chapters which are grouped into four parts. In Part I, consisting of six chapters, the spectral theory for monic matrix polynomials is developed beginning with the basic ideas of linearization and Jordan chains. We then go on to representation theorems for a monic polynomial and its inverse, followed by analysis of divisibility problems. These sections are illustrated by applications to differential and difference equations. The last three chapters are concerned with special varieties of factorization, perturbation and stability of divisors, and extension problems.

Part II consists of three chapters and is devoted to more general problems in which the "monic" condition is relaxed. The necessary extensions to the spectral theory are made and applied to differential and difference equations. Problems concerning least common multiples and greatest common divisors are then discussed in this context.

Further concepts are needed for the analysis of self-adjoint polynomials, which is presented in the four chapters of Part III. After developing the additional spectral theory required, including the introduction of the sign characteristic, some factorization problems are discussed. The last chapter provides illustrations of the theory of Part III in the case of matrix polynomials of second degree arising in the study of damped vibrating systems.

Finally, Part IV consists of six supplementary chapters added to make this work more self-contained. It contains topics not easily found elsewhere, or for which it is useful to have at hand a self-contained development of concepts and terminology.