Mathematics Monograph Series 3

Spectral and High-Order Methods with Applications

Jie Shen Tao Tang

（谱方法和高精度算法及其应用）

SCIENCE PRESS
Beijing

(O—2553.0101)

Spectral and High-Order Methods with Applications

This book expands lecture notes by the authors taught in the past few years in USA, Canada and China. The overall emphasis of these notes is to present basic algorithms together with some applications of spectral methods. The aim is to provide a sufficient background on the implementation and analysis of spectral and high-order methods so that the readers can approach the current research literature with the necessary tools and understanding.

It is expected that this book will be a useful supplement for people studying spectral methods on their own. This book is especially suited to students interested in high-order methods for PDEs, but it will appeal to numerical analysis and mathematically oriented students and researchers in engineering, physics, and related areas.

About the authors:
Jie Shen is a full Professor of Mathematics at Purdue University, USA.
Tao Tang is a Chair Professor of Mathematics at Hong Kong Baptist University.

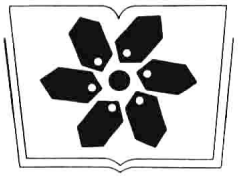
ISBN 7—03—017722—3
定 价：78.00 元

Mathematics Monograph Series

3

Spectral and High-Order Methods with Applications

Ji



中国科学院科学出版基金资助出版

Mathematics Monograph Series 3

Jie Shen Tao Tang

Spectral and High-Order Methods with Applications

SCIENCE PRESS
Beijjng

Responsible Editor: Chen Yuzhuo

Published by Science Press
16 Donghuangchenggen North Street
Beijing 100717, China

Printed in Beijing

ISBN 7-03-017722-3 (Beijing)

Preface

This book expands lecture notes by the authors for a course on *Introduction of Spectral Methods* taught in the past few years at Penn State University, Simon Fraser University, the Chinese University of Hong Kong, Hong Kong Baptist University, Purdue University and the Chinese Academy of Sciences. Our lecture notes were also used by Prof. Zhenhuan Teng in his graduate course at Peking University.

The overall emphasis of the present book is to present some basic spectral and high-order algorithms together with applications to some linear and nonlinear problems that one frequently encounters in practice. The algorithms in the book are presented in a *pseudocode* format or with MATLAB or FORTRAN codes that contain additional details beyond the mathematical formulas. The reader can easily write computer routines based on the pseudocodes in any standard computer language. We believe that the readers learn and understand numerical methods best by seeing how algorithms are developed from the mathematical theory and then writing and testing computer implementations of them. For those interested in the numerical analysis of the spectral methods, we have also provided self-contained error analysis for some basic spectral-Galerkin algorithms presented in the book. Our aim is to provide a sufficient background on the implementation and analysis of spectral and high-order methods so that the readers can approach the current research literature with the necessary tools and understanding.

We hope that this book will be useful for people studying spectral methods on their own. It may also serve as a textbook for advanced undergraduate/beginning graduate students. The only prerequisite for the present book is a standard course in Numerical Analysis.

This project has been supported by NSERC Canada, National Science Foundation, Research Grant Council of Hong Kong, and International Research Team of Complex System of the Chinese Academy of Sciences. In writing this book, we have received much help from our friends and students. In particular, we would like to thank Dr. Lilian Wang of Nanyang Technical University of Singapore for his many contributions throughout the book. We are grateful to the help provided by Zhongzhi Bai of the Chinese Academy of Sciences, Weizhu Bao of National University of Singapore, Raymond Chan of Chinese University of Hong Kong, Wai Son Don of Brown

University, Heping Ma of Shanghai University and Xuecheng Tai of Bergen University of Norway. Our gratitude also goes to Professor Hermann Brunner of Memorial University of Newfoundland, Dr. Zhengru Zhang of Beijing Normal University, and the following graduate students at Purdue, Qirong Fang, Yuen-Yick Kwan, Hua Lin, Xiaofeng Yang and Yanhong Zhao, who have read the entire manuscripts and provided many constructive suggestions. Last but not the least, we would like to thank our wives and children for their love and support.

A website relevant to this book can be found in

`http://www.math.hkbu.edu.hk/~ttang/PGteaching`

or

`http://lsec.cc.ac.cn/~ttang/PGteaching`

We welcome comments and corrections to the book. We can be reached by email to

`shen@math.purdue.edu` (Shen) and `ttang@math.hkbu.edu.hk` (Tang).

Jie Shen
Purdue University

Tao Tang
Hong Kong Baptist University

Contents

Chapter 1

Preliminaries

Contents

In this chapter, we present some preliminary materials which will be used throughout the book. The first section set the stage for the introduction of spectral methods. In Sections 1.2−1.4, we present some basic properties of orthogonal polynomials, which play an essential role in spectral methods, and introduce the notion of generalized Jacobi polynomials. Since much of the success and popularity of spectral methods can be attributed to the invention of Fast Fourier Transform (FFT), an algorithmic description of the FFT is presented in Section 1.5. In the next two sections, we collect some popular time discretization schemes and iterative schemes which will be frequently used in the book. In the last section, we present a concise error analysis for several projection operators which serves as the basic ingredients for the error analysis of spectral methods.

1.1 Some basic ideas of spectral methods

Comparison with the finite element method
Computational efficiency
Fourier spectral method
Phase error

Finite Difference (FD) methods approximate derivatives of a function by *local* arguments (such as $u'(x) \approx (u(x+h)-u(x-h))/2h$, where h is a small grid spacing) - these methods are typically designed to be exact for polynomials of low orders. This approach is very reasonable: since the derivative is a local property of a function, it makes little sense (and is costly) to invoke many function values far away from the point of interest.

In contrast, spectral methods are *global*. The traditional way to introduce them starts by approximating the function as a sum of very smooth basis functions:

$$u(x) \approx \sum_{k=0}^{N} a_k \Phi_k(x),$$

where the $\Phi_k(x)$ are polynomials or trigonometric functions. In practice, there are many feasible choices of the basis functions, such as:

$\Phi_k(x) = e^{ikx}$ (the Fourier spectral method);
$\Phi_k(x) = T_k(x)$ ($T_k(x)$ are the Chebyshev polynomials; the Chebyshev spectral method);
$\Phi_k(x) = L_k(x)$ ($L_k(x)$ are the Legendre polynomials; the Legendre spectral method).

In this section, we will describe some basic ideas of spectral methods. For ease of exposition, we consider the Fourier spectral method (i.e. the basis functions are chosen as e^{ikx}). We begin with the periodic heat equation, starting at time 0 from $u_0(x)$:

$$u_t = u_{xx}, \tag{1.1.1}$$

with a periodic boundary condition $u(x,0) = u_0(x) = u_0(x+2\pi)$. Since the exact solution u is periodic, it can be written as an infinite Fourier series. The approximate solution u^N can be expressed as a *finite* series. It is

$$u^N(x,t) = \sum_{k=0}^{N-1} a_k(t) e^{ikx}, \qquad x \in [0, 2\pi),$$

where each $a_k(t)$ is to be determined.

Comparison with the finite element method

We may compare the spectral method (before actually describing it) to the finite element method. One difference is this: the trial functions τ_k in the finite element method are usually 1 at the mesh-point, $x_k = kh$ with $h = 2\pi/N$, and 0 at the other mesh-points, whereas e^{ikx} is nonzero everywhere. That is not such an important distinction. We could produce from the exponentials an interpolating function like τ_k, which is zero at all mesh-points except at $x = x_k$:

$$F_k(x) = \frac{1}{N} \sin \frac{N}{2}(x - x_k) \cot \frac{1}{2}(x - x_k), \qquad N \text{ even}, \tag{1.1.2}$$

$$F_k(x) = \frac{1}{N} \sin \frac{N}{2}(x - x_k) \csc \frac{1}{2}(x - x_k), \qquad N \text{ odd}. \tag{1.1.3}$$

Of course it is not a piecewise polynomial; that distinction is genuine. A consequence of this difference is the following:

> Each function F_k spreads over the whole solution interval, whereas τ_k is zero in all elements not containing x_k. The stiffness matrix is sparse for the finite element method; in the spectral method it is full.

The computational efficiency

Since the matrix associated with the spectral method is full, the spectral method seems more time-consuming than finite differences or finite elements. In fact, the spectral method had not been used widely for a long time. The main reason is the expensive cost in computational time. However, the discovery of the Fast Fourier Transform (FFT) by Cooley and Tukey[33] solves this problem. We will describe the Cooley-Tukey algorithm in Chapter 5. The main idea is the following. Let $w_N = e^{2\pi i/N}$ and

$$(\mathcal{F}_N)_{jk} = w_N^{jk} = \cos \frac{2\pi jk}{N} + i \sin \frac{2\pi jk}{N}, \qquad 0 \leqslant j, \quad k \leqslant N - 1.$$

Then for any N-dimensional vector v_N, the usual N^2 operations in computing $\mathcal{F}_N v_N$ are reduced to $N \log_2 N$. The significant improvement can be seen from the following table:

N	N^2	$N\log_2 N$	N	N^2	$N\log_2 N$
16	256	64	256	65536	2048
32	1024	160	512	262144	4608
64	4096	384	1024	1048576	10240
128	16384	896	2048	4194304	22528

The Fourier spectral method

Unlike finite differences or finite elements, which replace the right-hand side u_{xx} by differences at nodes, the spectral method uses u_{xx}^N exactly. In the spectral method, there is no Δx. The derivatives with respect to space variables are computed explicitly and correctly.

The Fourier approximation u^N is a combination of oscillations e^{ikx} up to frequency $N-1$, and we simply differentiate them; hence

$$u_t^N = u_{xx}^N$$

becomes

$$\sum_{k=0}^{N-1} a_k'(t)e^{ikx} = \sum_{k=0}^{N-1} a_k(t)(ik)^2 e^{ikx}.$$

Since frequencies are uncoupled, we have $a_k'(t) = -k^2 a_k(t)$, which gives

$$a_k(t) = e^{-k^2 t} a_k(0),$$

where the values $a_k(0)$ are determined by using the initial function:

$$a_k(0) = \frac{1}{2\pi}\int_0^{2\pi} u_0(x)e^{-ikx}\mathrm{d}x.$$

It is an easy matter to show that

$$\begin{aligned} |u(x,t) - u^N(x,t)| &= \left|\sum_{k=N}^{\infty} a_k(0)e^{ikx}e^{-k^2 t}\right| \\ &\leqslant \max_k |a_k(0)| \sum_{k=N}^{\infty} e^{-k^2 t} \\ &\leqslant \max_{0\leqslant x\leqslant 2\pi} |u_0(x)| \int_N^{\infty} e^{-tx^2}\mathrm{d}x. \end{aligned}$$

Therefore, the error goes to zero very rapidly as N becomes reasonably large. The

convergence rate is determined by the integral term

$$J(t,N) := \int_N^\infty e^{-tx^2}\mathrm{d}x = \sqrt{\frac{\pi}{4t}}\mathrm{erfc}(\sqrt{t}N),$$

where `erfc(x)` is the complementary error function (both FORTRAN and MATLAB have this function). The following table lists the value of $J(t,N)$ at several values of t:

N	J(0.1, N)	J(0.5, N)	J(1, N)
1	1.8349e+00	3.9769e-01	1.3940e-01
2	1.0400e+00	5.7026e-02	4.1455e-03
3	5.0364e-01	3.3837e-03	1.9577e-05
4	2.0637e-01	7.9388e-05	1.3663e-08
5	7.1036e-02	7.1853e-07	1.3625e-12
6	2.0431e-02	2.4730e-09	1.9071e-17
7	4.8907e-03	3.2080e-12	3.7078e-23
8	9.7140e-04	1.5594e-15	9.9473e-30

In more general problems, the equation in time will not be solved exactly. It needs a difference method with time step Δt, as Chapter 5 will describe. For derivatives with respect to space variables, there are two ways:

(1) Stay with the harmonics e^{ikx} or $\sin kx$ or $\cos kx$, and use FFT to go between coefficients a_k and mesh values $u^N(x_j,t)$. Only the mesh values enter the difference equation in time.

(2) Use an expansion $U = \sum U_k(t)F_k(x)$, where $F_k(x)$ is given by (1.1.2) and (1.1.3), that works directly with values U_k at mesh points (where $F_k = 1$). There is a *differentiation matrix* D that gives mesh values of the derivatives, $D_{jk} = F_k'(x_j)$. Then the approximate heat equation becomes $U_t = D^2U$.

Phase error

The fact that x-derivatives are exact makes spectral methods free of phase error. Differentiation of the multipliers e^{ikx} give the right factor ik while finite differences lead to the approximate factor iK:

$$\frac{e^{ik(x+h)} - e^{ik(x-h)}}{2h} = iKe^{ikx}, \qquad K = \frac{\sin kh}{h}.$$

When kh is small and there are enough mesh points in a wavelength, K is close to k. When kh is large, K is significantly smaller than k. In the case of the heat

equation (1.1.1) it means a slower wave velocity. For details, we refer to Richtmyer and Morton[131] and LeVeque [101]. In contrast, the spectral method can follow even the nonlinear wave interactions that lead to turbulence. In the context of solving high Reynolds number flow, the low physical dissipation will not be overwhelmed by large numerical dissipation.

Exercise 1.1

Problem 1 Consider the linear heat equation (1.1.1) with homogeneous Dirichlet boundary conditions $u(-1,t) = 0$ and $u(1,t) = 0$. If the initial condition is $u(x,0) = \sin(\pi x)$, then the exact solution of this problem is given by $u(x,t) = e^{-\pi^2 t}\sin(\pi x)$. It has the infinite Chebyshev expansion

$$u(x,t) = \sum_{n=0}^{\infty} b_n(t) T_n(x),$$

where

$$b_n(t) = \frac{1}{c_n} J_n(\pi) e^{-\pi^2 t},$$

with $c_0 = 2$ and $c_n = 1$ if $n \geqslant 1$.

a. Calculate

$$J_n(\pi) = \int_{-1}^{1} \frac{1}{\sqrt{1-x^2}} T_n(x)\sin(\pi x)\mathrm{d}x$$

by some numerical method (e.g. Simpson's rule)①;

b. Plot $J_n(\pi)$ against n for $n \leqslant 25$. This will show that the truncation series converges at an exponential rate (a well-designed collocation method will do the same).

1.2 Orthogonal polynomials

Existence
Zeros of orthogonal polynomials
Polynomial interpolations
Quadrature formulas
Discrete inner product and discrete transform

① Hint: (a) Notice that $J_n(\pi) = 0$ when n is even; (b) a coordinate transformation like $x = \cos\theta$ may be used.

Orthogonal polynomials play a fundamental role in the implementation and analysis of spectral methods. It is thus essential to understand some general properties of orthogonal polynomials. Two functions f and g are said to be *orthogonal* in the weighted Sobolev space $L^2_\omega(a,b)$ if

$$\langle f,\, g\rangle := (f,g)_\omega := \int_a^b \omega(x)f(x)g(x)\mathrm{d}x = 0,$$

where ω is a fixed positive *weight function* in (a,b). It can be easily verified that $\langle\cdot,\,\cdot\rangle$ defined above is an inner product in $L^2_\omega(a,b)$.

A sequence of *orthogonal polynomials* is a sequence $\{p_n\}_{n=0}^\infty$ of polynomials with $\deg(p_n) = n$ such that

$$\langle p_i,\, p_j\rangle = 0 \qquad \text{for} \quad i \neq j. \tag{1.2.1}$$

Since orthogonality is not altered by multiplying a nonzero constant, we may normalize the polynomial p_n so that the coefficient of x^n is one, i.e.,

$$p_n(x) = x^n + a_{n-1}^{(n)}x^{n-1} + \cdots + a_0^{(n)}.$$

Such a polynomial is said to be *monic*.

Existence

Our immediate goal is to establish the existence of orthogonal polynomials. Although we could, in principle, determine the coefficients $a_j^{(n)}$ of p_n in the natural basis $\{x^j\}$ by using the orthogonality conditions (1.2.1), it is more convenient, and numerically more stable, to express p_{n+1} in terms of lower-order orthogonal polynomials. To this end, we need the following general result:

Let $\{p_n\}_{n=0}^\infty$ be a sequence of polynomials such that p_n is exactly of degree n. If

$$q(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0, \tag{1.2.2}$$

then q can be written uniquely in the form

$$q(x) = b_n p_n + b_{n-1}p_{n-1} + \cdots + b_0 p_0. \tag{1.2.3}$$

In establishing this result, we may assume that the polynomials $\{p_n\}$ are monic. We shall prove this result by induction. For $n = 0$, we have

$$q(x) = a_0 = a_0 \cdot 1 = a_0 p_0(x).$$

Hence we must have $b_0 = a_0$. Now assume that q has the form (1.2.2). Since p_n is the only polynomial in the sequence $p_n, p_{n-1}, \cdots, p_0$ that contains x^n and since p_n is monic, it follows that we must have $b_n = a_n$. Hence, the polynomial $q - a_n p_n$ is of degree $n-1$. Thus, by the induction hypothesis, it can be expressed uniquely in the form

$$q - a_n p_n = b_{n-1} p_{n-1} + \cdots + b_0 p_0,$$

which establishes the result.

A consequence of this result is the following:

Lemma 1.2.1 *If the sequence of polynomials $\{p_n\}_{n=0}^{\infty}$ is monic and orthogonal, then the polynomial p_{n+1} is orthogonal to any polynomial q of degree n or less.*

We can establish this by the following observation:

$$\langle p_{n+1},\, q\rangle = b_n\langle p_{n+1},\, p_n\rangle + b_{n-1}\langle p_{n+1},\, p_{n-1}\rangle + \cdots + b_0\langle p_{n+1},\, p_0\rangle = 0,$$

where the last equality follows from the orthogonality of the polynomials $\{p_n\}$.

We now prove the existence of orthogonal polynomials①. Since p_0 is monic and of degree zero, we have

$$p_0(x) \equiv 1.$$

Since p_1 is monic and of degree one, it must have the form

$$p_1(x) = x - \alpha_1.$$

To determine α_1, we use orthogonality:

$$0 = \langle p_1,\, p_0\rangle = \int_a^b \omega(x)x\mathrm{d}x - \alpha_1 \int_a^b \omega(x)\mathrm{d}x.$$

Since the weight function is positive in (a, b), it follows that

$$\alpha_1 = \int_a^b \omega(x)x\mathrm{d}x \Big/ \int_a^b \omega(x)\mathrm{d}x.$$

In general we seek p_{n+1} in the form $p_{n+1} = xp_n - \alpha_{n+1}p_n - \beta_{n+1}p_{n-1} - \gamma_{n+1}p_{n-2} - \cdots$. As in the construction of p_1, we use orthogonality to determine the coefficients above. To determine α_{n+1}, write

$$0 = \langle p_{n+1},\, p_n\rangle = \langle xp_n,\, p_n\rangle - \alpha_{n+1}\langle p_n,\, p_n\rangle - \beta_{n+1}\langle p_{n-1},\, p_n\rangle - \cdots.$$

① The procedure described here is known as Gram-Schmidt orthogonalization.