

AN INTRODUCTION TO
**NUMERICAL
METHODS** AND
ANALYSIS

Second Edition

JAMES F. EPPERSON

WILEY

**AN INTRODUCTION TO
NUMERICAL METHODS
AND ANALYSIS**

AN INTRODUCTION TO NUMERICAL METHODS AND ANALYSIS

Second Edition

JAMES F. EPPERSON

Mathematical Reviews

WILEY

Copyright © 2013 by John Wiley & Sons, Inc. All rights reserved.

Published by John Wiley & Sons, Inc., Hoboken, New Jersey.

Published simultaneously in Canada.

No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning or otherwise, except as permitted under Section 107 or 108 of the 1976 United States Copyright Act, without either the prior written permission of the Publisher, or authorization through payment of the appropriate per-copy fee to the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, (978) 750-8400, fax (978) 750-4470, or on the web at www.copyright.com. Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, (201) 748-6011, fax (201) 748-6008, or online at <http://www.wiley.com/go/permission>.

Limit of Liability/Disclaimer of Warranty: While the publisher and author have used their best efforts in preparing this book, they make no representation or warranties with respect to the accuracy or completeness of the contents of this book and specifically disclaim any implied warranties of merchantability or fitness for a particular purpose. No warranty may be created or extended by sales representatives or written sales materials. The advice and strategies contained herein may not be suitable for your situation. You should consult with a professional where appropriate. Neither the publisher nor author shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages.

For general information on our other products and services please contact our Customer Care Department within the United States at (800) 762-2974, outside the United States at (317) 572-3993 or fax (317) 572-4002.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print, however, may not be available in electronic formats. For more information about Wiley products, visit our web site at www.wiley.com.

Library of Congress Cataloging-in-Publication Data:

Epperson, James F., author.

An introduction to numerical methods and analysis / James F. Epperson, Mathematical Reviews. — Second edition.
pages cm

Includes bibliographical references and index.

ISBN 978-1-118-36759-9 (hardback)

1. Numerical analysis. I. Title.

QA297.E568 2013

518—dc23

2013013979

Printed in the United States of America.

10 9 8 7 6 5 4 3 2 1

To Mom (1920–1986) and Ed (1917–2012)
a story of love, faith, and grace

CONTENTS

Preface	xiii	
1	Introductory Concepts and Calculus Review	1
1.1	Basic Tools of Calculus	2
1.1.1	Taylor's Theorem	2
1.1.2	Mean Value and Extreme Value Theorems	9
1.2	Error, Approximate Equality, and Asymptotic Order Notation	14
1.2.1	Error	14
1.2.2	Notation: Approximate Equality	15
1.2.3	Notation: Asymptotic Order	16
1.3	A Primer on Computer Arithmetic	20
1.4	A Word on Computer Languages and Software	29
1.5	Simple Approximations	30
1.6	Application: Approximating the Natural Logarithm	35
1.7	A Brief History of Computing	37
1.8	Literature Review	40
	References	41
2	A Survey of Simple Methods and Tools	43
2.1	Horner's Rule and Nested Multiplication	43
2.2	Difference Approximations to the Derivative	48
		vii

2.3	Application: Euler's Method for Initial Value Problems	56
2.4	Linear Interpolation	62
2.5	Application—The Trapezoid Rule	68
2.6	Solution of Tridiagonal Linear Systems	78
2.7	Application: Simple Two-Point Boundary Value Problems	85
3	Root-Finding	89
3.1	The Bisection Method	90
3.2	Newton's Method: Derivation and Examples	97
3.3	How to Stop Newton's Method	103
3.4	Application: Division Using Newton's Method	106
3.5	The Newton Error Formula	110
3.6	Newton's Method: Theory and Convergence	115
3.7	Application: Computation of the Square Root	119
3.8	The Secant Method: Derivation and Examples	122
3.9	Fixed-Point Iteration	126
3.10	Roots of Polynomials, Part 1	136
3.11	Special Topics in Root-finding Methods	143
3.11.1	Extrapolation and Acceleration	143
3.11.2	Variants of Newton's Method	147
3.11.3	The Secant Method: Theory and Convergence	151
3.11.4	Multiple Roots	155
3.11.5	In Search of Fast Global Convergence: Hybrid Algorithms	159
3.12	Very High-order Methods and the Efficiency Index	165
3.13	Literature and Software Discussion	168
	References	168
4	Interpolation and Approximation	171
4.1	Lagrange Interpolation	171
4.2	Newton Interpolation and Divided Differences	177
4.3	Interpolation Error	187
4.4	Application: Muller's Method and Inverse Quadratic Interpolation	192
4.5	Application: More Approximations to the Derivative	195
4.6	Hermite Interpolation	198
4.7	Piecewise Polynomial Interpolation	202
4.8	An Introduction to Splines	210
4.8.1	Definition of the Problem	210
4.8.2	Cubic B-Splines	211
4.9	Application: Solution of Boundary Value Problems	223
4.10	Tension Splines	228
4.11	Least Squares Concepts in Approximation	234

4.11.1	An Introduction to Data Fitting	234
4.11.2	Least Squares Approximation and Orthogonal Polynomials	237
4.12	Advanced Topics in Interpolation Error	250
4.12.1	Stability of Polynomial Interpolation	250
4.12.2	The Runge Example	253
4.12.3	The Chebyshev Nodes	255
4.13	Literature and Software Discussion	261
	References	262
5	Numerical Integration	263
5.1	A Review of the Definite Integral	264
5.2	Improving the Trapezoid Rule	266
5.3	Simpson's Rule and Degree of Precision	271
5.4	The Midpoint Rule	282
5.5	Application: Stirling's Formula	286
5.6	Gaussian Quadrature	288
5.7	Extrapolation Methods	300
5.8	Special Topics in Numerical Integration	307
5.8.1	Romberg Integration	307
5.8.2	Quadrature with Non-smooth Integrands	312
5.8.3	Adaptive Integration	317
5.8.4	Peano Estimates for the Trapezoid Rule	322
5.9	Literature and Software Discussion	328
	References	328
6	Numerical Methods for Ordinary Differential Equations	329
6.1	The Initial Value Problem: Background	330
6.2	Euler's Method	335
6.3	Analysis of Euler's Method	339
6.4	Variants of Euler's Method	342
6.4.1	The Residual and Truncation Error	344
6.4.2	Implicit Methods and Predictor–Corrector Schemes	347
6.4.3	Starting Values and Multistep Methods	352
6.4.4	The Midpoint Method and Weak Stability	354
6.5	Single-Step Methods: Runge–Kutta	359
6.6	Multistep Methods	366
6.6.1	The Adams Families	366
6.6.2	The BDF Family	370
6.7	Stability Issues	372
6.7.1	Stability Theory for Multistep Methods	372
6.7.2	Stability Regions	376

6.8	Application to Systems of Equations	378
6.8.1	Implementation Issues and Examples	378
6.8.2	Stiff Equations	381
6.8.3	A-Stability	382
6.9	Adaptive Solvers	386
6.10	Boundary Value Problems	399
6.10.1	Simple Difference Methods	399
6.10.2	Shooting Methods	403
6.10.3	Finite Element Methods for BVPs	407
6.11	Literature and Software Discussion	414
	References	415
7	Numerical Methods for the Solution of Systems of Equations	417
7.1	Linear Algebra Review	418
7.2	Linear Systems and Gaussian Elimination	420
7.3	Operation Counts	427
7.4	The LU Factorization	430
7.5	Perturbation, Conditioning, and Stability	441
7.5.1	Vector and Matrix Norms	441
7.5.2	The Condition Number and Perturbations	443
7.5.3	Estimating the Condition Number	450
7.5.4	Iterative Refinement	453
7.6	SPD Matrices and the Cholesky Decomposition	457
7.7	Iterative Methods for Linear Systems: A Brief Survey	460
7.8	Nonlinear Systems: Newton's Method and Related Ideas	469
7.8.1	Newton's Method	469
7.8.2	Fixed-Point Methods	472
7.9	Application: Numerical Solution of Nonlinear Boundary Value Problems	474
7.10	Literature and Software Discussion	477
	References	477
8	Approximate Solution of the Algebraic Eigenvalue Problem	479
8.1	Eigenvalue Review	479
8.2	Reduction to Hessenberg Form	485
8.3	Power Methods	490
8.4	An Overview of the QR Iteration	509
8.5	Application: Roots of Polynomials, Part II	518
8.6	Literature and Software Discussion	519
	References	519

9	A Survey of Numerical Methods for Partial Differential Equations	521
9.1	Difference Methods for the Diffusion Equation	521
9.1.1	The Basic Problem	521
9.1.2	The Explicit Method and Stability	522
9.1.3	Implicit Methods and the Crank–Nicolson Method	527
9.2	Finite Element Methods for the Diffusion Equation	536
9.3	Difference Methods for Poisson Equations	539
9.3.1	Discretization	539
9.3.2	Banded Cholesky Solvers	542
9.3.3	Iteration and the Method of Conjugate Gradients	543
9.4	Literature and Software Discussion	553
	References	553
10	An Introduction to Spectral Methods	555
10.1	Spectral Methods for Two-Point Boundary Value Problems	556
10.2	Spectral Methods for Time-Dependent Problems	568
10.3	Clenshaw–Curtis Quadrature	577
10.4	Literature and Software Discussion	579
	References	579
	Appendix A: Proofs of Selected Theorems, and Additional Material	581
A.1	Proofs of the Interpolation Error Theorems	581
A.2	Proof of the Stability Result for ODEs	583
A.3	Stiff Systems of Differential Equations and Eigenvalues	584
A.4	The Matrix Perturbation Theorem	586
	Index	587

PREFACE

Preface to the Second Edition

This third version of the text is officially the Second Edition, because the second version was officially dubbed the Revised Edition. Now that the confusing explanation is out of the way, we can ask the important question: What is new?

- I continue to chase down typographical errors, a process that reminds me of herding cats. I'd like to thank everyone who has sent me information on this, especially Prof. Mark Mills of Central College in Pella, Iowa. I have become resigned to the notion that a typo-free book is the result of a (*slowly* converging) limiting process, and therefore is unlikely to be actually achieved. But I do keep trying.
- The text now assumes that the student is using MATLAB for computations, and many MATLAB routines are discussed and used in examples. I want to emphasize that this book is still a *mathematics* text, not a primer on how to use MATLAB.
- Several biographies were updated as more complete information has become widely available on the Internet, and a few have been added.
- Two sections, one on adaptive quadrature (§5.8.3) and one on adaptive methods for ODEs (§6.9) have been re-written to reflect the decision to rely more on MATLAB.
- Chapter 9 (A Survey of Numerical Methods for Partial Differential Equations) has been extensively re-written, with more examples and graphics.

- New material has been added:
 - Two sections on roots of polynomials. The first (§3.10) introduces the Durand–Kerner algorithm; the second (§8.5) discusses using the companion matrix to find polynomial roots as matrix eigenvalues.
 - A section (§3.12) on very high-order root-finding methods.
 - A section (§4.10) on splines under tension, also known as “taut splines;”
 - Sections on the finite element method for ODEs (§6.10.3) and some PDEs (§9.2);
 - An entire chapter (Chapter 10) on spectral methods¹.
- Several sections have been modified somewhat to reflect advances in computing technology.
- Later in this preface I devote some time to outlining possible chapter and section selections for different kinds of courses using this text.

It might be appropriate for me to describe how I see the material in the book. Basically, I think it breaks down into three categories:

- The fundamentals: All of Chapters 1 and 2, most of Chapters 3 (3.1, 3.2, 3.3, 3.5, 3.8, 3.9), 4 (4.1, 4.2, 4.3, 4.6, 4.7, 4.8, 4.11), and 5 (5.1, 5.2, 5.3, 5.4, 5.7); this is the basic material in numerical methods and analysis and should be accessible to any well-prepared students who have completed a standard calculus sequence.
- Second level: Most of Chapters 6, 7, and 8, plus much of the remaining sections from Chapters 3 (3.4, 3.6, 3.7, 3.10), 4 (4.4, 4.5), and 5 (5.5, 5.6), and some of 6 (6.8) and 7 (7.7); this is the more advanced material and much of it (from Chap. 6) requires a course in ordinary differential equations or (Chaps. 7 and 8) a course in linear algebra. It is still part of the core of numerical methods and analysis, but it requires more background.
- Advanced: Chapters 9 and 10, plus the few remaining sections from Chapters 3, 4, 5, 6, 7, and 8.
- It should go without saying that precisely *what* is considered “second level” or “advanced” is largely a matter of taste.

As always, I would like to thank my employer, *Mathematical Reviews*, and especially the Executive Editor, Graeme Fairweather, for the study leave that gave me the time to prepare (for the most part) this new edition; my editor at John Wiley & Sons, Susanne Steitz-Filler, who does a good job of putting up with me; an anonymous copy-editor at Wiley who saved me from a large number of self-inflicted wounds; and—most of all—my family of spouse Georgia, daughter Elinor, son James, and Border Collie mutts Samantha

¹The material on spectral methods may well not meet with the approval of experts on the subject, as I presented the material in what appears to be a very non-standard way, and I left out a lot of important issues that make spectral methods, especially for time dependent problems, practical. I did it this way because I wanted to write an introduction to the material that would be accessible to students taking a first course in numerical analysis/methods, and also in order to avoid cluttering up the exposition with what I considered to be “side issues.” I appreciate that these side issues have to be properly treated to make spectral methods practical, but since this tries to be an elementary text, I wanted to keep the exposition as clean as possible.

and Dylan. James was not yet born when I first began writing this text in 1997, and now he has finished his freshman year of high school; Elinor was in first grade at the beginning and graduated from college during the final editing process for this edition. I'm very proud of them both! And I can never repay the many debts that I owe to my dear spouse.

Online Material

There will almost surely be some online material to supplement this text. At a minimum, there will be

- MATLAB files for computing and/or reading Gaussian quadrature (§5.6) weights and abscissas for $N = 2^m$, $m = 0, 1, 2, \dots, 10$.
- Similar material for computing and/or reading Clenshaw-Curtis (§10.3) weights and abscissas.
- Color versions of some plots from Chapter 9.
- It is possible that there will be an entire additional section for Chapter 3.

To access the online material, go to

www.wiley.com/go/epperson2edition

The webpage should be self-explanatory.

A Note About the Dedication

The previous editions were dedicated to six teachers who had a major influence on the author's mathematics education: Frank Crosby and Ed Croteau of New London High School, New London, CT; Prof. Fred Gehring and Prof. Peter Duren of the University of Michigan Department of Mathematics; and Prof. Richard MacCamy and Prof. George J. Fix of the Department of Mathematics, Carnegie-Mellon University, Pittsburgh, PA. (Prof. Fix served as the author's doctoral advisor.) I still feel an unpayable debt of gratitude to these men, who were outstanding teachers, but I felt it appropriate to express my feelings about my parents for this edition, hence the new dedication to the memory of my mother and step-father.

Course Outlines

One can define several courses from this book, based on the level of preparation of the students and the number of terms the course runs, as well as the level of theoretical detail the instructor wishes to address. Here are some example outlines that might be used.

- A single semester course that does not assume any background in linear algebra or differential equations, and which does not emphasize theoretical analysis of methods:

- Chapter 1 (all sections²);
 - Chapter 2 (all sections³);
 - Chapter 3 (Sections 3.1–3.3, 3.8–3.10);
 - Chapter 4 (Sections 4.1–4.8);
 - Chapter 5 (Sections 5.1–5.7).
- A two-semester course which assumes linear algebra and differential equations for the second semester:
 - Chapter 1 (all sections);
 - Chapter 2 (all sections);
 - Chapter 3 (Sections 3.1–3.3, 3.8–3.10);
 - Chapter 4 (Sections 4.1–4.8);
 - Chapter 5 (Sections 5.1–5.7).
 - Semester break should probably come here.
 - Chapter 6 (6.1–6.6; 6.10 if time/preparation permits)
 - Chapter 7 (7.1–7.6)
 - Chapter 8 (8.1–8.4)
 - Additional material at the instructor’s discretion.
 - A two-semester course for well-prepared students:
 - Chapter 1 (all sections);
 - Chapter 2 (all sections);
 - Chapter 3 (Sections 3.1–3.10; 3.11 at the discretion of the instructor);
 - Chapter 4 (Sections 4.1–4.11, 4.12.1, 4.12.3; 4.12.2 at the discretion of the instructor);
 - Chapter 5 (Sections 5.1–5.7, 5.8.1; other sections at the discretion of the instructor).
 - Semester break should probably come here.
 - Chapter 6 (6.1–6.8; 6.10 if time/preparation permits; other sections at the discretion of the instructor)
 - Chapter 7 (7.1–7.8; other sections at the discretion of the instructor)
 - Chapter 8 (8.1–8.4)
 - Additional material at the instructor’s taste and discretion.

Some sections appear to be left out of all these outlines. Most textbooks are written to include extra material, to facilitate those instructors who would like to expose their students to different material, or as background for independent projects, etc.

I want to encourage anyone—teachers, students, random readers—to contact me with questions, comments, suggestions, or remaining typos. My professional email is still jfe@ams.org

²§§1.5 and 1.6 are included in order to expose students to the issue of approximation; if an instructor feels that the students in his or her class do not need this exposure, these sections can be skipped in favor of other material from later chapters.

³The material on ODEs and tridiagonal systems can be taught to students who have not had a normal ODE or linear algebra course.

Computer Access

Because the author no longer has a traditional academic position, his access to modern software is limited. Most of the examples were done using a very old and limited version of MATLAB from 1994. (Some were done on a Sun workstation, using FORTRAN code, in the late 1990s.) The more involved and newer examples were done using public access computers at the University of Michigan's Duderstadt Center, and the author would like to express his appreciation to this great institution for this.

A Note to the Student

(This is slightly updated from the version in the First Edition.) This book was written to be read. I am under no illusions that this book will compete with the latest popular novel for interest or thrilling narrative. But I have tried very hard to write a book on mathematics that can be read by students. So do not simply buy the book, work the exercises, and sell the book back to the bookstore at the end of the term. Read the text, think about what you have read, and ask your instructor questions about the things that you do not understand.

Numerical methods and analysis is a very different area of mathematics, certainly different from what you have seen in your previous courses. It is not harder, but the differentness of the material makes it seem harder. We worry about different issues than those in other mathematics classes. In a calculus course you are typically asked to compute the derivative or antiderivative of a given function, or to solve some equation for a particular unknown. The task is clearly defined, with a very concrete notion of "the right answer." Here, we are concerned with computing approximations, and this involves a slightly different kind of thinking. We have to understand what we are approximating well enough to construct a reasonable approximation, and we have to be able to think clearly and logically enough to analyze the accuracy and performance of that approximation. One former student has characterized this course material as "rigorously imprecise" or "approximately precise." Both are appropriate descriptions. Rote memorization of procedures is not of use here; it is vital in this course that the student learn the underlying concepts. Numerical mathematics is also *experimental* in nature. A lot can be learned simply by trying something out and seeing how the computation goes.

Preface to the Revised Edition

First, I would like to thank John Wiley for letting me do a Revised Edition of *An Introduction to Numerical Methods and Analysis*, and in particular I would like to thank Susanne Steitz and Laurie Rosatone for making it all possible.

So, what's new about this edition? A number of things. For various reasons, a large number of typographical and similar errors managed to creep into the original edition. These have been aggressively weeded out and fixed in this version. I'd like to thank everyone who emailed me with news of this or that error. In particular, I'd like to acknowledge Marzia Rivi, who translated the first edition into Italian and who emailed me with many typos, Prof. Nicholas Higham of Manchester University, Great Britain, and Mark Mills of Central College in Pella, Iowa. I'm sure there's a place or two where I did something silly like reversing the order of subtraction. If anyone finds any error of any sort, please email me at jfe@ams.org.

I considered adding sections on a couple of new topics, but in the end decided to leave the bulk of the text alone. I spent some time improving the exposition and presentation, but most of the text is the same as the first edition, except for fixing the typos.

I would be remiss if I did not acknowledge the support of my employer, the American Mathematical Society, who granted me a study leave so I could finish this project. Executive Director John Ewing and the Executive Editor of Mathematical Reviews, Kevin Clancey, deserve special mention in this regard. Amy Hendrikson of TeXnology helped with some L^AT_EX issues, as did my colleague at Mathematical Reviews, Patrick Ion. Another colleague, Maryse Brouwers, an extraordinary grammarian, helped greatly with the final copyediting process.

The original preface has the URL for the text website wrong; just go to www.wiley.com and use their links to find the book. The original preface also has my old professional email. The updated email is jfe@ams.org; anyone with comments on the text is welcome to contact me.

But, as is always the case, it is the author's immediate family who deserve the most credit for support during the writing of a book. So, here goes a big thank you to my wife, Georgia, and my children, Elinor and Jay. Look at it this way, kids: The end result will pay for a few birthdays.

Preface (To the First Edition)

This book is intended for introductory and advanced courses in numerical methods and numerical analysis, for students majoring in mathematics, sciences, and engineering. The book is appropriate for both single-term survey courses or year-long sequences, where students have a basic understanding of at least single-variable calculus and a programming language. (The usual first courses in linear algebra and differential equations are required for the last four chapters.)

To provide maximum teaching flexibility, each chapter and each section begins with the basic, elementary material and gradually builds up to the more advanced material. This same approach is followed with the underlying theory of the methods. Accordingly, one can use the text for a "methods" course that eschews mathematical analysis, simply by not covering the sections that focus on the theoretical material. Or, one can use the text for a survey course by only covering the basic sections, or the extra topics can be covered if you have the luxury of a full year course.

The objective of the text is for students to learn where approximation methods come from, why they work, why they sometimes don't work, and when to use which of many techniques that are available, and to do all this in a style that emphasizes readability and usefulness to the beginning student. While these goals are shared by other texts, it is the development and delivery of the ideas in this text that I think makes it different.

A course in numerical computation—whether it emphasizes the theory or the methods—requires that students think quite differently than in other mathematics courses, yet students are often not experienced in the kind of problem-solving skills and mathematical judgment that a numerical course requires. Many students react to mathematics problems by pigeon-holing them by category, with little thought given to the meaning of the answer. Numerical mathematics demands much more judgment and evaluation in light of the underlying theory, and in the first several weeks of the course it is crucial for students to adapt their way of thinking about and working with these ideas, in order to succeed in the course.

To enable students to attain the appropriate level of mathematical sophistication, this text begins with a review of the important calculus results, and why and where these ideas play an important role in this course. Some of the concepts required for the study of computational mathematics are introduced, and simple approximations using Taylor's theorem are treated in some depth, in order to acquaint students with one of the most common and basic tools in the science of approximation. Computer arithmetic is treated in perhaps more detail than some might think necessary, but it is instructive for many students to see the actual basis for rounding error demonstrated in detail, at least once.

One important element of this text that I have not seen in other texts is the emphasis that is placed on "cause and effect" in numerical mathematics. For example, if we apply the trapezoid rule to (approximately) integrate a function, then the error should go down by a factor of 4 as the mesh decreases by a factor of 2; if this is not what happens, then almost surely there is either an error in the code or the integrand is not sufficiently smooth. While this is obvious to experienced practitioners in the field, it is not obvious to beginning students who are not confident of their mathematical abilities. Many of the exercises and examples are designed to explore this kind of issue.

Two common starting points to the course are root-finding or linear systems, but diving in to the treatment of these ideas often leaves the students confused and wondering what the point of the course is. Instead, this text provides a second chapter designed as a "toolbox" of elementary ideas from across several problem areas; it is one of the important innovations of the text. The goal of the toolbox is to acclimate the students to the culture of numerical methods and analysis, and to show them a variety of simple ideas before proceeding to cover any single topic in depth. It develops some elementary approximations and methods that the students can easily appreciate and understand, and introduces the students, in the context of very simple methods and problems, to the essence of the analytical and coding issues that dominate the course. At the same time, the early development of these tools allows them to be used later in the text in order to derive and explain some algorithms in more detail than is usually the case.

The style of exposition is intended to be more lively and "student friendly" than the average mathematics text. This does not mean that there are no theorems stated and proved correctly, but it does mean that the text is not slavish about it. There is a reason for this: The book is meant to be read by the students. The instructor can render more formal anything in the text that he or she wishes, but if the students do not read the text because they are turned off by an overly dry regimen of definition, theorem, proof, corollary, then all of our effort is for naught. In places, the exposition may seem a bit wordier than necessary, and there is a significant amount of repetition. Both are deliberate. While brevity is indeed better mathematical style, it is not necessarily better pedagogy. Mathematical textbook exposition often suffers from an excess of brevity, with the result that the students cannot follow the arguments as presented in the text. Similarly, repetition aids learning, by reinforcement.

Nonetheless I have tried to make the text mathematically complete. Those who wish to teach a lower-level survey course can skip proofs of many of the more technical results in order to concentrate on the approximations themselves. An effort has been made—not always successfully—to avoid making basic material in one section depend on advanced material from an earlier section.

The topics selected for inclusion are fairly standard, but not encyclopedic. Emerging areas of numerical analysis, such as wavelets, are not (in the author's opinion) appropriate for a first course in the subject. The same reasoning dictated the exclusion of other, more mature areas, such as the finite element method, although that might change in future editions should there be sufficient demand for it. A more detailed treatment of

approximation theory, one of the author's favorite topics, was also felt to be poorly suited to a beginning text. It was felt that a better text would be had by doing a good job covering some of the basic ideas, rather than trying to cover everything in the subject.

The text is not specific to any one computing language. Most illustrations of code are made in an informal pseudo-code, while more involved algorithms are shown in a "macro-outline" form, and programming hints and suggestions are scattered throughout the text. The exercises assume that the students have easy access to and working knowledge of software for producing basic Cartesian graphs.

A diskette of programs is *not* provided with the text, a practice that sets this book at odds with many others, but which reflects the author's opinion that students must learn how to write and debug programs that implement the algorithms in order to learn the underlying mathematics. However, since some faculty and some departments structure their courses differently, a collection of program segments in a variety of languages is available on the text web site so that instructors can easily download and then distribute the code to their students. Instructors and students should be aware that these are program *segments*; none of them are intended to be ready-to-run complete programs. Other features of the text web site are discussed below. (*Note:* This material may be removed from the Revised Edition website.)

Exercises run the gamut from simple hand computations that might be characterized as "starter exercises" to challenging derivations and minor proofs to programming exercises designed to test whether or not the students have assimilated the important ideas of each chapter and section. Some of the exercises are taken from application situations, some are more traditionally focused on the mathematical issues for their own sake. Each chapter concludes with a brief section discussing existing software and other references for the topic at hand, and a discussion of material not covered in this text.

Historical notes are scattered throughout the text, with most named mathematicians being accorded at least a paragraph or two of biography when they are first mentioned. This not only indulges my interest in the history of mathematics, but it also serves to engage the interest of the students.

The web site for the text (<http://www.wiley.com/epperson>) will contain, in addition to the set of code segments mentioned above, a collection of additional exercises for the text, some application modules demonstrating some more involved and more realistic applications of some of the material in the text, and, of course, information about any updates that are going to be made in future editions. Colleagues who wish to submit exercises or make comments about the text are invited to do so by contacting the author at epperson@math.uah.edu.

Notation

Most notation is defined as it appears in the text, but here we include some commonplace items.

$\mathbb{R}$ — The real number line; $\mathbb{R} = (-\infty, \infty)$.

$\mathbb{R}^n$ — The vector space of real vectors of n components.

$\mathbb{R}^{n \times n}$ — The vector space of real $n \times n$ matrices.

$C([a, b])$ — The set of functions f which are defined on the interval $[a, b]$, continuous on all of (a, b) , and continuous from the interior of $[a, b]$ at the endpoints.

- $C^k([a, b])$ — The set of functions f such that f and its first k derivatives are all in $C([a, b])$.
- $C^{p,q}(Q)$ — The set of all functions u that are defined on the two-dimensional domain $Q = \{(x, t) \mid a \leq x \leq b, 0 < t \leq T\}$, and that are p times continuously differentiable in x for all t , and q times continuously differentiable in t for all x .
- $\approx$ — Approximately equal. When we say that $A \approx B$, we mean that A and B are approximations to each other. See §1.2.2.
- $\equiv$ — Equivalent. When we say that $f(x) = g(x)$, we mean that the two functions agree at the single point x . When we say that $f(x) \equiv g(x)$, we mean that they agree at all points x . The same thing is said by using just the function names, i.e., $f = g$.
- $\mathcal{O}$ — On the order of (“big O of”). We say that $A = B + \mathcal{O}(D(h))$ whenever $|A - B| \leq CD(h)$ for some constant C and for all h sufficiently small. See §1.2.3.
- $\mathbf{u}$ — Machine epsilon. The largest number such that, in computer arithmetic, $1 + \mathbf{u} = 1$. Architecture dependent, of course. See §1.3.
- sgn — Sign function. The value of $\text{sgn}(x)$ is 1, -1 , or 0, depending on whether or not x is positive, negative, or zero, respectively.

CHAPTER 1

INTRODUCTORY CONCEPTS AND CALCULUS REVIEW

It is best to start this book with a question: What do we mean by “Numerical Methods and Analysis”? What kind of mathematics is this book about?

Generally and broadly speaking, this book covers the mathematics and methodologies that underlie the techniques of *scientific computation*. More prosaically, consider the button on your calculator that computes the sine of the number in the display. Exactly how does the calculator know that correct value? When we speak of using the computer to solve a complicated mathematics or engineering problem, exactly what is involved in making that happen? Are computers “born” with the knowledge of how to solve complicated mathematical and engineering problems? No, of course they are not. Mostly they are *programmed* to do it, and the programs implement algorithms that are based on the kinds of things we talk about in this book.

Textbooks and courses in this area generally follow one of two main themes: Those titled “Numerical Methods” tend to emphasize the implementation of the algorithms, perhaps at the expense of the underlying mathematical theory that explains why the methods work; those titled “Numerical Analysis” tend to emphasize this underlying mathematical theory, perhaps at the expense of some of the implementation issues. The best approach, of course, is to properly mix the study of the algorithms and their implementation (“methods”) with the study of the mathematical theory (“analysis”) that supports them. This is our goal in this book.

Whenever someone speaks of using a computer to design an airplane, predict the weather, or otherwise solve a complex science or engineering problem, that person is talking about using numerical methods and analysis. The problems and areas of endeavor that use these

kinds of techniques are continually expanding. For example, *computational mathematics*—another name for the material that we consider here—is now commonly used in the study of financial markets and investment structures, an area of study that does not ordinarily come to mind when we think of “scientific” computation. Similarly, the increasingly frequent use of computer-generated animation in film production is based on a heavy dose of spline approximations, which we introduce in §4.8. And modern weather prediction is based on using numerical methods and analysis to solve the very complicated equations governing fluid flow and heat transfer between and within the atmosphere, oceans, and ground.

There are a number of different ways to break the subject down into component parts. We will discuss the derivation and implementation of the algorithms, and we will also analyze the algorithms, mathematically, in order to learn how best to use them and how best to implement them. In our study of each technique, we will usually be concerned with two issues that often are in competition with each other:

- *Accuracy*: Very few of our computations will yield the exact answer to a problem, so we will have to understand how much error is made, and how to control (or even diminish) that error.
- *Efficiency*: Does the algorithm take an inordinate amount of computer time? This might seem to be an odd question to concern ourselves with—after all, computers are fast, right?—but there are slow ways to do things and fast ways to do things. All else being equal (it rarely is), we prefer the fast ways.

We say that these two issues compete with each other because, generally speaking, the steps that can be taken to make an algorithm more accurate usually make it more costly, that is, less efficient.

There is a third issue of importance, but it does not become as evident as the others (although it is still present) until Chapter 6:

- *Stability*: Does the method produce similar results for similar data? If we change the data by a small amount, do we get vastly different results? If so, we say that the method is unstable, and unstable methods tend to produce unreliable results. It is entirely possible to have an accurate method that is efficiently implemented, yet is horribly unstable; see §6.4.4 for an example of this.

1.1 BASIC TOOLS OF CALCULUS

1.1.1 Taylor’s Theorem

Computational mathematics does not require a large amount of background, but it does require a good knowledge of that background. The most important single result in numerical computations, from all of the calculus, is Taylor’s Theorem,¹ which we now state:

¹Brook Taylor (1685–1731) was educated at St. John’s College of Cambridge University, entering in 1701 and graduating in 1709. He published what we know as Taylor’s Theorem in 1715, although it appears that he did not entirely appreciate its larger importance and he certainly did not bother with a formal proof. He was elected a member of the prestigious Royal Society of London in 1712.

Taylor acknowledged that his work was based on that of Newton and Kepler and others, but he did not acknowledge that the same result had been discovered by Johann Bernoulli and published in 1694. (But then Taylor discovered integration by parts first, although Bernoulli claimed the credit.)

Theorem 1.1 (Taylor's Theorem with Remainder) Let $f(x)$ have $n+1$ continuous derivatives on $[a, b]$ for some $n \geq 0$, and let $x, x_0 \in [a, b]$. Then,

$$f(x) = p_n(x) + R_n(x)$$

for

$$p_n(x) = \sum_{k=0}^n \frac{(x-x_0)^k}{k!} f^{(k)}(x_0), \quad (1.1)$$

and

$$R_n(x) = \frac{1}{n!} \int_{x_0}^x (x-t)^n f^{(n+1)}(t) dt. \quad (1.2)$$

Moreover, there exists a point ξ_x between x and x_0 such that

$$R_n(x) = \frac{(x-x_0)^{n+1}}{(n+1)!} f^{(n+1)}(\xi_x). \quad (1.3)$$

The point x_0 is usually chosen at the discretion of the user, and is often taken to be 0. Note that the two forms of the remainder are equivalent: the “pointwise” form (1.3) can be derived from the “integral” form (1.2); see Problem 23.

Taylor's Theorem is important because it allows us to represent, exactly, fairly general functions in terms of polynomials with a *known, specified, boundable* error. This allows us to replace, in a computational setting, these same general functions with something that is much simpler—a polynomial—yet at the same time we are able to bound the error that is made. No other tool will be as important to us as Taylor's Theorem, so it is worth spending some time on it here at the outset.

The usual calculus treatment of Taylor's Theorem should leave the student familiar with three particular expansions (for all three of these we have used $x_0 = 0$, which means we really should call them Maclaurin² series, but we won't):

$$\begin{aligned} e^x &= 1 + x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \dots + \frac{1}{n!}x^n + \frac{1}{(n+1)!}x^{n+1}e^{\xi_x} \\ &= \sum_{k=0}^n \frac{1}{k!}x^k + R_n(x), \\ \sin x &= x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 + \dots + \frac{(-1)^n}{(2n+1)!}x^{2n+1} + \frac{(-1)^{n+1}}{(2n+3)!}x^{2n+3} \cos \xi_x \\ &= \sum_{k=0}^n \frac{(-1)^k}{(2k+1)!}x^{2k+1} + R_n(x), \\ \cos x &= 1 - \frac{1}{2!}x^2 + \frac{1}{4!}x^4 + \dots + \frac{(-1)^n}{(2n)!}x^{2n} + \frac{(-1)^{n+1}}{(2n+2)!}x^{2n+2} \cos \xi_x \\ &= \sum_{k=0}^n \frac{(-1)^k}{(2k)!}x^{2k} + R_n(x). \end{aligned}$$

²Colin Maclaurin (1698–1746) was born and lived almost his entire life in Scotland. Educated at Glasgow University, he was professor of mathematics at Aberdeen from 1717 to 1725 and then went to Edinburgh. He worked in a number of areas of mathematics, and is credited with writing one of the first textbooks based on Newton's calculus, *Treatise of fluxions* (1742). The Maclaurin series appears in this book as a special case of Taylor's series.

(Strictly speaking, the indices on the last two remainders should be $2n + 1$ and $2n$, because those are the exponents in the last terms of the expansion, but it is commonplace to present them as we did here.) In fact, Taylor's Theorem provides us with our first and simplest example of an approximation and an error estimate. Consider the problem of approximating the exponential function on the interval $[-1, 1]$. Taylor's Theorem tells us that we can represent e^x using a polynomial with a (known) remainder:

$$e^x = \underbrace{1 + x + \frac{1}{2!}x^2 + \frac{1}{3!}x^3 + \dots + \frac{1}{n!}x^n}_{p_n(x), \text{ polynomial}} + \underbrace{\frac{1}{(n+1)!}x^{n+1}e^{c_x}}_{R_n(x), \text{ remainder}},$$

where c_x is an unknown point between x and 0. Since we want to consider the most general case, where x can be any point in $[-1, 1]$, we have to consider that c_x can be any point in $[-1, 1]$, as well. For simplicity, let's denote the polynomial by $p_n(x)$, and the remainder by $R_n(x)$, so that the equation above becomes

$$e^x = p_n(x) + R_n(x).$$

Suppose that we want this approximation to be accurate to within 10^{-6} in absolute error, i.e., we want

$$|e^x - p_n(x)| \leq 10^{-6}$$

for *all* x in the interval $[-1, 1]$. Note that if we can make $|R_n(x)| \leq 10^{-6}$ for all $x \in [-1, 1]$, then we will have

$$|e^x - p_n(x)| = |R_n(x)| \leq 10^{-6},$$

so that the error in the approximation will be less than 10^{-6} . The best way to proceed is to create a *simple* upper bound for $|R_n(x)|$, and then use that to determine the number of terms necessary to make this upper bound less than 10^{-6} .

Thus we proceed as follows:

$$\begin{aligned} |R_n(x)| &= \frac{|x^{n+1}e^{c_x}|}{(n+1)!}, \\ &= \frac{|x|^{n+1}e^{c_x}}{(n+1)!}, \quad \text{because } e^z > 0 \text{ for all } z, \\ &\leq \frac{e^{c_x}}{(n+1)!}, \quad \text{because } |x| \leq 1 \text{ for all } x \in [-1, 1], \\ &\leq \frac{e}{(n+1)!}, \quad \text{because } e^{c_x} \leq e \text{ for all } x \in [-1, 1]. \end{aligned}$$

Thus, if we find n such that

$$\frac{1}{(n+1)!}e \leq 10^{-6},$$

then we will have

$$|e^x - p_n(x)| = |R_n(x)| \leq \frac{1}{(n+1)!}e \leq 10^{-6}$$

and we will *know* that the error is less than the desired tolerance, for *all* values of x of interest to us, i.e., all $x \in [-1, 1]$. A little bit of exercise with a calculator shows us that we need to use $n = 9$ to get the desired accuracy. Figure 1.1 shows a plot of the

exponential function e^x , the approximation $p_9(x)$, as well as the less accurate $p_2(x)$; since it is impossible to distinguish by eye between the plots for e^x and $p_9(x)$, we have also provided Figure 1.2, which is a plot of the error $e^x - p_9(x)$; note that we begin to lose accuracy as we get away from the interval $[-1, 1]$. This is not surprising. Since the Taylor polynomial is constructed to match f and its first n derivatives at $x = x_0$, it ought to be the case that p_n is a good approximation to f only when x is near x_0 .

Here we are in the first section of the text, and we have already constructed our first approximation and proved our first theoretical result. Theoretical result? Where? Why, the error estimate, of course. The material in the previous several lines is a proof of Proposition 1.1.

Proposition 1.1 *Let $p_9(x)$ be the Taylor polynomial of degree 9 for the exponential function. Then, for all $x \in [-1, 1]$, the error in the approximation of $p_9(x)$ to e^x is less than 10^{-6} , i.e.,*

$$|e^x - p_9(x)| \leq 10^{-6}$$

for all $x \in [-1, 1]$.

Although this result is not of major significance to our work—the exponential function is approximated more efficiently by different means—it does illustrate one of the most important aspects of the subject. The result tells us, ahead of time, that we can approximate the exponential function to within 10^{-6} accuracy using a specific polynomial, and this accuracy holds for all x in a specified interval. That is the kind of thing we will be doing throughout the text—constructing approximations to difficult computations that are accurate in some sense, and we *know* how accurate they are.

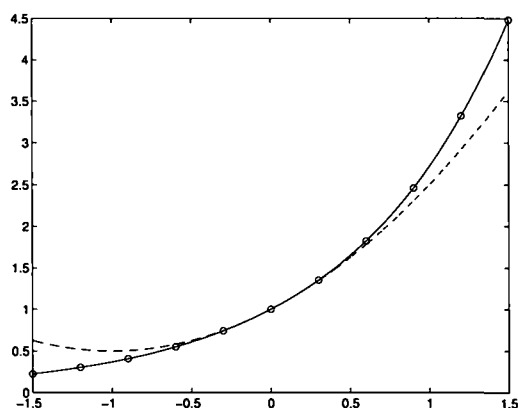


Figure 1.1 Taylor approximation: e^x (solid line), $p_9(x) \approx e^x$ (circles), and $p_2(x) \approx e^x$ (dashed line). Note that e^x and $p_9(x)$ are indistinguishable on this plot.

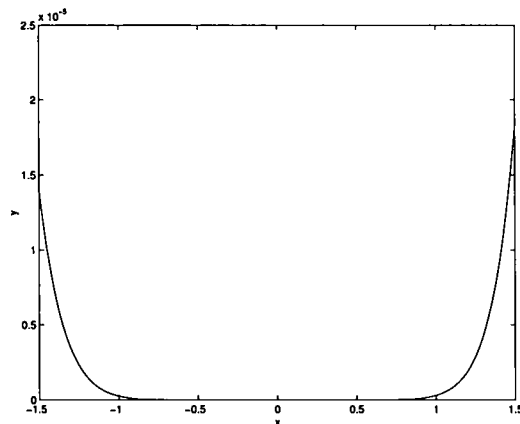


Figure 1.2 Error in Taylor approximation: $e^x - p_9(x)$.

■ EXAMPLE 1.1

Let $f(x) = \sqrt{x+1}$; then the second-order Taylor polynomial (computed about $x_0 = 0$) is computed as follows:

$$f(x_0) = f(0) = 1;$$

$$f'(x) = \frac{1}{2}(x+1)^{-1/2} \Rightarrow f'(x_0) = \frac{1}{2};$$

$$f''(x) = -\frac{1}{2} \times \frac{1}{2}(x+1)^{-3/2} \Rightarrow f''(x_0) = -\frac{1}{4};$$

$$p_2(x) = f(x_0) + (x - x_0)f'(x_0) + \frac{1}{2}(x - x_0)^2 f''(x_0) = 1 + \frac{1}{2}x - \frac{1}{8}x^2.$$

The error in using p_2 to approximate $\sqrt{x+1}$ is given by $R_2(x) = \frac{1}{3!}(x-x_0)^3 f'''(\xi_x)$, where ξ_x is between x and x_0 . We can simplify the error as follows:

$$\begin{aligned} |R_2(x)| &= \left| \frac{1}{3!}(x-x_0)^3 f'''(\xi_x) \right|, \\ &= \frac{1}{6}|x|^3 \left| \frac{3}{2} \times \frac{1}{2} \times \frac{1}{2}(\xi_x + 1)^{-5/2} \right|, \\ &= \frac{1}{16}|x|^3 |\xi_x + 1|^{-5/2}. \end{aligned}$$

If we want to consider this for all $x \in [0, 1]$, then, we observe that $x \in [0, 1]$ and ξ_x between x and 0 imply that $\xi_x \in [0, 1]$; therefore

$$|\xi_x + 1|^{-5/2} \leq |0 + 1|^{-5/2} = 1,$$

so that the upper bound on the error becomes $|R_3(x)| \leq 1/16 = 0.0625$, for all $x \in [0, 1]$. If we are only interested in $x \in [0, \frac{1}{2}]$, then the error is much smaller:

$$\begin{aligned} |R_2(x)| &= \frac{1}{16}|x|^3 |\xi_x + 1|^{-5/2}, \\ &\leq \frac{1}{16}(1/2)^3, \\ &= \frac{1}{128}, \\ &= 0.0078125. \end{aligned}$$

■ EXAMPLE 1.2

Consider the problem of finding a polynomial approximation to the function $f(x) = \sin \pi x$ that is accurate to within 10^{-4} for all $x \in [-\frac{1}{2}, \frac{1}{2}]$, using $x_0 = 0$. A direct computation with Taylor's Theorem gives us (note that the indexing here is a little different, because we want to take advantage of the fact that the Taylor polynomial here has only odd terms)

$$p_n(x) = \pi x - \frac{1}{6}\pi^3 x^3 + \frac{1}{120}\pi^5 x^5 + \cdots + (-1)^n \frac{1}{(2n+1)!} \pi^{2n+1} x^{2n+1},$$