Preface

Functions of matrices have been studied for as long as matrix algebra itself. Indeed, in his seminal *A Memoir on the Theory of Matrices* (1858), Cayley investigated the square root of a matrix, and it was not long before definitions of $f(A)$ for general f were proposed by Sylvester and others. From their origin in pure mathematics, matrix functions have broadened into a subject of study in applied mathematics, with widespread applications in science and engineering. Research on matrix functions involves matrix theory, numerical analysis, approximation theory, and the development of algorithms and software, so it employs a wide range of theory and methods and promotes an appreciation of all these important topics.

My first foray into $f(A)$ was as a graduate student when I became interested in the matrix square root. I have worked on matrix functions on and off ever since. Although there is a large literature on the subject, including chapters in several books (notably Gantmacher [203, 1959], Horn and Johnson [296, 1991], Lancaster and Tismenetsky [371, 1985], and Golub and Van Loan [224, 1996]), there has not previously been a book devoted to matrix functions. I started to write this book in 2003. In the intervening period interest in matrix functions has grown significantly, with new applications appearing and the literature expanding at a fast rate, so the appearance of this book is timely.

This book is a research monograph that aims to give a reasonably complete treatment of the theory of matrix functions and numerical methods for computing them, as well as an overview of applications. The theory of matrix functions is beautiful and nontrivial. I have strived for an elegant presentation with illuminating examples, emphasizing results of practical interest. I focus on three equivalent definitions of $f(A)$, based on the Jordan canonical form, polynomial interpolation, and the Cauchy integral formula, and use all three to develop the theory. A thorough treatment is given of problem sensitivity, based on the Fréchet derivative. The applications described include both the well known and the more speculative or recent, and differential equations and algebraic Riccati equations underlie many of them.

The bulk of the book is concerned with numerical methods and the associated issues of accuracy, stability, and computational cost. Both general purpose methods and methods for specific functions are covered. Little mention is made of methods that are numerically unstable or have exorbitant operation counts of order n^4 or higher; many methods proposed in the literature are ruled out for at least one of these reasons.

The focus is on theory and methods for general matrices, but a brief introduction to functions of structured matrices is given in Section 14.1. The problem of computing a function of a matrix times a vector, $f(A)b$, is of growing importance, though as yet numerical methods are relatively undeveloped; Chapter 13 is devoted to this topic.

One of the pleasures of writing this book has been to explore the many connections between matrix functions and other subjects, particularly matrix analysis and numerical analysis in general. These connections range from the expected, such as divided differences, the Kronecker product, and unitarily invariant norms, to the unexpected, which include the Mandelbrot set, the geometric mean, partial isometries, and the role of the Fréchet derivative beyond measuring problem sensitivity.

I have endeavoured to make this book more than just a monograph about matrix functions, and so it includes many useful or interesting facts, results, tricks, and techniques that have a (sometimes indirect) $f(A)$ connection. In particular, the book contains a substantial amount of matrix theory, as well as many historical references, some of which appear not to have previously been known to researchers in the area. I hope that the book will be found useful as a source of statements and applications of results in matrix analysis and numerical linear algebra, as well as a reference on matrix functions.

Four main themes pervade the book.

Role of the sign function. The matrix sign function has fundamental theoretical and algorithmic connections with the matrix square root, the polar decomposition, and, to a lesser extent, matrix pth roots. For example, a large class of iterations for the matrix square root can be obtained from corresponding iterations for the matrix sign function, and Newton's method for the matrix square root is mathematically equivalent to Newton's method for the matrix sign function.

Stability. The stability of iterations for matrix functions can be effectively defined and analyzed in terms of power boundedness of the Fréchet derivative of the iteration function at the solution. Unlike some earlier, more ad hoc analyses, no assumptions are required on the underlying matrix. General results (Theorems 4.18 and 4.19) simplify the analysis for idempotent functions such as the matrix sign function and the unitary polar factor.

Schur decomposition and Parlett recurrence. The use of a Schur decomposition followed by reordering and application of the block form of the Parlett recurrence yields a powerful general algorithm, with f-dependence restricted to the evaluation of f on the diagonal blocks of the Schur form.

Padé approximation. For transcendental functions the use of Padé approximants, in conjunction with an appropriate scaling technique that brings the matrix argument close to the origin, yields an effective class of algorithms whose computational building blocks are typically just matrix multiplication and the solution of multiple right-hand side linear systems. Part of the success of this approach rests on the several ways in which rational functions can be evaluated at a matrix argument, which gives the scope to find a good compromise between speed and stability.

In addition to surveying, unifying, and sometimes improving existing results and algorithms, this book contains new results. Some of particular note are as follows.

- Theorem 1.35, which relates $f(\alpha I_m + AB)$ to $f(\alpha I_n + BA)$ for $A \in \mathbb{C}^{m\times n}$ and $B \in \mathbb{C}^{n\times m}$ and is an analogue for general matrix functions of the Sherman–Morrison–Woodbury formula for the matrix inverse.

- Theorem 4.15, which shows that convergence of a scalar iteration implies convergence of the corresponding matrix iteration when applied to a Jordan block, under suitable assumptions. This result is useful when the matrix iteration can be block diagonalized using the Jordan canonical form of the underlying matrix, A. Nevertheless, we show in the context of Newton's method for the matrix square root that analysis via the Jordan canonical form of A does not always give the strongest possible convergence result. In this case a stronger result, Theorem 6.9, is obtained essentially by reducing the convergence analysis to the consideration of the behaviour of the powers of a certain matrix.

- Theorems 5.13 and 8.19 on the stability of essentially all iterations for the matrix sign function and the unitary polar factor, and the general results in Theorems 4.18 and 4.19 on which these are based.
- Theorems 6.14–6.16 on the convergence of the binomial, Pulay, and Visser iterations for the matrix square root.
- An improved Schur–Parlett algorithm for the matrix logarithm, given in Section 11.6, which makes use of improved implementations of the inverse scaling and squaring method in Section 11.5.

The Audience

The book's main audience is specialists in numerical analysis and applied linear algebra, but it will be of use to anyone who wishes to know something of the theory of matrix functions and state of the art methods for computing them. Much of the book can be understood with only a basic grounding in numerical analysis and linear algebra.

Using the Book

The book can be used as the basis for a course on functions of matrices at the graduate level. It is also a suitable reference for an advanced course on applied or numerical linear algebra, which might cover particular topics such as definitions and properties of $f(A)$, or the matrix exponential and logarithm. It can be used by instructors at all levels as a supplementary text from which to draw examples, historical perspective, statements of results, and exercises. The book, and the subject itself, are particularly well suited to self-study.

To a large extent the chapters can be read independently. However, it is advisable first to become familiar with Sections 1.1–1.3, the first section of Chapter 3 (*Conditioning*), and most of Chapter 4 (*Techniques for General Functions*).

The *Notes and References* are an integral part of each chapter. In addition to containing references, historical information, and further details, they include material not covered elsewhere in the chapter and should always be consulted, in conjunction with the index, to obtain the complete picture.

This book has been designed to be as easy to use as possible and is relatively self-contained. Notation is summarized in Appendix A, while Appendix B (*Background: Definitions and Useful Facts*) reviews basic terminology and needed results from matrix analysis and numerical analysis. When in doubt about the meaning of a term the reader should consult the comprehensive index. Appendix C provides a handy summary of operation counts for the most important matrix computation kernels. Each bibliography entry shows on which pages the item is cited, which makes browsing through the bibliography another route into the book's content.

The exercises, labelled "problems", are an important part of the book, and many of them are new. Solutions, or occasionally a reference to where a solution can be found, are given for almost every problem in Appendix E. Research problems given at the end of some sets of problems highlight outstanding open questions.

A Web page for the book can be found at

`http://www.siam.org/books/ot104`

It includes

- The Matrix Function Toolbox for MATLAB, described in Appendix D. This toolbox contains basic implementations of the main algorithms in the book.
- Updates relating to material in the book.
- A BibTeX database `functions-of-matrices.bib` containing all the references in the bibliography.

Acknowledgments

A number of people have influenced my thinking about matrix functions. Discussions with Ralph Byers in 1984, when he was working on the matrix sign function and I was investigating the polar decomposition, first made me aware of connections between these two important tools. The work on the matrix exponential of Cleve Moler and Charlie Van Loan has been a frequent source of inspiration. Beresford Parlett's ideas on the exploitation of the Schur form and the adroit use of divided differences have been a guiding light. Charles Kenney and Alan Laub's many contributions to the matrix function arena have been important in my own research and are reported on many pages of this book. Finally, Nick Trefethen has shown me the importance of the Cauchy integral formula and has offered valuable comments on drafts at all stages.

I am grateful to several other people for providing valuable help, suggestions, or advice during the writing of the book:

> Rafik Alam, Awad Al-Mohy, Zhaojun Bai, Timo Betcke, Rajendra Bhatia, Tony Crilly, Philip Davies, Oliver Ernst, Andreas Frommer, Chun-Hua Guo, Gareth Hargreaves, Des Higham, Roger Horn, Bruno Iannazzo, Ilse Ipsen, Peter Lancaster, Jörg Liesen, Lijing Lin, Steve Mackey, Roy Mathias, Volker Mehrmann, Thomas Schmelzer, Gil Strang, Françoise Tisseur, and Andre Weideman.

Working with the SIAM staff on the publication of this book has been a pleasure. I thank, in particular, Elizabeth Greenspan (acquisitions), Sara Murphy (acquisitions), Lois Sellers (design), and Kelly Thomas (copy editing).

Research leading to this book has been supported by the Engineering and Physical Sciences Research Council, The Royal Society, and the Wolfson Foundation.

Manchester
December 2007

Nicholas J. Higham