## NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensuie the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

# THE PSEUDOSPECTRAL CHEBYSHEV METHOD FOR TWO-POINT BOUNDARY VALUE PROBLEMS 

by<br>Yunyun Liu<br>B.Sc. Shanghai University of Science and Technolugy 1982<br>A thesis submitted in partial fulfillment<br>of the requirements for the degree of<br>Master of Science<br>in the Department of Mathematics and Statistics<br>of<br>Simon Fraser University<br>(C) Yunyun Liu 1992<br>SIMON FRASER UNIVERSITY<br>April, 1992

All rights reserved. This work may not be reproduced in whole or in part, by photocopy or other means, without the permission of the author.

The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

L'auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

L'auteur conserve la propriété du droit d'auteur qui protège sá thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

## APPROVAL

```
Name: Yunyun Liu
Degree: Master of Science
Title of thesis:
THE PSEUDOSPECTRAL CHEBYSHEV METHOD
FOR TWO-POINT BOUNDARY VALUE PROBLEMS
```

Examining Committee:

Dr. G. A. C. Graham

Chairman

Dr. Manfred Trummer, Senior Supervisor

Dr. Robert D. Russell

Dr. George Bojadziev

Dr. Tao Tang, External Examiner

Date Approved:

I hereby grant to Simon Fraser University tho right lo lend my thesis, project or extended essay (the title of which is shown below) to users of tho Simon Fraser University Library, and to make partial or single copies only for such users or in response to a roquest from the library of any other university, or other educational institution, on its own behalf or for ono of its users. I further agree that permission for multiple copying of this work for scholarly purposes may be granted by me or the Dean of Graduate Studies. It is understood that copying or publication of this work for financial gain shall not be allowed without my written permission.

Title of Thesis/Rrojoct/Extended Essay
THE PSEUDOSPECTRAL CHEBYSHEV METHOD
FOR TWO-PUINT BOUNDARY VALUE PROBLEM
$\qquad$

Author:
(signature)
$\frac{\text { YunYun } L 1 L E}{\text { (name) }}$
$\frac{\text { May } 20,1992}{\text { (date). }}$

## Abstract

Although spectral methods have attracted much of the attention in current research of numerical methods for solving differential equations, little experience is available in applying spectral methods to solve (two - point) boundary value problems. In this thesis, two spectral methods (the pseudospectral Chebyshev method and the method of transformation to the circle) have been studied to solve second order boundary value problems (BVPs).

Emphasis is given to singularly perturbed BVPs that are controlled by a singularity parameter $\epsilon$. Three kinds of singularities are investigated: mild ( $\epsilon=10^{-2}$ ), stiff $\left(\epsilon=10^{-4}\right)$ and very stiff $\left(\epsilon=10^{-6}\right)$, and problems with interior or boundary layers are considered.

Numerical performances of the above mentioned two spectral methods are judged by condition numbers, average errors and cpu times. It has been found that the pseudospectral Chebyshev method is more suitable to solve non-stiff problems than to stiff problems; that the method of transformation to the circle can decrease the condition numbers of the relevant matrices when applied to solve non-stiff problems. However, for stiff problems, it is nevertheless seldom successful. A brief comparison of the above two methods to the Chebyshev collocation method is also presented in this thesis. It is found that once sufficient resolution is achieved, numerical solutions are accurate despite the large condition numbers.

## Acknowledgements

First of all, I want to thank my senior supervisor, Dr. Manfred K. Trummer for his guidance, encouragement and patience during the preparation of this thesis.

My thanks also go to Dr. Robert D. Russell, Dr. Margorzata Dubiel and Dr. Kathy Heinrich for their support.

Finally, I want to thank Graduate Secretary Mrs. Sylvia Holmes and other department staff. It was their kindness and help that made my stay at SFU comfortable.

A special thank is to the Department of Mathematics and Statistics of Simon Fraser University for its financial support.

## Dedication

## To My Parents

To the Memory of My Grandparents
To Great Chen

## Contents

Abstract ..... iii
Acknowledgements ..... iv
Dedication ..... v
Contents ..... vii
List of Tables ..... vii
List of Figures ..... vii
1 Introduction ..... 1
1.1. Pseudospectral Chebyshev Methods ..... 1
1.2 Collocation Method ..... 4
1.3 Thesis Plan ..... 7
2 A Pseudospectral Chebyshev Method ..... 9
2.1 Introduction ..... 9
2.2 Chebyshev Polynomial Expansions ..... 10
2.3 A Pseudospectral Chebyshev Method ..... 12
2.4 Linear System Solver ..... 15
2.5 Numerical Results ..... 17
3 Preliminary Transform To The Circle ..... 39
3.1 Introduction ..... 39
3.2 The Method of Transformation to the Circle ..... 40
3.3 Iterative Methods ..... 44
3.3.1 CGNR (The Conjugate Gradient Iteration Applied to the Normal Equations) ..... 44
3.3.2 CGS (The Conjugate Gradient Squared) ..... 45
3.4 Numerical Results ..... 47
4 Chebyshev Collocation Method ..... 63
4.1 Chebyshev Collocation Method ..... 63
4.2 Numerical Results ..... 66
5 Discussion ..... 74
Bibliography ..... 85

## List of Tables

2.1 Example 1, pseudospectral Chebyshev method, $\epsilon=10^{-2}$ ..... 25
2.2 Example 1, pseudospectral Chebyshev method, $\epsilon=10^{-4}$ ..... 25
2.3 Example 1, pseudospectral Chebyshev method, $\epsilon=10^{-6}$ ..... 25
2.4 Example 2, pseudospectral Chebyshev method, $\epsilon=10^{-2}$ ..... 30
2.5 Example 2, pseudospectral Chebyshev method, $\epsilon=10^{-4}$ ..... 30
2.6 Example 2, pseudospectral Chebyshev method, $\epsilon=10^{-6}$ ..... 31
2.7 Example 3, pseudospectral Chebyshev method, $\epsilon=10^{-2}$ ..... 35
2.8 Example 3, pseudospectral Chebyshev method, $\epsilon=10^{-4}$ ..... 35
2.9 Example 3, pseudospectral Chebyshev method, $\epsilon=10^{-6}$ ..... 35
3.1 Example 1, transformation to the circle method, $\epsilon=10^{-2}$ ..... 54
3.2 Example 1, transformation to the circle method, $\epsilon=10^{-4}$ ..... 55
3.3 Example 1, transformation to the circle method, $\epsilon=10^{-6}$ ..... 55
3.4 Example 2, transformation to the circle method, $\epsilon=10^{-2}$ ..... 60
3.5 Example 2, transformation to the circle method, $\epsilon=10^{-4}$ ..... 60
3.6 Example 2, transformation to the circie method, $\epsilon=10^{-6}$ ..... 60
4.1 Example 2, Chebyshev collocation method, $\epsilon=10^{-2}$. ..... 66
4.2 Example 2, Chebyshev collocation method, $\epsilon=10^{-4}$. ..... 66
4.3 Example 2, Chebyshev collocation method, $\epsilon=10^{-6}$. ..... 66
5.1 Example 1, the condition number for two methods, $\epsilon=10^{-2}$ ..... 79
5.2 Example 1, the condition number for two methods, $\epsilon=10^{-4}$. ..... 79
5.3 Example 1, the condition number for two methods, $\epsilon=10^{-6}$ ..... 79
5.4 Example 2, the condition number for two methods, $\epsilon=10^{-2}$ ..... 79
5.5 Example 2, the condition number for two methods, $\epsilon=10^{-4}$ ..... 80
5.6. Example 2, the condition number for two methods, $\epsilon=10^{-6}$ ..... 80

## List of Figures

1.1 Cubic $B$-spline functions ..... 6
2.1 Distribution of Chebyshev collocation points, $N=32$ ..... 13
2.2 Example 1, pseudospectral Chebyshev method, $\epsilon=10^{-2}, N=32$ ..... 19
2.3 Example 1, pseudospectral Chebyshev method, $\epsilon=10^{-2}, N=64$ ..... 20
2.4 Example 1, pseudospectral Chebyshev method, $\epsilon=10^{-2}, N=96$ ..... 20
2.5 Example 1, pseudospectral Chebyshev method, $\epsilon=10^{-4}, N=32$ ..... 21
2.6 Example 1, pseudospectral Chebyshev method, $\epsilon=10^{-4}, N=64$ ..... 21
2.7 Example 1, pseudospectral Chebyshev method, $\epsilon=10^{-4}, N=96$ ..... 22
2.8 Example 1, pseudospectral Chebyshev method, $\epsilon=10^{-4}, N=128$ ..... 22
2.9 Example 1, pseudospectral Chebyshev method, $\epsilon=10^{-6}, N=32$ ..... 23
2.10 Example 1, pseudospectral Chebyshev method, $\epsilon=10^{-\hat{6}}, N=64$ ..... 23
2.11 Example 1, pseudospectral Chebyshev method, $\epsilon=17^{-6}, N=96$ ..... 24
2.12 Example 1, pseudospectral Chebyshev method, $\epsilon=10^{-6}, N=128$ ..... 24
2.13 Example 2, pseudospectral Chebyshev method, $\epsilon=10^{-2}, N=32$ ..... 27
2.14 Example 2, pseudospectral Chebyshev method, $\epsilon=10^{-2}, N=64$ ..... 27
2.15 Example 2, pseudospectral Chebyshev method, $\epsilon=10^{-4}, N=128$ ..... 28
2.16 Example 2, pseudospectral Chebyshev method, $\epsilon=10^{-4}, N=256$ ..... 28
2.17 Example 2, pseudospectral Chebyshev method, $\epsilon=10^{-4} ; N=512$ ..... 29
2.18 Example 2, pseudospectral Chebyshev method, $\epsilon=10^{-6}, N=128$ ..... 29
2.19 Example 2, pseudospectral Chebyshev method, $\epsilon=10^{-6}, N=256$ ..... 30
2.20 Example 3, pseudospectral Chebyshev method, $\epsilon=10^{-2}, N=32$ ..... 32
2.21 Example 3, pseudospectral Chebyshev method, $\epsilon=10^{-2}, N=64$ ..... 32
2.22 Example 3, pseudospectral Chebyshev method, $\epsilon=10^{-4}, N=128$ ..... 33
2.23 Example 3, pseudospectral Chebyshev method, $\epsilon=10^{-4}, N=256$ ..... 33
2.24 Example 3, pseudospectral Chebyshev method, $\epsilon=10^{-6}, N=256$ ..... 34
2.25 Example 3; pseudospectral Chebyshev method, $\epsilon=10^{-6}, N=512$ ..... 34
2.26 Example 1, average error ( $* 10^{12}$ ) for $\epsilon=10^{-2}, N=32,64,96,128$ ..... 36
2.27 Example 1, average error $\left(* 10^{2}\right)$ for $\epsilon=10^{-4}, N=32,64,96,128$ ..... 36
2.28 Example 1, average error ( $* 10^{2}$ ) for $\epsilon=10^{-6}, N=32,64,96,128$ ..... 37
2.29 Example 2, average error ( $* 10^{3}$ ) for $\epsilon=10^{-2}, N=32,64,96,128$ ..... 37
2.30 Example 2, average error for $\epsilon=10^{-4}, N=128,256,512$ ..... 38
2.31 Example 2, average error for $\epsilon=10^{-6}, N=128,256,512$ ..... 38
3.1 Example 1, transformation to the circle method, $\epsilon=10^{-2}, N=32$ ..... 49
3.2 Example 1, transformation to the circle method, $\epsilon=10^{-2}, N=64$ ..... 49
3.3 Example 1, transformation to the circle method, $\epsilon=10^{-2}, N=96$ ..... 50
3.4 Example 1, transformation to the circle method, $\epsilon=10^{-4}, N=32$ ..... 50
3.5 Example 1, transformation to the circle method, $\epsilon=10^{-4}, N=64$ ..... 51
3.6 Example 1, transformation to the circle method, $\epsilon=10^{-4}, N=96$ ..... 51
3.7 Example 1, transformation to the circle method, $\epsilon=10^{-4}, N=128$ ..... 52
3.8 Example 1, transformation to the circle method, $\epsilon=10^{-6}, N=32$ ..... 52
3.9 Example 1, transformation to the circle method, $\epsilon=10^{-6}, N=64$ ..... 53
3.10 Example 1, transformation to the circle method, $\epsilon=10^{-6}, N=96$ ..... 53
3.11 Example 1, transformation to the circle method, $\epsilon=10^{-6}, N=128$ ..... 54
3.12 Example 2, transformation to the circle method, $\epsilon=10^{-2}, N=64$ ..... 57
3.13 Example 2, transformation to the circle method, $\epsilon=10^{-4}, N=128$ ..... 57
3.14 Example 2, transformation to the circle method, $\epsilon=10^{-4}, N=256$ ..... 58
3.15 Example 2, transformation to the circle method, $\epsilon=10^{-6}, N=128$ ..... 58
3.16 Example 2, transformation to the circle method, $\epsilon=10^{-6}, N=256$ ..... 59
3.17 Example 1, the pseudospectral method, $N=96, \epsilon=10^{-4}$ ..... 61
3.18 Example 1, the pseudospectral method, $N=96, \epsilon=10^{-4}$ ..... 61
3.19 Example 1, the circle technique, $N=96, \epsilon=10^{-4}$ ..... 62
3.20 Example 1, the circle technique, $N=96, \epsilon=10^{-4}$ ..... 62
4.1 Example 2, Chebyshev collocation method, $\epsilon=10^{-2}, N=32$ ..... 68
4.2 Example 2, Chebyshev collocation method, $\epsilon=10^{-2}, N=64$ ..... 68
4.3 Example 2, Chebyshev collocation method, $\epsilon=10^{-2}, N=96$ ..... 69
4.4 Example 2, Chebyshev collocation method, $\epsilon=10^{-2}, N=128$ ..... 69
4.5 Example 2, Chebyshev collocation method, $\epsilon=10^{-2}, N=256$ ..... 70
4.6 Example 2, Chebyshev collocation method, $\epsilon=10^{-2}, N=512$ ..... 70
4.7 Example 2, Chebyshev collocation method, $\epsilon=10^{-4}, N=96$ ..... 71
4.8 Example 2, Chebyshev collocation method, $\epsilon=10^{-4}, N=128$ ..... 71
4.9 Example 2, Chebyshev collocation method, $\epsilon=10^{-4}, N=256$ ..... 72
4.10 Example 2, Chebyshev collocation method, $\epsilon=10^{-6}, N=96$ ..... 72
4.11 Example 2, Chebyshev collocation method, $\epsilon=10^{-6}, N=128$ ..... 73
4.12 Example 2, Chebyshev collocation method, $\epsilon=10^{-6}, N=256$ ..... 73
5.1 Example 1, singular vector ( $u_{1}$ ) versus Chebyshev points, $\epsilon=10^{-2}, N=96$ ..... 80
5.2 Example 1 , singular vector ( $u_{2}$ ) versus Chebyshev points, $\epsilon=10^{-2}, N=96$ ..... 81
5.3 Example 1, singular vector ( $u_{3}$ ) versus Chebyshev points, $\epsilon=10^{-2}, N=96$. ..... 81
5.4 Example 1, singular vector ( $u_{93}$ ) versus Chebyshev points, $\epsilon=10^{-2}, N=96$. ..... 82
5.5 Example 1, singular vector ( $u_{94}$ ) versus Chebyshev points, $\epsilon=10^{-2}, N=96$. ..... 82
5.6 Example 1, singular vector ( $u_{95}$ ) versus Chebyshev points, $\epsilon=10^{-2}, N=96$. ..... 83
5.7 Example 1, singular values $\epsilon=10^{-2}, N=96$ ..... 83
5.8 Example 1, coefficients $\left\{b_{i}\right\}$ of the equation (5.1) $\epsilon=10^{-2}, N=96 \ldots$. . . 84

## Chapter 1

## Introduction

Numerical methods for solving ordinary and partial differential equations have always been important in scientific investigations. With the advent of computers, the use of numerical methods has been popularized, and more importantly, people are now able to attack those problems which are fundamental to our understanding of scientific phenomenon, but were so much more difficult to study in the past.

There are a large number of numerical methods available for use now, and among them, spectral methods have attracted much of the attention in current research on numerical methods. In this thesis, we plan to study two specific spectral methods for solving boundary value problems. In this Chapter, we first introduce the pseudospectral Chebyshev method, and then briefly mention a quite popular numerical method for solving boundary value problems, viz., the collocation method.

### 1.1 Pseudospectral Chebyshev Methods

Spectral methods are today not as widely used in the numerical solution of (two-point) boundary value problems (BVPs), even though they are very popular for time-dependent PDE's. However, for a variety of BVPs, they are competitive with the classical finite
difference, shooting, and collocation methods (see [15]). In spectral methods, the solution is assumed to be a finite linear combination of some set of global analytic basis functions, for example, Chebyshev polynomials. The differential equation yields then a system of equations for the coefficients.

These schemes can be very efficient because the rate of convergence or the order of accuracy as the number of modes increases depends only on the smoothness of the solution.

In particular, for an analytic solution of the differential equation, the error decays exponentially (see [15]). By contrast, for example, in finite difference methods, the order of accuracy is fixed by the scheme.

For the pseudospectral Chebyshev methods, the solution is discretized at the Chebyshev collocation points. The approximate solution is forced to satisfy the equation only at the Chebyshev collocation points. This method has the advantage of being able to deal more easily with nonlinear terms than the spectral method. The basic idea of pseudospectral Chebyshev methods consists of replacing exact derivatives by derivatives of interpolating polynomials at the Chebyshev points. Unfortunately; the pseudospectral Chebyshev methods lead to very ill-conditioned matrices ([4], [17]). The condition number of the first derivative matrix is proportional to $N^{2}$, while the condition number of the second derivative related matrix increases like $N^{4}$, where N is the number of collocation points.

In order to overcome the ill-conditioning an efficient preconditioner was suggested by Orszag [17]. The resulting condition number is nearly constant with respect to N [8]. Canuto and Quarteroni [6] proposed two classes of preconditioning matrices: one arising from finite difference, the other from finite elements. Further, Orszag [17] chose a different mesh for the preconditioner. An interesting alternative procedure was proposed in [8]: the preconditioning operator computes the first derivative at the intermediate grid points, then shifts the value to the original grid points.

Berrut [4] transfers the problem into an equivalent one on the circle, and then solves the problem on the circle with the pseudospectral Fourier method so that the matrix has
condition $\mathrm{O}\left(N^{2}\right)$ for second order ordinary differential equations. We will study this "circle technique" in Chapter 3.

In this section, we will present the pseudospectral Chebyshev method. We consider a smooth function $u(x)$ in the domain $x \in[-1,1]$. The Chebyshev collocation points are

$$
x_{i}=\cos \left(\frac{\pi i}{N}\right), \quad i=0, \cdots, N
$$

which are the extrema of the Nth order Chebyshev polynomial

$$
T_{N}(x)=\cos \left(N \cos ^{-1} x\right)
$$

The function $\mathrm{u}(\mathrm{x})$ is interpolated by a polynomial $P(x),\left(P\left(x_{i}\right)=u\left(x_{i}\right)=u_{i}\right)$, of degree $\leq N$,

$$
P(x)=\sum_{j=0}^{N} u_{j} L_{j}(x)
$$

where $L_{j}$ is the polynomial of degree N with

$$
L_{j}\left(x_{k}\right)=\delta_{j k}
$$

It can be shown [14] that

$$
L_{j}(x)=\frac{(-1)^{j+1}\left(1-x^{2}\right) T_{N}^{\prime}(x)}{c_{j} N^{2}\left(x-x_{j}\right)}, \quad j=0, \cdots, N
$$

where

$$
c_{j}= \begin{cases}2, & j=0, N \\ 1, & j=1, \cdots, N-1\end{cases}
$$

The derivative of $u(x)$ at the collocation points $x_{j}$ can be approximated in many different ways. The most obvious way to compute the derivative is via matrix-vector multiplication [7]. The entries of the Chebyshev derivative matrix $D_{x}$ are computed by taking the analytical derivative of $L_{j}(x)$ and evaluating it at the collocation points $x_{k}$ for $\mathrm{j}, k=0, \cdots, i v$, i.e., $d_{k j}=L_{j}^{\prime}\left(x_{k}\right)$. Then the entries of the matrix are

$$
d_{k j}=\frac{c_{k}}{c_{j}} \frac{(-1)^{j+k}}{\left(x_{k}-x_{j}\right)}, \quad j \neq k
$$

$$
\begin{gathered}
d_{k k}=-\frac{1}{2} \frac{x_{k}}{\left(1-x_{k}^{2}\right)}, \quad k \neq 0, N \\
d_{00}=-d_{N N}=\frac{2 N^{2}+1}{6}
\end{gathered}
$$

and now the derivative of $u\left(x_{i}\right)$ becomes

$$
u_{i}^{\prime}=\sum_{j=0}^{N} d_{i j} u_{j}, \quad i=0, \cdots, N
$$

Some researchers have used other methods [21]. Variation of this matrix - vector multiplication algorithm are described in [21]; the other popular method uses the FFT (Fast Fourier Transform), and is asymptotically faster ( $\mathrm{O}(\mathrm{N} \log \mathrm{N})$ operations) than a matrix vector multiply $\left(\mathrm{O}\left(N^{2}\right)\right)$.

### 1.2 Collocation Method

In this section, we will discuss a collocation method which in some sense is equivalent to the Runge-Kutta method; for an extensive treatment (see [1]). The important difference between this method and the pseudospectral method is that it uses basis functions with local support and limited differentiability.

We are given the linear equation

$$
\begin{equation*}
L x=y \tag{1.2.1}
\end{equation*}
$$

where y is a given function. Approximating the solution $x(t)$ of (1.2.1) by the method of collocation consists of finding a function $x_{N}(t) \in \operatorname{span}\left\{\phi_{1}, \cdots, \phi_{N}\right\}$.

$$
x_{N}(t)=a_{1} \phi_{1}(t)+a_{2} \phi_{2}(t)+\cdots+a_{N} \phi_{N}(t)
$$

by solving the $N \times N$ system of linear equations

$$
\begin{equation*}
L x_{N}\left(t_{i}\right)=\sum_{j=1}^{N} a_{j} L \phi_{j}\left(t_{i}\right)=y\left(t_{i}\right) \quad 1 \leq i \leq N \tag{1.2.2}
\end{equation*}
$$

Here $t_{1}, t_{2}, \cdots, t_{N}$ are N distinct points of the domain at which all the terms of (1.2.2) are defined. The function $x_{N}(t)$, if it exists, is said to collocate $x(t)$ at the points $t_{1}, \cdots, t_{N}$.

Any function $x_{N}(t)$ so obtained is referred to as an approximate solution obtained by the method of collocation.

For example, let us consider the linear boundary value problem

$$
\begin{gathered}
L x(t)=x^{\prime \prime}(t)+p(t) x^{\prime}(t)+q(t) x(t)=f(t), \quad 0 \leq t \leq 1 \\
x(0)=x(1)=0
\end{gathered}
$$

and we assume that this problem has a unique solution $x(t)$, and that $p(t)$ and $q(t)$ are continuous on $[0,1]$. To approximate $x(t)$ by collocation using the cubic B-splines $\left\{B_{i}\right\}_{i=-1}^{n+1}$, let $\pi:=0=t_{0}<t_{1}<\cdots<t_{n}=1$. Let $x_{N} \in \operatorname{span}\left\{B_{-1}, B_{0}, \cdots, B_{n+1}\right\}$. Using collocation with these approximating functions, we seek

$$
\begin{equation*}
x_{N}=a_{-1} B_{-1}(t)+a_{0} B_{0}(t)+\cdots+a_{n+1} B_{n+1}(t) \tag{1.2.3}
\end{equation*}
$$

such that

$$
\left\{\begin{array}{l}
x_{N}(0)=0 \\
L x_{N}\left(t_{i}\right)=f\left(t_{i}\right), \quad 0 \leq i \leq n \\
x_{N}(1)=0
\end{array}\right.
$$

Notice that we are coliocating at $n+1$ knots and we force $x_{N}(t)$ to satisfy the same boundary data as $x(t)$. However, since we have $n+3$ basis functions, we have the same number of unknowns as equations.

To compute $x_{N}(t)$, we first use the linearity of $L$.
Thus for $0 \leq i \leq n$,

$$
\begin{equation*}
L x_{N}\left(t_{i}\right)=\sum_{j=-1}^{n+1} a_{j} L B_{j}\left(t_{i}\right)=f\left(t_{i}\right) \tag{1.2.4}
\end{equation*}
$$

where

$$
\begin{equation*}
L B_{j}\left(t_{i}\right)=B_{j}^{\prime \prime}\left(t_{i}\right)+p\left(t_{i}\right) B^{\prime}\left(t_{i}\right)+q\left(t_{i}\right) B_{j}\left(t_{i}\right) \tag{1.2.5}
\end{equation*}
$$

while

$$
x_{N}(0)=a_{-1}+4 a_{0}+a_{1}
$$



Figure 1.1: Cubic B-spline functions
and

$$
x_{N}(1)=a_{n-1}+4 a_{n}+a_{n+1}
$$

Recalling that $B_{i}(t)=B_{i}^{\prime}(t)=B_{i}^{\prime \prime}(t)=0$ when $t \geq t_{i+2}$ and $t \leq t_{i-2}$ for each i , we see that (1.2.4) reduces to the linear system

$$
\begin{equation*}
C_{n} a=b \tag{1.2.6}
\end{equation*}
$$

where the coefficient matrix $C_{n}$ is

$$
\left[\begin{array}{cccccccc}
1 & 4 & 1 & & & & & \\
L B_{-1}\left(t_{0}\right) & L B_{0}\left(t_{0}\right) & L B_{1}\left(t_{0}\right) & & & & & \\
& L B_{0}\left(t_{1}\right) & L B_{1}\left(t_{1}\right) & L B_{2}\left(t_{1}\right) & & & & \\
& & \ddots & \ddots & \ddots & & & \\
& & & L B_{j-1}\left(t_{j}\right) & L B_{j}\left(t_{j}\right) & L B_{j+1}\left(t_{j}\right) & & \\
& & & & \ddots & \ddots & \ddots & \\
& & & & & L B_{n-1}\left(t_{n}\right) & L B_{n}\left(t_{n}\right) & L B_{n+1}\left(t_{n}\right) \\
& & & & & 1 & 4 & 1
\end{array}\right]
$$

and $a=\left(a_{-1}, a_{0}, a_{1}, \cdots, a_{n+1}\right)^{T}$ and $b=\left(0, f\left(x_{0}\right), f\left(x_{1}\right), \cdots, f\left(x_{n}\right), 0\right)^{T} . C_{n}$ is almost tridiagonal. Eliminating $a_{-1}$ and $a_{n+1}$ leads to an $(n+1)$ by $(n+1)$ tridiagonal system. It is clear that $x_{N}(t)$ exists and is unique if and only if $C_{n}$ is nonsingular.

### 1.3 Thesis Plan

In this thesis, we will study three numerical methods for solving singularly perturbed twopoint boundary value problems of order 2. The plan of the thesis is as follows. In Chapter 2, we will study a spectral method called the pseudospectral Chebyshev method. The necessary background about Chebyshev polynomials and expansions in terms of Chebyshev polynomials is reviewed, then the pseudospectral Chebyshev method is derived and the numerical procedure for carrying out this method is explained: This is followed by an application of the pseudospectral Chebyshev method to solve three carefully chosen BVPs.

In Chapter 3, we will first study a spectral method with preliminary transformation to the circle. This method starts with transforming a problem defined on $x \in[-1,1]$ into an equivalent problem defined on the unit circle through $x=\cos \phi$; then the method proceeds to apply the Fourier spectral method to solve the transformed problem. Next, we will study two
iterative methods. The first iterative method is the conjugate gradient iteration applied to the normal equations, and the second iterative method is the biorthogonalization algorithm adapted from the biconjugate gradient iteration. We then apply the transformation method and the iterative methods to solve the three problems studied in Chapter 2.

In Chapter 4, we briefly discuss the Chebyshev collocation method and compare with the pseudospectral Chebyshev method discussed in Chapter 2.

In Chapter 5, we will discuss the numerical results obtained in the previous three Chapters comprehensively. Both merits and shortcomings with each method will be considered, together with some explanations and a mention of future research.

Ernphasis in this thesis has been given to singularly perturbed BVPs controlled by a singularity parameter $\epsilon$. Three kinds of singularities have been investigated: mild ( $\epsilon=10^{-2}$ ), stiff $\left(\epsilon=10^{-4}\right)$ and very stiff $\left(\epsilon=10^{-6}\right)$, and three types of layers in exact solutions have been considered: one layer at the center of the domain; two layers on the boundaries of the domain; and one layer on the left boundary of the domain. Numerical performance of the above mentioned methods have been judged in terms of condition numbers, average errors and cpu times. It is shown that numerical solutions are accurate despite the large condition number.

## Chapter 2

## A Pseudospectral Chebyshev

## Method

### 2.1 Introduction

There are many different numerical methods for solving differential equations, and spectral methods are now widely used. However, for boundary value problems (BVPs), the use of spectral method has not been studied in great detail. In fact, standard textbooks usually do not cover the spectral method or its variants for BVPs, although in the sixties polynomial collocation methods had received some attention for BVPs (see for example [23]). However, spline collocation methods became more popular, and renewed interest in spectral methods is more recent. In this Chapter, we study a specific variant of spectral methods, namely, the pseudospectral Chebyshev method.

The basic idea of pseudospectral Chebyshev methods consists of replacing exact derivatives by derivatives of interpolating polynomials at the Chebyshev points in the domain. Improving computational efficiency of spectral methods is one of the main purposes of the current investigation into these methods. Discrete Chebyshev differentiation can be represented by matrices, and in practice, these matrices are severely ill-conditioned and full.

Thus, it is important to find efficient techniques for solving the algebraic systems. The condition number of the matrix related to the N -point Chebyshev pseadospectral approximation of the first derivative operator is proportional to $N^{2}$, while for the matrix related to the second derivative operator, the condition number increases like $N^{4}$. Iterative methods and preconditioning techniques have been proposed to solve these spectral algebraic systems by Orszag [17], Canuto and Quarteroni [6] and other authors. We will discuss iterative methods in Chapter 3. In this Chapter, we consider the direct method, namely, the Gaussian elimination method.

### 2.2 Chebyshev Polynomial Expansions

First, we review several results from approximation theory. The Chebyshev polynomial of degree $\mathrm{k}(k=0,1, \cdots)$ on $[-1,1]$ is defined by the formula

$$
\begin{equation*}
T_{k}=\cos \left(k \cos ^{-1} x\right) \tag{2.2.1}
\end{equation*}
$$

Clearly, $\left|T_{k}(x)\right| \leq 1$ for $x \in[-1,1]$. The $T_{k}$ are indeed polynomials in $x$. For example,

$$
\begin{equation*}
T_{0}(x)=1, \quad T_{1}(x)=x \tag{2.2.2}
\end{equation*}
$$

by definition, and using elementary trigonometric identities, we can obtain the recursion

$$
\begin{equation*}
T_{k+1}(x)=2 x T_{k}(x)-T_{k-1}(x) \quad k \geq 1 \tag{2.2.3}
\end{equation*}
$$

Let $L_{[-1,1]}^{2}$ be the space of square integrable functions defined on $[-1,1]$. Then the furictions $T_{k}$ constitute an orthonormal basis with respect to the inner product

$$
\begin{equation*}
(f, g)=\int_{-1}^{1} f(x) g(x)\left(1-x^{2}\right)^{-1 / 2} d x \tag{2.2.4}
\end{equation*}
$$

in $L_{[-1,1]}^{2}$. The convergence theory of Chebyshev polynomial expansions is very similar to that of Fourier cosine series. In fact, suppose for $f(x) \in L_{[-1,1]}^{2}$, we write formally

$$
\begin{equation*}
g(x)=\sum_{k=0}^{\infty} a_{k} T_{k}(x) \tag{2.2.5}
\end{equation*}
$$

the so-called the Chebyshev series associated with $f(x)$, where $a_{k}$ is to be determined, then $G(\theta):=g(\cos \theta)$ is the Fourier cosine series of $F(\theta):=f(\cos \theta)$ for $0 \leq \theta \leq \pi$. This result follows from the definition of $T_{k}$, because $T_{k}(\cos \theta)=\cos (k \theta), G(\theta)=g(\cos \theta)=$ $\sum_{k=0}^{\infty} a_{k} \cos (k \theta)$. Thus,

$$
\begin{equation*}
a_{k}=\frac{2}{\pi c_{k}} \int_{0}^{\pi} f(\cos \theta) \cos (k \theta) d \theta=\frac{2}{\pi c_{k}} \int_{-1}^{1} f(x) T_{k}(x)\left(1-x^{2}\right)^{-1 / 2} d x \tag{2.2.6}
\end{equation*}
$$

where $c_{0}=2, \quad c_{k}=1(k>0)$.
It follows from this close relation between Chebyshev series and Fourier cosine series that if $\mathrm{f}(\mathrm{x})$ is piecewise continuous and if $\mathrm{f}(\mathrm{x})$ is of bounded total variation for $-1 \leq x \leq 1$ then $g(x)=\frac{1}{2}[f(x+)+f(x-)]$ for each $x(-1<x<1)$ and $g(1)=f(1-), g(-1)=f(-1+)$. Also, if $f^{(p)}(x)$ is continuous for all $|x| \leq 1$ for $p=0,1, \ldots \ldots, n-1$, and $f^{(n)}(x)$ is integrable, then

$$
\begin{equation*}
a_{k}=O\left(k^{-n}\right) \tag{2.2.7}
\end{equation*}
$$

Since $\left|T_{k}(x)\right| \leq 1$ for $|x| \leq 1$, it follows that the remainder after k terms of the Chebyshev series (2.2.5) is asymptotically much smaller than $k^{-(n-1)}$ as $k \rightarrow \infty$. If $f(x)$ is infinitely differentiable for $|x| \leq 1$, the error in the Chebyshev series goes to zero more rapidly than any finite power of $k^{-1}$ as $k \rightarrow \infty$ [15].

The most important feature of Chebyshev series is that their convergence properties are not affected by the values of $\mathrm{f}(\mathrm{x})$ or its derivatives at the boundaries $x= \pm 1$, but only by the smoothness of $f(x)$ and its derivatives throughout $-1<x<1$. In contrast, the Gibbs phenomenon shows that the rate of convergence of Fourier series depends on the value of $f(x)$ and its derivatives at the boundaries in addition to the smoothness of $f(x)$ and its derivative in the interior of the interval. The reason for the absence of a Gibbs phenomenon for the Chebyshev series of $f(x)$ at $x= \pm 1$ is due to the fact that $F(\theta)=f(\cos \theta)$ satisfies $F_{2 p+1}(0)=F_{2 p+1}(\pi)=0$ provided only that all derivatives of $\mathrm{f}(\mathrm{x})$ of order at most $2 p+1$ exist at $x= \pm 1$.

An important consequence of the rapid convergence of Chebyshev polynomial expansions
of smooth functions is that Chebyshev expansions may normally be differentiated termwise. Since

$$
\begin{equation*}
\frac{d^{p}}{d x^{p}} T_{k}(x)=O\left(k^{2 p}\right), \quad k \rightarrow \infty \tag{2.2.8}
\end{equation*}
$$

uniformly for $|x| \leq 1$, if $a_{k} \rightarrow 0$ faster than any finite power of $k^{-1}$ as $k \rightarrow \infty$, then (2.2.5) may be differentiated formally [15]:

$$
\begin{equation*}
\frac{d^{p} g(x)}{d x^{p}}=\sum_{k=0}^{\infty} a_{k} \frac{d^{p} T_{k}(x)}{d x^{p}} \tag{2.2.9}
\end{equation*}
$$

### 2.3 A Pseudospectral Chebyshev Method

In this section, we consider a pseudospectral Chebyshev method for the linear second order two-point boundary value problem,

$$
\left\{\begin{array}{l}
\epsilon U^{\prime \prime}(x)+p(x) U^{\prime}+q(x) U(x)=f(x)  \tag{2.3.10}\\
U(1)=U_{0}, \quad U(-1)=U_{N}
\end{array}\right.
$$

where $\epsilon$ is a parameter that controls the singularity of the problem. Our discussion will follow Berrut [4], who formulated the algorithm for $\epsilon=1$. Suppose that the problem is well-posed and in particular that it has a unique solution. The Chebyshev interpolation polynomial can be written as

$$
\begin{equation*}
\bar{U}(x)=\sum_{j=0}^{N} \bar{U}_{j} L_{j}(x) \tag{2.3.11}
\end{equation*}
$$

where

$$
x_{j}=\cos \frac{j \pi}{N}, \quad j=0,1, \cdots, N
$$

are the interpolating points (also called collocation points), $\left\{\bar{U}_{j}\right\}_{j=1}^{N-1}$ are the unknown coefficients to be determined, and $L_{j}(x)$ is the Lagrange interpolation polynomial associated with $x_{j}$, that is,

$$
\begin{equation*}
L_{j}(x)=\prod_{i=0, i \neq j}^{N} \frac{\left(x-x_{i}\right)}{\left(x_{j}-x_{i}\right)} \tag{2.3.12}
\end{equation*}
$$

Since $L_{j}\left(x_{k}\right)=\delta_{j k}\left(\delta_{j k}\right.$ is the Kronecker delta), it follows that $\bar{U}\left(x_{j}\right)=\bar{U}_{j}$.


Figure 2.1: Distribution of Chebyshev collocation points, $N=32$

Here we analyze the relation between the number of the collocation points and the width $O(\epsilon)$ of the boundary layer.

Since

$$
\begin{aligned}
\Delta x_{1} & =\left|x_{1}-x_{0}\right|=|\cos (\pi / N)-1| \\
& =\left|1-\sin (0)(\pi / N)-\frac{1}{2} \cos (0)(\pi / N)^{2}+\cdots-1\right| \\
& \approx \frac{1}{2} \frac{\pi^{2}}{N^{2}} \approx \frac{5}{N^{2}}, \\
\Delta \bar{x} & =\left|\cos \left(\frac{\pi}{N}+\frac{\pi}{2}\right)-0\right| \approx \frac{\pi}{N} .
\end{aligned}
$$

Therefore we can see that the spacing between the collocation points near the boundary is $\Delta x_{1} \approx O\left(\frac{1}{N^{2}}\right)$ and the spacing near the center of the interval is $\Delta \tilde{x} \approx O\left(\frac{1}{N}\right)$, while the width of the boundary layer is $O(\epsilon)$. For good resolution of our mumerical solutions at least one of the collocation points should lie in the boundary layer (or any transition layer). Therefore we need $5 / N^{2}<\epsilon$, i.e., $N>\sqrt{5} / \sqrt{\epsilon}$. If the width of the boundary layer was exactly $\epsilon$, we should use $N \approx 2200,220$ and 22 for $\epsilon=10^{-6}, 10^{-4}$ and $10^{-2}$ respectively.

In our numerical computations we never used N larger than 1024 , and therefore, were not able to resolve the boundary layers for $\epsilon=10^{-6}$. See [24] for a method which achieves resolution for smaller values of $N$ in the case of small $\epsilon$.

The Chebyshev method substitutes the interpolation polynomial (2.3.11) into (2.3.10), and then replaces $\widetilde{U}\left(x_{0}\right)$ by $U_{0}$ and $\bar{U}\left(x_{N}\right)$ by $U_{N}$. Doing this, we have

$$
\left\{\begin{array}{l}
\epsilon \bar{U}^{\prime \prime}(x)+p(x) \bar{U}^{\prime}(x)+q(x) \bar{U}(x)=f(x)  \tag{2.3.13}\\
\bar{U}(1)=U_{0}, \quad \bar{U}(-1)=U_{N}
\end{array}\right.
$$

Using collocation points $\left\{x_{j}\right\}_{j=1}^{N-1}$, we have collocation equations

$$
\left\{\begin{array}{l}
\epsilon \bar{U}^{\prime \prime}\left(x_{j}\right)+p\left(x_{j}\right) \bar{U}^{\prime}\left(x_{j}\right)+q\left(x_{j}\right) \vec{U}\left(x_{j}\right)=f\left(x_{j}\right)  \tag{2.3.14}\\
\bar{U}(1)=U_{0}, \quad \bar{U}(-1)=U_{1} .
\end{array}\right.
$$

Denoting by $D_{x}$ the matrix of the first derivatives $L_{i}^{\prime}\left(x_{j}\right)(i, j=0,1, \cdots, N)$, and letting $\bar{U}^{\prime}\left(x_{j}\right)=\bar{U}_{j}^{\prime}(j=0,1, \cdots, N), \bar{U}\left(x_{j}\right)=\bar{U}_{j}(j=0,1, \cdots, N)$ gives

$$
\left[\begin{array}{c}
\bar{U}_{0}^{\prime} \\
\vdots \\
\bar{U}_{N}^{\prime}
\end{array}\right]=D_{x}\left[\begin{array}{c}
\bar{U}_{0} \\
\vdots \\
\bar{U}_{N}
\end{array}\right]
$$

where

$$
\begin{aligned}
D_{x} & =\left(L_{i}^{\prime}\left(x_{j}\right)\right)=\left(d_{i j}\right) \\
d_{i j} & =\frac{(-1)^{j+i}}{x_{i}-x_{j}} \cdot \frac{c_{i}}{c_{j}}, \quad i \neq j \\
d_{00} & =\frac{2 N^{2}+1}{6}=-d_{N N} \\
d_{i i} & =\frac{-x_{i}}{2\left(1-x_{i}^{2}\right)}, \quad 1 \leq i \leq N-1 \\
c_{0} & =c_{N}=2, \\
c_{i} & =1, \quad 1 \leq i \leq N-1
\end{aligned}
$$

In view of the fact that the matrix of the second derivatives $L_{j}^{\prime \prime}\left(x_{i}\right)$ is the square of $D_{x}$ (see [4]), we obtain the system of linear equations

$$
\begin{align*}
\sum_{j=1}^{N-1}\left[\epsilon\left(D_{x}^{2}\right)_{i j}+p_{i}\left(D_{x}\right)_{i j}+q_{i} \delta_{i j}\right] \bar{U}\left(x_{i}\right)= & f\left(x_{i}\right)-\left[\epsilon\left(D_{x}^{2}\right)_{i 0}+p_{i}\left(D_{x}\right)_{i 0}\right] U_{\mathrm{C}} \\
& -\left[\epsilon\left(D_{x}^{2}\right)_{i N}+p_{i}\left(D_{x}\right)_{i N}\right] U_{N} \tag{2.3.15}
\end{align*}
$$

for the $\bar{U}\left(x_{i}\right), i=1,2, \cdots, N-1$, with $p_{i}:=p\left(x_{i}\right), q_{i}:=q\left(x_{i}\right)$. The largest eigenvalue modulus of $D_{x}^{2}$ is known to grow with N like $O\left(N^{4}\right)$ (see [4]), and in the example we consider in section 2.5 , the numerical value of that modulus grows about that fast.

For simplicity, we write equation (2.3.15) in matrix form

$$
\begin{equation*}
A \bar{U}=b \tag{2.3.16}
\end{equation*}
$$

where

$$
A=\left(a_{i j}\right) \in R^{(N-1) \times(N-1)}
$$

with

$$
\begin{aligned}
a_{i j} & =\epsilon\left(D_{x}^{2}\right)_{i j}+p_{i}\left(D_{x}\right)_{i j}+q_{i} \delta_{i j} . \\
\bar{U} & =\left(\bar{U}_{1}, \cdots, \bar{U}_{N-1}\right)^{T} \in R^{N-1} \\
b & =\left(b_{1}, b_{2}, \cdots, b_{N-1}\right)^{T} \in R^{N-1}
\end{aligned}
$$

with

$$
b_{i}=f\left(x_{i}\right)-\left[\epsilon\left(D_{x}^{2}\right)_{i 0}+p_{i}\left(D_{x}\right)_{i 0}\right] U_{0}-\left[\epsilon\left(D_{x}^{2}\right)_{i N}+p_{i}\left(D_{x}\right)_{i N}\right] U_{N}
$$

For a fixed integer $N$, solving system (2.3.16) for $\bar{U}$ and substituting $\bar{U}, U_{0}$ and $U_{N}$ in to (2.3.11) will lead to our numerical solution $\bar{U}(x)$ to the BVP (2.3.10).

### 2.4 Linear System Solver

The LINPACK [7] subroutines operate on general square nonsymmetric matrices. The operations performed include the triangular factorization of matrices, the estimation of the
matrix condition number, the solution of simultaneous linear equations and the calculation of determinants and inverses. Let $A$ be a real or complex square matrix of order $n$. There is an upper triangular matrix U and a matrix L which is the product of elementary lower triangular and permutation matrices such that $A=L U$. This factorization can be used to solve linear equations $A x=b$ by solving successively $L(U x)=b$ to compute the inverse of A as $A^{-1}=U^{-1} L^{-1}$.

The condition number $k(A)$ is a quantity which measures the sensitivity of the solution $x$ to errors in the matrix $A$ and the right hand side $b$. If the relative error in $A$ is of size $\epsilon$, then the resulting relative error in x can be as large as $k(A) \epsilon$. Error in A can arise in many ways. In particular, the effect of the roundoff error introduced by the subroutines in this Chapter can usually be assessed by taking $\epsilon$ to be a small multiple of the rounding unit.

It is possible to efficiently compute a quantity RCOND which is an estimate of the reciprocal condition, $1 / k(A)$. If exact singularity is detected, RCOND may be set to 0 .

If $A$ is badly scaled, then the interpretation of RCOND is more delicate.
We use the double precision, general subroutines DGECO, DGEFA to solve linear system $A x=b$. DGECO is usually called first to factor the matrix and estimate its condition. The actual factorization is done by DGEFA which can be called in place of DGECO if the condition estimate is not needed. The time required by DGECO is roughly $(1+9 / N)$ times the time required by DGEFA. Thus when $N=9$, DGECO costs twice as much as DGEFA, but when $N=90$, DGECO costs only 10 percent more.

Roundoff errors in floating point arithmetic operations usually cause the quantities computed by the subroutines in this Chapter to be somewhat inaccurate. The following fairly vague statements give a general idea of the extent of these inaccuracies.

DGEFA produces matrices $L$ and $U$ for which the product $L U$ is almost always within roundoff error of $A$, no matter how close $A$ is to being singular.

DGECO has experimentally been shown to produce an estimate RCOND for which $1 / R C O N D$ is usually the same order of magnitude as the actual condition number.

To make these statements more precise, we introduce the following notation. let $a_{j}$ be the columns of A. For $x \in R^{n}$, define

$$
\begin{gathered}
\|x\|_{1}=\sum_{j=1}^{n}\left|x_{j}\right| \\
\|A x\|_{1}=\sum_{i=1}^{n}\left|\sum_{j=1}^{n} a_{i j} x_{j}\right| \leq \sum_{i=1}^{n} \max _{j}\left|a_{i j}\right| \sum_{j=1}^{n}\left|x_{j}\right| \leq\left(\sum_{i=1}^{n} \max _{j}\left|a_{i j}\right|\right)\|x\|_{1} \\
\|A\|_{1}=\max _{x} \frac{\|A x\|_{1}}{\|x\|_{1}}=\max _{j}\left\|a_{j}\right\|_{1}
\end{gathered}
$$

Then the condition number of $A$ with respect to $\|\cdot\|_{1}$ is

$$
k_{1}(A)=\|A\|_{1}\left\|A^{-1}\right\|_{1}
$$

If errors are measured in the usual Euchidean norm,

$$
\|x\|_{2}=\left(\sum_{j=1}^{n} x_{j}^{2}\right)^{1 / 2}
$$

then the condition number is

$$
k_{2}(A)=\|A\|_{2}\left\|A^{-1}\right\|_{2}=\frac{\sigma_{1}(A)}{\sigma_{n}(A)}
$$

Here $\sigma_{1}(A)$ and $\sigma_{n}(A)$ are the largest and smallest singular value of $\Lambda . k_{1}(A)$ and $k_{2}(A)$ are different, but are usually of the same order of magnitude.

### 2.5 Numerical Results

The pseudospectral Chebyshev method discussed in section 2.3 was seldom applied to solve singularly perturbed BVPs. For such problems, we need to consider two regions: narrow regions of very fast variation (so-called boundary or interior layers), and wider regions of slow variation. The concept of singularly perturbed BVP relates to the concept of stiff BVP in numerical analysis. All the computations in this thesis were done on a Sparc station. In our Tables, the average errors are computed by

$$
\text { Average Error }=\frac{1}{N} \sum_{j=1}^{N}\left|u\left(x_{j}\right)-u_{e x a c t}\left(x_{j}\right)\right|
$$

where N is the number of terms used in the Chebyshev series, $\mathrm{u}(\mathrm{x})$ is the numerical solution and $u_{\text {exact }}$ is an exact solution. The condition number (RCOND) of A is obtained by LINPACK [7]

$$
k_{1}(A)=\|A\|_{1}\left\|A^{-1}\right\|_{1}
$$

## Example 1.

Consider

$$
\left\{\begin{array}{l}
\epsilon U^{\prime \prime}(x)+x U^{\prime}(x)=-\epsilon \pi^{2} \cos (\pi x) \sin (\pi x), \quad-1 \leq x \leq 1  \tag{2.5.17}\\
U(-1)=-2, \quad U(1)=0
\end{array}\right.
$$

whose solution is

$$
\begin{equation*}
U(x)=\cos (\pi x)+\operatorname{erf}(x / \sqrt{2 \epsilon}) / \operatorname{erf}(1 / \sqrt{2 \epsilon}) \tag{2.5.18}
\end{equation*}
$$

where

$$
\operatorname{erf}(x)=\frac{2}{\sqrt{\pi}} \int_{0}^{x} \exp \left(-t^{2}\right) d t
$$

is the error function. For $0 \leq \epsilon \ll 1$ this solution has a rapid transition layer at $x=0$.
COLSYS [1] has no problem for $\epsilon=10^{-4}$ to reach error $O\left(10^{-9}\right)$ with mesh points $\mathrm{N}=$ 132, but if the problem is run using no mesh selection, comparable accuracy is not achieved until $\mathrm{N}=1280$ [1]. The interior layer gets resolved well, because one of the collocation points is exactly at the location of the layer. Usually interior layers pose more difficulty.

We first tested $\epsilon=10^{-2}$. For $N=32,64$ and 96 , we plotted our numerical solutions (the solid-dotted lines) and the exact solution (the dotted line) in Figure 2.1 to Figure 2.3. As can be seen, our numerical solutions captured all the properties of the exact solution for $N$ as small as 64 . See Table 2.1 for a summary of the related condition numbers, average errors and cpu times. It is interesting that for $\epsilon=10^{-2}$ the condition numbers of the matrix are much smaller than the theoretical $O\left(N^{4}\right)$.

Next, we tested $\epsilon=10^{-4}$. In this case, the rapid transition of the exact solution at $x=0$ showed up moderately. We plotted our numerical solutions in Figure 2.4 to Figure 2.7 for $N=32,64,96$ and 128 , respectively. The approximation to the exact solution improves as
$N$ increases. However, even when $N=128$, there is noticeable oscillation near $x=0$. Table 2.2 contains the relevant numerical summary. Still, the condition numbers are smaller than $O\left(N^{4}\right)$.

Finally, we tested $\epsilon=10^{-6}$. In this case, the exact solution has a sharp change at $x=0$, and it took $N$ as large as 128 to reach a fairly good numerical solution. For this stiff problem (i.e. $\epsilon=10^{-6}$ ), see Figure 2.8 to Figure 2.11 and Table 2.3.


Figure 2.2: Example 1, pseudospectral Chebyshev method, $\epsilon=10^{-2}, N=32$


Figure 2.3: Example 1, pseudospectral Chebyshev method, $\epsilon=10^{-2}, N=64$


Figure 2.4: Example 1, pseudospectral Chebyshev method, $\epsilon=10^{-2}, N=96$


Figure 2.5: Example 1, pseudospectral Chebyshev method, $\epsilon=10^{-4}, N=32$


Figure 2.6: Example 1, pseudospectral Chebyshev method, $\epsilon=10^{-4}, N=64$


Figure 2.7: Example 1, pseudospectral Chebyshev method, $\epsilon=10^{-4}, N=96$


Figure 2.8: Example 1, pseudospectral Chebyshev method, $\epsilon=10^{-4}, N=128$


Figure 2.9: Example 1, pseudospectral Chebyshev method, $\epsilon=10^{-6}, N=32$


Figure 2.10: Example 1, pseudospectral Chebyshev method, $\epsilon=10^{-6}, N=64$


Figure 2.11: Example 1, pseudospectral Chebyshev method, $\epsilon=10^{-6}, N=90$


Figure 2.12: Example 1, pseudospectral Chebyshev method, $\epsilon=10^{-6}, N=128$

| N | Condition Number | Average Error | Cpu-Time |
| :---: | :---: | :---: | :---: |
| 32 | $6.4 \times 10^{2}$ | $1.23 \times 10^{-4}$ | 0.3 |
| 64 | $8.3 \times 10^{3}$ | $4.18 \times 10^{-11}$ | 1.58 |
| 96 | $1.4 \times 10^{5}$ | $4.59 \times 10^{-14}$ | 5.62 |
| 128 | $1.2 \times 10^{5}$ | $6.42 \times 10^{-14}$ | 12.53 |

Table 2.1: Example 1, pseudospectral Chebyshev method, $\epsilon=10^{-2}$

| N | Condition Number | Average Error | Cpu-Time |
| :---: | :---: | :--- | :---: |
| 32 | $4.5 \times 10^{4}$ | $4.38 \times 10^{-2}$ | 0.28 |
| 64 | $7.7 \times 104$ | $2.15 \times 10^{-2}$ | 1.58 |
| 96 | $1.3 \times 10^{5}$ | $1.17 \times 10^{-2}$ | 5.68 |
| 128 | $1.4 \times 10^{5}$ | $6.19 \times 10^{-3}$ | 12.66 |

Table 2.2: Example 1, pseudospectral Chebyshev method, $\epsilon=10^{-4}$

| N | Condition Number | Average Error | Cpu-Time |
| :---: | :---: | :---: | :---: |
| 32 | $5.8 \times 10^{6}$ | $4.65 \times 10^{-2}$ | 0.41 |
| 64 | $1.2 \times 10^{7}$ | $2.69 \times 10^{-2}$ | 1.88 |
| 96 | $1.8 \times 10^{7}$ | $1.95 \times 10^{-2}$ | 5.68 |
| 128 | $2.4 \times 10^{7}$ | $1.54 \times 10^{-2}$ | 12.64 |

Table 2.3: Example 1, pseudospectral Chebyshev method, $\epsilon=10^{-\epsilon}$

## Example 2.

As the second example, we considered the BVP

$$
\left\{\begin{array}{l}
\epsilon U^{\prime \prime}(x)=x U^{\prime}(x)+\frac{1}{2} U(x), \quad-1 \leq x \leq 1  \tag{2.5.19}\\
U(-1)=1, \quad U(1)=2
\end{array}\right.
$$

which has the exact solution

$$
\begin{equation*}
u(x)=e^{-(x+1) / \epsilon}+2 e^{(x-1) / \epsilon}+O(\epsilon) \tag{2.5.20}
\end{equation*}
$$

whese $O(\epsilon)$ denotes an quantity which can be bounded by a constant times $\epsilon$. This solution has two boundary layers at the ends, and is smooth near the "turning point" $x=0$. The numerical error propagates from the boundaries to the middle of the interval [-1, 1]. This equation has tendency for an interior layer. Maybe oscillations in interior are related to this fact. COLSYS succeeded in solving this problem for $\epsilon=10^{-6}$ although the error is not too small (See [1]).

Again, we considered three different cases corresponding to $\epsilon=10^{-2}, \epsilon=10^{-4}$, and $\epsilon=10^{-6}$, respectively. When $\epsilon=10^{-2}$, accurate numerical approximation to the true solution was easy to obtain; see Figure 2.12, Figure 2.13 and Table 2.4. When $\epsilon=10^{-4}$, it required a relatively large $N$ to obtain a reasonable numerical solution; see Figure 1.14 to Figure 2.16 for $n=128,256,512$, respectively and Table 2.5. However, for $N=512$, a fairly good numerical solution was achieved. When $\epsilon=10^{-6}$, it was difficult to reach good numerical solutions even for large $N$. It is shown that the pseudospectral Chebyshev method fails to solve this two boundary layers problem with small $\epsilon$ (e.g. $10^{-6}$ ), while it succeeds in solving it with $\epsilon=10^{-4}$. See Figures 2.17 and 2.18. and Table 2.6. However, it appears reasonable to expect highly accurate solution for sufficiently large $\mathbf{N}$, even for small $\epsilon\left(\epsilon=10^{-6}\right)$.


Figure 2.13: Example 2, pseudospectral Chebyshev method, $\epsilon=10^{-2}, N=32$


Figure 2.14: Example 2, pseudospectral Chebyshev method, $\epsilon=10^{-2}, N=64$


Figure 2.15: Example 2, pseudospectral Chebyshev method, $\epsilon=10^{-4}, N=128$


Figure 2.16: Example 2, pseudospectral Chebyshev method, $\epsilon=10^{-4}, N=256$


Figure 2.17: Example 2, pseudospectral Chebyshev method, $\epsilon=10^{-4}, N=512$


Figure 2.18: Example 2, pseudospectral Chebyshev method, $\epsilon=10^{-6}, N=128$


Figure 2.19: Example 2, pseudospectral Chebyshev method, $\epsilon=10^{-6}, N=256$

| N | Condition Number | Average Error | Cpu-Time |
| :---: | :---: | :--- | :---: |
| 32 | $3.0 \times 10^{3}$ | $2.10 \times 10^{-3}$ | 0.39 |
| 64 | $5.2 \times 10^{4}$ | $7.89 \times 10^{-4}$ | 1.89 |
| 96 | $2.7 \times 10^{5}$ | $7.85 \times 10^{-4}$ | 5.64 |
| 128 | $8.8 \times 10^{5}$ | $7.83 \times 10^{-4}$ | 12.77 |

Table 2.4: Example 2, pseudospectral Chebyshev method, $\epsilon=10^{-2}$

| N | Condition Number | Average Error | Cpu-Time |
| :---: | :---: | :--- | :---: |
| 128 | $3.3 \times 10^{4}$ | $2.10 \times 10^{0}$ | 14.57 |
| 256 | $2.5 \times 10^{5}$ | $2.43 \times 10^{-2}$ | 102.58 |
| 512 | $3.9 \times 10^{6}$ | $1.30 \times 10^{-6}$ | 804.09 |

Table 2.5: Example 2, pseudospectral Chebyshev method, $\epsilon=10^{-4}$

| $N$ | Condition Number | Average Error | Cpu-Time |
| :---: | :---: | :--- | :---: |
| 128 | $7.1 \times 10^{3}$ | $1.62 \times 10^{0}$ | 12.71 |
| 256 | $2.3 \times 10^{4}$ | $2.43 \times 10^{0}$ | 93.08 |
| 512 | $1.1 \times 10^{5}$ | $1.32 \times 10^{0}$ | 730.58 |

Table 2.6: Example 2, pseudospectral Chebyshev method, $\epsilon=10^{-6}$

## Example 3.

As the third example, we considered a problem with one boundary layer

$$
\left\{\begin{array}{l}
2 \epsilon U^{\prime \prime}(x)+U^{\prime}(x)=0, \quad-1 \leq x \leq 1  \tag{2.5.21}\\
U(-1)=0, \quad U(1)=1
\end{array}\right.
$$

which has the exact solution

$$
U(x)=\left\{1-\exp \left(-\frac{x+1}{2 \epsilon}\right)\right\} /\left\{1-\exp \left(-\frac{1}{\epsilon}\right)\right\} .
$$

Similar to the previous two examples, when $\epsilon=10^{-2}$, our numerical solutions became very accurate for $N$ as small as 64 ; see Figure 2.19, Figure 2.20 and Table 2.7. When $\epsilon=10^{-4}$, a large $N$ was needed to produce a fairly good numerical solution; see Figure 2.21, Figure 2.22 and Table 2.8. When $\epsilon=10^{-6}$, our numerical solutions were not close to the exact solution for $N$ as large as 512; see Figure 2.23, Figure 2.24 and Table 2.9. For this problem, the condition numbers are much smaller than $O\left(N^{4}\right)$, because there is a small $\epsilon$ multiplying the matrix $D_{x}^{2}$.

Finally, we plot some pictures of the error versus N for Example 1 and 2, as shown in Figure 2.25 to Figure 2.30. The errors become smaller as N is larger.


Figure 2.20: Example 3, pseudospectral Chebyshev method, $\epsilon=10^{-2}, N=32$


Figure 2.21: Example 3, pseudospectral Chebyshev method, $\epsilon=10^{-2}, N=64$


Figure 2.22: Example 3, pseudospectral Chebyshev method, $\epsilon=10^{-4}, N=128$


Figure 2.23: Example 3, pseudospectral Chebyshev method, $\epsilon=10^{-4}, N=256$


Figure 2.24: Example 3, pseudospectral Chebyshev method, $\epsilon=10^{-6}, N=256$


Figure 2.25: Example 3, pseudospectral Chebyshev method, $\epsilon=10^{-6}, N=512$

| N | Condition Number | Average Error | Cpu-Time |
| :---: | :---: | :---: | :---: |
| 32 | $1.7 \times 10^{3}$ | $4.52 \times 10^{-6}$ | 0.36 |
| 64 | $2.4 \times 10^{4}$ | $4.36 \times 10^{-15}$ | 1.90 |
| 96 | $1.2 \times 10^{5}$ | $3.25 \times 10^{-14}$ | 5.59 |
| 128 | $3.7 \times 10^{5}$ | $4.96 \times 10^{-14}$ | 12.69 |

Table 2.7: Example 3, pseudospectral Chebyshev method, $\epsilon=10^{-2}$

| N | Condition Number | Average Error | Cpu-Time |
| :---: | :---: | :--- | :---: |
| 32 | $1.4 \times 10^{3}$ | $0.25 \times 10^{1}$ | 0.39 |
| 64 | $2.0 \times 10^{3}$ | $0.59 \times 10^{0}$ | 1.86 |
| 96 | $3.7 \times 10^{3}$ | $0.22 \times 10^{0}$ | 5.58 |
| 128 | $7.5 \times 10^{3}$ | $0.81 \times 10^{-1}$ | 12.47 |
| 256 | $7.0 \times 10^{4}$ | $0.30 \times 10^{-3}$ | 93.52 |

Table 2.8: Example 3, pseudospectral Chebyshev method, $\epsilon=10^{-4}$

| N | Condition Number | Average Error | Cpu-Time |
| :---: | :---: | :--- | :---: |
| 256 | $2.3 \times 10^{4}$ | $1.54 \times 10^{0}$ | 94.07 |
| 512 | $1.3 \times 10^{5}$ | $1.32 \times 10^{0}$ | 749.28 |

Table 2.9: Example 3, pseudospectral Chebyshev method, $\epsilon=10^{-6}$


Figure 2.26: Example 1, average error $\left(* 10^{12}\right)$ for $\epsilon=10^{-2}, N=32,64,96,128$


Figure 2.27: Example 1, average error $\left(* 10^{2}\right)$ for $\epsilon=10^{-4}, N=32,64,96,128$


Figure 2.28: Example 1, average error $\left(* 10^{2}\right)$ for $\epsilon=10^{-6}, N=32,64,96,128$


Figure 2.29: Example 2, average error $\left(* 10^{3}\right)$ for $\epsilon=10^{-2}, N=32,64,96,128$


Figure 2.30: Example 2, average error for $\epsilon=10^{-4}, N=128,256,512$


Figure 2.31: Example 2, average error for $\epsilon=10^{-6}, N=128,256,512$

## Chapter 3

## Preliminary Transform To The

## Circle

### 3.1 Introduction

In our study of the pseudospectral Chebyshev method for linear ODEs of order 2 in Chapter 2, we must solve a system of linear equations whose condition number grows with the number of collocation points N like $O\left(N^{4}\right)$. This property arises from the differentiation of the Chebyshev interpolating polynomials. When a relatively large $N$ is needed to produce a fairly good numerical solution, the related condition zumber is large, indicating relatively low precision in the numerical solution. Therefore, various efforts have been made to reduce the condition number whenever possible.

In this Chapter, we will first discuss a variant of the peudospectral Chebyshev method studied in Chapter 2, which was suggested by Berrut . The idea is simple: first, the problem is replaced by an equivalent one on the circle using a transformation; secondly, the transformed problem on the circle is solved with the pseudospectral Fourier method. It is hoped that the related matrices resulted from doing this will have smaller ondition numbers like $O\left(N^{2}\right)$; see [4]. Following this, we will discuss two iterative methods briefly.

### 3.2 The Method of Transformation to the Circle

For simplicity, we consider a first order problem: Consider the initial value problem

$$
\left\{\begin{array}{l}
U^{\prime}(x)=F(x)  \tag{3.2.1}\\
U(1)=0
\end{array}\right.
$$

If the pseudospectral Chebyshev method is used to find the Chebyshev interpolation polynomials

$$
\tilde{U}(x)=\sum_{j=0}^{N} \tilde{U}\left(x_{j}\right) L_{j}(x)
$$

and if the Chebyshev points

$$
x_{j}=\cos \left(\frac{j \pi}{N}\right), \quad j=0,1, \cdots, N
$$

are used as collocation points, one will reach a system of linear equations

$$
\begin{equation*}
A \bar{U}=b \tag{3.2.2}
\end{equation*}
$$

where A is $D_{x}$ defined in section $2.3, b=\left(F\left(x_{1}\right), \cdots, F\left(x_{N}\right)\right)^{T}, \bar{U}=\left(\bar{U}\left(x_{1}\right), \cdots, \bar{U}\left(x_{N}\right)\right)^{T}$. The direct solution to (2.2.2) requires $O\left(N^{3}\right)$ operations. The condition number of the matrix A grows with N like $O\left(N^{2}\right)$ (see [2]).

We now define the process of transformation to the circle with $\phi$ denoting the angle; we have

$$
\begin{aligned}
x & =\cos \phi \\
u(\phi) & :=U(\cos \phi), \\
f(\phi) & :=F(\cos \phi) .
\end{aligned}
$$

Thus $u(\phi)$ and $f(\phi)$ are $2 \pi$-periodic even functions on the full unit circle by defining $f(2 \pi-\phi)=f(\phi)$.

With these definitions, problem (3.2.1) now becomes

$$
\left\{\begin{array}{l}
-u^{\prime}(\phi) / \sin (\phi)=f(\phi)  \tag{3.2.3}\\
u(0)=0
\end{array}\right.
$$

or, multiplying both sides by $\sin 0$.

$$
\left\{\begin{array}{l}
u^{\prime}(\phi)=-f(\phi) \sin \phi=: h(\phi)  \tag{3.2.4}\\
u(0)=0
\end{array}\right.
$$

It is clear from the definitions that all functions appearing in (3.2.4) are $2 \pi$-periodic and inherit the differentiability properties of the corresponding functions in the original problem. More importantly, problem (3.2.4) can be solved by the pseudospectral Fourier method as described below.

Interpolating the trigonometric polynomials of degree N between the equidistant points

$$
\phi_{j}=j \frac{2 \pi}{2 N}, \quad j=0,1, \cdots, 2 N-1
$$

and collocating at the $\phi_{j}$ yield the linear system

$$
\begin{equation*}
\sum_{n=0}^{2 N-1}\left(D_{\phi}\right)_{j n} \bar{u}\left(\phi_{n}\right)=h\left(\phi_{j}\right) \tag{3.2.5}
\end{equation*}
$$

for the approximate values $\bar{u}\left(\varphi_{0}\right), \cdots, \bar{u}\left(\phi_{2 N-1}\right)$. The coefficients $\left(D_{\phi}\right)_{j n}$ are the values at $\oint_{j}$ of the derivative of the coefficient [4]

$$
l_{n}(\phi)=\frac{1}{2 N} \sin \left[N\left(\phi-\phi_{n}\right)\right] \cot \frac{\phi-\phi_{n}}{2} .
$$

This result comes from the general Lagrangian form

$$
\begin{equation*}
\sum_{n=0}^{2 N-1} f\left(\phi_{n}\right) l_{n}(\phi) \tag{3.2.6}
\end{equation*}
$$

of the trigonometric polynomial interpolating an arbitrary function $f(\phi)$ between the $\phi_{j}$. From the definition of the derivative and Bernoulli rule, one can easily compute [4]

$$
\left(D_{\phi}\right)_{j n}=l_{n}^{\prime}\left(\phi_{j}\right)= \begin{cases}\frac{1}{2}(-1)^{j+n} \cot \frac{\dot{\phi}_{2}-\phi_{n}}{2}=\frac{1}{2}(-1)^{j-n} \cot \frac{(j-n) \pi}{2 N}, & j \neq n \\ 0, & j=n\end{cases}
$$

The matrix $D_{\phi}$ of the system (3.2.5) is therefore circulant (and skew symmetric). The $\phi_{j}$ are the images of the Chebyshev points under the transformation $\phi=\arccos (x)$; therefore,
the values of $F(\phi)$ used in the computation of $h\left(\phi_{j}\right)$ are the same as those used in $A \bar{U}=b$ of (3.2.2).

The transformation to the circle technique can be applied to higher order problems. We now consider the (singularly perturbed) second order linear two-point boundary value problem

$$
\left\{\begin{array}{l}
\epsilon U^{\prime \prime}(x)+P(x) U^{\prime}(x)+Q(x) U(x)=F(x)  \tag{3.2.7}\\
U(1)=u_{0}, \quad U(-1)=u_{N}
\end{array}\right.
$$

Now, instead of solving (2.3.15) directly, we propose to apply the circle transformation to obtain an equivalent problem which is defined on the circle. To do this, let $\phi(x)=\arccos (x)$ be the inverse of $x=\cos \phi$ and define $u(\phi):=U(\cos \phi), p(\phi):=P(\cos \phi), q(\phi):=Q(\cos \phi)$, and $f(\phi):=F(\cos \phi) ;$ then

$$
\begin{aligned}
\phi^{\prime}(x) & =-(\sin \phi(x))^{-1} \\
\phi^{\prime \prime}(x) & =[\sin \phi(x)]^{-2} \cdot \cos \phi(x) \cdot \phi^{\prime}(x)=-\frac{\cos \phi(x)}{\sin ^{3} \phi(x)} \\
\frac{d}{d x} U(x) & =\frac{d}{d \phi} U(\cos \phi) \phi^{\prime}(x)=u^{\prime}(\phi) \phi^{\prime}(x) \\
\frac{d^{2}}{d x^{2}} U(x) & =\frac{d}{d x}\left[u^{\prime}(\phi) \phi^{\prime}(x)\right] \\
& \left.=u^{\prime \prime}(\phi) \phi^{\prime}(x)^{2}+u^{\prime}(\phi) \phi^{\prime \prime}(x)\right)
\end{aligned}
$$

Substituting the above expressions into (3.2.7) leads to

$$
\epsilon\left[u^{\prime \prime}(\phi) \frac{1}{\sin ^{2} \phi}+u^{\prime}(\phi)\left(-\frac{\cos \phi}{\sin ^{3} \phi}\right)\right]+p(\phi)\left(-\frac{1}{\sin \phi}\right) u^{\prime}(\phi)+q(\phi) u(\phi)=f(\phi) .
$$

Multiplying both sides by $\sin ^{3} \phi$, we get

$$
\begin{equation*}
\epsilon u^{\prime \prime}(\phi) \sin (\phi)-\epsilon u^{\prime}(\phi) \cos (\phi)-p(\phi) \sin ^{2}(\phi) u^{\prime}(\phi)+q(\phi) \sin ^{3}(\phi) u(\phi)=f(\phi) \sin ^{3}(\phi) \tag{3.2.8}
\end{equation*}
$$

The boundary conditions in (3.2.6) are transformed into $u\left(\phi_{0}\right)=u_{0}$ and $u\left(\phi_{N}\right)=u_{N}$.
The pseudospectral Fourier method applied to the above equation then leads to the linear system $(j=0,1, \cdots, 2 N-1)$

$$
\begin{equation*}
\sum_{n=0}^{2 N-1} a_{j n} u\left(\phi_{n}\right)=b_{j} \tag{3.2.9}
\end{equation*}
$$

with

$$
\begin{aligned}
a_{j n} & :=\epsilon \sin \left(\phi_{j}\right)\left(D_{\phi}^{2}\right)_{j n}-\left[\epsilon \cos \left(\phi_{j}\right)+p\left(\phi_{j}\right) \sin ^{2}\left(\phi_{j}\right)\right]\left(D_{\phi}\right)_{j n}+q\left(\phi_{j}\right) \sin ^{3}\left(\phi_{j}\right) \delta_{j n} \\
b_{j} & :=f\left(\phi_{j}\right) \sin ^{3}\left(\phi_{j}\right)
\end{aligned}
$$

Similar to the pseudospectral Chebyshev method, the matrix of the second derivatives $l_{n}^{\prime \prime}\left(\phi_{j}\right)$ is the square of the matrix $D_{\phi}$ of the first derivatives. And there even exists a simple formula [4]:

$$
\left(D_{\phi}^{2}\right)_{j n}= \begin{cases}\frac{1}{2}(-1)^{j+n+1} \csc ^{2} \frac{\phi_{j}-\phi_{n}}{2}=\frac{1}{2}(-1)^{j-n+1} \csc ^{2} \frac{(j-n) \pi}{2 N}, & j \neq n \\ -\frac{2 N^{2}+1}{6}, & j=n\end{cases}
$$

Since $\mathrm{p}(\phi), \mathrm{q}(\phi), \mathrm{f}(\phi)$ and $\bar{u}(\phi)$ are even functions of $\phi$, and moreover, since $D_{\phi}$ is antisymmetric and $D_{\phi}^{2}$ is symmetric, only $N-1$ of the $2 N$ equations (3.2.10) are linearly independent: equations number $j=0$ and N reduce to $0\left(\sin \phi_{j}=0\right.$ and $\left.\left(D_{\phi}\right)_{j n} \bar{u}\left(\phi_{n}\right)=0\right)$, and for $n=1, \cdots, N-1$, equation number $2 N-n$ is the negative of equation number $n$. If we substitute $\bar{u}\left(\phi_{2 N-n}\right)=\bar{u}\left(\phi_{n}\right)$ into (3.2.9), we only need to solve the system composed of the equations numbered 1 to $N-1$ of (3.2.9), namely

$$
\begin{equation*}
\sum_{n=1}^{N-1}\left(a_{j n}+a_{j, 2 N-n}\right) u\left(\phi_{n}\right)=b_{j}-a_{j 0} u_{0}-a_{j N} u_{N} \tag{3.2.10}
\end{equation*}
$$

by Gaussian elimination. Our computations later show that the condition number of (3.2.10) grows like $O\left(N^{2}\right)$ for non-stiff problems (3.2.7)[4]. But for stiff problems, this is not true. The principle is the transformation of the original problem to the circle and the consequent elimination of the endpoint singularities of the factor of the derivatives $(\csc (\phi)$ or an odd power of it). The Fourier method is then applied to the regularized equation; it makes use of the same values of the functions appearing in the differential problem as the classical Chebyshev method.

### 3.3 Iterative Methods

It is well-known that algebraic systems resulting from spectral methods are full and rather ill-conditioned. For problem (3.2.6), the condition number of the matrix $D_{x}^{2}$ grows like $O\left(N^{4}\right)$. Considerable attention has been devoted to the use of iterative methods. In this section, we briefly discuss two iterative methods [16].

### 3.3.1 CGNR (The Conjugate Gradient Iteration Applied to the Normal Equations)

CGNR is the name for implementation of the conjugate gradient iteration method applied to the normal equations [16]. We apply the conjugate gradient iteration to the nonsymmetric system

$$
A \bar{U}=b
$$

The normal equations are

$$
A^{T} A U=A^{T} b
$$

This algorithm constructs a sequence

$$
r_{0}, A^{T} r_{0},\left(A^{T} A\right) A^{T} r_{0}, \cdots,\left(A^{T} A\right)^{n-1} A^{T} r_{0}
$$

We want to find the unique sequence of vectors

$$
u_{n} \in \operatorname{span}\left\{A^{T} r_{0},\left(A^{T} A\right) A^{T} r_{0}, \cdots,\left(A^{T} A\right)^{n-1} A^{T} r_{0}\right\}
$$

with minimal residual at each step:

$$
\left\|r_{n}\right\|=\text { minimum }
$$

where $r_{n}=b-A u_{n}$. This is equivalent to the orthogonality condition:

$$
\tau_{n} \perp \operatorname{span}\left\{A A^{T} r_{0},\left(A A^{T}\right)^{2} r_{0}, \cdots,\left(A A^{T}\right)^{n} r_{0}\right\}
$$

Now, $u_{n}$ can be formed by a three-term recurrence relation in the conjugate gradient iteration.

The algorithm of CGNR:

1. Given an initial guess $u_{0}$ and initial residual

$$
r_{0}=b-A u_{0}
$$

2. Given $\beta_{0}:=0 ; p_{0}:=0$,
3. Do $n:=1,2, \cdots$,

$$
\begin{aligned}
p_{n} & :=A^{T} r_{n-1}+\beta_{n-1} p_{n-1}, \\
\alpha_{n} & :=\left\|A^{T} r_{n-1}\right\|_{2}^{2} /\left\|A p_{n}\right\|_{2}^{2} \\
u_{n} & :=u_{n-1}+\alpha_{n} p_{n}, \\
r_{n} & :=r_{n-1}-\alpha_{n} A p_{n}, \\
\beta_{n} & :=\left\|A^{T} r_{n}\right\|_{2}^{2} /\left\|A^{T} r_{n-1}\right\|_{2}^{2}
\end{aligned}
$$

The convergence of CGNR is determined solely by the singular values of A. Any two matrices with the same singular values have identical worst-case convergence rates. If A is normal, the moduli of the eigenvalues are equal to the singular values. For details see [16].

### 3.3.2 CGS (The Conjugate Gradient Squared)

CGS is a biorthogonalization algorithm adapted from the biconjugate gradient iteration [16].

First, we introduce biconjugate gradient iteration (BCG) which constructs non-optimal approximations.

BCG constructs a sequence of vectors

$$
u_{n} \in \operatorname{span}\left\{r_{0}, A r_{0}, \cdots, A_{n-1} r_{0}\right\}
$$

which implies

$$
\begin{equation*}
e_{n}=p_{n}(A) e_{0}, \quad r_{n}=p_{n}(A) r_{0} \tag{3.3.11}
\end{equation*}
$$

for some polynomial $p_{n}$ of degree n , where $e_{0}=A^{-1} b-u_{0}$, and $r_{0}=b-A u_{0}$. The value $p_{n}$ is now determined by the orthogonality condition

$$
r_{n} \perp \operatorname{span}\left\{\tilde{r}_{0}, A^{T} \tilde{r}_{0}, \cdots, A_{n-1}^{T} \tilde{r}_{0}\right\}
$$

where $\tilde{r}_{0} \in R^{n-1}$ is a vector often taken equal to $r_{0}$. BCG computes its choice of $u_{n}$ by three-term recurrence relations.

CGS, which stands for "CG-squared", is a modification of BCG by Sonneveld [20] who replaces equation (3.3.11) by

$$
\begin{equation*}
e_{n}=p_{n}^{2}(A) e_{0}, \quad r_{n}=p_{n}^{2}(A) r_{0} \tag{3.3.12}
\end{equation*}
$$

The polynomial $p_{n}$ remains the same and there is no increase in the amount of work per step.

Furthermore, whereas BCG requires vector multiplications by both A and $A^{T}$, CGS only requires multiplications by A .

CGS algorithm:

1. Given an initial guess $u_{0}$ and initial residual

$$
r_{0}=b-A u_{0}
$$

2. Given $q_{0}:=p_{0}:=0 ; \rho_{0}:=1 ; \tilde{r}_{0}=r_{0}$,
3. For $n=1,2, \cdots$,

$$
\begin{aligned}
\rho_{n} & :=\tilde{r}_{0}^{T} r_{n-1} \\
\beta_{n} & :=\rho_{n} / \rho_{n-1} \\
w_{n} & :=r_{n-1}+\beta_{n} q_{n-1} \\
p_{n} & :=w_{n}+\beta_{n}\left(q_{n-1}+\beta_{n} p_{n-1}\right) \\
v_{n} & :=A p_{n} \\
\sigma_{n} & :=\tilde{r}_{0}^{T} v_{n}
\end{aligned}
$$

$$
\begin{aligned}
\alpha_{n} & :=\rho_{n} / \sigma_{n} \\
q_{n} & :=w_{n}-\alpha_{n} v_{n} \\
r_{n} & :=r_{n-1}-\alpha_{n} A\left(w_{n}+q_{n}\right) \\
u_{n} & :=u_{n-1}+\alpha_{n}\left(w_{n}+q_{n}\right)
\end{aligned}
$$

Although the convergence of CGS is governed by eigenvalues of A [16], its convergence properties are less well understood than CGNR. It is still unknown how frequently CGS turns out to be effective.

### 3.4 Numerical Results

We look again at the three examples studied in Chapter 2, using the method of transformation to the circle of section 3.2. We emphasize the comparison of the methods of this Chapter and the pseudospectral Chebyshev method of Chapter 2. The reader is reminded that in all the figures to follow, the exact solutions are plotted as dotted lines, and the numerical solutions are plotted as solid-dotted lines.

## Example 1 Revisited.

For the three different values of $\epsilon\left(10^{-2}, 10^{-4}\right.$, and $\left.10^{-6}\right)$, the method of transformation to the circle was applied to solve the BVP of Example 1. Figure 3.1 to Figure 3.11 and Table 3.1 to Table 3.3 show the results of the same kind of numerical work that was done in Example 1 of Chapter 2.

The pictures here and the pictures in Chapter 2 look similar, however, the condition numbers here are much smaller than those of Chapter 2 (except for the case where $N=128$, $\epsilon=10^{-4}$ ), and the average errors here tend to be smaller for $\epsilon=10^{-2}$, and slightly larger for $\epsilon=10^{-4}$ and $\epsilon=10^{-6}$. The cpu times here tend to be larger than those of Chapter 2 (except for the case where $N=128, \epsilon=10^{-6}$ ). When we applied the iterative methods above, we found that CGNR converged reasonably fast, but CGS did not. Also CGNR appears to work better when applied to the original problems versus the transformation to
the circle. This came as a surprise, since the matrices associated with the circle thechnique are better conditioned. Figure 3.18 also shows that quite a few iterations of CGNR are required, making this rather expensive. Further study, especially on the use of efficient preconditioners is needed. We have chosen $N=96$ and computed $(2 \times N)$ iterations of the CGNR method. The results for the pseudospectral method are shown in Figure 3.17 and Figure 3.18. The Figure 3.19 and Figure 3.20 are the results for the transformation to the circle technique.


Figure 3.1: Example 1, transformation to the circle method, $\epsilon=10^{-2}, N=32$


Figure 3.2: Example 1, transformation to the circle method, $\epsilon=10^{-2}, N=64$


Figure 3.3: Example 1, transformation to the circle method, $\epsilon=10^{-2}, N=96$


Figure 3.4: Example 1, transformation to the circle method, $\epsilon=10^{-4}, N=32$


Figure 3.5: Example 1, transformation to the circle method, $\epsilon=10^{-4}, N=64$


Figure 3.6: Example 1, transformation to the circle method, $\epsilon=10^{-4}, N=96$


Figure 3.7: Example 1, transformation to the circle method, $\epsilon=10^{-4}, N=128$


Figure 3.8: Example 1, transformation to the circle method, $\epsilon=10^{-6}, N=32$


Figure 3.9: Example 1, transformation to the circle method, $\epsilon=10^{-6}, N=64$


Figure 3.10: Example 1, transformation to the circle method, $\epsilon=10^{-6}, N=96$


Figure 3.11: Example 1, transformation to the circle method, $\epsilon=10^{-6}, N=128$

| N | Condition Number | Average Error | Cpu-Time |
| :---: | :---: | :---: | :---: |
| 32 | $7.1 \times 10^{2}$ | $1.23 \times 10^{-4}$ | 0.97 |
| 64 | $2.3 \times 10^{3}$ | $4.18 \times 10^{-11}$ | 3.26 |
| 96 | $4.7 \times 10^{3}$ | $1.92 \times 10^{-14}$ | 7.17 |
| 128 | $7.9 \times 10^{3}$ | $2.10 \times 10^{-14}$ | 13.66 |

Table 3.1: Example 1, transformation to the circle method, $\epsilon=10^{-2}$

| N | Condition Number | Average Error | Cpu-Time |
| :---: | :---: | :--- | :---: |
| 32 | $3.2 \times 10^{3}$ | $4.39 \times 10^{-2}$ | 0.92 |
| 64 | $2.2 \times 10^{4}$ | $2.15 \times 10^{-2}$ | 3.29 |
| 96 | $7.7 \times 10^{4}$ | $1.17 \times 10^{-2}$ | 7.17 |
| 128 | $1.7 \times 10^{5}$ | $6.20 \times 10^{-3}$ | 12.74 |

Table 3.2: Example 1, transformation to the circle method, $\epsilon=10^{-4}$

| N | Condition Number | Average Error | Cpu-Time |
| :---: | :---: | :--- | :---: |
| 32 | $5.3 \times 10^{5}$ | $4.65 \times 10^{-2}$ | 0.91 |
| 64 | $6.3 \times 10^{5}$ | $2.97 \times 10^{-2}$ | 2.82 |
| 96 | $5.8 \times 10^{5}$ | $1.95 \times 10^{-2}$ | 6.06 |
| 128 | $5.3 \times 10^{5}$ | $1.54 \times 10^{-2}$ | 11.02 |

Table 3.3: Example 1, transformation to the circle method, $\epsilon=10^{-6}$

## Example 2 Revisited.

The same kind of numericai work was carried out as we did in Example 2 of Chapter 2. The results are displayed in Figure 3.15 to Figure 3.20 and Table 3.4 to Table 3.6.

The pictures here look close to the pictures in Chapter 2 again, but the condition numbers and the average errors show a crossed pattern. For $\epsilon=10^{-2}$, the method of transformation to the circle produced much smaller condition numbers to reach the same average errors when compared with the pseudosnectral Chebyshev method, however, for $\epsilon=10^{-4}$ and $\epsilon=10^{-6}$, this is no longer true. For $\epsilon=10^{-4}$, the average errors are still the same for the two methods, but the condition numvers are larger for the first method when $N$ is small, and then become smaller when $N$ is large. For $\epsilon=10^{-6}$, the condition numbers with the first method are always larger, while the average errors with the first method are always smaller. The cup times also show a crossed pattern.

We wonder why the condition numbers related with circle technique for smaller $\epsilon\left(10^{-4}\right.$ and $10^{-6}$ ) are much larger than the condition numbers obtained by pseudospectral Chebyshev method. It seems that the circle technique is only successful in solving non-singularly perturbed boundary value problems, but not so successful for solving singularly perturbed boundary value problems.

BVPs. We


Figure 3.12: Example 2, transformation to the circle method, $\epsilon=10^{-2}, N=64$


Figure 3.13: Example 2, transformation to the circle method, $\epsilon=10^{-4}, N=128$


Figure 3.14: Example 2, transformation to the circle method, $\epsilon=10^{-4}, N=256$


Figure 3.15: Example 2, transformation to the circle method, $\epsilon=10^{-6}, N=128$


Figure 3.16: Example 2, transformation to the circle method, $\epsilon=10^{-6}, N=256$

| N | Condition Number | Average Error | Cpu-Time |
| :---: | :---: | :--- | :---: |
| 32 | $9.2 \times 10^{1}$ | $2.10 \times 10^{-3}$ | 0.91 |
| 64 | $3.3 \times 10^{2}$ | $7.89 \times 10^{-4}$ | 3.2 |
| 96 | $6.5 \times 10^{2}$ | $7.85 \times 10^{-4}$ | 7.18 |
| 128 | $1.1 \times 10^{3}$ | $7.83 \times 10^{-4}$ | 12.87 |

Table 3.4: Example 2, transformation to the circle method, $\epsilon=10^{-2}$

| N | Condition Number | Average Error | Cpu-Time |
| :---: | :---: | :--- | :---: |
| 128 | $1.5 \times 10^{6}$ | $0.21 \times 10^{1}$ | 12.78 |
| 256 | $3.7 \times 10^{5}$ | $0.24 \times 10^{-1}$ | 49.12 |
| 512 | $1.2 \times 10^{5}$ | $0.13 \times 10^{-5}$ | 343.65 |

Table 3.5: Example 2, transformation to the circle method, $\epsilon=10^{-4}$

| N | Condition Number | Average Error | Cpu-Time |
| :---: | :---: | :--- | :---: |
| 128 | $4.0 \times 10^{5}$ | $1.62 \times 10^{-1}$ | 12.91 |
| 256 | $3.9 \times 10^{6}$ | $1.54 \times 10^{-1}$ | 54.63 |

Table 3.6: Example 2, transformation to the circle method, $\epsilon=10^{-6}$


Figure 3.17: Example 1, the pseudospectral method, $N=96, \epsilon=10^{-4}$


Figure 3.18: Example 1, the pseudospectral method, $N=96, \epsilon=10^{-4}$


Figure 3.19: Example 1, the circle technique, $N=96, \epsilon=10^{-4}$


Figure 3.20: Example 1, the circle technique, $N=96, \epsilon=10^{-4}$

## Chapter 4

## Chebyshev Collocation Method

### 4.1 Chebyshev Collocation Method

In this Chapter, we will consider the collocation method with Chebyshev polynomials to solve two point boundary value pr blems. This method is discussed by Boyd [5] and it is a "nontraditional" spectral method, in the sense that we formulate the problem in terms of the (Chebyshev) expansion coefficients, but still use collocation. Usually these coefficients are determined by a Galerkin approach. We will compare the collocation method with the pseudospectral Chebyshev method discussed in Chapter 2. The pseudospectral Chebyshev method has a feature that is very attractive for certain types of problems, but leads to difficulties with others. This feature is very high resolution near the boundaries. For example, the collocation points for Chebyshev polynomial pseudospectral methods for problems on $-1 \leq x \leq 1$ are usually chosen such that $x_{j}=\cos (\pi j / N)(j=0,1, \cdots, N)$. The collocation points $x_{1}$ and $x_{N-1}$ are within approximately $\pi^{2} / 2 N^{2}$ of the boundary points $x_{0}$ and $x_{N}$, respectively, so that the boundary resolution is $\Delta x=O\left(1 / N^{2}\right)$.

This leads to extremely good resolution properties of spectral methods for boundarylayer problems. While resolution of a problem with a boundary layer of thickness $\epsilon \ll 1$ would require $O(1 / \epsilon)$ uniformly spaced grid points, it requires only $O\left(1 / \epsilon^{1 / 2}\right)$ terms in the

Chebyshev spectral series. Nonuniform grids could be used in many of these problems, as they can be implemented fairly efficiently in spectral methods using coordinate transformation. We will consider linear two point boundary value problems of the form:

$$
\left\{\begin{array}{l}
\epsilon u^{\prime \prime}(x)+p(x) u^{\prime}(x)+q(x) u(x)=f(x)  \tag{4.1.1}\\
u(-1)=u_{N}, \quad u(1)=u_{0}
\end{array}\right.
$$

Using the Chebyshev collocation method, we seek a solution

$$
\begin{equation*}
u(x)=\sum_{k=0}^{N} a_{k} T_{k}(x) \tag{4.1.2}
\end{equation*}
$$

where $\left\{T_{k}\right\}$ is the Chebyshev polynomial of degree k on $[-1,1]$ defined by formula

$$
T_{k}(\cos \theta)=\cos (k \theta)
$$

Clearly,

$$
\begin{gathered}
\left|T_{k}(x)\right| \leq 1 \quad \text { for } x \in[-1,1] \\
T_{0}(x)=1, \quad T_{1}(x)=x
\end{gathered}
$$

and, using elementary trigonometric identities

$$
T_{k+1}(x)=2 x T_{k}(x)-T_{k-1}(x) \quad \text { for } k \geq 1
$$

We introduce $N+1$ suitable collocation points which are the extreme of the Chebyshev polynomial of order N. i.e., $x_{j}=\cos (\pi j / N)$ for $j=0,1, \cdots, N$. Then, the approximate solution (4.1.2) must satisfy the ODE (4.1.1) and its boundary conditions at the collocation points $\left\{x_{j}\right\}$. Moreover, the following three steps are done:
(1) Determination of $N+1$ coefficients $a_{k}(k=0,1, \cdots, N)$, so that

$$
\begin{equation*}
u\left(x_{j}\right)=\sum_{k=0}^{N} a_{k} T_{k}\left(x_{j}\right) \quad(j=0,1, \cdots, N) . \tag{4.1.3}
\end{equation*}
$$

and; (2) Evaluation of $u^{\prime}\left(x_{j}\right)$ by

$$
\begin{equation*}
u^{\prime}\left(x_{j}\right)=\sum_{k=0}^{N} a_{k} T_{k}^{\prime}\left(x_{j}\right) \quad(j=0,1, \cdots, N) \tag{4.1.4}
\end{equation*}
$$

Finally, (3) we evaluate $u^{\prime \prime}\left(x_{j}\right)$ by

$$
\begin{equation*}
u^{\prime \prime}\left(x_{j}\right)=\sum_{k=0}^{N} a_{k} T_{k}^{\prime \prime}\left(x_{j}\right) \quad(j=0,1, \cdots, N) \tag{4.1.5}
\end{equation*}
$$

Here, the collocation equations are

$$
\left\{\begin{array}{l}
\epsilon u^{\prime \prime}\left(x_{j}\right)+p(x) u^{\prime}\left(x_{j}\right)+q(x) u\left(x_{j}\right)=f\left(x_{j}\right) \quad(j=1, \cdots, N-1)  \tag{4.1.6}\\
u(-1)=u_{N}, \quad u(1)=u_{0}
\end{array}\right.
$$

and substituting (4.1.3) (4.1.4) and (4.1.5) into the collocation equations, we have a linear system for determining coefficients $\left\{a_{k}\right\}_{k=0}^{N}$, i.e.,

$$
\left\{\begin{array}{l}
\sum_{k=0}^{N} a_{k} T_{k}(1)=u_{0} \\
\sum_{k=0}^{N}\left[\epsilon T_{k}^{\prime \prime}\left(x_{j}\right)+p\left(x_{j}\right) T^{\prime}\left(x_{j}\right)+q\left(x_{j}\right) T_{k}\left(x_{j}\right)\right] a_{k}=f\left(x_{j}\right) \quad(j=1, \cdots, N-1) \\
\sum_{k=1}^{N} a_{k} T_{k}(-1)=u_{N}
\end{array}\right.
$$

We write it in matrix form

$$
\begin{equation*}
A x=b \tag{4.1.7}
\end{equation*}
$$

where $A \in R^{(N+1) \times(N+1)}, \quad x, b \in R^{N+1}$,

$$
\begin{gathered}
x=\left(a_{0}, a_{1}, \cdots, a_{N}\right)^{T}, b=\left(u_{0}, f\left(x_{1}\right), \cdots, f\left(x_{N-1}\right), u_{N}\right)^{T}, \\
A=\left(a_{i j}\right),
\end{gathered}
$$

with

$$
\begin{gathered}
a_{1 k}=T_{k}(1) \quad(k=0, \cdots, N), \\
a_{j k}=\epsilon T_{k}^{\prime \prime}\left(x_{j}\right)+p\left(x_{j}\right) T_{k}^{\prime}\left(x_{j}\right)+q\left(x_{j}\right) T_{k}\left(x_{j}\right) \quad(k=0,1, \cdots, N, \text { and } j=1, \cdots, N-1), \\
a_{N k}=T_{k}(-1) \quad(k=0,1, \cdots, N) .
\end{gathered}
$$

Chapter 2 indicates that the matrix A is a very ill-condition matrix. Its condition number is $O\left(N^{4}\right)$. Some researchers consider a preconditioning scheme to solve this system(4.1.7) [6]. However, we use Linpack subroutines (Gaussian elimination) to solve (4.1.7).

$$
\begin{equation*}
A x=b \tag{4.1.8}
\end{equation*}
$$

### 4.2 Numerical Results

The method described in section 4.1 is applied to Example 2. Table 4.1 to Table 4.3 contain the summaries of the related results.

| N | Condition Number | Average Error | Cpu-Time |
| :---: | :---: | :--- | :---: |
| 32 | $1.6 \times 10^{3}$ | $2.30 \times 10^{-2}$ | 0.73 |
| 64 | $1.7 \times 10^{4}$ | $6.31 \times 10^{-3}$ | 2.12 |
| 96 | $6.2 \times 10^{4}$ | $5.16 \times 10^{-3}$ | 4.63 |
| 128 | $1.8 \times 10^{5}$ | $4.95 \times 10^{-3}$ | 8.43 |
| 256 | $1.9 \times 10^{6}$ | $4.76 \times 10^{-3}$ | 40.42 |
| 512 | $2.5 \times 10^{7}$ | $4.76 \times 10^{-3}$ | 206.77 |

Table 4.1: Example 2, Chebyshev collocation method, $\epsilon=10^{-2}$.

| N | Condition Number | Average Error | Cpu-Time |
| :---: | :---: | :--- | :---: |
| 96 | $6.1 \times 10^{4}$ | $1.27 \times 10^{0}$ | 4.59 |
| 128 | $2.2 \times 10^{5}$ | $2.02 \times 10^{0}$ | 8.38 |
| 256 | $6.0 \times 10^{5}$ | $8.31 \times 10^{-2}$ | 38.98 |
| 512 | $2.9 \times 10^{6}$ | $6.43 \times 10^{-3}$ | 207.17 |

Table 4.2: Example 2, Chebyshev collocation method, $\epsilon=10^{-4}$.

| N | Condition Number | Average Error | Cpu-Time |
| :---: | :---: | :--- | :---: |
| 96 | $5.0 \times 10^{4}$ | $1.28 \times 10^{0}$ | 4.59 |
| 128 | $1.0 \times 10^{5}$ | $1.29 \times 10^{0}$ | 8.38 |
| 256 | $6.0 \times 10^{5}$ | $1.21 \times 10^{0}$ | 39.92 |
| 512 | $2.9 \times 10^{6}$ | $1.14 \times 10^{\circ}$ | 208.74 |

Table 4.3: Example 2, Chebyshev collocation method, $\epsilon=10^{-6}$.

For the large parameter $\left(\epsilon=10^{-2}\right)$, the numerical results are gradually better with increasing N. See Figure 4.1 to Figure 4.6 for $N=32,64,96,128,256,512$ respectively.

For stiff problem, $\epsilon=10^{-4}$, we first try $\mathrm{N}=96$ and $\mathrm{N}=128$, respectively. Figure 4.7 and Figure 4.8 show that there are large oscillations. A large N is needed to produce a fairly good numerical solution, see Figure 4.9 and Figure 4.10 and Table 4.2.

For the very stiff problem (viz., $\epsilon=10^{-6}$ ), the method becomes difficult to solve it with the large average errors. See Figure 4.10, Figure 4.11 , Figure 4.12 and Table 4.3.

We compare the present method with pseudospectral Chebyshev method we studied in Chapter 2. The Chebyshev collocation method for solving the non-stiff problem ( $\epsilon=10^{-2}$ ) is less efficient than the pseudospectral Chebyshev method. For the stiff problem $\epsilon=10^{-4}$, the Chebyshev collocation method can compete with the pseudospectral Chebyshev method. For the very stiff problem $\left(\epsilon=10^{-6}\right)$, the Chebyshev collocation method is not quite as bad as the pseudospectral Chebyshev method. It is interesting the oscillation that occurred in the Chebyshev collocation method is still reasonable, while the oscillation that occurred in the pseudospectral method is unacceptable in some sense. However, our basic conclusion is that the pseudospectral is better than the collocation approach. Also, the condition numbers resulting from the collocation approach appear to be worse.


Figure 4.1: Example 2, Chebyshev collocation method, $\epsilon=10^{-2}, N=32$


Figure 4.2: Example 2, Chebyshev collocation method, $\epsilon=10^{-2}, N=64$


Figure 4.3: Example 2, Chebyshev collocation method, $\epsilon=10^{-2}, N=96$


Figure 4.4: Example 2, Chebyshev collocation method, $\epsilon=10^{-2}, N=128$


Figure 4.5: Example 2, Chebyshev collocation method, $\epsilon=10^{-2}, N=256$


Figure 4.6: Example 2, Chebyshev collocation method, $\epsilon=10^{-2}, N=512$


Figure 4.7: Example 2, Chebyshev collocation method, $\epsilon=10^{-4}, N=96$


Figure 4.8: Example 2, Chebyshev collocation method, $\epsilon=10^{-4}, N=128$


Figure 4.9: Example 2, Chebyshev collocation method, $\epsilon=10^{-4}, N=256$


Figure 4.10: Example 2, Chebyshev collocation method, $\epsilon=10^{-6}, N=96$


Figure 4.11: Example 2, Chebyshev collocation method, $\epsilon=10^{-6}, N=128$


Figure 4.12: Example 2, Chebyshev collocation method, $\epsilon=10^{-6}, N=256$

## Chapter 5

## Discussion

In this thesis we have studied three numerical methods for solving singularly perturbed boundary value problems of order 2. The pseudospectral Chebyshev method was little applied to solve singularly perturbed boundary value problems, therefore, three problems have been chosen for our study-the first problem has a layer in the middle of the interval $[-1,1]$; the second problem has two boundary layers at $x=-1$ and $x=1$; and the third problem has one boundary layer at $x=-1$. With each problem and with each method, we have considered three cases of singularity-mild $\left(\epsilon=10^{-2}\right)$, stiff $\left(\epsilon=10^{-4}\right)$, and very stiff $\left(\epsilon=10^{-6}\right)$. Our attention in the study has been given to three areas: condition numbers, average errors and cpu times.

The pseudospectral Chebyshev method was studied in Chapter 2. The numerical results in this Chapter show that the condition number of the matrix related to the $N$-point pseudospectral Chebyshev approximation is smaller than $N^{4}$, but it depends on the singularity paraneter $\epsilon$. In general, the smaller the $\epsilon$, the larger the condition number.

For $\epsilon=10^{-2}$ and $\epsilon=10^{-4}$, the pseudospectral Chebyshev method can be successfully applied to solve the three singularly perturbed BVPs, even though for $\epsilon=10^{-4}$, a relatively large $N(N=256$ or 512$)$ is needed to reach a fairly good solution. For $\epsilon=10^{-6}$, the pseudospectral Chebyshev method succeeds in solving the first problem. When applied to
solve the second and the third problems, the method fails to produce converging solutions for $N$ as large as 512 . The failure in these cases is most serious in the middle of the interval $[-1,1]$, where the exact solutions are flat and smooth, while the numerical solutions have large oscillations.

The method of transformation to the circle was studied in Chapter 3. By transforming the problem defined on $x \in[-1,1]$ into an equivalent problem defined on the unit circle through $x=\cos \phi$, this method proceeds to apply the Fourier spectral method. The numerical results in this chapter also show that the condition number of the N-point approximation increases as the the singularity parameter $\epsilon$ decreases.

The numerical solutions produced with the method of transformation to the circle are of similar quality when compared with those of the numerical solutions produced with the pseudospectral Chebyshev method. However, the transformation method does outperform the pseudospectral Chebyshev method for the mildly singularly perturbed problems $\left(\epsilon=10^{-2}\right)$. For the stiff $\left(\epsilon=10^{-4}\right)$ and very stiff $\left(\epsilon=10^{-6}\right)$ problems, there is no dominance in performance; sometimes the transformation method does better; sometimes the pseudospectral Chebyshev method does better. We compare the condition number for above two methods as shown in Table 5.1 to Table 5.6.

It is noticed that when $\epsilon=10^{-6}$, both the pseudospectral Chebyshev method and the transformation method failed to produce acceptable numerical solutions in all the cases related to the second and the third examples, for $N$ as large as 512 . The failure is in the middle of the interval $[-1,1]$, where large oscillations occur in the numerical solutions even when $N=512$. One possible reason for this failure is that there are always fewer collocation points near $x=0$ than near $x=-1$ and $x=1$ in the above two methods. (For the transformation method, think in terms of the inverse transformation $\phi=\arccos (\mathrm{x})$. .

The two iterative methods, the conjugate gradient iteration applied to the normal equations (CGNR) and the biorthogonalization algorithm adapted from the biconjugate gradient iteration (CGS) were also studied in Chapter 3. We had hoped to see a good performance
of these two iterative methods, but only the CGNR method converged.
Other effort has also been made to study the possibility of improving the quality of our numerical solutions. For instance, in the study of the pseudospectral Chebyshev method, we also tried to use Chebyshev polynomials to construct approximations (known as the Chebyshev collocation method) in addition to the standard use of Lagrange polynomials. The condition number appears to be worse.

Finally, we will briefly analyze the numerical errors. We consider the linear system $A x=b$, then the relative error in $x[10]$ is

$$
\frac{\|\tilde{x}-x\|}{\|x\|} \leq k_{2}(A)\left(\frac{\|\tilde{b}-b\|}{\|b\|}\right)
$$

where $\tilde{x}$ and $\tilde{b}$ are the numerical results, i.e. $A \tilde{x}=\tilde{b}$. Thus, the relative error in x can be the condition number $k_{2}(A)$ times the relative error in A and b . In this sense, the condition number $k_{2}(A)$ quantifies the sensitivity of the $A x=b$ problem. On the other hand, by a singular value decomposition (SVD) analysis, the SVD of A is

$$
A=U \Sigma V^{T}=\sum_{i=1}^{n} \sigma_{i} u_{i} v_{i}^{T}
$$

where $U=\left(u_{1}, u_{2}, \cdots, u_{n}\right), V=\left(v_{1}, v_{2}, \cdots, v_{n}\right)$ and $\sum=\operatorname{diag}\left(\sigma_{1}, \sigma_{2}, \cdots, \sigma_{n}\right)$ with $\sigma_{1} \geq$ $\sigma_{2} \geq \cdots \geq \sigma_{n} \geq 0$. Then

$$
\begin{gathered}
x=A^{-1} b=\left(U \Sigma V^{T}\right)^{-1} b=\sum_{i=1}^{n} \frac{u_{i}^{T} b}{\sigma_{i}} v_{i}, \\
\text { and } \quad \tilde{x}=A^{-1} \tilde{b}=\sum_{i=1}^{n} \frac{u_{i}^{T} \tilde{b}}{\sigma_{i}} v_{i} .
\end{gathered}
$$

The absolute error in $x$ is

$$
x-\tilde{x}=\sum_{i=1}^{n} \frac{u_{i}^{T}(b-\tilde{b})}{\sigma_{i}} v_{i}
$$

Let

$$
\begin{align*}
b & =\sum_{i=1}^{n} b_{i} u_{i}  \tag{5.1}\\
\delta b=b-\tilde{b} & =\sum_{i=1}^{n} \epsilon_{i} b_{i} u_{i} \quad\left(\text { where } \quad \epsilon_{i}= \pm \epsilon\right)
\end{align*}
$$

Here we assume that every component in $\tilde{b}$ has the same relative error

$$
\delta x=x-\tilde{x}=\sum_{i=1}^{n} \frac{u_{i}^{T}(b-\tilde{b})}{\sigma_{i}} v_{i}
$$

Therefore, we have

$$
\begin{gathered}
\|x\|_{2}^{2} \approx \sum_{i=1}^{n} \frac{b_{i}^{2}}{\sigma_{i}^{2}} \\
\|b\|_{2}^{2} \approx \sum_{i=1}^{n} b_{i}^{2} \\
\|\delta x\|_{2}^{2} \approx \epsilon^{2} \sum_{i=1}^{n} \frac{b_{i}^{2}}{\sigma_{i}^{2}} \\
\|\delta b\|_{2}^{2} \approx \epsilon^{2} \sum_{i=1}^{n} b_{i}^{2}
\end{gathered}
$$

The relative errors in x and b have the same order,

$$
\begin{equation*}
\frac{\|\delta x\|_{2}^{2} /\|x\|_{2}^{2}}{\|\delta \delta\|_{2}^{2} /\|b\|_{2}^{2}} \approx \frac{\epsilon}{\epsilon}=1 \tag{5.2}
\end{equation*}
$$

We have plotted the singular vectors versus Chebyshev points for Example 1 (with $\epsilon=10^{-2}$ and $N=96$ ) as shown in Figure 5.1 to Figure 5.6. The behaviour of the singular values and the coefficients $\left\{b_{i}\right\}$ of the equation (5.1) are shown in Figure 5.7 and Figure 5.8. Based on our numerical experiments, we can see that the numerical solutions are accurate despite the large condition numbers. This needs to be contrasted with the worst case scenario, namely that $b=b_{1} u_{1}$ and $\delta b=\beta_{n} u_{n}$, i.e. the right hand side is the direction of the singular vector associated with the largest singular value $\sigma_{1}$, but the perturbation is only in the direction of $u_{n}$ associated with the smallest singular value $\sigma_{n}$. Our investigation shows that the singular vectors associated with large singular values are highly oscillatory, whereas the ones associated with small $\sigma_{k}^{\prime} s$ are smooth. This is expected since A describes a differentiation process. Now, although the right hand side of our BVPs may be smooth, the boundary conditions are incorporated in the vector $b$, making it a nonsmooth vector (at least for problems which generate interior or boundary layers); as is illustrated by Figure 5.8. However, the full effect of the large condition number $k_{2}(A)=\sigma_{1} / \sigma_{n}$ could only be felt, if the perturbation in $b$ were very smooth (associated with the small singular values), which
is extremely unlikely for perturbation generated by finite precision arithmetic. Thus it is plausible that (5.2) holds. This also explains why the transformation to the circle, although decreasing the condition numbers, does not promote higher accuracy.

| N | Direct Method | Circle Technique |
| :---: | :---: | :---: |
| 32 | $6.4 \times 10^{2}$ | $7.1 \times 10^{2}$ |
| 64 | $8.3 \times 10^{3}$ | $2.3 \times 10^{3}$ |
| 96 | $1.4 \times 10^{5}$ | $4.7 \times 10^{3}$ |
| 128 | $1.2 \times 10^{5}$ | $7.9 \times 10^{3}$ |

Table 5.1: Example 1, the condition number for two methods, $\epsilon=10^{-2}$

| N | Direct Method | Circle Technique |
| :---: | :---: | :---: |
| 32 | $4.5 \times 10^{4}$ | $3.2 \times 10^{3}$ |
| 64 | $7.7 \times 10^{4}$ | $2.2 \times 10^{4}$ |
| 96 | $1.3 \times 10^{5}$ | $7.7 \times 10^{4}$ |
| 128 | $1.4 \times 10^{5}$ | $1.7 \times 10^{5}$ |

Table 5.2: Example 1, the condition number for two methods, $\epsilon=10^{-4}$

| N | Direct Method | Circle Technique |
| :---: | :---: | :---: |
| 32 | $5.8 \times 10^{6}$ | $5.3 \times 10^{5}$ |
| 64 | $1.2 \times 10^{7}$ | $6.3 \times 10^{5}$ |
| 96 | $1.8 \times 10^{7}$ | $5.8 \times 10^{5}$ |
| 128 | $2.4 \times 10^{7}$ | $5.3 \times 10^{5}$ |

Table 5.3: Example 1, the condition number for two methods, $\epsilon=10^{-6}$

| N | Direct Method | Circle Technique |
| :---: | :---: | :---: |
| 32 | $3.0 \times 10^{3}$ | $9.2 \times 10^{1}$ |
| 64 | $5.2 \times 10^{4}$ | $3.3 \times 10^{2}$ |
| 96 | $2.7 \times 10^{5}$ | $6.5 \times 10^{2}$ |
| 128 | $8.8 \times 10^{5}$ | $1.1 \times 10^{3}$ |

Table 5.4: Example 2, the condition number for two methods, $\epsilon=10^{-2}$

| N | Direct Method | Circle Technique |
| :---: | :---: | :---: |
| 128 | $3.3 \times 10^{4}$ | $1.5 \times 10^{6}$ |
| 256 | $2.5 \times 10^{5}$ | $3.7 \times 10^{5}$ |
| 512 | $3.9 \times 10^{6}$ | $1.2 \times 10^{5}$ |

Table 5.5: Example 2, the condition number for two methods, $\epsilon=10^{-4}$

| N | Direct Method | Circle Technique |
| :---: | :---: | :---: |
| 128 | $7.1 \times 10^{3}$ | $4.0 \times 10^{5}$ |
| 256 | $2.3 \times 10^{4}$ | $3.9 \times 10^{6}$ |

Table 5.6: Example 2, the condition number for two methods, $\epsilon=10^{-6}$


Figure 5.1: Example 1, singular vector $\left(u_{1}\right)$ versus Chebyshev points, $\epsilon=10^{-2}, N=96$


Figure 5.2: Example 1, singular vector ( $u_{2}$ ) versus Chebyshev points, $\epsilon=10^{-2}, N=96$


Figure 5.3: Example 1, singular vector ( $u_{3}$ ) versus Chebyshev points, $\epsilon=10^{-2}, N=96$


Figure 5.4: Example 1, singular vector ( $u_{93}$ ) versus Chebyshev points, $\epsilon=10^{-2}, N=96$


Figure 5.5: Example 1, singular vector ( $u_{94}$ ) versus Chebyshev points, $\epsilon=10^{-2}, N=96$


Figure 5.6: Example 1, singular vector ( $u_{95}$ ) versus Chebyshev points, $\epsilon=10^{-2}, N=96$


Figure 5.7: Example 1, singular values $\epsilon=10^{-2}, N=96$


Figure 5.8: Example 1, coefficients $\left\{b_{i}\right\}$ of the equation (5.1) $\epsilon=10^{-2}, N=96$

## Bibliography

[1] Ascher, U. M., R. M. M. Mattheij, R. M. M. and Russell, R. D., Numerical Solution of Boundary Value Problems for Ordinary Differential Equations, Prentice-Hall Inc., 1988.
[2] Atkinson, K. E., A Survey of Numerical Methods for the Solution of Fredholm Integral Equations of the Second Kind, Society for Industrial and Applied Mathematics. 1976.
[3] Becker, R. A., Chambers, J. M. and Wilks, A. R., The New S Language: A Programming Environment for Data Analysis and Graphics, Wadsworth \& Brooks/Cole Advanced Books \& Software, Pacific Grove, California, 1988.
[4] Berrut, J. P., A Pseudospectral Chebyshev Method with Preliminary Transform to the Circle: O. D. E., Manuscript, 1990.
[5] Boyd, J. P., Chebyshev 8 Fourier Spectral Methods, Lecture Notes in Enginerring, 49, Brebbia, C. A. and Orszag, S. A. ed., Springer-Verlag, 1989.
[6] Canuto, C. and Quarteroni, A., Preconditioned Minimal Residul Methods for Chebyshev Spectral Calculations, J. Comput. Phys., 60, 315-337, 1985.
[7] Dongarra, J. J., Moler, C. B. and Bunch, J. R., Stewart, G. W., Linpack users' Guide, SLAM Philadelphia, 1979.
[8] Funaro, D., A Preconditioning Matrix for The Chebyshev Differencing Operator,SIAM J. Numer. Anal.,24, 1024-1031, 1987.
[9] Funaro, D., Some Results about the Spectrum of Chebyshev Differencing Operator, In the proceedings of the International Symposium of Numerical Analysis (Madrid, Sept. 1985).
[10] Golub, G. H. and Van Loan, C. F., Matrix Computations, second edition, The Johns Hopkins Universty Press, Baltimore and London, 1989.
[11] Greengard, L., Spectral Integration and Two - Point Boundary Value, SIAM J. Numer. Anal., 28, 1071-1080, 1991.
[12] Gottlieb, D., Hussaini, M. T., and Orszag, S. A., Theory and Applications of Spectral Methods, Report No. 83-66, 1983.
[13] Gutknecht, M. H., Variants of BICGSTAB for Matrices with Complx Spectrum, IPS Research report No. 91-14, ETH-Zentrum CH-8092 Zurich, 1991.
[14] Hussaini, M. Y., and Zang, T. A., Iterative Spectral Method and Spectral Solution to Compressible Flows, Spectral Methods for Partial Differential Equations (R. Voigt, D. Gottlieb, Y. Hussaini Eds.), SLAM, Philadelphia, 1984.
[15] Gottlieb, D., and Orszag, S. A., Numerical Analysis of Spectral Methods: Theory and Applications, CBMS regional Conferance Series in Applied Math., 26, SIAM Philadelphia, 1977.
[16] Nachtigal, N. M., Reddy, S. C., and Trefethen, L. N., How Fast Are Nonsymmetric Matrix Iterations?, Manuscript, Massachusetts Institute of Technology, Cambridge, 1991.
[17] Orszag, S. A., Spectral Methods for Problems in Complex Geometries, J. Comput. Phys., 37, 70-92, 1980.
[18] Prenter, P. M., Splines and Variational Methods, John Wiley \& Sons, New York, 1975.
[19] Smith, R. C., Bogar, G. A., Bower, K. L., and Lund, J., The Sinc-Galerkin Method for Fourth-Order Differantial Equations, SIAM J. Numer. Anal., 28, 760-788, 1991.
[20] Sonneveld, P., CGS, A Fast Lanczos - Type Solver for Nonsymmetric Linear Systems, SIAM J. Sci. Stat. Comp., 10, 36-52, 1989.
[21] Don, W. S. and Solomonoff, A., Accuracy and Speed in Computing the Chebyshev Collocation Derivative, Division of Applied Mathematics, Brown University, Providence, 1991.
[22] Yamasaki, A., New Preconditioners Based on Low-Rank Elimination, Numerical Analysis Report 89-10, Massachusetts Institute of Technology, Cambrige, 1989.
[23] Wright, K., Chebyshev Collocation methods for Ordinary Differential Equations, J. Comput., 16, 358-365, 1964.
[24] Huang, W., and Sloan, D. M., A New Pseudospectral Method with Upwind Features. Manuscript, 1992.

