NOTICE WARNING CONCERNING COPYRIGHT RESTRICTIONS:
The copyright law of the United States (title 17, U.S. Code) governs the making of photocopies or other reproductions of copyrighted material. Any copying of this document without permission of its author may be prohibited by law.

NAMT
96-013

Pseudospectral vs Finite Difference Methods for Initial Value Problems with Discontinuous Coefficients.

Erding Luo
Carnegie Mellon University

Heinz-Otto Kreiss
University of California at Los Angeles

Research Report No. 96-NA-013

August 1996

Sponsors

U.S. Army Research Office
Research Triangle Park
NC 27709

National Science Foundation
1800 G Street, N.W.
Washington, DC 20550

University Libraries
Carnegie Mellon University
Pittsburgh PA 15213-3890

TMAN
E10-JP

Pseudospectral vs Finite Difference Methods for Initial Value Problems with Discontinuous Coefficients

Erding Luo * Heinz-Otto Kreiss†

First version: March, 1995 This version: March, 1996

Abstract

An initial value problem with piecewise constant coefficients is considered. The accuracies for both finite difference methods and the pseudospectral method are analyzed, and a modification of the initial value problem is suggested. The modified problem is shown to have the same temporal period as the original problem does, and a second order accuracy is obtained for the pseudospectral method at integral multiples of the temporal period.

Key Words. pseudospectral, finite difference, initial value problem

AMS subject classifications. 65M06, 65M70

1 Introduction

The pseudospectral method is very powerful for periodic initial value problems with smooth coefficients. Well known results include its convergence at a spectral rate, [2, 3, 4, 8, 9]. This is largely because the fourier modes of the pseudospectral method can be considered as the limit of higher and higher order difference approximations. For equations with discontinuous coefficients, the accuracy of the pseudospectral method is far from being clear, even in the one dimensional case, [1, 4, 5, 6, 7].

*Department of Mathematics, Carnegie Mellon University, Pittsburgh, PA 15213. Research supported by the NSF through grants to IMA and by the ARO and the NSF through grants to the Center for Nonlinear Analysis. Email: eluo@andrew.cmu.edu.

†Department of Mathematics, University of California at Los Angeles, LA, CA90024. Email: kreiss@math.ucla.edu.

University Libraries
Carnegie Mellon University
Pittsburgh PA 15213-3890

The purpose of this paper is to compare the pseudospectral method with finite difference methods in this case. We will focus on the one dimensional scalar wave equation with periodic boundary conditions. The coefficient of the problem is assumed to be piecewise constant. We study the accuracies both for finite difference methods and the pseudospectral method using eigenfunction analysis. Our approach is to determine how many eigenmodes are needed to well represent the initial condition with a given accuracy, and to analyze the accuracies of these eigenmodes. For finite difference methods, local truncation error analysis is used to obtain a global error estimate. For the pseudospectral method, the error estimate is generated by numerical computations. Since the coefficient is discontinuous, the error depends on the relation of the grids to the points of discontinuity. If the points of discontinuity align with the grid points, and we use an average harmonic values of the coefficient, both finite difference methods and the pseudospectral method give us an accuracy of order $O(h\sqrt{h})$ in l_2 norm and $O(h)$ in l_∞ norm. In general, the points of discontinuity are not aligned with the meshes. Both finite difference methods and the pseudospectral method give us an approximation with a shift, and we cannot expect more than first order accuracy, [1]. To improve this accuracy, we suggest a modified problem by truncating the fourier expansion of the inverse of the coefficient. This modified problem retains the same temporal period as the original one. It turns out that the pseudospectral method still has $O(h)$ accuracy in l_∞ norm, but $O(h\sqrt{h})$ in l_2 norm. Moreover, at integral multiples of the temporal period, the pseudospectral method always has $O(h^2)$ accuracy in l_∞ norm.

The rest of the paper is organized as follows: our model problem is presented in section 2. In section 3, Numerical schemes are introduced. The accuracies of both numerical methods are demonstrated in section 4 and some numerical results are presented in section 5.

2 Model Problem

Consider the following equation

$$\frac{\partial u(x,t)}{\partial t} = -a(x)\frac{\partial u(x,t)}{\partial x}, \quad x \in (-1,1) \tag{2.1}$$

with periodic boundary condition, where the coefficient $a(x)$ is also periodic with two discontinuous points at $x = -1/2, 1/2$, such that

$$a(x) = \begin{cases} 0.5, & x \in (-0.5, 0.5), \\ 1, & x \in [-1,1]/[-0.5,0.5]. \end{cases} \tag{2.2}$$

The initial condition takes the following form,

$$f(x) = \exp(-Cx^2) = \exp(-(\sqrt{C}x)^2), \quad x \in [-1,1], \tag{2.3}$$

where $C >> 1$.

2.1 Analytic Solution

Denote the kth ($k \in Z$) eigenfunction of (2.1) by ϕ_k, which satisfies the periodic boundary condition. Then,

$$-a(x)\frac{d\phi_k(x)}{dx} = \lambda\phi_k(x).$$

Thus,

$$\phi_k(x) = \exp(-\lambda \int_{-1}^{x} \frac{1}{a(\xi)} d\xi),$$

where

$$\lambda = 2\pi i\mu k, \quad \mu = (\int_{-1}^{1} \frac{1}{a(\xi)} d\xi)^{-1} = \frac{1}{3}.$$

Let

$$\lambda_k = 2\pi ik, \quad k \in Z.$$

Then, ϕ_k can be rewritten as

$$\phi_k(x) = \exp(-\lambda_k\mu \int_{-1}^{x} \frac{1}{a(\xi)} d\xi), \quad x \in [-1, 1]. \tag{2.4}$$

The sequence $\{\phi_k(x)\}$ forms an orthogonal base with the weighing function $\frac{1}{a(x)}$ in the domain $[-1, 1]$. The solution of (2.1) can thus be expanded in terms of the eigenfunctions as

$$u(x,t) \sim \sum_k c_k \exp(\lambda_k\mu t)\phi_k(x), \tag{2.5}$$

where $c_k (k \in Z)$ is a constant that is determined by the initial condition (2.3), i.e.,

$$f(x) = \sum_k c_k\phi_k(x).$$

2.2 Expansion of Initial Condition

Denote the eigenmodes by

$$\varphi_k(x) = \exp(-\lambda_k x), \quad k \in Z. \tag{2.6}$$

The initial condition (2.3) can also be expanded in terms of a certain number of the fourier eigenmodes (2.6) within a given accuracy.

For $-\frac{1}{2} \leq x \leq \frac{1}{2}$, the eigenmodes are of the form

$$\phi_k(x) = const \exp(\frac{4}{3}\pi i x).$$

Thus they are the usual Fourier modes for the interval $-\frac{3}{4} \leq x \leq \frac{3}{4}$. Since $C >> 1$, we can assume that the initial data $f(x) \equiv 0$ for $1 > |x| \geq \frac{1}{2}$ and therefore the eigen expansion of f is nothing else but the Fourier expansion of a smooth periodic function. If $\frac{\sqrt{C}}{M} << 1$, then,

$$|f(x) - \sum_{k=-M}^{M} c_k \phi_k| \leq \eta \leq 1.$$

Numerical calculations show that for $C = 60$, the error $\eta \leq 0.01$ for $M \geq 32$.

2.3 Modified Problem

We modify equation (2.1) as follows

$$\frac{\partial v(x,t)}{\partial t} = -\frac{1}{b(x)} \frac{\partial v(x,t)}{\partial x}, \quad x \in (-1,1), \tag{2.7}$$

where the coefficient $b(x)$ is the truncated fourier series of $\frac{1}{a(x)}$. The fourier expansion of $\frac{1}{a(x)}$ is

$$\frac{1}{a(x)} \sim \sum_{\omega=-\infty}^{\infty} \hat{b}_\omega \exp(i\omega\pi(x+1)),$$

where

$$\hat{b}_\omega = \frac{1}{2}\int_{-1}^{1} \frac{1}{a(x)} \exp(-i\omega\pi(x+1))dx = \begin{cases} -\frac{1}{\omega\pi}\sin\frac{\omega\pi}{2}, & \omega \neq 0; \\ \frac{1}{2}\int_{-1}^{1} \frac{1}{a(x)}dx, & \omega = 0. \end{cases}$$

Let

$$R_N(x) = \sum_{\omega<-N/2}^{\omega\geq N/2} \hat{b}_\omega \exp(i\omega\pi(x+1)),$$

then $b(x)$ is given by

$$b(x) = \frac{1}{a(x)} - R_N(x) = \sum_{\omega\geq -N/2}^{\omega<N/2} \hat{b}_\omega \exp(i\omega\pi(x+1)). \tag{2.8}$$

The kth eigenfunction $\psi_k(x)$ of (2.7) is the solution of

$$-\frac{1}{b(x)}\frac{d\psi_k(x)}{dx} = \lambda\psi_k(x),$$

i.e.,

$$\psi_k(x) = \exp(-\lambda_k \mu \int_{-1}^{x} b(\xi) d\xi), \quad x \in [-1, 1]. \tag{2.9}$$

Observe that

$$\int_{-1}^{1} b(x) dx = \int_{-1}^{1} \frac{1}{a(x)} dx.$$

Therefore the eigenvalues of the modified equation are the same as of (2.1). The sequence $\{\psi_k(x)\}$ forms an orthogonal base with the weighing function $b(x)$ in the domain $[-1, 1]$. Subtracting (2.4) from (2.9),

$$\begin{aligned} \phi_k(x) - \psi_k(x) &= \phi_k(x)(1 - \exp(\lambda_k \mu \int_{-1}^{x} R_N(\xi) d\xi)) \\ &= \phi_k(x)(-\lambda_k \mu \int_{-1}^{x} R_N(\xi) d\xi) + O(\frac{k^2}{N^2}) = O(\frac{k}{N}). \end{aligned} \tag{2.10}$$

Theorem 2.1 *Assume T is the temporal period of equation (2.1). Then, the difference between the solutions for equations (2.1) and (2.7) is of the following order,*

$$u(x,t) - v(x,t) = \begin{cases} O(\frac{1}{N}), & for \quad t \neq kT, \quad k \in Z, \\ 0, & for \quad otherwise. \end{cases}$$

Proof. Since the eigenvalues do not change, we have that at $t = kT$,

$$u(x,t) = v(x,t) = f(x).$$

□

3 Numerical Scheme

Divide the domain $[-1, 1]$ with equal step size h by grid points $\{x_j\}$, and denote by u_j the approximation of the solution of equation (2.1) at x_j for $j = 0, 1, \cdots, N$. Since the coefficient in the equation is strictly positive, finite difference methods and the pseudospectral method are stable, [7]. In this section, we add dissipations both to finite difference methods and to the pseudospectral method, in order to reduce the high frequency modes caused by the discontinuity of the coefficient.

3.1 Finite Difference Methods

The difference approximations of (2.1) are given by

$$\frac{du_j}{dt} = -a(x_j) Q_p u_j, \quad j = 0, 1, \cdots, N-1 \tag{3.11}$$

where Q_p is the centered difference operator that approximates $\partial/\partial x$ with accuracy of order p. Thus,

$$Q_p = D_0 \sum_{\gamma=0}^{p/2-1} (-1)^\gamma \alpha_\gamma (h^2 D_+ D_-)^\gamma, \tag{3.12}$$

with the coefficients determined as follows,

$$\alpha_0 = 1, \tag{3.13}$$

$$\alpha_\gamma = \frac{\gamma}{4\gamma+2}\alpha_{\gamma-1}, \quad \gamma = 1, 2, \cdots, p/2-2. \tag{3.14}$$

Assuming the 4-th order Runge-Kutta method with artificial dissipation is used for time discretization, we determine the approximation u_j^{n+1} at time level $n+1$ from the previous approximation u_j^n in the following way,

$$\begin{aligned} u_j^{(n,1)} &= u_j^n, \\ u_j^{(n,2)} &= u_j^n - \frac{\Delta t}{2} a(x_j) Q_p u_j^{(n,1)}, \\ u_j^{(n,3)} &= u_j^n - \frac{\Delta t}{2} a(x_j) Q_p u_j^{(n,2)}, \\ u_j^{(n,4)} &= u_j^n - \Delta t a(x_j) Q_p u_j^{(n,3)}, \end{aligned}$$

and

$$\begin{aligned} u_j^{n+1} &= (1 + \gamma h^p (-1)^{p/2+1} Q_{p,q}) u_j^n \\ &- \frac{\Delta t}{6} a(x_j) Q_p (u_j^{(n,1)} + 2u_j^{(n,2)} + 2u_j^{(n,3)} + u_j^{(n,4)}), \quad j = 0, 1, \cdots, N-1, \end{aligned} \tag{3.15}$$

where $Q_{p,q}$ denotes the centered difference operator that approximates $\frac{\partial^p}{\partial x^p}$ with accuracy of order q.

3.2 Pseudospectral Method

Let U denote the vector with components $u_j, j = 0, \cdots, N-1$. Then, the pseudospectral method has the form

$$\frac{dU}{dt} = ASU. \tag{3.16}$$

Here S presents the standard FFT operator which approximates $\frac{\partial}{\partial x}$. A is the diagonal matrix with entry $a(x_j), \quad j = 0, \cdots, N-1$. If the 4-th order Runge-Kutta method is used

for time discretization with artificial dissipation, then the approximation U^{n+1} at time level $n+1$ is determined from the approximation U^n at time level n by

$$\begin{aligned} U^{(n,1)} &= U^n, \\ U^{(n,2)} &= U^n + \frac{\Delta t}{2} ASU^{(n,1)}, \\ U^{(n,3)} &= U^n + \frac{\Delta t}{2} ASU^{(n,2)}, \\ U^{(n,4)} &= U^n + \Delta t ASU^{(n,3)}, \end{aligned}$$

and

$$\begin{aligned} U^{n+1} &= (1 + \frac{2\gamma h^{10}}{\pi^9} \frac{d^{10}}{dx^{10}}) U^n \\ &+ \frac{\Delta t}{6} AS(U^{(n,1)} + 2U^{(n,2)} + 2U^{(n,3)} + U^{(n,4)}). \end{aligned} \tag{3.17}$$

4 Accuracy Analysis

4.1 Truncation Error for Finite Difference

In this section, we analyze the local truncation error of the eigenfunction for the 2nd order finite difference method. We first establish

Theorem 4.1 *Consider (2.1) with coefficient (2.2) and a partition of* $[-1,1]$ *with equal step* $h = \frac{2}{N}$ *by grid points* $x_j = \frac{2j}{N} - 1, \quad j = 0, \cdots, N$. *Assume the two discontinuous points* $-1/2$ *and* $1/2$ *align with two grid points, and at these two points, we use the harmonic average* $a(-\frac{1}{2}) = a(\frac{1}{2}) = \frac{2}{3}$ *as coefficients. Then,*

$$T_j^k = a(x_j)(D_0\phi_k(x_j) - \frac{d\phi_k(x_j)}{dx}) = \begin{cases} O(\lambda_k^2 h), & x_j = -1/2, 1/2, \\ O(\lambda_k^3 h^2), & otherwise, \end{cases} \tag{4.18}$$

where $\phi_k(x)$ *is defined in (2.4).*

Proof. Since

$$\phi_k(x) = \exp(-\lambda_k \mu \int_{-1}^{x} \frac{1}{a(\xi)} d\xi),$$

we have

$$\begin{aligned} \phi_k(x-h) &= \exp(-\lambda_k \mu \int_{-1}^{x-h} \frac{1}{a(\xi)} d\xi) = \phi_k(x) \exp(-\lambda_k \mu \int_{x}^{x-h} \frac{1}{a(\xi)} d\xi), \\ \phi_k(x+h) &= \exp(-\lambda_k \mu \int_{-1}^{x+h} \frac{1}{a(\xi)} d\xi) = \phi_k(x) \exp(-\lambda_k \mu \int_{x}^{x+h} \frac{1}{a(\xi)} d\xi). \end{aligned}$$

Simple calculation using Taylor expansion shows

$$2hD_0\phi_k(x_j) = 2h\phi_k(x_j)(-\frac{\lambda_k\mu}{a_j} + O(\lambda_k^3h^2)) = 2h\phi_k(x_j)(-\frac{\lambda_k\mu}{a_j} + R_j)$$

for a continuous point x_j of $a(x)$, and for $x_j = -\frac{1}{2}$ or $\frac{1}{2}$,

$$2hD_0\phi_k(x_j) = 2h\phi_k(x_j)(-\frac{\lambda_k\mu}{a_j} + O(\lambda_k^2h)) = 2h\phi_k(x_j)(\frac{-\lambda_k\mu}{a_j} + R_j),$$

where

$$R_j = \begin{cases} O(\lambda_k^2h), & x_j = -1/2, 1/2, \\ O(\lambda_k^3h^2), & other \quad x_j. \end{cases}$$

□

Denote now by $(\phi_{k,j}^h, \Lambda_k)$ the kth eigenpair of the following discrete eigen problem

$$-a_jD_0\phi_{k,j}^h = \Lambda_k\phi_{k,j}^h, \quad k, j = 1, \cdots, N, \tag{4.19}$$

where the discrete eigenfunction is orthonormal with a discrete weighting function $\{1/a_j\}_{j=1}^N$,

$$\sum_{j=1}^{N} \frac{1}{a_j}\phi_{k,j}^h\bar{\phi}_{s,j}^h h = \delta_{k,s}.$$

We are ready to show the following

Theorem 4.2 *Under the assumptions in Theorem 4.1, the L_2 norm of the difference between the continuous and the discretized eigenfunction defined in (4.19) is bounded by that of the corresponding truncation error. That is,*

$$\|\phi_k - \phi_k^h\|_2 \leq C\|T^k\|_2,$$

where C is a constant independent of h, k.

Proof. Note first that

$$\bar{\Lambda}_k = -\Lambda_k, \forall k.$$

We then introduce by $\|\cdot\|_a$ a energy norm. For any vector $u = (u_1, \cdots, u_N)^T$,

$$\|u\|_a^2 = \sum_{j=1}^{N} \frac{1}{a_j}u_j\bar{u_j}h.$$

Since $\{a_j\}_{j=1}^N$ is bounded positive, $\|u\|_a$ is equivalent to $\|u\|_2$. Recall that

$$-a(x)\frac{d\phi_k(x)}{dx} = \lambda_k \phi_k(x)$$

with $\lambda_k = \frac{2}{3}\pi ik$, $\phi_k(x) = \exp(-\frac{2}{3}\pi ik \int_{-1}^{x} \frac{1}{a(s)} ds)$, we have

$$T_j^k = -a_j D_0 \phi_k(x_j) + a_j \frac{d\phi_k(x_j)}{dx} = -a_j D_0 \phi_k(x_j) - \lambda_k \phi_k(x_j).$$

Write into a matrix form,

$$(L_h - \lambda_k I)(\phi_k(x_1), \cdots, \phi_k(x_N))^T = (T_1^k, \cdots, T_N^k)^T,$$

where L_h is a matrix with entries generated from $-a_j D_0$ and I is the identical matrix, we can establish the following

$$|\Lambda_k - \lambda_k| \leq \frac{\|T^k\|_2}{\|\phi_k\|_2} = \|T^k\|_2.$$

Denote $e_{k,j}^h = \phi_k(x_j) - \phi_{k,j}^h$,

$$\begin{aligned}
\Lambda_k e_{k,j}^h &= \Lambda_k(\phi_k(x_j) - \phi_{k,j}^h) \\
&= (\Lambda_k - \lambda_k)\phi_k(x_j) + \lambda_k \phi_k(x_j) - \Lambda_k \phi_{k,j}^h \\
&= (\Lambda_k - \lambda_k)\phi_k(x_j) - a_j(\frac{d\phi_k(x_j)}{dx} - D_0 \phi_{k,j}^h) \\
&= (\Lambda_k - \lambda_k)\phi_k(x_j) - a_j D_0(\phi_k(x_j) - \phi_{k,j}^h) - a_j(\frac{d\phi_k(x_j)}{dx} - D_0\phi_k(x_j)) \\
&= (\Lambda_k - \lambda_k)\phi_k(x_j) - a_j D_0 e_{k,j}^h - T_j^k.
\end{aligned}$$

Taking inner product (denoted by $< \cdot, \cdot >$) by $\phi_{\xi,j}^h$ on both sides,

$$< D_0 e_{k,j}^h, \phi_{\xi,j}^h >= -\Lambda_k < \frac{e_{k,j}^h}{a_j}, \phi_{\xi,j}^h > -(\Lambda_k - \lambda_k) < \frac{\phi_k(x_j)}{a_j}, \phi_{\xi,j}^h > - < \frac{T_j^k}{a_j}, \phi_{\xi,j}^h > .$$

That is,

$$< e_{k,j}^h, \frac{\Lambda_\xi - \Lambda_k}{a_j} \phi_{\xi,j}^h >= (\Lambda_k - \lambda_k) < \frac{\phi_k(x_j)}{a_j}, \phi_{\xi,j}^h > - < \frac{T_j^k}{a_j}, \phi_{\xi,j}^h > .$$

When $\xi \neq k$,

$$< e_{k,j}^h, \frac{\phi_{\xi,j}^h}{a_j} >= \frac{\lambda_k - \Lambda_k}{\Lambda_\xi - \Lambda_k} < \frac{\phi_k(x_j)}{a_j}, \phi_{\xi,j}^h > - \frac{1}{\Lambda_k - \Lambda_\xi} < \frac{T_j^k}{a_j}, \phi_{\xi,j}^h > .$$

Since $e^h_{k,j} = \sum_\xi < e^h_{k,j}, \frac{\phi^h_{\xi,j}}{a_j} > \phi^h_{\xi,j}$,

$$\begin{aligned}
\|e^h_{k,j}\|^2_a &= \sum_{\xi \neq k} | < e^h_{k,j}, \frac{\phi^h_{k,j}}{a_j} > |^2 h + | < e^h_{k,j}, \frac{\phi^h_{k,j}}{a_j} > |^2 h \\
&\leq \frac{1}{\min_{\xi \neq k} |\Lambda_k - \Lambda_\xi|^2} (|\lambda_k - \Lambda_k|^2 \|\phi_k\|^2_a + \|T^k\|^2_a) + | < e^h_{k,j}, \frac{\phi^h_{k,j}}{a_j} > |^2 h \\
&\leq \frac{C}{\min_{\xi \neq k} |\Lambda_k - \Lambda_\xi|^2} \|T^k\|^2_a + | < e^h_{k,j}, \frac{\phi^h_{k,j}}{a_j} > |^2 h.
\end{aligned}$$

Note that

$$| < e^h_{k,j}, \frac{\phi^h_{k,j}}{a_j} > |^2 h \leq \|e^h_k\|^2_a \|\phi^h_k\|^2_a h = \|e^h_k\|^2_a h.$$

Therefore,

$$\|e^h_k\|_a \leq \frac{C\|T^k\|_a}{\min_{\xi \neq k} |\Lambda_k - \Lambda_\xi|^2 \sqrt{1-h}}.$$

Without any other notation, C always denotes a constant independent of h, k. Since Λ_k is pure imaginary for all k, and we can show that $|\Lambda_1| < |\Lambda_2| < \cdots < |\Lambda_N|$ for small h,

$$\min_{\xi \neq k} |\Lambda_k - \Lambda_\xi| = \min\{|\Lambda_k - \Lambda_{k+1}|, |\Lambda_k - \Lambda_{k-1}|\}.$$

Without loss of generality, we assume $\|T^{k+1}\|_2 \geq \|T^k\|_2$ and

$$\min_{\xi \neq k} |\Lambda_k - \Lambda_\xi| = |\Lambda_k - \Lambda_{k+1}|.$$

$$\begin{aligned}
|\Lambda_k - \Lambda_{k+1}| &\geq |\lambda_k - \lambda_{k+1}| - |\Lambda_k - \lambda_k| - |\Lambda_{k+1} - \lambda_{k+1}| \\
&\geq C - |\Lambda_k - \lambda_k| - |\Lambda_{k+1} - \lambda_{k+1}| \\
&\geq C - \|T^k\|_2 - \|T^{k+1}\|_2 \\
&\geq C - 2\|T^{k+1}\|_2 > 0.
\end{aligned}$$

Therefore,

$$\|e^h_k\|_a \leq \frac{\|T^{k+1}\|_a}{C - 2\|T^{k+1}\|_2} \frac{1}{\sqrt{1-h}} \leq C\|T^{k+1}\|_2 \leq C\|T^k\|_2.$$

□

Remark. If the discontinuous points align with grid points, we may take the harmonic average values at those points. In so doing, the global error can be improved to be $O(h\sqrt{h})$

in l_2 norm. If the discontinuous points do not align with grid points, assigning the harmonic average values at those points is not possible. By taking the exact point values, the global error we can obtain is in general $O(h)$. However, either one of them would give us an accuracy $O(h)$ in l_∞ norm.

The order of numerical methods are usually determined by the number of eigenfunctions of the methods that can well approximate those of analytic ones. In our case, since we can explicitly write all the operators of the numerical methods, we are able to evaluate all the corresponding eigenpairs.

4.2 Eigenvalues

The positive eigenvalues of the analytic problem are

$$\lambda_k = 2\pi k/3.$$

By taking $N = 32$ and $N = 64$ for the number of grid points, the eigenvalues of the different approximations are calculated and presented in Table 11 and Table 12. Throughout the rest of the paper, the notation "PS" stands for the pseudospectral method, "PS-mod" denotes the pseudospectral method with the coefficient $b(x)$. "FD2", $\cdots$, "FD12" denotes 2nd to 12th order finite difference method, respectively. The values in the tables are divided by $2\pi/3$. The values at points of discontinuity for coefficient $a(x)$ are as in (5.22). From the comparison of eigenvalues, the pseudospectral method is not better than finite difference methods. And the higher order finite difference method is no longer superior to lower order ones either. It is evident that if $k << N$ then the eigenvalues $\tilde{\lambda}_k$ of PS-mod are much closer to the analytic ones than those of finite difference methods and PS. For a given required accuracy and a given number of eigenvalues, PS-mod is the best. However, as the number of the grid points increases, the accuracies of all the methods is seemingly the same.

Remark. In Table 11 and Table 12, we arrange the eigenvalues with respect to the frequencies of the corresponding eigenfunctions. In the columns corresponding to the analytic and pseudospectral methods, the number of the eigenvalues increases with the frequencies of the corresponding eigenfunctions. For the finite difference methods, the case is totally different. To illustrate this in more detail, consider an arbitrary fourier eigenmode denoted by $e^{i\omega x}$ and the operator of the 2nd order finite difference method. Then,

$$\frac{de^{i\omega x}}{dx} = i\omega e^{i\omega x},$$

and

$$D_0 e^{i\omega x} = i\frac{\sin(\omega h)}{h} e^{i\omega x}.$$

When the frequency ω goes from 0 to π/h, the eigenvalue of the finite difference method is $i\frac{\sin(\omega h)}{h}$ behaving exactly as a *sin* function. At $\omega = \frac{\pi}{h}$, the corresponding eigenvalue is 0. However, the functions for the analytic and pseudospectral are $i\omega$ increasing with ω.

4.3 Eigenvector

4.3.1 First Eigenvector

Our estimates tell us that the approximation error for the eigenfunctions is $O(h)$. This is confirmed by numerical experiments, using both 8th order finite difference method and the pseudospectral method in Table 1, Table 2 and Table 3.

norm	N=	16	32	64	128	256
2	FD8	0.1568	0.0778	0.0388	0.0194	0.0097
	PS	0.1246	0.0639	0.0323	0.0163	0.0082
∞	FD8	0.2014	0.0992	0.0494	0.0247	0.0123
	PS	0.1367	0.0709	0.0361	0.0182	0.0091

Table 1: Accuracy of the first eigenfunction with (5.21)

norm	N=	16	32	64	128	256
2	FD8	0.0986	0.0507	0.0259	0.0131	0.0066
	PS	0.0570	0.0282	0.0141	0.0070	0.0035
∞	FD8	0.1032	0.0531	0.0270	0.0137	0.0069
	PS	0.0868	0.0476	0.0248	0.0126	0.0064

Table 2: Accuracy of the first eigenfunction with (5.22)

The values of the coefficient at points of discontinuity in Table 1, Table 2 and Table 3 are as in (5.21), (5.22) and (5.23) respectively (see section 5). In all these tables, we get $O(h)$ accuracy in l_∞ norm. This is also true in l_2 norm in Table 1 and Table 2. In Table 3 where we take the harmonic values at the discontinuous points and consider the pseudospectral method for the modified problem (2.7), the error in l_2 norm is increased from $O(h)$ to $O(h\sqrt{h})$. But, the pseudospectral method for the modified problem still gives us $O(h)$ accuracy in l_∞ norm, due to the difference between the two analytic solutions in

norm	N=	16	32	64	128	256
2	FD8	0.0287	0.0089	0.0029	0.0010	3.4743E-4
	PS	0.0251	0.0089	0.0031	0.0011	3.9153E-4
	PS-mod	0.0144	0.0049	0.0017	5.9968E-4	2.1125E-4
∞	FD8	0.0476	0.0227	0.0111	0.0055	0.0027
	PS	0.0498	0.0249	0.0124	0.0062	0.0031
	PS-mod	0.0273	0.0134	0.0067	0.0033	0.0017

Table 3: Accuracy of the first eigenfunction with (5.23)

norm	N=	16	32	64	128	256
2	PS-mod	0.0039	8.9968e-04	2.1335e-04	5.1556e-05	1.2613e-05
∞	PS-mod	0.0042	0.0011	2.9377e-04	8.0042e-05	2.1793e-05

Table 4: Accuracy of the first eigenfunction for modified problem

(2.1) and (2.7). But for points away from discontinuity, the accuracy should be $O(h^2)$. If we only consider the accuracy of the approximation for (2.7) (instead of (2.1)), we can get $O(h^2)$ both in l_2 and l_∞ norms (see Table 4), which is determined by the construction of $b(x)$ for $a(x)$.

4.3.2 Comparison of Other Eigenvectors

We now compare other eigenfunctions of finite difference methods and the pseudospectral method with the analytic ones. To do so, we take N equal to 64, and the values at discontinuous points to be (5.22). The k-th eigenfunction corresponds to the k-th eigenvalue, for $k = -N/2, \cdots, N/2 - 1$. We define the kth eigenfunction of the numerical methods by checking if its corresponding eigenvalue is the closest one to the kth analytic eigenvalue. As we can see in Figures 1 to 4, the pseudospectral method gives a good representation of the analytic eigenfunctions for large eigenvalues.

5 Numerical Tests

In this section approximations of the soution of equation (2.1) are determined with the following initial value,

$$u(x,0) = f(x) = \exp(-60x^2), \quad x \in [-1,1]. \tag{5.20}$$

The errors in l_2 and l_∞ norms of the pseudospectral method, 2nd to 12th order finite difference methods and Fourier Garlerkin methods are compared after 1 analytic temporal period.

5.1 Jumps Not Aligning With Grid Points

When the jumps do not align with grid points, directly taking the values of the coefficient at the grid points usually gives an accuracy of $O(h)$. To see this, consider the original problem (2.1) with a slightly shift ϵ(less than h) for coefficient $a(x)$

$$a(x) = \begin{cases} 0.5, & x \in (-0.5-\epsilon, 0.5-\epsilon), \\ 1, & x \in [-1,1]/(-0.5-\epsilon, 0.5-\epsilon). \end{cases}$$

We have

$$a(-0.5) = 0.5, \quad a(0.5) = 1. \tag{5.21}$$

The error of the approximation is presented in Table 5 and Fourier Garlerkin method, denoted by "Garlerkin", is also tested.

As shown in Table 5, the accuracy is $O(h)$ for the finite difference methods and for the Fourier Garlerkin method as well. FD6 is faster than FD8 here, which is caused by the discontinuity of the coefficient $a(x)$. Table 6 is obtained by adding dissipation to the schemes. Faster convergence can be obtained by reducing the high frequency modes coming from the discontinuity of $a(x)$. Note that as the order of the finite difference method increases, the accuracy converges to PS's. In Table 5 and 6, the pseudospectral method has almost $O(h^2)$. We explain the result as the symmetric structure of $a(x)$. In Table 7, we test the pseudospectral method by considering the case where

$$a(x) = \begin{cases} 0.5, & x \in (-0.5+\epsilon, 0.5-\epsilon), \\ 1, & x \in [-1,1]/(-0.5+\epsilon, 0.5-\epsilon). \end{cases}$$

Here

$$a(-0.5) = 1, \quad a(0.5) = 1. \tag{5.22}$$

In the calculations, we take $\epsilon = 1/1600$. Hence, the analytic period of time is $3 - 2\epsilon = 2.99875$. The usual accuracy $O(h)$ for the pseudospectral method holds under these new values of the coefficient.

norm	N=	32	64	128	256	512
2	FD2	0.2718	0.1196	3.8801E-02	1.3454E-02	5.6908E-03
	FD4	0.1161	3.0405E-02	1.2286E-02	5.9391E-03	2.9510E-03
	FD6	4.6414E-02	8.8467E-03	4.9821E-03	2.5456E-03	1.2763E-03
	FD8	4.7689E-02	2.4720E-02	1.2522E-02	6.2406E-03	3.1101E-03
	FD10	6.7000E-02	3.1623E-02	1.5444E-02	7.6833E-03	3.8386E-03
	FD12	6.5851E-02	2.8301E-02	1.3816E-02	6.8813E-03	3.4397E-03
	PS	3.1714E-03	5.0574E-04	1.0228E-04	2.3101E-05	5.5019E-06
	Galerkin	1.5034E-02	6.7207E-03	3.2234E-03	1.5855E-3	7.8779E-4
∞	FD2	0.5338	0.2923	8.3895E-02	2.0928E-02	6.5619E-03
	FD4	0.1792	5.6035E-02	2.5319E-02	1.2245E-02	6.0712E-03
	FD6	8.3810E-02	1.6688E-02	9.2034E-03	4.7101E-03	2.3550E-03
	FD8	7.9152E-02	3.8458E-02	1.8202E-02	8.8561E-03	4.3748E-03
	FD10	8.7014E-02	3.5695E-02	1.7159E-02	8.4787E-03	4.2285E-03
	FD12	0.1083	3.3149E-02	1.5119E-02	7.4288E-03	3.6909E-03
	PS	7.6521E-03	1.0566E-03	1.9869E-04	4.1411E-05	9.3782E-06
	Galerkin	3.1276E-02	1.4339E-02	7.0307E-03	3.4195E-3	1.6858E-3

Table 5: $\Delta t = 0.01h$, $\gamma = 0$, final time equals 3.

5.2 Jumps Aligning With Grid Points

Consider now the case (2.2) where the discontinuous points align with grid points. For this case we are able to assign the values of $a(x)$ at the discontinuous points to be the harmonic average. That is,

$$a(x) = \begin{cases} 2/3, & x = -0.5 \\ 2/3, & x = 0.5. \end{cases} \tag{5.23}$$

From Table 8, it's clear that the finite difference schemes has an accuracy of $O(h^2)$, and the accuracy for the pseudospectral method is even better. This happens partially because of the specific structure of the coefficient and initial condition. We should also note that some cancellations caused by the symmetric data occur here.

5.3 Pseudospectral Method With Coefficient $b(x)$

Under the pseudospectral method with coefficient $b(x)$, the accuracy is improved from $O(h)$ to $O(h^2)$ as shown in Table 9. The approximation for (2.7) is calculated after 3 analytic

norm	N=	32	64	128	256	512
2	FD2	0.2077	0.1008	$2.8127E-2$	$7.0527E-3$	$1.8300E-3$
	FD4	$6.0897E-2$	$9.8626E-3$	$4.8435E-3$	$1.4237E-3$	$2.3856E-4$
	FD6	$3.0626E-2$	$5.7495E-3$	$2.7663E-3$	$4.0663E-4$	$3.2385E-5$
	FD8	$2.9323E-2$	$5.8767E-3$	$2.6648E-3$	$3.5896E-4$	$2.7255E-5$
	FD10	$1.2623E-2$	$1.7685E-3$	$6.9677E-4$	$8.4135E-5$	$5.4129E-5$
	FD12	$1.1173E-2$	$1.6729E-3$	$6.4647E-4$	$7.8288E-5$	$5.1680E-5$
	PS	$3.7246E-3$	$5.7748E-4$	$1.3565E-4$	$3.9678E-5$	$1.3960E-5$
∞	FD2	0.4082	0.2464	$7.2503E-2$	$1.7682E-2$	$4.5447E-3$
	FD4	0.1377	$2.3232E-2$	$1.0362E-2$	$3.0188E-3$	$5.0795E-4$
	FD6	$7.1106E-2$	$1.3652E-2$	$5.9420E-3$	$8.7226E-4$	$7.5566E-5$
	FD8	$7.1813E-2$	$1.3604E-2$	$5.7510E-3$	$7.7225E-4$	$6.4344E-5$
	FD10	$3.2796E-2$	$4.0848E-3$	$1.5495E-3$	$2.4314E-4$	$1.1352E-4$
	FD12	$2.9959E-2$	$3.9670E-3$	$1.4437E-3$	$2.0535E-4$	$8.4518E-5$
	PS	$8.0702E-3$	$1.3275E-3$	$2.9103E-4$	$8.2365E-5$	$2.9038E-5$

Table 6: Error with dissipations. $\Delta_t = 0.01h$. Final time equals 3.

norm	N=	32	64	128	256	512
2	PS	9.7319E-2	4.7086E-2	2.2470E-2	1.0241E-2	4.1424E-3
∞	PS	0.2033	1.0021E-1	4.7921E-2	2.1838E-2	8.8329E-3

Table 7: $\Delta t = 0.01h$, $\gamma = 0$, final time equals 2.99875.

temporal period. There is no difference between the analytic solutions of (2.1) and (2.7), since both have the same temporal period.

Acknowledgment. We want to express our gratitude to Bengt Fornberg, whose short course given at IMA motivated our interest in finding an explanation for the problem considered in this paper.

References

[1] D. L. BROWN, *A Note on the Numerical Solution of the Wave Equation With Piecewise Smooth Coefficients*, Mathematics of Computation, Vol. 42, No. 166, 1984, pp.369-

norm	N=	32	64	128	256	512
2	FD2	0.3868	0.2354	9.0539D-2	2.4744D-2	6.2381D-3
	FD4	0.1809	3.3606D-2	4.8228D-3	1.0429D-3	2.5622D-4
	FD6	8.8369D-2	5.2536D-3	6.7254D-4	1.6959D-4	4.2194D-5
	FD8	5.0305D-2	8.7696D-3	8.5339D-4	1.4887D-4	3.4516D-5
	FD10	6.6675D-2	1.9340D-2	4.5957D-3	1.1029D-3	2.7275D-4
	FD12	4.1083D-2	4.0960D-3	1.2958D-3	3.3263D-4	8.3718D-5
	PS	1.2916D-3	1.0546D-4	2.3944D-6	1.3839D-7	1.5182D-8
∞	FD2	0.5379	0.4703	0.2349	6.4592D-2	1.5800D-2
	FD4	0.3014	7.6635D-2	1.0146D-2	1.3509D-3	2.6898D-4
	FD6	0.1601	1.5550D-2	1.2512D-3	2.3931D-4	5.6745D-5
	FD8	8.2814D-2	1.2156D-2	1.2001D-3	2.2164D-4	4.7912D-5
	FD10	1.0220D-1	2.7691D-2	5.3558D-3	1.2161D-3	2.9632D-4
	FD12	7.1897D-2	4.7769D-3	1.3942D-3	3.3306D-4	8.2900D-5
	PS	3.2982D-3	4.6778D-5	5.0230D-6	2.7638D-7	6.4413D-8

Table 8: $\Delta t = 0.01h$, $\gamma = 0$, final time equals 9.

norm	N=	32	64	128	256	512
2	PS-mod	5.5522D-3	1.1314D-3	2.4840D-4	3.3470D-5	1.3399D-5
∞	PS-mod	1.3857D-2	2.5183D-3	5.2898D-4	6.1897D-5	1.4727D-5

Table 9: $\Delta t = 0.01h$, $\gamma = 0$, final time equals 9.

391.

[2] C. CANUTO, M. Y. HUSSAINI, A. QUARTERONI, AND T. A. ZANG, *Spectral Methods in Fluid Dynamics*, Springer-Verlag, 1988.

[3] B. GUSTAFSSON, H. O. KREISS, AND J. OLIGER, *Time Dependent Problems and Difference Methods*, 1994.

[4] H. O. KREISS AND J. OLIGER, *Methods for the Approximate Solution of Time Dependent Problems*, GARP, No. 10, 1973.

[5] B. FORNBERG, *The pseudospectral Methods: Accurate Representation of Interfaces in Elastic Wave Calculations*, Geophysics, Vol. 53, No. 5. (1988), pp. 625-637.

N=	32	64	128	256	512
FD2	γ_1	γ_1	γ_1	γ_2	γ_3
FD4	γ_1	γ_1	γ_1	γ_2	γ_3
FD6	γ_2	γ_2	γ_2	γ_3	γ_4
FD8	γ_2	γ_2	γ_2	γ_3	γ_4
FD10	γ_3	γ_3	γ_3	γ_4	γ_5
FD12	γ_3	γ_3	γ_3	γ_4	γ_5
PS	γ_6	γ_6	γ_6	γ_6	γ_6

Table 10: The dissipation parameter used in Table 6. $\gamma_1 = E - 4$, $\gamma_2 = E - 5, \gamma_3 = E - 6$, $\gamma_4 = E - 7$, $\gamma_5 = E - 8$, $\gamma_6 = E - 3$. The dissipation term is $Q_{12,2}$ for the finite difference methods, except the two cases with γ_5 where it is $Q_{14,2}$.

[6] B. FORNBERG, *The pseudospectral Method: Comparisons with Finite Differences for the Elastic Wave Equation*, Geophysics, Vol. 52, No. 4. (1987), pp. 483-501.

[7] B. FORNBERG, *On a Fourier Method For the Integration of Hyperbolic Equations*, SIAM J. Numer. Anal. Vol. 12, No. 4. (1975), pp. 509-528.

[8] B. FORNBERG, *A practical Guide to Pseudospectral methods*, Cambridge University Press, (to appear).

[9] D. GOTTLIEB, AND S. A. ORSZAG, *Numerical Analysis of Spectral Methods: Theory and Application*, SIAM, Philadelphia, Vol. 26, 1977.

Analytic	PS	PS-mod	FD8	FD6	FD4	FD2
16.0000	22.2489	22.0982	0	0	0	0
15.0000	19.8029	19.5177	2.6527	2.2173	1.6759	1.0263
14.0000	17.4273	17.0101	4.9130	4.2152	3.2261	1.9592
13.0000	15.1367	14.6156	6.3823	5.5055	4.4369	2.7529
12.0000	13.0216	12.4933	12.9685	11.9026	4.7574	3.4542
11.0000	11.3617	11.0578	12.2427	11.2827	10.3159	3.4824
10.0000	10.2847	10.0401	11.1287	10.3139	9.8257	4.2914
9.0000	9.2229	9.0226	9.7456	9.0826	9.0452	6.7009
8.0000	8.1913	8.0094	8.2286	7.6978	8.0275	7.5323
7.0000	7.1625	7.0059	6.7813	6.3443	6.8479	7.2152
6.0000	6.1339	6.0038	5.9395	5.8380	5.6357	6.0120
5.0000	5.1102	5.0014	5.0706	4.9884	5.1309	5.1851
4.0000	4.0871	4.0006	4.0749	4.0077	3.9837	3.7644
3.0000	3.0644	3.0003	3.0634	3.0579	3.0198	2.8855
2.0000	2.0428	2.0001	2.0400	2.0317	2.0438	1.9834
1.0000	1.0213	1.0000	1.0213	1.0212	1.0208	0.9985

Table 11: Positive Eigenvalues for $N = 32$.

JUN 2 8 2004

Carnegie Mellon University Libraries
3 8482 01423 7339

Analytic	PS	PS-mod	FD8	FD6	FD4	FD2
32.0000	45.8393	45.7490	0	0	0	0
31.0000	43.1450	42.9778	2.6699	2.2155	1.6777	1.0144
30.0000	40.4984	40.2504	5.2430	4.3690	3.3225	2.0115
29.0000	37.8717	37.5410	7.6152	6.4005	4.8838	2.9412
28.0000	35.2647	34.8519	9.6960	8.2634	6.3602	3.8733
27.0000	32.6853	32.1943	11.4302	9.7182	7.6821	4.7500
26.0000	30.1493	29.5898	12.4716	10.7553	8.6475	5.5133
25.0000	27.6884	27.0828	12.7621	11.4436	9.4698	5.6393
24.0000	25.3807	24.7927	25.8441	11.8132	9.7214	6.3016
23.0000	23.4933	23.1295	26.2908	23.7289	10.2605	6.809
22.0000	22.3705	22.1027	25.1175	24.1053	10.4219	7.3052
21.0000	21.2986	21.0612	24.1367	23.1129	20.5742	7.4147
20.0000	20.2743	20.0370	22.9340	22.2738	20.8676	7.6005
19.0000	19.2494	19.0314	21.5465	21.2332	20.0913	14.2920
18.0000	18.2227	18.0229	20.0134	20.0172	19.4275	14.7205
17.0000	17.2066	17.0147	18.3740	18.6546	18.5950	15.2164
16.0000	16.1905	16.0122	16.6685	17.1769	17.6087	15.0297
15.0000	15.1732	15.0095	14.9462	15.6182	16.4865	13.7485
14.0000	14.1600	14.0061	13.3468	14.0214	15.2491	13.0957
13.0000	13.1467	13.0048	13.1206	12.4823	13.9209	12.3410
12.0000	12.1330	12.0038	11.8293	12.0224	12.5341	11.4935
11.0000	11.1211	11.0024	10.9809	11.0113	11.1498	10.5648
10.0000	10.1093	10.0017	10.0566	9.9927	10.0261	9.5718
9.0000	9.0973	9.0013	9.0729	8.9930	8.8724	8.5442
8.0000	8.0861	8.0008	8.0825	8.0018	7.8848	7.6175
7.0000	7.0749	7.0005	7.0722	7.0529	6.9595	6.9590
6.0000	6.0638	6.0003	6.0634	6.0450	6.0126	6.2809
5.0000	5.0531	5.0002	5.0522	5.0500	5.0407	4.8030
4.0000	4.0423	4.0001	4.0414	4.0364	4.0370	3.9267
3.0000	3.0316	3.0000	3.0320	3.0314	3.0279	3.0047
2.0000	2.0211	2.0000	2.0204	2.0188	2.0221	1.9960
1.0000	1.0105	1.0000	1.0105	1.0105	1.0105	1.0022

Table 12: Positive Eigenvalues for $N = 64$