A Course in Approximation Theory

Ward Cheney
Will Light

Graduate Studies in Mathematics
Volume 101

American Mathematical Society

A Course in Approximation Theory

Ward Cheney
Will Light

Graduate Studies in Mathematics
Volume 101

American Mathematical Society
Providence, Rhode Island

2000 *Mathematics Subject Classification.* Primary 41–01.

For additional information and updates on this book, visit
www.ams.org/bookpages/gsm-101

Library of Congress Cataloging-in-Publication Data

Cheney, E. W. (Elliott Ward), 1929–
A course in approximation theory / Ward Cheney, Will Light.
p. cm. — (Graduate studies in mathematics ; v. 101)
Originally published: Pacific Grove : Brooks/Cole Pub. Co., c2000.
Includes bibliographical references and index.
ISBN 978-0-8218-4798-5 (alk. paper)
1. Approximation theory—Textbooks. I. Light, W. A. (William Allan), 1950– II. Title.

QA221.C44 2009
511′.4—dc22 2008047417

∞ The paper used in this book is acid-free and falls within the guidelines
established to ensure permanence and durability.
Visit the AMS home page at http://www.ams.org/

10 9 8 7 6 5 4 3 2 1 14 13 12 11 10 09

A Course in Approximation Theory

Titles in This Series

102 **Mark A. Pinsky,** Introduction to Fourier analysis and wavelets, 2009
101 **Ward Cheney and Will Light,** A course in approximation theory, 2009
100 **I. Martin Isaacs,** Algebra: A graduate course, 2009
99 **Gerald Teschl,** Mathematical methods in quantum mechanics: With applications to Schrödinger operators, 2009
98 **Alexander I. Bobenko and Yuri B. Suris,** Discrete differential geometry: Integrable structure, 2008
97 **David C. Ullrich,** Complex made simple, 2008
96 **N. V. Krylov,** Lectures on elliptic and parabolic equations in Sobolev spaces, 2008
95 **Leon A. Takhtajan,** Quantum mechanics for mathematicians, 2008
94 **James E. Humphreys,** Representations of semisimple Lie algebras in the BGG category $\mathcal{O}$, 2008
93 **Peter W. Michor,** Topics in differential geometry, 2008
92 **I. Martin Isaacs,** Finite group theory, 2008
91 **Louis Halle Rowen,** Graduate algebra: Noncommutative view, 2008
90 **Larry J. Gerstein,** Basic quadratic forms, 2008
89 **Anthony Bonato,** A course on the web graph, 2008
88 **Nathanial P. Brown and Narutaka Ozawa,** C^*-algebras and finite-dimensional approximations, 2008
87 **Srikanth B. Iyengar, Graham J. Leuschke, Anton Leykin, Claudia Miller, Ezra Miller, Anurag K. Singh, and Uli Walther,** Twenty-four hours of local cohomology, 2007
86 **Yulij Ilyashenko and Sergei Yakovenko,** Lectures on analytic differential equations, 2007
85 **John M. Alongi and Gail S. Nelson,** Recurrence and topology, 2007
84 **Charalambos D. Aliprantis and Rabee Tourky,** Cones and duality, 2007
83 **Wolfgang Ebeling,** Functions of several complex variables and their singularities (translated by Philip G. Spain), 2007
82 **Serge Alinhac and Patrick Gérard,** Pseudo-differential operators and the Nash–Moser theorem (translated by Stephen S. Wilson), 2007
81 **V. V. Prasolov,** Elements of homology theory, 2007
80 **Davar Khoshnevisan,** Probability, 2007
79 **William Stein,** Modular forms, a computational approach (with an appendix by Paul E. Gunnells), 2007
78 **Harry Dym,** Linear algebra in action, 2007
77 **Bennett Chow, Peng Lu, and Lei Ni,** Hamilton's Ricci flow, 2006
76 **Michael E. Taylor,** Measure theory and integration, 2006
75 **Peter D. Miller,** Applied asymptotic analysis, 2006
74 **V. V. Prasolov,** Elements of combinatorial and differential topology, 2006
73 **Louis Halle Rowen,** Graduate algebra: Commutative view, 2006
72 **R. J. Williams,** Introduction the the mathematics of finance, 2006
71 **S. P. Novikov and I. A. Taimanov,** Modern geometric structures and fields, 2006
70 **Seán Dineen,** Probability theory in finance, 2005
69 **Sebastián Montiel and Antonio Ros,** Curves and surfaces, 2005

For a complete list of titles in this series, visit the
AMS Bookstore at **www.ams.org/bookstore/**.

To our wives, Victoria and Anita.

Preface

This book offers a graduate-level exposition of selected topics in modern approximation theory. A large portion of the book focuses on multivariate approximation theory, where much recent research is concentrated. Although our own interests have influenced the choice of topics, the text cuts a wide swath through modern approximation theory, as can be seen from the table of contents. We believe the book will be found suitable as a text for courses, seminars, and even solo study. Although the book is at the graduate level, it does not presuppose that the reader already has taken a course in approximation theory.

Topics of This Book

A central theme of the book is the problem of interpolating data by smooth multivariable functions. Several chapters investigate interesting families of functions that can be employed in this task; among them are the polynomials, the positive definite functions, and the radial basis functions. Whether these same families can be used, in general, for approximating functions to arbitrary precision is a natural question that follows; it is addressed in further chapters.

The book then moves on to the consideration of methods for concocting approximations, such as by convolutions, by neural nets, or by interpolation at more and more points. Here there are questions of limiting behavior of sequences of operators, just as there are questions about interpolating on larger and larger sets of nodes.

A major departure from our theme of multivariate approximation is found in the two chapters on univariate wavelets, which comprise a significant fraction of the book. In our opinion wavelet theory is so important a development in recent times—and is so mathematically appealing—that we had to devote some space to expounding its basic principles.

The Style of This Book

In style, we have tried to make the exposition as simple and clear as possible, electing to furnish proofs that are complete and relatively easy to read without the reader needing to resort to pencil and paper. Any reader who finds this style too prolix can proceed quickly over arguments and calculations that are routine. To paraphrase Shaw: We have done our best to avoid conciseness! We have also made considerable efforts to find simple ways to introduce and explain each topic. We hope that in doing so, we encourage readers to delve deeper into some areas. It should be borne in mind that further exploration of some topics may require more mathematical sophistication than is demanded by our treatment.

Organization of the Book

A word about the general plan of the book: we start with relatively elementary matters in a series of about ten short chapters that do not, in general, require more of the reader than undergraduate mathematics (in the American university system). From that point on, the gradient gradually increases and the text becomes more demanding, although still largely self-contained. Perhaps the most significant demands made on the technical knowledge of the reader fall in the areas of measure theory and the Fourier transform. We have freely made use of the Lebesgue function spaces, which bring into play such measure-theoretic results as the Fubini Theorem. Other results such as the Riesz Representation Theorem for bounded linear functionals on a space of continuous functions and the Plancherel Theorem for Fourier transforms also are employed without compunction; but we have been careful to indicate explicitly how these ideas come into play. Consequently, the reader can simply accept the claims about such matters as they arise. Since these theorems form a vital part of the equipment of any applied analyst, we are confident that readers will want to understand for themselves the essentials of these areas of mathematics. We recommend Rudin's *Real and Complex Analysis* (McGraw-Hill, 1974) as a suitable source for acquiring the necessary measure theoretic ideas, and the book *Functional Analysis* (McGraw-Hill, 1973) by the same author as a good introduction to the circle of ideas connected with the Fourier transform.

Additional Reading

We call the reader's attention to the list of books on approximation theory that immediately precedes the main section of references in the bibliography. These books, in general, are concerned with what we may term the "classical" portion of approximation theory—understood to mean the parts of the subject that already were in place when the authors were students. As there are very few textbooks covering recent theory, our book should help to fill that "much needed gap," as some wag phrased it years ago. This list of books emphasizes only the systematic textbooks for the subject as a whole, not the specialized texts and monographs.

Acknowledgments

It is a pleasure to have this opportunity of thanking three agencies that supported our research over the years when this book was being written: the Division of Scientific Affairs of NATO, the Deutsche Forschungsgemeinschaft, and the Science and Engineering Research Council of Great Britain. For their helpful reviews of our manuscript, we thank Robert Schaback, University of Göttingen, and Larry Schumaker, Vanderbilt University. We acknowledge also the contribution made by many students, who patiently listened to us expound the material contained in this book and who raised incisive questions. Students and colleagues in Austin, Leicester, Würzburg, Singapore, and Canterbury (NZ) all deserve our thanks. Professor S. L. Lee was especially helpful.

A special word of thanks goes to Ms. Margaret Combs of the University of Texas Mathematics Department. She is a superb technical typesetter in the modern sense of the word, that is, an expert in TEX. She patiently created the TEX files for lecture notes, starting about six years ago, and kept up with the constant editing of these notes, which were to become the backbone of the book.

The staff of Brooks/Cole Publishing has been most helpful and professional in guiding this book to its publication. In particular, we thank Gary Ostedt, sponsoring editor; Ragu Raghavan, marketing representative; and Janet Hill, production editor, for their personal contact with us during this project.

How to Reach Us

Readers are encouraged to bring errors and suggestions to our attention. E-mail is excellent for this purpose: our addresses are cheney@math.utexas.edu and pwl@mcs.le.ac.uk. A web site for the book is maintained at http:www.math.utexas.edu/user/cheney/ATBOOK.

Ward Cheney

Will Light

A Course in

Approximation Theory

Contents

1

Introductory Discussion of Interpolation

We shall be concerned with real-valued functions defined on a domain X, which need not be specified at this moment. (It will often be a subset of $\mathbb{R}$, $\mathbb{R}^2$, ..., but can be more general.) In the domain X a set of n distinct points is given:

$$\mathcal{N} = \{x_1, x_2, \dots, x_n\}$$

These points are called **nodes**, and $\mathcal{N}$ is the **node set**. For each node x_i an **ordinate** $\lambda_i \in \mathbb{R}$ is given. (Each λ_i is a real number.) The problem of **interpolation** is to find a suitable function $F : X \to \mathbb{R}$ that takes these prescribed n values. That is, we want

$$F(x_i) = \lambda_i \qquad (1 \le i \le n)$$

When this occurs, we say that F **interpolates** the given data $\{(x_i, \lambda_i)\}_{i=1}^n$. Usually F must be chosen from a preassigned family of functions on X.

A wide variety of functions F may be suitable. Figures 1.1 and 1.2 show 12 different interpolation functions for a single data set. The nodes are 5 real numbers. They and the specified ordinates are given in this table:

x	1.2	3.6	4.3	6.1	7.8
y	1.2	3.1	−1.3	2.7	1.4

In Figure 1.1a, the raw data points are shown. In Figures 1.1b to 1.1f, F has the form $F(x) = \sum_1^5 c_j u(x - x_j)$, in which u is a function of our choosing. First we took a B-spline of degree 0. To avoid the discontinuous nature of this example, we then took u to be a B-spline of degree 1, as shown in Figure 1.1c. To avoid discontinuities in the first two derivatives, we then let u be a cubic B-spline, as in Figure 1.1d. In Figure 1.1e we show the interpolant when $u(x) = |x|$, and in Figure 1.1f we used $u(x) = |x|^{1/2}$.

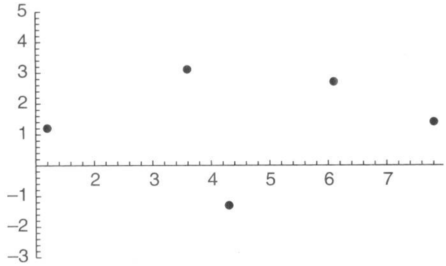

Figure 1.1a

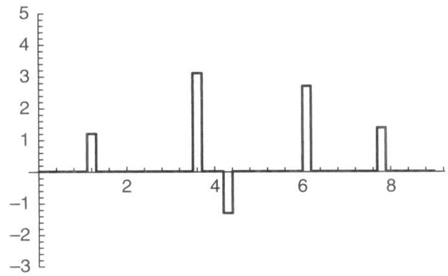

Figure 1.1b

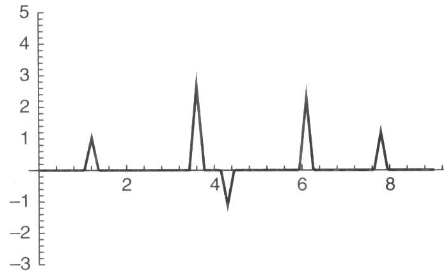

Figure 1.1c

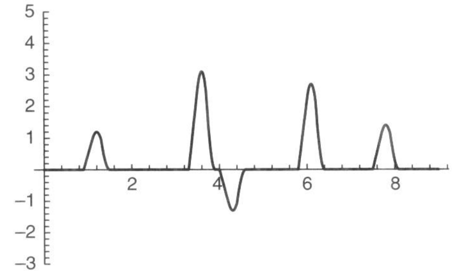

Figure 1.1d

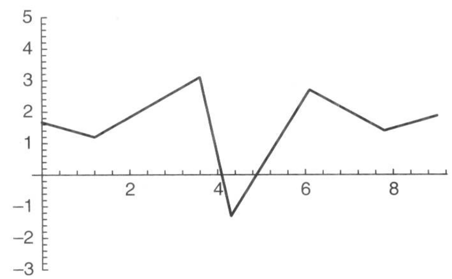

Figure 1.1e

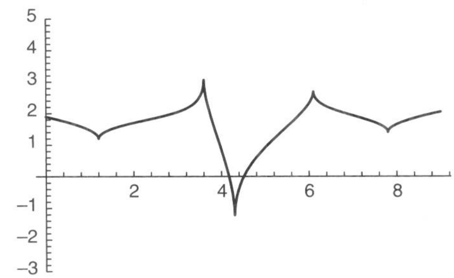

Figure 1.1f

Further examples are shown in Figure 1.2. Here we have used the same data as in Figure 1.1, but a different choice of interpolating functions. Specifically, 1.2a employs a fourth-degree polynomial; 1.2b employs a natural cubic spline; 1.2c is given by the Interpolation command in Mathematica and is also a cubic spline. In 1.2d, we used a cubic *B*-spline, B^3, determined by integer knots, and interpolated with $\sum_1^5 c_i B^3(x - x_i)$. In 1.2e, we used $\sum c_i e^{-(x-x_i)^2}$, and in 1.2f we used, in the same manner, a 0-degree *B*-spline. Some variations in scaling are noticeable in the figures.

The examples in Figures 1.1 and 1.2 suggest the great diversity among different types of interpolating functions. The selection of an appropriate type of interpolant must be made according to further criteria, above and beyond the basic requirement of inter-

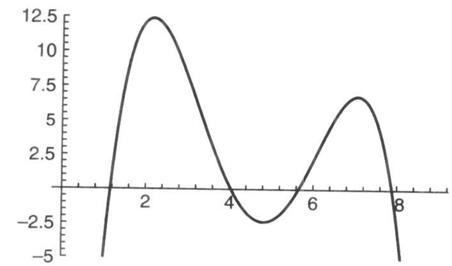

Figure 1.2a

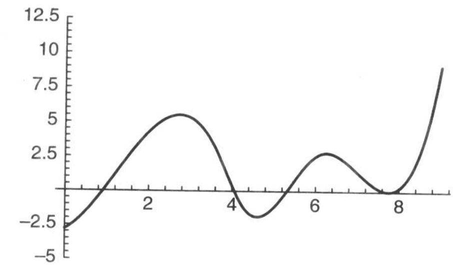

Figure 1.2b

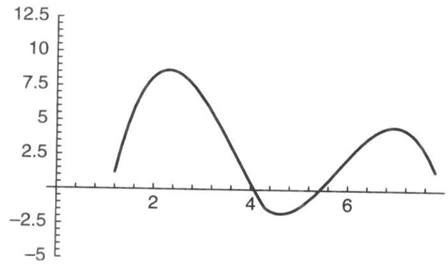

Figure 1.2c

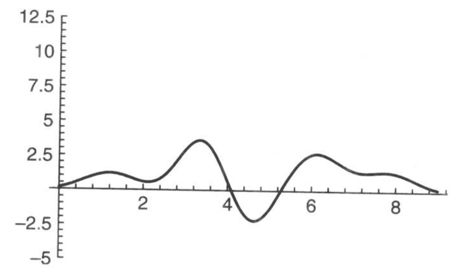

Figure 1.2d

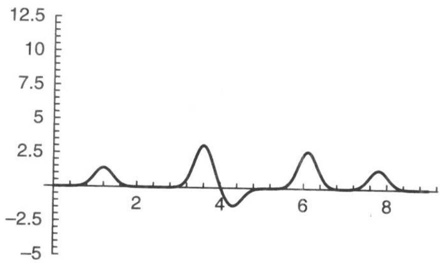

Figure 1.2e

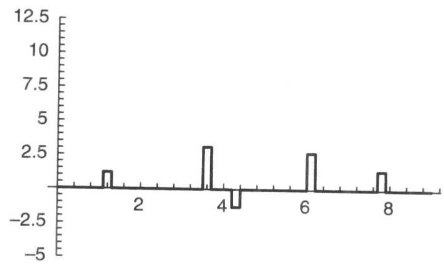

Figure 1.2f

polation. For example, in a specific application we may want the interpolating function to have a continuous first derivative. (That requirement would disqualify most of the functions in Figure 1.1.)

The **linear interpolation problem** is a special case that arises when F is to be chosen from a prescribed n-dimensional vector space of functions on X. Suppose that U is this vector space and that a basis for U is $\{u_1, u_2, \ldots, u_n\}$. The function F that we seek must have the form

$$F = \sum_{j=1}^{n} c_j u_j$$

When the interpolation conditions are imposed on F, we obtain

$$\lambda_i = F(x_i) = \sum_{j=1}^{n} c_j u_j(x_i) \qquad (1 \leq i \leq n)$$

This is a system of n linear equations in n unknowns. It can be written in matrix form as $Ac = y$, or in detail as

$$\begin{bmatrix} u_1(x_1) & u_2(x_1) & \cdots & u_n(x_1) \\ u_1(x_2) & u_2(x_2) & \cdots & u_n(x_2) \\ \cdots & \cdots & \cdots & \cdots \\ u_1(x_n) & u_2(x_n) & \cdots & u_n(x_n) \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \cdots \\ c_n \end{bmatrix} = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \cdots \\ \lambda_n \end{bmatrix}$$

The $n \times n$ matrix A appearing here is called the **interpolation matrix.** In order that our problem be solvable for any choice of ordinates λ_i it is necessary and sufficient that the interpolation matrix be nonsingular. The ideal situation is that this matrix be nonsingular for all choices of n distinct nodes.

THEOREM 1. *Let U be an n-dimensional linear space of functions on X. Let $x_1, x_2, \ldots, x_n$ be n distinct nodes in X. In order that U be capable of interpolating arbitrary data at the nodes it is necessary and sufficient that zero data be interpolated only by the zero-element in U.*

Proof. The space U can furnish an interpolant for arbitrary data if and only if the interpolation matrix A is nonsingular. An equivalent condition on the matrix A is that the equation $Ac = 0$ can be true only if $c = 0$. ■

Example. Let $X = \mathbb{R}$ and let $u_j(x) = x^{j-1}$, for $j = 1, 2, \ldots, n$. An $n \times n$ interpolation matrix in this special case is called a **Vandermonde matrix.** It looks like this:

$$V = \begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{bmatrix}$$

The determinant of V is given by the formula

$$\det V = \prod_{1 \leq j < i \leq n} (x_i - x_j)$$

This is obviously nonzero if and only if the nodes are distinct. Hence the interpolation problem has a unique solution for any choice of distinct nodes. We can also use Theorem 1 to see that V is nonsingular. Thus, we consider the "homogeneous" linear problem, in which we attempt to interpolate zero data. The solution will be a polynomial of degree at most $n - 1$ that takes the value 0 at each of the n nodes. Since a nonzero polynomial of degree at most $n-1$ can have at most $n-1$ zeros, we conclude that the zero polynomial is the only possible solution. ∎

The Vandermonde matrix occurs often in mathematics. Refer to Rushanan [Rush], Grosof and Taiani [GT], Cheney [C1], for example. It is ill-conditioned for numerical work. See Gautschi, [Gau1, Gau2].

An n-dimensional vector space U of functions on a domain X is said to be a **Haar space** if the only element of U which has more than $n-1$ roots in X is the zero element. The next theorem provides some properties equivalent to the Haar property. In the theorem, we refer to **point-evaluation functionals.** If V is a vector space of functions on a set X, and if x is a point of X, then the point-evaluation functional corresponding to x is denoted by x^* and is defined on V by

$$x^*(f) = f(x) \qquad (f \in V)$$

Obviously x^* is linear, because

$$x^*(\alpha f + \beta g) = (\alpha f + \beta g)(x) = \alpha f(x) + \beta g(x) = \alpha x^*(f) + \beta x^*(g)$$

THEOREM 2. *Let U have the basis $\{u_1, u_2, \ldots, u_n\}$. These properties are equivalent:*

a. *U is a Haar space*
b. $\det(u_j(x_i)) \neq 0$ *for any set of distinct points $x_1, x_2, \ldots, x_n$ in X*
c. *For any distinct points $x_1, x_2, \ldots, x_n$ in X, the set of point-evaluation functionals $x_1^*, x_2^*, \ldots, x_n^*$ spans the algebraic dual space U^**
d. *If $x_1, x_2, \ldots, x_m$ are distinct in X and if $\sum_{i=1}^m \lambda_i u_j(x_i) = 0$ for $j = 1, 2, \ldots, n$ then either at least $n+1$ of the coefficients λ_i are nonzero, or $\sum_{i=1}^m |\lambda_i| = 0$*

Proof. To show that **a** implies **b**, suppose **b** false. Since the determinant of $(u_j(x_i))$ is zero, the matrix is singular, and there exists a nonzero vector $(c_1, c_2, \ldots, c_n)$ such that $\sum_{j=1}^n c_j u_j(x_i) = 0$, $(1 \le i \le n)$. Put $u = \sum_{j=1}^n c_j u_j$. Since $\{u_1, u_2, \ldots, u_n\}$ is linearly independent, $u \neq 0$. But $u(x_i) = 0$ for $1 \le i \le n$. Hence **a** is false.

To show that **b** implies **c**, suppose **b** true. Then the set $\{x_1^*, x_2^*, \ldots, x_n^*\}$ is linearly independent when these functionals are restricted to U. Indeed, if $\sum_{i=1}^n a_i x_i^* | U = 0$, then $\sum_{i=1}^n a_i x_i^*(u_j) = 0$ for $1 \le j \le n$, and by **b**, $\sum_{i=1}^n |a_i| = 0$. Since U^* is of dimension n, the functionals span U^*.

To show that **c** implies **d**, assume **c**. Let $x_1, \ldots, x_m$ be distinct points that satisfy $\sum_{i=1}^m \lambda_i u_j(x_i) = 0$ for $1 \le j \le n$. If $m \le n$, then by **c** we can take additional points and obtain a basis $\{x_1^*, \ldots, x_n^*\}$ for U^*. Then the subset $\{x_1^*, \ldots, x_m^*\}$ is linearly independent on U and all λ_i are zero. If $m > n$ and $\sum_{i=1}^m |\lambda_i| \neq 0$ then at least $n+1$ of the λ_i are nonzero, for otherwise we will have a nontrivial linear combination of n (or fewer) x_i^* that vanishes on U, contrary to **c**.

To prove that **d** implies **a**, assume **d** and take $m = n$. Then the equation $\sum_{i=1}^n \lambda_i u_j(x_i) = 0$ for $1 \le j \le n$ implies $\sum_{i=1}^n |\lambda_i| = 0$. Hence the matrix $(u_j(x_i))$ is nonsingular. Thus if $(c_1, c_2, \ldots, c_n) \neq 0$, we cannot have $\sum_{j=1}^n c_j u_j(x_i) = 0$. In other words, a nonzero member of U cannot have n zeros. ∎