

IMPLICIT SOLUTIONS TO THE WAVE EQUATION BASED ON THE METHOD OF
LINES TRANSPOSE

By

Gerard Lee Van Groningen

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Mathematics

2012

ABSTRACT

IMPLICIT SOLUTIONS TO THE WAVE EQUATION BASED ON THE METHOD OF LINES TRANSPOSE

By

Gerard Lee Van Groningen

We present a numerical method for computing the wave equation implicitly. The approach discretizes the wave equation in time using the method of lines transpose, also known as the transverse method of lines or Rothe's method. This differs from conventional methods in that we solve the resulting system of ODEs using boundary integral methods. We then analyze the fully discretized solution resulting from the midpoint and trapezoidal quadrature rules and show that convergent and unconditionally stable schemes result. We also show that the choice of discretization in time can lead to various schemes of prescribed accuracy, which may or may not introduce numerical dissipation into the approximate solution.

We start with the simplest case of solving the wave equation in one dimension using either a free space solution or Dirichlet or Neumann Boundary conditions. We then analyze the stability and consistency of the method, as well as investigating the dispersion relations and deriving the phase error. Next, some numerical examples are presented which give validation to the error estimates. Further, the method is adapted for both outflow boundary conditions, using either one-way waves or a perfectly matched absorbing layer, as well as the implementation of a soft source. Finally, we utilize an ADI scheme to explore solving the wave equation in higher dimensions.

Since this method of lines transpose approach is implicit, it removes the usual CFL stability limit inherent in explicit time stepping methods for solving the wave equation, and thus the algorithm will be more efficient than these explicit methods.

TABLE OF CONTENTS

List of Tables	v
List of Figures	vi
1 Introduction	1
1.1 Literature Review	1
1.2 Prerequisites	8
1.2.1 Eigenvalues of Perturbed Tri-diagonal Toeplitz Matrices	8
1.2.2 Schur Polynomial Test	12
2 MOL^T Formulation and Integral Solution	15
2.1 Formulation	15
2.1.1 Dissipative Schemes	15
2.1.2 Purely Dispersive Schemes	16
2.2 Integral Solution	17
2.3 Numerical Approach	22
3 Analysis of the Free Space Solution	27
3.1 Convergence	27
3.1.1 Consistency of the MOL ^T method	28
3.1.1.1 Non-dissipative Scheme	28
3.1.1.2 Dissipative Schemes	30
3.1.1.3 Fully Discrete Schemes	33
3.1.2 Stability of the MOL ^T method	36
3.1.2.1 The Eigenvalues of A_p	38
3.1.2.2 Dissipative Schemes	42
3.1.2.3 Non-dissipative Scheme	45
3.2 Dispersion	46
3.2.1 Semi-discrete Non-dissipative Scheme	47
3.2.2 Semi-discrete Dissipative Schemes	50
3.2.3 Fully Discrete Non-dissipative Scheme	54
3.2.4 Fully Discrete Dissipative Schemes	59

4	Analysis of the Solution on a Bounded Domain	64
4.1	Consistency	64
4.1.1	Non-dissipative Scheme	64
4.1.2	Dissipative Schemes	71
4.2	Stability	72
4.3	Numerical Results	83
4.4	Fast Implementation	88
5	Non-Reflecting Boundary Conditions	90
5.1	Perfectly Matched Layers	92
5.1.1	Constant Damping Function	95
5.1.1.1	Implementation	95
5.1.1.2	Implicit Update	99
5.1.1.3	Performance	106
5.1.2	Polynomial Damping Functions	108
5.1.2.1	Implementation	109
5.1.2.2	Optimization	129
5.2	One-Way Waves	130
5.2.1	Performance	139
6	MOL^T in Higher Dimensions Utilizing an Alternating Directly Implicit Scheme	141
6.1	Freespace Solution	142
6.2	Finite Domain	144
7	Soft Sources	148
7.1	Implementation	151
7.2	Higher Dimensions	153
8	Future Directions	155
8.1	Further Problems	155
8.2	Numerical Work	157
	Bibliography	159

LIST OF TABLES

Table 5.1	Reflection and Damping for the Dissipative 1st Order Scheme with a Piecewise Constant Damping Function	108
Table 5.2	Reflection and Damping for the Dissipative 2nd Order Scheme with a Piecewise Constant Damping Function	108
Table 5.3	Optimization of the Dissipative 1st Order Quadratic Damping Function	130
Table 5.4	Performance of the Dissipative 1st Order Scheme with a Non-reflecting Boundary	140

LIST OF FIGURES

Figure 3.1	Relative Phase Error for the Non-dissipative Scheme	60
Figure 4.1	Contour Plot of the Minimum Eigenvalue of the Dissipative Schemes for Zero Dirichlet Boundary Conditions with Midpoint Quadrature .	84
Figure 4.2	Contour Plot of the Minimum Eigenvalue of the Dissipative Schemes for Zero Neumann Boundary Conditions with Midpoint Quadrature	84
Figure 4.3	Convergence of the First Order Dissipative Scheme with $\nu = \Delta x/(c\Delta t) = 0.1$	85
Figure 4.4	Convergence of the Second Order Dissipative Scheme with $\nu/\sqrt{2} = \Delta x/(c\Delta t) = 0.05$	86
Figure 4.5	Convergence of the Second Order Non-dissipative Scheme with $\nu/\sqrt{2} = \Delta x/(c\Delta t) = 0.05$	87
Figure 5.1	Damping Profiles for Low Order Polynomials	115
Figure 6.1	Propagation of the First Order Dissipative ADI Scheme at $t = 0$. .	146
Figure 6.2	Propagation of the First Order Dissipative ADI Scheme at $t = 0.5$.	147
Figure 6.3	Propagation of the First Order Dissipative ADI Scheme at $t = 1.25$	147
Figure 7.1	Wave Introduction via Soft Source	152
Figure 7.2	Long Term Time Behavior of the Soft Source	153

Chapter 1

Introduction

1.1 Literature Review

The wave equation is an important hyperbolic Partial Differential Equation (PDE) that arises in acoustics, electromagnetics, and fluid dynamics. The focus of this dissertation is computing numerical solutions to the wave equation

$$\begin{aligned}v_{tt} &= c^2 \nabla^2 v, \quad x \in \Omega, \quad t \geq 0, \\v(x, 0) &= f(x), \quad v_t(x, 0) = g(x),\end{aligned}\tag{1.1}$$

where Ω may either consist of the entire region of $\mathbb{R}^n$, $n \in \{1, 2, 3\}$, or of a subset of $\mathbb{R}^n$, in which case boundary conditions must be applied.

We are particularly interested in the setting of Maxwell's equations as they arise in non-relativistic plasma situations. Maxwell's equations may be expressed in terms of a scalar

potential Ψ and a vector potential A as

$$E = -\nabla\Psi - \frac{\partial A}{\partial t}, \quad B = \nabla \times A, \quad (1.2)$$

where E represents the electric field and B represents the magnetic field. If one chooses the Lorenz gauge condition

$$\nabla \cdot A + \frac{1}{c^2} \frac{\partial \Psi}{\partial t} = 0,$$

(1.2) may be expressed in the form of a wave equation as

$$\frac{1}{c^2} \frac{\partial^2 \Psi}{\partial t^2} - \nabla^2 \Psi = \frac{\rho}{\epsilon}, \quad \frac{1}{c^2} \frac{\partial^2 A}{\partial t^2} - \nabla^2 A = \mu J,$$

where ρ is the charge density, ϵ is the free space permittivity, μ is the free space permeability and J is the current. In this situation, we wish to develop a numerical solver which will be able to handle the time scale gap between the electron plasma frequency and the speed of light.

There currently exist several well established approaches to computing a numerical solution to the wave equation. One such approach is to use a separable solution. That is, we assume the ansatz

$$v(x, t) = \psi(x)e^{-i\omega t},$$

which leads to the oscillatory Helmholtz equation

$$\nabla^2 \psi + k^2 \psi = 0,$$

where $k = \omega/c$ is the wave number and ω is the frequency of an eigenmode. While the resulting Helmholtz equation can be solved efficiently using fast summation methods [22, 41], this approach is only useful for generating time harmonic solutions to the wave equation.

A fairly standard approach for hyperbolic PDEs is the Method of Lines (MOL) discretization. The prevalent idea behind the Method of Lines was introduced in 1949 by Courant and Lax [28] as an adaptation for using the method of characteristics as a numerical solver for hyperbolic PDEs. Courant and Lax began with semi-linear problems in which the characteristics are known a priori and the solution to the PDE is then found using an iterative method over the resulting integral equations. However, when this approach was extended to non-linear problems, problems arose since the characteristics were not known beforehand. Therefore, the resulting integral equations might no longer be easily solved using the same iterative methods. Courant and Lax overcame this by replacing the characteristics with lines upon which the spatial variable x is held constant. This gave rise to a PDE which is discretized only in space, resulting in a system of Ordinary Differential Equations (ODEs), which can then be solved using whichever Initial Value Problem (IVP) solver is desired. Each individual ODE then represents a solution at some point on the grid as a function of time.

Many studies have been done on the numerical properties of the MOL discretization. Courant et al. [27] studied convergence based on the ratio of step sizes in different dimensions. Peter Lax [52] used his MOL discretization design to prove the existence of solutions to non-analytic initial value problems that result from quasi-linear hyperbolic PDEs. In

addition to its original purpose in solving hyperbolic PDEs, the method of lines has also been studied for both parabolic [71] and elliptic [49] equations as well. Much work has been done in establishing the accuracy and stability of the MOL discretization, including Verwer and Sanz-Serna [69], who showed convergence of both the space discretization based on a logarithmic matrix norm and convergence of the time integration based on C-stability, as well as Reddy and Trefethen [60], who used ϵ -pseudo-eigenvalues to obtain conditions that were both necessary and sufficient for the stability of the method of lines.

Another technique whose roots trace back to the method of characteristics are the semi-Lagrangian methods, first introduced by Courant et al. [27]. In a purely Lagrangian solver for the advection equation, the exact value of the solution at a given time is found by tracing back along its characteristic to find the value at the previous time step. However, this requires that the mesh be flexible and be able to move in time. As this is frequently impractical to implement, semi-Lagrangian methods require a fixed mesh and then trace back the characteristic to find its value at the previous time step. In the case of more complicated PDEs where the characteristic may not be easily found, the characteristic must instead be numerically approximated. In the likely event that this characteristic does not lie on a mesh point at the previous time step, polynomial interpolation will be used to determine the value of the point at which the characteristic does cross. This essentially reduces the problem to solving a set of independent ODEs which will approximate the characteristic and then determining the function value at those points using interpolation.

While originally based on advection equations (e.g. [62]), semi-Lagrangian methods are now very commonly used for a range of problems, such as the Vlasov equation [65], but especially in the atmospheric sciences [63, 66] due to its beneficial numerical properties, such

as the lack of restriction based on a Courant number [19]. Since the actual semi-Lagrangian method can vary in order based on both the approximation of the characteristic and the order of the interpolation used, a variety of studies have been done on the stability, convergence and cost effectiveness of these various implementations [10, 19, 34, 53, 66].

Another common solver for hyperbolic PDEs is the finite volume methods, beginning in 1959 with Godunov's method [38]. Godunov divided his domain into cells and found the average of the numerical solution over each cell. He then constructed a Riemann problem which consisted of the PDE whose initial conditions were given by the piecewise constant function corresponding to the average numerical value over each cell. Next, he evolved the solution to this Riemann problem for a time step and then recalculated the average value over each cell. This process was then repeated until the desired final time was reached. The most significant benefit of Godunov type schemes was a greater ability to handle shocks compared to other contemporary numerical solvers, due to procuring the average value of the numerical solution over each cell. The further study of the convergence and other numerical properties of Godunov-type schemes have been studied for a range of problems [29, 45, 54, 59]. Current examples of finite volume methods whose roots trace back to Godunov-type schemes include the weighted essentially non-oscillatory (WENO) method, which uses an adaptive stencil and a combination of lower order reconstructions to obtain a higher order solution to the PDE, as well as the discontinuous Galerkin (DG) method, which constructs a solution over each element using lower order polynomials before updating the inter-element terms.

Godunov-type schemes are not without their drawbacks, however. Due to the Riemann problem's initial condition of a piecewise constant function, the scheme was limited to first order. This order may be made higher though, if the initial condition uses piecewise poly-

nomials instead of piecewise constant functions. However, making this change will increase the complexity of the problem and higher order schemes may still be hard to achieve. Additionally, the scheme is restricted in that the size of the time step is bounded by half the size of the spatial step multiplied by the fastest signal in the solution.

Integral equation methods, such as [6, 39, 40, 51], may also be used to solve hyperbolic problems. The benefits of integral equation methods, in comparison to traditional finite difference and finite element methods, include a much simpler incorporation of complex boundary geometries and the easier attainment of higher order. The biggest drawback to integral equation methods, however, is that they possess a large computational complexity, with a time independent problem requiring a dense $N^2 \times N^2$ matrix for an $N \times N$ grid. Advances are being made in reducing the amount of necessary operations, most notably the fast multipole method (FMM) [25, 61] for the wave scattering problem. In FMM, if particles are located within each other's near-field, their interaction is calculated directly. Otherwise, a particle's contribution to its nearest multipole is calculated and the interaction between multipoles is then calculated by traversing up and down a tree.

A further example of an integral equation method that is of particular interest comes from Alpert, Greengard and Hagstrom [5]. Here, the authors take the wave equation and use a Laplace transform to create the equivalent PDE in the frequency domain. They then take a centered finite difference of the temporal derivative followed by an inverse Laplace transform to revert to the time domain. This gives that the finite difference approximation of the temporal derivative is equivalent to the integral of the spatial derivative multiplied by the appropriate Green's function. This evolution scheme can be seen as an integral form of the Lax-Wendroff scheme and the numerical properties of the scheme will depend upon the

technique used to compute the integral.

A less common approach is the Method of Lines Transpose (MOL^T), also known as the transverse methods of lines or Rothe’s method or the horizontal line method, where the time derivative is replaced by an algebraic approximation and the resulting system of Boundary Value Problems (BVPs) must be solved. Here, each individual BVP represents the solution at a given point in time. This approach has been relatively sparsely considered [8, 46, 56], in part because the numerical solution to BVPs may be more challenging to find than for the ODEs that result from the standard method of lines, especially in the case where the BVP is stiff or contains boundary layers. While the MOL^T may not be as widely used as the MOL, it does not possess all of the same drawbacks and thus has been used in a variety of applications such as solving the heat equation [21, 51], defect correction methods [48], evolution equations [50], etc. With the continued development of boundary integral methods [36, 55], the method of lines transpose approach may be very beneficial for certain classes of problems. For example, Li et al. [55] constructed recurrence relations for efficiently evaluating derivatives to the Green’s kernel for the non-oscillatory Helmholtz equation in $\mathbb{R}^3$.

In considering the vast array of previously developed temporal discretizations, two classes of numerical schemes emerge: those which introduce numerical diffusion (i.e. dissipative schemes), and those that do not (i.e. non-dissipative schemes). The non-dissipative schemes will be useful for long time simulations that must maintain the strength of the wave amplitudes. The diffusive schemes are also of interest, for instance, in asymptotic preserving applications in which the steady state approximation must be recovered.

1.2 Prerequisites

The following results that will be useful in showing the convergence of the MOL^T scheme.

1.2.1 Eigenvalues of Perturbed Tri-diagonal Toeplitz Matrices

In section 3.1.2.1, we will use the inverse of a Toeplitz matrix to help determine stability of the method. This inverse comes from Yueh and Cheng's 2007 paper [70] on finding the eigenvalues and inverses of a Tri-Diagonal Toeplitz matrix with perturbed corners. The relevant matrix will be of the form

$$A_n = \begin{bmatrix} b + \gamma & a & 0 & \dots & 0 \\ a & b & a & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & a & b & a \\ 0 & \dots & 0 & a & b + \gamma \end{bmatrix}.$$

Let λ be an eigenvalue of A_n with u its corresponding eigenvector. $A_n u = \lambda u$ can then be written as

$$\begin{aligned} bu_1 + au_2 &= \lambda u_1 - \gamma u_1 \\ au_1 + bu_2 + au_3 &= \lambda u_2 \\ &\vdots \\ au_{n-2} + bu_{n-1} + au_n &= \lambda u_{n-1} \\ au_{n-1} + bu_n &= \lambda u_n - \gamma u_n \end{aligned}$$

which is equivalent to

$$au_{k-1} + (b - \lambda)u_k + au_{k+1} = 0, \quad k = 1, 2, \dots, n, \quad (1.3)$$

$$au_0 = \gamma u_1,$$

$$au_{n+1} = \gamma u_n.$$

Let $u_i, i \in \{0, 1, \dots, n+1\}$ represent the first $n+1$ terms of an infinite complex sequence $u = \{u_i\}_{i=1}^\infty$. The equations in (1.3) then satisfy the recurrence relation

$$au_{k-1} + bu_k + au_{k+1} = \lambda u_k + f_k, \quad k \in \mathbb{N}$$

under the condition

$$au_0 = au_{n+1} = 0$$

with

$$f_k = \begin{cases} -\gamma u_1, & k = 1 \\ -\gamma u_n, & k = n \\ 0, & \text{elsewise.} \end{cases}$$

This relation is equivalent to

$$\left(a\bar{h}^2 + (b - \lambda)\bar{h} + \bar{a} \right) u = (a\bar{u}_1 + f)\bar{h}$$

where

$$\bar{h} = \{0, 1, 0, \dots\},$$

$$\bar{h}^2 = \{0, 0, 1, 0, \dots\},$$

$$\bar{a} = \{a, 0, 0, \dots\},$$

$$\bar{u}_1 = \{u_1, 0, 0, \dots\},$$

$$f = \{f_k\}_{k=1}^{\infty}.$$

Note that $\bar{h}^2 = \bar{h} * \bar{h}$, where $*$ is a convolution operator.

Since $a \neq 0$, this gives that

$$u = \frac{(a\bar{u}_1 + f)\bar{h}}{(a\bar{h}^2 + (b - \lambda)\bar{h} + \bar{a})}. \quad (1.4)$$

Consider the denominator as a quadratic in terms of $\bar{h}$, whose roots are given by

$$\delta_{\pm} = \frac{-(b - \lambda) \pm \sqrt{\omega}}{2a}, \quad \omega = (b - \lambda)^2 - 4a^2. \quad (1.5)$$

Then $\delta_+\delta_- = 1$ and thus,

$$\delta_{\pm} = \cos(\theta) \pm i \sin(\theta)$$

for some θ in the strip $\{z \in \mathbb{C} | 0 \leq \operatorname{Re}(z) < \pi\}$.

Consider first the case where $\theta \neq 0$. This will give that

$$\cos(\theta) = \frac{\lambda - b}{2a} \implies \lambda = b + 2a \cos(\theta). \quad (1.6)$$

Then, by partial fractions, equation (1.4) becomes

$$\begin{aligned} u &= \frac{1}{\omega} \left(\frac{1}{\delta_- - \bar{h}} - \frac{1}{\delta_+ - \bar{h}} \right) (a\bar{u}_1 + f)\bar{h} \\ &= \frac{2i}{\omega} \{\sin(j\theta)\} * (a\bar{u}_1 + f)\bar{h}. \end{aligned}$$

Evaluating this convolution product will give

$$u_j = \frac{2i}{\omega} \{au_1 \sin(j\theta) - \delta u_1 \sin((j-1)\theta) - H_j^n \delta u_n \sin((j-n)\theta)\}$$

where $H_j^n = 1$ if $j \geq n$ and $H_j^n = 0$ if $j < n$.

This will then give

$$a \sin(\theta) u_n = au_1 \sin(n\theta) - \delta u_1 \sin((n-1)\theta) \text{ and}$$

$$0 = au_1 \sin((n+1)\theta) - \delta u_1 \sin(n\theta) - \delta u_n \sin(\theta).$$

This can be rewritten as

$$Au_1 + Bu_n = 0$$

$$Cu_1 + Du_n = 0.$$

Since u_1 and u_n both can't be zero, this implies that

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix} = 0$$

which gives the necessary condition

$$a^2 \sin((n+1)\theta) + \delta^2 \sin((n-1)\theta) - 2a\delta \sin(n\theta) = 0.$$

For any θ that meets this qualification, the eigenvalues may then be found based on (1.6).

The case of $\theta = 0$ will be similar, but now instead with equation (1.5) giving

$$\begin{aligned} 1 &= \frac{-(b-\lambda) \pm \sqrt{(b-\lambda)^2 - 4a^2}}{2a} \\ \implies 2a &= -(b-\lambda) \pm \sqrt{(b-\lambda)^2 - 4a^2} \\ \implies (2a + (b-\lambda))^2 &= (b-\lambda)^2 - 4a^2 \\ \implies 4a^2 + 4a(b-\lambda) &= -4a^2 \\ \implies b-\lambda = -2a &\implies \lambda = b+2a. \end{aligned}$$

1.2.2 Schur Polynomial Test

The Schur polynomial test will be used in section 3.1.2.2 to guarantee that all the roots of a polynomial are in the unit disc. A polynomial $\phi_d(z) = \sum_{i=0}^d a_i z^i$ is said to be a Schur polynomial of exact degree d if $\phi(z)$ has exactly d roots, counting multiplicities, each of which has magnitude less than one. The roots of this polynomial may be hard to find, but the Schur polynomial test [67] may determine if all the roots are within the unit circle. To

check these conditions, define the corresponding polynomials

$$\phi_d^*(z) = \sum_{i=0}^d a_{d-i} z^i \text{ and}$$

$$\phi_{d-1}(z) = \frac{\phi_d^*(0)\phi_d(z) - \phi_d(0)\phi_d^*(z)}{z}.$$

Theorem 1 (Schur Polynomial Test). $\phi_d(z)$ is a Schur polynomial of exact degree d if $\phi_{d-1}(z)$ is a Schur polynomial of exact degree $d-1$ and $|\phi_d(0)| < |\phi_d^*(0)|$.

Proof. Assume that $\phi_{d-1}(z)$ is a Schur polynomial of exact degree $d-1$ and that $|\phi_d(0)| < |\phi_d^*(0)|$.

First, note that on the unit circle itself,

$$|\phi_d^*(z)| = \left| \phi_d(\bar{z}^{-1}) \right| = |\phi_d(z)|$$

since $\bar{z}^{-1} = z$ on the unit circle.

Define

$$\psi(z) = \frac{z\phi_{d-1}(z)}{\phi_d^*(0)}.$$

Then

$$\begin{aligned} |\phi_d(z) - \psi(z)| &= \left| \phi_d(z) - \frac{\phi_d^*(0)\phi_d(z) - \phi_d(0)\phi_d^*(z)}{\phi_d^*(0)} \right| \\ &= \left| \frac{\phi_d(0)}{\phi_d^*(0)} \phi_d^*(z) \right| < |\phi_d^*(z)| = |\phi_d(z)| \text{ on the unit circle.} \end{aligned}$$

Thus, by Rouché's theorem, $\phi_d(z)$ and $\psi(z)$ must have the same number of zeros inside the

unit circle. Since $\phi_{d-1}(z)$ is a Schur polynomial of exact degree $d - 1$, $\psi(z)$ must have d zeros, each of which lie within the unit circle, which implies that $\phi_d(z)$ also has d roots, each of which lies within the unit circle.

Therefore, $\phi_d(z)$ is a Schur polynomial of exact degree d . □

Chapter 2

MOL^T Formulation and Integral Solution

2.1 Formulation

In the MOL^T discretizations for the wave equation, the PDE is first discretized in time. Both dissipative and non-dissipative classes of discretization schemes may be derived. The following derivations are done in the context of one dimension but are easily extended to higher dimensions.

2.1.1 Dissipative Schemes

To construct the dissipative schemes, first observe that the one-sided first and second order finite difference approximations to the operator $v_{tt}^{n+1} = v_{tt}(x, t^{n+1})$ are:

$$v_{tt}^{n+1} = \frac{v^{n+1} - 2v^n + v^{n-1}}{\Delta t^2} + O(\Delta t) \text{ and} \tag{2.1}$$

$$v_{tt}^{n+1} = \frac{2v^{n+1} - 5v^n + 4v^{n-1} - v^{n-2}}{\Delta t^2} + O\left(\Delta t^2\right). \tag{2.2}$$

Now let $u^n = u(x, t^n)$ be the semi-discrete approximation to $v(x, t^n)$, which satisfies the wave equation (1.1). If the Laplacian term is evaluated implicitly, with either the substitution (2.1) or (2.2), we obtain, respectively, the semi-discrete equations

$$\frac{u^{n+1} - 2u^n + u^{n-1}}{\Delta t^2} = c^2 u_{xx}^{n+1} \text{ and} \quad (2.3)$$

$$\frac{2u^{n+1} - 5u^n + 4u^{n-1} - u^{n-2}}{\Delta t^2} = c^2 u_{xx}^{n+1}. \quad (2.4)$$

Here and below, in accordance with context, we will reserve the variable v to mean the continuous solution of the wave equation and u its semi-discrete approximation. It should be emphasized that equations (2.3) and (2.4) produce numerical dissipation, which will be shown in section 3.1. While not the focus of this work, this property is effective in recovering magnetostatic limits and will be the subject of future work.

2.1.2 Purely Dispersive Schemes

A purely dispersive scheme may be constructed by carefully considering the choice of finite difference approximations for both terms in $v_{tt}^n = c^2 \partial_{xx} v^n$, where the Laplacian term is replaced with an appropriate time-averaged approximation. For example, since

$$\partial_{xx} v^n = \partial_{xx} \left(\frac{v^{n+1} + v^{n-1}}{2} \right) + O(\Delta t^2),$$

a second order accurate semi-discrete approximation to equation (1.1) is

$$\frac{u^{n+1} - 2u^n + u^{n-1}}{(c\Delta t)^2} = \partial_{xx} \left(\frac{u^{n+1} + u^{n-1}}{2} \right). \quad (2.5)$$

As will be shown in section 3.1, this formulation removes numerical dissipation. This purely dispersive scheme may be more favorable for long time simulations, which wish to avoid the spurious effects of numerical diffusion.

2.2 Integral Solution

This work differs from other MOL^T methods in that we consider integral solutions to the boundary value problem, rather than utilizing finite difference or finite elements methods for the spatial derivatives. Specifically, we solve the resulting BVP through a free space Green's function [46, 48, 56], which is then corrected with a boundary integral method.

Collecting u^{n+1} in any of equations (2.3) - (2.5) results in

$$\mathcal{L}_{\beta_j} [u^{n+1}] (x) = -\frac{\beta_j}{(c\Delta t)^2} f_{\beta_j} (u^n, u^{n-1}, u^{n-2}, x), \quad (2.6)$$

where the Helmholtz operator is

$$\mathcal{L}_{\beta_j} [u] := \left(\partial_{xx} - \frac{\beta_j}{(c\Delta t)^2} \right) u, \quad (2.7)$$

where β_j is dependent upon the scheme chosen.

Equation (2.3) will correspond to choosing $\beta_{1a} = 1$, for which the Helmholtz operator is

$$\mathcal{L}_{\beta_{1a}} = \partial_{xx} - \frac{1}{(c\Delta t)^2},$$

and

$$f_{\beta_{1a}} = 2u^n(x) - u^{n-1}(x).$$

Equation (2.4) will correspond to choosing $\beta_{1b} = 2$, for which the Helmholtz operator is

$$\mathcal{L}_{\beta_{1b}} = \partial_{xx} - \frac{2}{(c\Delta t)^2}$$

and

$$f_{\beta_{1b}} = 5u^n(x) - 4u^{n-1}(x) + u^{n-2}(x).$$

Similarly, equation (2.5) will also correspond to $\beta_{2a} = 2$, so that the Helmholtz operator is

$$\mathcal{L}_{\beta_{2a}} = \mathcal{L}_{\beta_{1b}} \text{ and}$$

$$f_{\beta_{2a}} = 2u^n(x) - \frac{(c\Delta t)^2}{\beta_{2a}} \mathcal{L}_2 \left[u^{n-1} \right] (x).$$

The subscripts will be dropped when discussing the general case.

Equation (2.6) can be solved formally by inverting the Helmholtz operator (2.7). The resulting integral solution will consist of two components: a particular solution, $u_p(x)$, which solves the free space problem, and a homogeneous solution $u_h(x)$, which acts as a correction to enforce the boundary conditions. This will produce a solution in the form

$$u^{n+1}(x) = \mathcal{L}_{\beta}^{-1} \left[-\alpha_j^2 f_{\beta_j} \right] = u_p^{n+1}(x) + u_h^{n+1}(x), \quad (2.8)$$

where

$$\alpha_j = \frac{\sqrt{\beta_j}}{c\Delta t},$$

and the respective particular and homogeneous solutions are given as

$$u_p^{n+1}(x) = -\alpha_j^2 \int_{-L}^L f_{\beta_j} \left(u^n(y), u^{n-1}(y), u^{n-2}(y) \right) G(x|y) dy \quad (2.9)$$

and

$$u_h^{n+1}(x) = \left[u^{n+1}(y) \partial_y G(x|y) - \partial_y u^{n+1}(y) G(x|y) \right]_{-L}^L. \quad (2.10)$$

The Green's function for the Helmholtz operator (2.7) is

$$\begin{cases} G(x|y) = -\frac{1}{2\alpha} e^{-\alpha|x-y|}, & x \in \mathbb{R}^1 \\ G(x|y) = -\frac{1}{4} K_0(\alpha|x-y|), & x \in \mathbb{R}^2 \\ G(x|y) = \frac{e^{-\alpha|x-y|}}{4\pi|x-y|}, & x \in \mathbb{R}^3, \end{cases}$$

where K_0 is a modified Bessel function.

Additionally, in the case where $x \in \mathbb{R}^1$, the homogeneous solution (2.10) may be replaced with the Ansatz

$$u_h^{n+1}(x) = c_1 e^{-\alpha(L-x)} + c_2 e^{-\alpha(L+x)}, \quad (2.11)$$

where c_1 and c_2 are determined by the imposed boundary conditions.

In the case of zero Dirichlet boundary conditions, this produces the system

$$\begin{aligned} u_p^{n+1}(L) + c_1 + c_2 e^{-2\alpha L} &= 0, \\ u_p^{n+1}(-L) + c_1 e^{-2\alpha L} + c_2 &= 0 \end{aligned}$$

which has a solution of

$$\begin{aligned} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} &= \begin{bmatrix} 1 & e^{-2\alpha L} \\ e^{-2\alpha L} & 1 \end{bmatrix}^{-1} \begin{bmatrix} -u_p^{n+1}(L) \\ -u_p^{n+1}(-L) \end{bmatrix} \\ &= \frac{1}{1 - e^{-4\alpha L}} \begin{bmatrix} 1 & -e^{-2\alpha L} \\ -e^{-2\alpha L} & 1 \end{bmatrix} \begin{bmatrix} -u_p^{n+1}(L) \\ -u_p^{n+1}(-L) \end{bmatrix}. \end{aligned}$$

Substituting in c_1 and c_2 in (2.11) then gives

$$u_h^{n+1}(x) = -\frac{\sinh(\alpha(L-x))}{\sinh(2\alpha L)} u_p^{n+1}(-L) - \frac{\sinh(\alpha(L+x))}{\sinh(2\alpha L)} u_p^{n+1}(L).$$

In the case of zero Neumann boundary conditions, there is now the system

$$\begin{aligned} \frac{d}{dx} u_p^{n+1}(L) + \alpha c_1 - \alpha c_2 e^{-2\alpha L} &= 0, \\ \frac{d}{dx} u_p^{n+1}(-L) + \alpha c_1 e^{-2\alpha L} - \alpha c_2 &= 0 \end{aligned}$$

which has a solution of

$$\begin{aligned} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} &= \begin{bmatrix} \alpha & -\alpha e^{-2\alpha L} \\ \alpha e^{-2\alpha L} & -\alpha \end{bmatrix}^{-1} \begin{bmatrix} -\frac{d}{dx} u_p^{n+1}(L) \\ -\frac{d}{dx} u_p^{n+1}(-L) \end{bmatrix} \\ &= \frac{1}{\alpha(1 - e^{-4\alpha L})} \begin{bmatrix} 1 & -e^{-2\alpha L} \\ e^{-2\alpha L} & -1 \end{bmatrix} \begin{bmatrix} -\frac{d}{dx} u_p^{n+1}(L) \\ -\frac{d}{dx} u_p^{n+1}(-L) \end{bmatrix}. \end{aligned}$$

Note then that since $u(x)$ is defined above in (2.9) as

$$u_p^{n+1}(x) = -\alpha^2 \int_{-L}^L f_\beta \left(u^n(y), u^{n-1}(y), u^{n-2}(y) \right) G(x|y) dy.$$

This yields

$$\begin{aligned} \frac{d}{dx} u_p^{n+1} &= -\alpha^2 \int_{-L}^L f_\beta \left(u^n(y), u^{n-1}(y), u^{n-2}(y) \right) \frac{d}{dx} G(x|y) dy \\ &= -\alpha^2 \int_{-L}^L f_\beta \left(u^n(y), u^{n-1}(y), u^{n-2}(y) \right) (-\alpha \operatorname{sgn}(x - y) G(x|y)) dy \\ &= -\alpha^2 \int_{-L}^L -\alpha \operatorname{sgn}(x - y) f_\beta \left(u^n(y), u^{n-1}(y), u^{n-2}(y) \right) G(x|y) dy \end{aligned}$$

and thus,

$$\begin{aligned} \frac{d}{dx} u_p^{n+1}(L) &= -\alpha^2 \int_{-L}^L -\alpha \operatorname{sgn}(L - y) f_\beta \left(u^n(y), u^{n-1}(y), u^{n-2}(y) \right) G(L|y) dy \\ &= -\alpha^3 \int_{-L}^L f_\beta \left(u^n(y), u^{n-1}(y), u^{n-2}(y) \right) G(L|y) dy \\ &= -\alpha u_p^{n+1}(L) \end{aligned}$$

while

$$\begin{aligned}
\frac{d}{dx}u_p^{n+1}(-L) &= -\alpha^2 \int_{-L}^L -\alpha \operatorname{sgn}(-L-y) f_\beta \left(u^n(y), u^{n-1}(y), u^{n-2}(y) \right) G(-L|y) dy \\
&= \alpha^3 \int_{-L}^L f_\beta \left(u^n(y), u^{n-1}(y), u^{n-2}(y) \right) G(-L|y) dy \\
&= \alpha u_p^{n+1}(-L).
\end{aligned}$$

This gives that

$$\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \frac{1}{1 - e^{-4\alpha L}} \begin{bmatrix} 1 & -e^{-2\alpha L} \\ e^{-2\alpha L} & -1 \end{bmatrix} \begin{bmatrix} u_p^{n+1}(L) \\ -u_p^{n+1}(-L) \end{bmatrix}.$$

Substituting in for c_1 and c_2 in equation (2.11) gives this then as

$$u_h^{n+1}(x) = \frac{\cosh(\alpha(L-x))}{\sinh(2\alpha L)} u_p^{n+1}(-L) + \frac{\cosh(\alpha(L+x))}{\sinh(2\alpha L)} u_p^{n+1}(L).$$

2.3 Numerical Approach

Consider first the discretization of the particular solution (2.9). Divide the domain into non-overlapping regions Ω_i , so that

$$u_p^{n+1}(x) = \sum_{i=1}^N -\alpha^2 \int_{\Omega_i} f_\beta \left(u^n(y), u^{n-1}(y), u^{n-2}(y) \right) G(x|y) dy. \quad (2.12)$$

Generally, these regions may be of varying size, but in this dissertation, we will consider a uniform partition for the composite midpoint rule, so that $\Omega_i = [x_{i-1/2}, x_{i+1/2}]$, $i \in \{1, 2, \dots, N\}$, where $x_i = -L + (i - 1/2) \Delta x$ and $L = N\Delta x/2$. A midpoint quadrature

approximation to equation (2.12), evaluated at $x = x_j$, gives for each $j = 1, \dots, N$

$$u_p^{n+1}(x_j) = \sum_{i=1}^N A_{ij} f_\beta \left(u^n(x_i), u^{n-1}(x_i), u^{n-2}(x_i) \right) + E_M(x_j), \quad (2.13)$$

where the remainder term satisfies

$$E_M(x_j) = \Delta x \sum_{i=1}^N B_{ij} \partial_x f_\beta \left(u^n(\xi_i), u^{n-1}(\xi_i), u^{n-2}(\xi_i) \right)$$

with $\xi_i \in \Omega_i$ and

$$A_{ij} = -2\alpha^2 \int_{x_{i-1/2}}^{x_{i+1/2}} G(x_j|y) dy, \quad B_{ij} = -2\alpha^2 \int_{x_{i-1/2}}^{x_{i+1/2}} (y - x_i) G(x_j|y) dy.$$

Evaluating elements of the matrix A using the Green's function gives

$$A_{ij} = \alpha \int_{x_{i-1/2}}^{x_{i+1/2}} e^{-\alpha|x_j-y|} dy = 2 \begin{cases} e^{-|i-j|\nu} \sinh\left(\frac{\nu}{2}\right), & i \neq j \\ 1 - e^{-\nu/2}, & i = j \end{cases} \quad (2.14)$$

where

$$\nu = \alpha \Delta x = \frac{\sqrt{\beta} \Delta x}{c \Delta t}.$$

Likewise, the elements of the matrix B satisfy

$$B_{ij} = \alpha \int_{x_{i-1/2}}^{x_{i+1/2}} (y - x_i) e^{-\alpha|x_j-y|} dy = \begin{cases} c(\nu) e^{-|i-j|\nu}, & i \neq j \\ 0, & i = j \end{cases}$$

where $c(\nu) = \cosh\left(\frac{\nu}{2}\right) - \frac{2}{\nu}\sinh\left(\frac{\nu}{2}\right)$. In practice, we are developing implicit methods, and as such, the case where $\Delta x < c\Delta t$, corresponding to $\nu < 1$, is of particular interest. Taking a Taylor expansion for small ν shows

$$\begin{aligned} c(\nu) &= \cosh\left(\frac{\nu}{2}\right) - \frac{2}{\nu}\sinh\left(\frac{\nu}{2}\right) = \sum_{n=0}^{\infty} \frac{(\nu/2)^{2n}}{(2n)!} - \frac{2}{\nu} \sum_{n=0}^{\infty} \frac{(\nu/2)^{2n+1}}{(2n+1)!} \\ &= 1 + \frac{(\nu/2)^2}{2} + \frac{(\nu/2)^4}{24} - \frac{2}{\nu} \left(\frac{\nu}{2} + \frac{(\nu/2)^3}{6} + \frac{(\nu/2)^5}{120} \right) + O(\nu^6) \\ &= \frac{\nu^2}{12} + \frac{\nu^4}{480} + O(\nu^6). \end{aligned}$$

Thus, the error is bounded by

$$\begin{aligned} |E_M(x_j)|_{\infty} &\leq 2\Delta x c(\nu) \|\partial_x f_{\beta}\|_{\infty} \sum_{i=1}^{\infty} e^{-i\nu} \\ &= \Delta x c(\nu) \|\partial_x f_{\beta}\|_{\infty} \frac{2e^{-\nu}}{1 - e^{-\nu}} \\ &\leq \frac{\nu\Delta x}{6} \|\partial_x f_{\beta}\|_{\infty}. \end{aligned} \tag{2.15}$$

Here, we must note that each term in f_{β} is a numerical solution to the wave equation at a previous time step. Thus, since $\frac{\partial}{\partial x}f = \pm \frac{1}{c} \frac{\partial}{\partial t}f$ and we have already discretized the temporal derivative, we have that $\partial_x f_{\beta}$ behaves like $f_{\beta}\Delta t^{-1}$, giving that the error from equation (2.15) is actually $O(\nu^2)$.

For a trapezoidal approximation, consider the points $x_i = -L + i\Delta x$ with $\Omega_i = [x_{i-1}, x_i]$.

This approximation then gives for each $j = 0, 1, \dots, N$

$$u_p^{n+1}(x_j) = \sum_{i=0}^N C_{ij} f_{\beta} \left(u^n(x_i), u^{n-1}(x_i), u^{n-2}(x_i) \right) + E_T(x_j), \tag{2.16}$$

where the error term is

$$E_T(x_j) = \frac{\Delta x^2}{2} \sum_{i=0}^N D_{ij} \partial_{xx} f_\beta \left(u^n(\xi_i), u^{n-1}(\xi_i), u^{n-2}(\xi_i) \right).$$

Analogous evaluation of the matrix terms gives

$$\begin{aligned} C_{ij} &= \alpha \int_{x_{i-1}}^{x_i} \left(\frac{y - x_{i-1}}{\Delta x} \right) e^{-\alpha|x_j - y|} dy + \alpha \int_{x_i}^{x_{i+1}} \left(\frac{x_{i+1} - y}{\Delta x} \right) e^{-\alpha|x_j - y|} dy \\ &= \frac{2}{\nu} \begin{cases} 2 \sinh^2(\nu/2) e^{-|i-j|\nu}, & i \neq j \\ e^{-\nu} - 1 + \nu, & i = j \end{cases} \end{aligned} \quad (2.17)$$

and

$$D_{ij} = \alpha \int_{x_{i-1}}^{x_i} (y - x_{i-1})(y - x_i) e^{-\alpha|x_j - y|} dy = \left(\frac{2}{\nu} \right) c(\nu) e^{-|i-j+1/2|\nu}$$

where again $c(\nu) = \cosh\left(\frac{\nu}{2}\right) - \frac{2}{\nu} \sinh\left(\frac{\nu}{2}\right)$. The corresponding error bound is

$$\begin{aligned} |E_T(x_j)| &\leq \left(\frac{\Delta x^2}{2} \right) \|\partial_{xx} f_\beta\|_\infty \left(\frac{2}{\nu} \right) c(\nu) \sum_{i=-\infty}^{\infty} e^{-|i-j+1/2|\nu} \\ &= \left(\frac{\Delta x^2}{\nu} \right) c(\nu) \|\partial_{xx} f_\beta\|_\infty 2 \sum_{i=0}^{\infty} e^{-|i-j|\nu} e^{-\nu/2} \\ &\leq \left(\frac{\Delta x^2}{\nu} \right) c(\nu) \|\partial_{xx} f_\beta\|_\infty \frac{2e^{-\nu/2}}{1 - e^{-\nu}} \\ &\leq \frac{\Delta x^2}{3} \|\partial_{xx} f_\beta\|_\infty. \end{aligned} \quad (2.18)$$

Similar to the proof of the error estimate of the midpoint quadrature, $\partial_{xx} f_\beta$ behaves like $f_\beta \Delta t^{-2}$, showing that the error in equation (2.18) is truly $O(\nu^2)$.

Since both of the error estimates for the midpoint and trapezoidal quadrature rule are $O(\nu^2)$, this will demand that we must satisfy that our spatial step size goes to zero at an exponentially faster rate than the temporal step size. As we are creating an implicit scheme, this is much less of a restriction than for an explicit scheme. However, in section 3.1.1.3, we will show that the method of undetermined coefficients may be used to find weights that will change the error from $O(\nu^2)$ to $O(\Delta x^2)$, as is desired.

Chapter 3

Analysis of the Free Space Solution

3.1 Convergence

The main result of this chapter is the following convergence theorem.

Theorem 2. *Let $u^n(x)$ be the semi-discrete MOL^T solution satisfying equations (2.8), (2.9), (2.10), with initial and boundary conditions specified by equation (1.1). Then the semi-discrete solutions satisfying equations (2.3), (2.4) and (2.5) will converge to $v(x, t^n)$ as $O(\Delta t^\beta)$, with $\beta = 1, 2, 2$ respectively. Further, suppose the fully discrete solution u_j^n is computed using either the midpoint (2.13) or trapezoidal (2.16) quadrature rules. Then, u_j^n will converge to $v(x_j, t^n)$ as $O(\Delta t^\beta + \nu^2)$ and will be stable for $0 \leq t \leq T$.*

The proof of this theorem follows from the Lax–Richtmyer equivalence theorem upon establishing consistency and stability. Consistency is established in section 3.1.1 and stability in 3.1.2. In section 3.1.1.3, we will determine quadrature weights that will create a fully discrete scheme that is $O(\Delta t^\beta + \Delta x^2)$. In section 3.2, the dispersion of these methods is discussed.

3.1.1 Consistency of the MOL^T method

3.1.1.1 Non-dissipative Scheme

Let v be the continuous solution satisfying the wave equation (1.1) with $L = \infty$. Now, the integral solution for the semi-discrete solution for (2.5) in free space is given by

$$u^{n+1}(x) = I[u^n](x) - u^{n-1}(x),$$

where $I[v^n]$ is defined by

$$I[v^n] = \int_{-\infty}^{\infty} 2\alpha^2 G(x|y) v^n(y) dy = \alpha \int_{-\infty}^{\infty} e^{-\alpha|z|} v^n(x+z) dz$$

after applying the change of variables $y - x = z$.

Inserting v into this equation, we can define the global truncation error as

$$\tau^n = \frac{1}{(c\Delta t)^2} \left(v^{n+1} + v^{n-1} - I[v^n] \right), \quad (3.1)$$

The integral can be split at $z = 0$, giving

$$I = (I_+) + (I_-) = \alpha \int_0^{\infty} e^{-\alpha z} v^n(x+z) dz + \alpha \int_0^{\infty} e^{-\alpha z} v^n(x-z) dz.$$

For each integral, repeatedly apply integration by parts to get

$$\begin{aligned}
I_{\pm} &= \alpha \int_0^{\infty} e^{-\alpha z} v^n(x \pm z) dz \\
&= \left[-(\pm 1) v^n(x \pm z) e^{-\alpha z} \right]_0^{\infty} + \int_0^{\infty} \pm \partial_x v^n(x \pm z) e^{-\alpha z} dz \\
&\vdots \\
&= \left[-\alpha \sum_{j=0}^k \left(\frac{\pm 1}{\alpha} \partial_x \right)^j v^n(x \pm z) e^{-\alpha z} \right]_0^{\infty} \\
&\quad + \alpha \int_0^{\infty} \left(\frac{\pm 1}{\alpha} \partial_x \right)^{k+1} v^n(x \pm z) e^{-\alpha z} dz.
\end{aligned}$$

Notice that when evaluating at $z = 0$, the alternating signs cause the odd derivatives to cancel and the even derivatives to combine for $I = I_+ + I_-$. Furthermore, the upper limit makes no contribution due to the dominance of the negative exponential.

Truncating the integration by parts at the leading order for the error, corresponding here to $k = 3$, gives

$$\begin{aligned}
I[v^n] &= \left[\left(1 + \frac{1}{\alpha} \partial_x + \frac{1}{\alpha^2} \partial_{xx} + \frac{1}{\alpha^3} \partial_{xxx} \right) v^n(x+z) e^{-\alpha z} \right]_0^{\infty} \\
&\quad + \left[\left(1 - \frac{1}{\alpha} \partial_x + \frac{1}{\alpha^2} \partial_{xx} - \frac{1}{\alpha^3} \partial_{xxx} \right) v^n(x-z) e^{-\alpha z} \right]_0^{\infty} \\
&\quad + \alpha \int_0^{\infty} \frac{1}{\alpha^4} (v_{xxxx}^n(x+z) + v_{xxxx}^n(x-z)) e^{-\alpha z} dz \\
&= 2 \left(v^n + \frac{1}{\alpha^2} v_{xx}^n \right) + \alpha \int_0^{\infty} \frac{1}{\alpha^4} (v_{xxxx}^n(x+z) + v_{xxxx}^n(x-z)) e^{-\alpha z} dz \\
&= 2v^n + (c\Delta t)^2 v_{xx}^n + \frac{(c\Delta t)^4}{4} (2v_{xxxx}^n(\xi)) \alpha \int_0^{\infty} e^{-\alpha z} dz. \tag{3.2}
\end{aligned}$$

Note that we use the mean value theorem for integrals in the last line. The truncation

error (3.1) then becomes

$$\begin{aligned}
\tau^n &= \frac{1}{(c\Delta t)^2} \left(v^{n+1} + v^{n-1} - 2v^n - (c\Delta t)^2 v_{xx}^n - \frac{(c\Delta t)^4}{2} v_{xxxx}^n(\xi) \right) \\
&= \frac{1}{c^2} \left(v_{tt}^n + \frac{\Delta t^2}{12} v_{tttt}^n + O(\Delta t^4) \right) - v_{xx}^n - \frac{(c\Delta t)^2}{2} v_{xxxx}^n(\xi) \\
&= \frac{1}{c^2} \left(\frac{\Delta t^2}{12} v_{tttt}^n + O(\Delta t^4) \right) - \frac{(c\Delta t)^2}{2} v_{xxxx}^n(\xi) \\
&= \frac{-5(c\Delta t)^2}{12} v_{xxxx}^n(\xi) + O(\Delta t^4)
\end{aligned}$$

where the terms $v^{n\pm 1}$ have been expanded about t^n and differentiated in the context of the wave equation to determine that

$$v_{tttt} = c^2 v_{ttxx} = c^4 v_{xxxx}.$$

Hence, the semi-discrete error is second order in time.

3.1.1.2 Dissipative Schemes

The same ideas may be applied to the dissipative free space solver (2.6). However, the expansion is now about t^{n+1} , where now the first order semi-discrete solution is given by

$$u^{n+1}(x) = I[u^n](x) + \frac{1}{2}I[u^{n-1}](x)$$

where $I[u]$ is defined, as before, by

$$I[u] = \int_{-\infty}^{\infty} 2\alpha^2 G(x|y)u(y)dy = \alpha \int_{-\infty}^{\infty} e^{-\alpha|z|}u(x+z) dz.$$

Substituting in the continuous solution v and once again changing variables to $z = y - x$, the global truncation error is defined as

$$\tau^{n+1} = \frac{1}{(c\Delta t)^2} \left(v^{n+1} - I[v^n] - \frac{1}{2}I[v^{n-1}] \right).$$

Repeating the integral splitting and integration by parts done just above, it becomes easy to see that truncating after three steps gives

$$\begin{aligned} & I[v^n] - \frac{1}{2}I[v^{n-1}] \\ &= \left[\left(1 + \frac{1}{\alpha}\partial_x + \frac{1}{\alpha^2}\partial_{xx} + \frac{1}{\alpha^3}\partial_{xxx} \right) \left(v^n - \frac{1}{2}v^{n-1} \right) (x+z)e^{-\alpha z} \right]_0^\infty \\ &+ \left[\left(1 - \frac{1}{\alpha}\partial_x + \frac{1}{\alpha^2}\partial_{xx} - \frac{1}{\alpha^3}\partial_{xxx} \right) \left(v^n - \frac{1}{2}v^{n-1} \right) (x-z)e^{-\alpha z} \right]_0^\infty \\ &+ \alpha \int_0^\infty \frac{1}{\alpha^4} \left(\left(v^n - \frac{1}{2}v^{n-1} \right)_{xxxx} (x+z) + \left(v^n - \frac{1}{2}v^{n-1} \right)_{xxxx} (x-z) \right) e^{-\alpha z} dz \\ &= 2 \left(\left(v^n - \frac{1}{2}v^{n-1} \right) + \frac{1}{\alpha^2} \left(v^n - \frac{1}{2}v^{n-1} \right)_{xx} \right) \\ &+ \alpha \int_0^\infty \frac{1}{\alpha^4} \left(\left(v^n - \frac{1}{2}v^{n-1} \right)_{xxxx} (x+z) + \left(v^n - \frac{1}{2}v^{n-1} \right)_{xxxx} (x-z) \right) e^{-\alpha z} dz \\ &= 2 \left(v^n - \frac{1}{2}v^{n-1} \right) + (c\Delta t)^2 \left(v^n - \frac{1}{2}v^{n-1} \right)_{xx} \\ &+ \frac{(c\Delta t)^4}{4} \left(2 \left(v^n - \frac{1}{2}v^{n-1} \right)_{xxxx} (\xi) \right) \alpha \int_0^\infty e^{-\alpha z} dz. \end{aligned}$$

Thus, the global truncation error is

$$\begin{aligned}
\tau^{n+1} &= \frac{1}{(c\Delta t)^2} \left(v^{n+1} - 2v^n + v^{n-1} - (c\Delta t)^2 (2v^n - v^{n-1})_{xx} + O(\Delta t^4) \right) \\
&= \frac{1}{c^2} \left(v_{tt}^{n+1} - \Delta t v_{ttt}^{n+1} \right) - v_{xx}^{n+1} + O(\Delta t^2) \\
&= -\frac{1}{c^2} \Delta t v_{ttt}^{n+1}(\zeta) + O(\Delta t^2) \\
&= -\Delta t v_{txx}^{n+1}(\zeta) + O(\Delta t^2),
\end{aligned}$$

which shows that this method is first order in time.

The second order semi-discrete solution is given by

$$u^{n+1}(x) = \frac{5}{4}I[u^n](x) + I[u^{n-1}](x) - \frac{1}{4}I[u^{n-2}](x),$$

and thus

$$\tau^{n+1} = \frac{1}{(c\Delta t)^2} \left(v^{n+1} - \frac{5}{4}I[v^n] + I[v^{n-1}] - \frac{1}{4}I[v^{n-2}] \right).$$

Using the same calculations as above,

$$\begin{aligned}
I[v^n] &= 2 \left(v^n + \frac{1}{\alpha^2} v_{xx}^n \right) + \alpha \int_0^\infty \frac{1}{\alpha^4} (v_{xxxx}^n(x+z) + v_{xxxx}^n(x-z)) e^{-\alpha z} dz \\
&= 2v^n + (c\Delta t)^2 v_{xx}^n + \frac{(c\Delta t)^4}{4} (2v_{xxxx}^n(\xi)) \alpha \int_0^\infty e^{-\alpha z} dz.
\end{aligned}$$

and therefore, after utilizing the linearity of I , this shows that

$$\begin{aligned}
\tau^{n+1} &= \frac{1}{(c\Delta t)^2} \left(v^{n+1} - \frac{5}{2}v^n - \frac{5}{4}(c\Delta t)^2 v_{xx}^n + 2v^{n-1} + (c\Delta t)^2 v_{xx}^{n-1} \right. \\
&\quad \left. - \frac{1}{2}v^{n-2} - (c\Delta t)^2 \frac{1}{4}v_{xx}^{n-2} + O(\Delta t^4) \right) \\
&= \frac{1}{c^2} \left(\frac{1}{2}v_{tt}^{n+1} \right) - \frac{1}{2}v_{xx}^{n+1} - \frac{3\Delta t^2}{4}v_{xxtt}^{n+1} + O(\Delta t^4) \\
&= -\frac{3(c\Delta t)^2}{4}v_{xxxx}^{n+1} + O(\Delta t^4),
\end{aligned}$$

which gives that the method is indeed second order.

3.1.1.3 Fully Discrete Schemes

In the case of the fully discrete implementations, the results from sections 3.1.1.1 and 3.1.1.2 will still follow in the exact same manner, with the addition that $I[u]$ now contributes an $O(\nu^2)$ error caused by the quadrature estimate.

A quick observation from equations (2.14) and (2.17) will show that the terms that come from integrating equation (2.12) using quadrature will be of the form

$$I[u](x_j) = w_0 u(x_j) + w_1 \sum_{k=1}^{\infty} e^{-|k-j|\nu} (u(x_{j-k}) + u(x_{j+k})).$$

Consider now the truncation error given by the non-dissipative scheme, shown in (3.1), adapted for the fully discrete scheme as

$$\tau^n(x_j) = \frac{1}{(c\Delta t)^2} \left(u^{n+1}(x_j) + u^{n-1}(x_j) - I[u^n](x_j) \right). \quad (3.3)$$

Let

$$w_0 = 2 - \frac{2w_1}{e^\nu - 1} = 2 - w_1 \sum_{k=1}^{\infty} 2e^{-k\nu}$$

(3.3) then becomes

$$\begin{aligned} \tau^n(x_j) &= \frac{u^{n+1}(x_j) - 2u^n(x_j) + u^{n-1}(x_j)}{\Delta t^2} - \frac{w_1}{\Delta t^2} \sum_{k=1}^{\infty} e^{-k\nu} (u^n(x_{j+k}) - 2u^n(x_j) + u^n(x_{j-k})) \\ &= u_{tt}^n(x_j) + \frac{\Delta t^2}{12} u_{tttt}(\xi_1) - \frac{w_1}{\Delta t^2} \sum_{k=1}^{\infty} e^{-k\nu} \left((k\Delta x)^2 u_{xx}^n(x_j) + \frac{(k\Delta x)^4}{12} u_{xxxx}^n(\xi_2) \right) \\ &= u_{tt}^n(x_j) + \frac{\Delta t^2}{12} u_{tttt}(\xi_1) - w_1 \sum_{k=1}^{\infty} e^{-k\nu} \left(\frac{(k\nu)^2}{2} u_{tt}^n(x_j) + \frac{(k\nu)^4}{48} u_{tttt}^n(\xi_2) \right) \\ &= \eta_1 u_{tt}^n(x_j) + \Delta t^2 \eta_2 u_{tttt}^n(\xi) \end{aligned} \tag{3.4}$$

where

$$\begin{aligned} \eta_1 &= 1 - \frac{w_1 \nu^2}{2} \sum_{k=1}^{\infty} e^{-k\nu} k^2, \\ \eta_2 &= \frac{1}{12} - \frac{w_1 \nu^4}{48} \sum_{k=1}^{\infty} e^{-k\nu} k^4. \end{aligned}$$

Note that

$$\sum_{k=1}^{\infty} e^{-k\nu} k^j = (-1)^j \left(\frac{d}{d\nu} \right)^j \sum_{k=1}^{\infty} e^{-k\nu} = (-1)^j \left(\frac{d}{d\nu} \right)^j \frac{1}{e^\nu - 1},$$

most notably with

$$\sum_{k=1}^{\infty} e^{-k\nu} k^2 = \frac{1}{4} \coth\left(\frac{\nu}{2}\right) \operatorname{csch}^2\left(\frac{\nu}{2}\right) \text{ and}$$

$$\sum_{k=1}^{\infty} e^{-k\nu} k^4 = \frac{1}{4} \coth\left(\frac{\nu}{2}\right) \operatorname{csch}^2\left(\frac{\nu}{2}\right) \left(3 \operatorname{csch}^2\left(\frac{\nu}{2}\right) + 1\right).$$

We must choose w_1 such that $\eta_1 = 0$ if we wish to have the order be $O(\Delta x^2)$. Thus,

$$\begin{aligned} \eta_1 = 0 &\implies 1 - \frac{w_1 \nu^2}{2} \frac{1}{4} \coth\left(\frac{\nu}{2}\right) \operatorname{csch}^2\left(\frac{\nu}{2}\right) = 0 \\ &\implies w_1 = \frac{8}{\nu^2} \tanh\left(\frac{\nu}{2}\right) \sinh^2\left(\frac{\nu}{2}\right). \end{aligned}$$

Then,

$$\begin{aligned} \eta_2 &= \frac{1}{12} - \frac{8\nu^4}{48} \tanh\left(\frac{\nu}{2}\right) \sinh^2\left(\frac{\nu}{2}\right) \frac{1}{4} \coth\left(\frac{\nu}{2}\right) \operatorname{csch}^2\left(\frac{\nu}{2}\right) \left(3 \operatorname{csch}^2\left(\frac{\nu}{2}\right) + 1\right) \\ &= \frac{1}{12} - \frac{\nu^2}{24} \left(3 \operatorname{csch}^2\left(\frac{\nu}{2}\right) + 1\right). \end{aligned}$$

Using the expansion

$$\operatorname{csch}^2(x) = \frac{1}{x^2} - \frac{1}{3} + \left(\frac{1}{36} + \left(\frac{7}{360}\right)^2\right) x^2 + O(x^4)$$

gives that

$$\eta_2 = \frac{-5}{12} - \frac{3\nu^2}{24} + O(\nu^4)$$

and thus, substituting in η_2 into (3.4) gives that the truncation error is now

$$\tau^n = \Delta t^2 \left(\frac{-5}{12} - \frac{3\nu^2}{24} + O(\nu^4) \right) = O(\Delta t^2 + \Delta x^2).$$

Similar results will hold for the same w_0 and w_1 for the dissipative schemes.

3.1.2 Stability of the MOL^T method

To establish the stability of the MOL^T method, we treat the fully discretized system as a one step method. To that extent, define

$$\bar{u}^n = [u^n(x_1), \dots, u^n(x_N)]^T, \quad n \geq 1,$$

where $x_j = -L + (j - 0.5)\Delta x$ in the case of the midpoint rule and $x_j = -L + j\Delta x$ when considering the trapezoidal rule. Then, for the dissipative scheme corresponding to equation (2.3),

$$\bar{u}^{n+1} = A_t \left(\bar{u}^n - \frac{1}{2} \bar{u}^{n-1} \right). \quad (3.5)$$

For the second order dissipative scheme in accordance with equation (2.4), we have

$$\bar{u}^{n+1} = \frac{1}{4} A_t \left(5\bar{u}^n - 4\bar{u}^{n-1} + \bar{u}^{n-2} \right) \quad (3.6)$$

and for the non-dissipative scheme, corresponding to equation (2.5),

$$\bar{u}^{n+1} = A_t \bar{u}^n - \bar{u}^{n-1}. \quad (3.7)$$

In each of the cases, the matrix A_t will possess the same structure, although the individual matrix elements, which are based upon ν , will differ. A_t will consist of a particular and homogeneous contribution, in analog to the free space and boundary terms in the integral solution of the continuous problem. It also contains the information of the discretization scheme in use; in this case, the midpoint or trapezoidal quadrature rules. If we write $A_t = A_p + A_h$, then for the midpoint rule, A_p is given by the matrix A in equation (2.14), and the matrix C of equation (2.17) in the case of the trapezoidal rule. The boundary matrix entries are

$$(A_h)_{ij} = -2a_0 \sinh\left(\frac{\nu}{2}\right) \left(\frac{\sinh(\alpha(L+x_i))}{\sinh(2\alpha L)} e^{-\alpha(L-x_j)} + \frac{\sinh(\alpha(L-x_i))}{\sinh(2\alpha L)} e^{-\alpha(L+x_j)} \right), \quad (3.8)$$

where the coefficient a_0 distinguishes between the midpoint and trapezoidal rules with

$$a_0 = \begin{cases} 1, & \text{Midpoint Rule} \\ \frac{2}{\nu} \sinh\left(\frac{\nu}{2}\right), & \text{Trapezoidal Rule.} \end{cases} \quad (3.9)$$

We will establish stability for these schemes by studying the eigenvalues of A_p and A_h . Since A_h will be exponentially small away from the boundaries, the main part of this work will focus on analyzing the eigenvalues of the free-space problem, which would neglect A_h altogether. Once the free space problem is understood, we will proceed with the boundary correction matrix.

3.1.2.1 The Eigenvalues of A_p

In order to simplify the proceeding analysis, which is for two different methods of discretization, write the matrices $A_p = A$ of equation (2.14) and $A_p = C$ of equation (2.17) in the form

$$A_p = 2\tilde{a}_0 I_N + 2\tilde{a}_1 \sinh\left(\frac{\nu}{2}\right) T, \quad (3.10)$$

where I_N is the $N \times N$ identity matrix, and the constants $\tilde{a}_0$ and $\tilde{a}_1$ distinguish the quadrature rule. They are given by

$$\tilde{a}_0, \tilde{a}_1 = \begin{cases} 1 - e^{-\nu/2} - \sinh\left(\frac{\nu}{2}\right), & 1, & \text{Midpoint Rule} \\ \frac{1}{\nu} \left(e^{-\nu} - 1 + \nu - 2 \sinh^2\left(\frac{\nu}{2}\right) \right), & \frac{2}{\nu} \sinh\left(\frac{\nu}{2}\right), & \text{Trapezoidal Rule.} \end{cases} \quad (3.11)$$

Note that these expressions can be simplified to $1 - \cosh\left(\frac{\nu}{2}\right)$ and $1 - \frac{1}{\nu} \sinh(\nu)$, respectively.

Finally, the matrix T is a Toeplitz matrix and is given by

$$T_{ij} = x^{|i-j|}, \quad x = e^{-\nu}.$$

Notice that the eigenvalues of A_p can be found from the eigenvalues t_k of T , since A_p differs from T by a scaling factor and a multiple of the identity matrix. Therefore, the eigenvalues λ_k of A_p are given by

$$\lambda_k = 2\tilde{a}_0 + 2\tilde{a}_1 \sinh\left(\frac{\nu}{2}\right) t_k. \quad (3.12)$$

Now, due to its structural symmetry, the inverse matrix T^{-1} has a special form, and the eigenvalues may be found analytically. In fact, T^{-1} is a nearly-Toeplitz, tri-diagonal matrix,

$$T^{-1} = \frac{1}{1-x^2} \begin{bmatrix} 1 & -x & 0 & \dots & 0 \\ -x & 1+x^2 & -x & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & -x & 1+x^2 & -x \\ 0 & \dots & 0 & -x & 1 \end{bmatrix}, \quad x = e^{-\nu}.$$

Since this is a tri-diagonal Toeplitz matrix with permuted corners, recall that based on the results in section 1.2.1, the eigenvalues of this matrix are given by

$$t_k^{-1} = \frac{x^2 - 2x \cos(\theta_k) + 1}{1 - x^2},$$

where θ_k is a root of

$$\sin((N+1)\theta) + x^2 \sin((N-1)\theta) - x \sin(N\theta) = 0, \quad \theta \in (0, \pi).$$

This gives that the eigenvalues of our Toeplitz matrix are given by

$$t_k = \frac{1 - x^2}{x^2 - 2x \cos(\theta_k) + 1},$$

Making use of $x = e^{-\nu}$, we can find an explicit form for the eigenvalues by solving

$$\begin{aligned}
t_k &= \frac{1 - x^2}{x^2 - 2x \cos(\theta_k) + 1} \\
&= \frac{1 - e^{-2\nu}}{e^{-2\nu} - 2e^{-\nu} \cos(\theta_k) + 1} \\
&= \frac{e^\nu - e^{-\nu}}{e^{-\nu} - 2\cos(\theta_k) + e^\nu} \\
&= \frac{2\sinh(\nu)}{2\cosh(\nu) - 2\cos(\theta_k)} \\
&= \frac{\sinh(\nu)}{\cosh(\nu) - \cos(\theta_k)} \\
&= \frac{2\sinh(\nu/2) \cosh(\nu/2)}{1 + 2\sinh^2(\nu/2) - 1 + 2\sin^2(\theta_k/2)} \\
&= \frac{\sinh(\nu/2) \cosh(\nu/2)}{\sinh^2(\nu/2) + \sin^2(\theta_k/2)}.
\end{aligned}$$

Thus, the eigenvalues are given by

$$t_k = \cosh(\nu/2) \left(\frac{\sinh(\nu/2)}{\sinh^2(\nu/2) + \sin^2(\theta_k/2)} \right), \quad 0 < \theta_k < \pi.$$

This leads to the following lemma.

Lemma 3. *The eigenvalues of the matrix given by equations (2.14) and (2.17) are contained in the interval $(0, 2)$.*

Proof. This is a consequence of equation (3.12), using the coefficients from (3.11). In the

former case, the eigenvalues for $k = 1, 2, \dots, N$, are

$$\begin{aligned}\lambda_k &= 2 \left(1 - \cosh(\nu/2) + \cosh(\nu/2) \left(\frac{\sinh^2(\nu/2)}{\sinh^2(\nu/2) + \sin^2(\theta_k/2)} \right) \right) \\ &= 2 \left(1 - \cosh(\nu/2) \left(\frac{\sin^2(\theta_k/2)}{\sinh^2(\nu/2) + \sin^2(\theta_k/2)} \right) \right), \quad 0 < \theta_k < \pi.\end{aligned}\quad (3.13)$$

We obtain the bounds by letting $\theta_k = 0, \pi$, and therefore

$$2 - 2\operatorname{sech}\left(\frac{\nu}{2}\right) < \lambda_k < 2, \quad (3.14)$$

where $\theta_k = 0$ corresponds to 2 and $\theta_k = \pi$ corresponds with $2 - 2\operatorname{sech}\left(\frac{\nu}{2}\right)$. Similarly, for the trapezoidal rule, the eigenvalues are for $k = 1, 2, \dots, N$,

$$\begin{aligned}\lambda_k &= 2 \left(1 - \frac{1}{\nu} \sinh(\nu) + \frac{2}{\nu} \sinh(\nu/2) \cosh(\nu/2) \left(\frac{\sinh^2(\nu/2)}{\sinh^2(\nu/2) + \sin^2(\theta_k/2)} \right) \right) \\ &= 2 \left(1 - \frac{1}{\nu} \sinh(\nu) \left(\frac{\sin^2(\theta_k/2)}{\sinh^2(\nu/2) + \sin^2(\theta_k/2)} \right) \right), \quad 0 < \theta_k < \pi.\end{aligned}\quad (3.15)$$

Now, the bounds obtained by letting $\theta_k = 0, \pi$, are

$$2 - \frac{4}{\nu} \tanh\left(\frac{\nu}{2}\right) < \lambda_k < 2, \quad (3.16)$$

where $\theta_k = 0$ corresponds to 2 and $\theta_k = \pi$ corresponds with $2 - \frac{4}{\nu} \tanh\left(\frac{\nu}{2}\right)$.

The lower bound in (3.14) and (3.16) will be identically zero only for $\nu = 0$ and positive otherwise. Therefore, the lemma is proven. $\square$

Lemma 4. *The matrices A and C given by equation (2.14) and (2.17) respectively, are*

symmetric positive definite.

Proof. Note first that A_p is symmetric since T is symmetric as $T_{ij} = x^{|i-j|} = x^{|j-i|} = T_{ji}$.

Also,

$$\begin{aligned}
& \frac{\sinh(2\alpha L)}{-2a_0 \sinh(\nu/2)} (A_h)_{ij} = \\
& \sinh(\alpha(L+x_i))e^{-\alpha(L-x_j)} + \sinh(\alpha(L-x_i))e^{-\alpha(L+x_j)} \\
& = \left(e^{\alpha(L+x_i)} - e^{-\alpha(L+x_i)}\right)e^{-\alpha(L-x_j)} + \left(e^{\alpha(L-x_i)} - e^{-\alpha(L-x_i)}\right)e^{-\alpha(L+x_j)} \\
& = e^{\alpha(x_i+x_j)} - e^{-2\alpha L}e^{\alpha(-x_i+x_j)} + e^{-\alpha(x_i+x_j)} - e^{-2\alpha L}e^{\alpha(x_i-x_j)} \\
& = \left(e^{\alpha(L+x_j)} - e^{-\alpha(L+x_j)}\right)e^{-\alpha(L-x_i)} + \left(e^{\alpha(L-x_j)} - e^{-\alpha(L-x_j)}\right)e^{-\alpha(L+x_i)} \\
& = \sinh(\alpha(L+x_j))e^{-\alpha(L-x_i)} + \sinh(\alpha(L-x_j))e^{-\alpha(L+x_i)} \\
& = \frac{\sinh(2\alpha L)}{-2a_0 \sinh(\nu/2)} (A_h)_{ji}.
\end{aligned}$$

Thus, the matrices A and C are symmetric. They are positive definite because of the previous lemma which proved that all of the eigenvalues were positive. $\square$

3.1.2.2 Dissipative Schemes

We are now in a position to show stability for the free-space problem using Von Neumann analysis. Let the numerical solution be given by $\bar{u}^n = \mu^n \bar{u}^0$ for $n \geq 0$. Inserting this expression into equations (3.5) and (3.6) with $A_t = A_p$, and canceling the common factor produces the stability equations

$$\begin{aligned}
\mu^2 \bar{u}^0 &= \left(\mu - \frac{1}{2}\right) A_p \bar{u}^0 \text{ and} \\
\mu^3 \bar{u}^0 &= \frac{1}{4} (5\mu^2 - 4\mu + 1) A_p \bar{u}^0.
\end{aligned} \tag{3.17}$$

Now, since A_p is a symmetric positive definite matrix, it is diagonalizable as $A_p = Q\Lambda Q^{-1}$, where Q is orthonormal and Λ is the diagonal matrix with entries given by (3.13) or (3.15). Thus, solvability of the expression (3.17) leads to the degree $2N$ Von-Neumann polynomial

$$\begin{aligned} p(\mu) &= \left| \mu^2 I_N - \left(\mu - \frac{1}{2} \right) A_p \right| \\ &= \left| \mu^2 Q I_N Q^{-1} - \left(\mu - \frac{1}{2} \right) Q \Lambda Q^{-1} \right| \\ &= \left| \mu^2 I_N - \left(\mu - \frac{1}{2} \right) \Lambda \right| = 0. \end{aligned}$$

Since this latter form is expressed as the determinant of a diagonal matrix, the Von-Neumann polynomial can be written in factored form, and the amplification factors $\mu_k^\pm$ will be determined in pairs by the expression

$$\mu_k^2 - \lambda_k \mu_k + \frac{\lambda_k}{2} = 0, \quad k = 1, 2, \dots, N.$$

Solving these resulting quadratic equations gives

$$\mu_k^\pm = \frac{\lambda_k \pm \sqrt{\lambda_k^2 - 2\lambda_k}}{2}.$$

Since $\lambda_k \in (0, 2)$, it is immediately apparent that the amplification factors form complex conjugate pairs. Further, since each θ_k will be distinct, this means that for each k , the roots of the Von-Neumann polynomial will also be distinct. If it is further shown that $|\mu_k| \leq 1$, then stability of the numerical scheme will be established. Taking the modulus of the roots,

we have

$$|\mu_k^\pm| = \sqrt{\left(\frac{\lambda_k}{2}\right)^2 + \frac{\lambda_k}{2} - \left(\frac{\lambda_k}{2}\right)^2} = \sqrt{\frac{\lambda_k}{2}} < 1.$$

Thus, the first order system will be stable, and, since $|\mu| < 1$, dissipative.

Repeating the process for the second order formulation gives

$$4\mu_k^3 - 5\lambda_k\mu_k^2 + 4\lambda_k\mu_k - \lambda_k = 0.$$

Lemma 5. *The magnitude of the roots of $4\mu_k^3 - 5\lambda_k\mu_k^2 + 4\lambda_k\mu_k - \lambda_k$ are all less than one.*

Proof. This will be shown using the Schur polynomial test given in section [1.2.2](#).

Define

$$\phi_0(z) = 4z^3 - 5\lambda_k z^2 + 4\lambda_k z - \lambda_k \text{ and}$$

$$\phi_0^*(z) = -\lambda_k z^3 + 4\lambda_k z^2 - \lambda_k z + 4.$$

Note that $|\phi_0(0)| = \lambda_k < 4 = |\phi_0^*(0)|$.

Then,

$$\begin{aligned} \phi_1(z) &= \frac{4(4z^3 - 5\lambda_k z^2 + 4\lambda_k z - \lambda_k) + \lambda_k(-\lambda_k z^3 + 4\lambda_k z^2 - \lambda_k z + 4)}{z} \\ &= (16 - \lambda_k)z^2 + (4\lambda_k^2 - 20\lambda_k)z + 16\lambda_k - 5\lambda_k^2 \text{ and thus} \\ \phi_1^*(z) &= (16\lambda_k - 5\lambda_k^2)z^2 + (4\lambda_k^2 - 20\lambda_k)z + 16 - \lambda_k. \end{aligned}$$

This gives that $\phi_1(0) = 16\lambda_k - 5\lambda_k^2 > 0$ and $\phi_1^*(0) = 16 - \lambda_k > 0$. Thus, we can check the

condition of $|\phi_1(0)| \leq |\phi_1^*(0)|$ by finding the roots of $\phi_m(0) - \phi_m^*(0) = -5\lambda_k^2 + 17\lambda_k - 16$. Solving this quadratic shows that the roots are imaginary and thus the equation is negative everywhere since using $\lambda_k = 0$ gives -16 .

Taking this further by one step yields

$$\begin{aligned}\phi_2(z) &= \frac{1}{z} \left((16 - \lambda_k) \left((16 - \lambda_k) z^2 + (4\lambda_k^2 - 20\lambda_k) z + 16\lambda_k - 5\lambda_k^2 \right) \right. \\ &\quad \left. - (16\lambda_k - 5\lambda_k^2) \left((16\lambda_k - 5\lambda_k^2) z^2 + (4\lambda_k^2 - 20\lambda_k) z + (16 - \lambda_k) \right) \right) \\ &= \left((16 - \lambda_k)^2 - (16\lambda_k - 5\lambda_k^2)^2 \right) z + (16 - 17\lambda_k + 5\lambda_k^2) (4\lambda_k^2 - 20\lambda_k),\end{aligned}$$

which is obviously a polynomial of degree 1.

Furthermore, this yields that

$$\tilde{\phi}_2(z) = \frac{(16 - \lambda_k)^2 - (16\lambda_k - 5\lambda_k^2)^2}{(16 - 17\lambda_k + 5\lambda_k^2) (4\lambda_k^2 - 20\lambda_k)} = \frac{4\lambda_k (5\lambda_k^3 - 42\lambda_k^2 + 85\lambda_k - 80)}{(16 - 17\lambda_k + 5\lambda_k^2) (4\lambda_k^2 - 20\lambda_k)}.$$

Mathematica can then be used to show that $\tilde{\phi}_2(z)$ has a critical point at 0 and no other critical points in $[0, 2]$. Thus, since $|\tilde{\phi}(1)| \approx 0.4848$, I have that ϕ_2 is a Schur polynomial and thus subsequently, so are ϕ_1 and ϕ_0 .

□

3.1.2.3 Non-dissipative Scheme

The process of the previous section is now repeated on (3.7). This gives rise to a polynomial of the form

$$\mu_k^2 - \lambda_k \mu_k + 1 = 0, \quad k = 1, 2, \dots, N.$$

Thus, the amplification factors are now given as

$$\mu_k^\pm = \frac{\lambda_k \pm \sqrt{\lambda_k^2 - 4}}{2}.$$

This time, the modulus of each complex conjugate pair is

$$|\mu_k^\pm| = \sqrt{\left(\frac{\lambda_k}{2}\right)^2 + 1 - \left(\frac{\lambda_k}{2}\right)^2} = 1.$$

Thus, since the eigenvalues are all distinct, the method is stable, and, since $|\mu| = 1$, the scheme is non-dissipative.

3.2 Dispersion

We now analyze the phase error of the free-space approximation. The continuous dispersion relation results from looking at sinusoidal solutions of the wave equation

$$v(x, t) = e^{i(kx - \omega t)} \text{ with } \omega^2 = c^2 k^2.$$

We now analyze the semi-discrete dispersion relations and define the phase error for the free space dissipative and non-dissipative schemes.

3.2.1 Semi-discrete Non-dissipative Scheme

For the semi-discrete equation, define $u^n = \hat{u}e^{ikx-i\tilde{\omega}(n\Delta t)}$, where $\tilde{\omega}$ denotes the discrete temporal frequency. Substituting this ansatz into the non-dissipative scheme yields

$$\left(e^{-i\tilde{\omega}\Delta t} + e^{i\tilde{\omega}\Delta t}\right) \hat{u}e^{ikx-i\tilde{\omega}(n\Delta t)} = \alpha \int_{-\infty}^{\infty} e^{-\alpha|x-y|} \hat{u}e^{iky-i\tilde{\omega}(n\Delta t)} dy.$$

Cancel the common term $\hat{u}e^{-i\omega n\Delta t}$ to get

$$\begin{aligned} \cos(\tilde{\omega}\Delta t) &= \frac{\alpha}{2} \int_{-\infty}^{\infty} e^{-\alpha|x-y|-ik(x-y)} dy \\ &= \frac{\alpha}{2} \int_0^{\infty} e^{-\alpha z} \left(e^{ikz} + e^{-ikz}\right) dz \end{aligned} \quad (3.18)$$

after the change of variables $z = y - x$. Evaluation of the integral gives

$$\cos(\tilde{\omega}\Delta t) = \frac{\alpha^2}{\alpha^2 + k^2} = \frac{1}{1 + \frac{(kc\Delta t)^2}{2}}. \quad (3.19)$$

This is the semi-discrete analog to the continuous dispersion relation. To avoid aliasing, consider only the wave numbers in the region $0 \leq \tilde{\omega}\Delta t \leq \pi$. The phase error can then be defined as

$$\Phi(\Delta t) = \left| \frac{\tilde{\omega}}{kc} - 1 \right|.$$

For a fixed frequency k , we can analyze the convergence of this approximation using a Taylor series approximation for small $kc\Delta t$. First, make use of the trigonometric identity

$$\cos(2\theta) = 1 - 2\sin^2(\theta)$$

in (3.19) to obtain

$$\begin{aligned} 1 - 2\sin^2\left(\frac{\omega\Delta t}{2}\right) &= \frac{1}{1 + \frac{(kc\Delta t)^2}{2}} \\ \implies 2\sin^2\left(\frac{\omega\Delta t}{2}\right) &= 1 - \frac{1}{1 + \frac{(kc\Delta t)^2}{2}} \\ \implies 2\sin^2\left(\frac{\omega\Delta t}{2}\right) &= \frac{(kc\Delta t)^2}{2 + (kc\Delta t)^2}, \end{aligned}$$

which gives

$$\sin^2\left(\frac{\tilde{\omega}\Delta t}{2}\right) = \frac{(kc\Delta t)^2}{4} \left(\frac{1}{1 + \frac{(kc\Delta t)^2}{2}} \right) = \frac{z^2}{1 + 2z^2},$$

where $z = \frac{kc\Delta t}{2}$. Taking the square root and arcsin of both sides of the equation now yields

$$\tilde{\omega} = \frac{2}{\Delta t} \arcsin\left(\frac{z}{\sqrt{1 + 2z^2}}\right). \quad (3.20)$$

The first terms of the binomial expansion for the denominator are given by

$$\begin{aligned}\left(1 + 2z^2\right)^{-1/2} &= 1 - \frac{1}{2}2z^2 + \frac{3}{8}\left(2z^2\right)^2 \\ &= 1 - \frac{(kc\Delta t)^2}{4} + \frac{3(kc\Delta t)^4}{32}\end{aligned}$$

which, in turn, gives that

$$\frac{z}{\sqrt{1 + 2z^2}} = \frac{kc\Delta t}{2} \left(1 - \frac{(kc\Delta t)^2}{4}\right) + O\left(\Delta t^5\right).$$

Substituting this into the arcsin expansion

$$\arcsin(z) = z + \frac{z^3}{6} + \frac{3z^5}{40} + O\left(z^7\right)$$

now turns (3.20) into

$$\begin{aligned}\tilde{\omega} &= \frac{2}{\Delta t} \left(\frac{kc\Delta t}{2} \left(1 - \frac{(kc\Delta t)^2}{4}\right) + \frac{\left(\frac{kc\Delta t}{2} \left(1 - \frac{(kc\Delta t)^2}{4}\right)\right)^3}{6} + O\left(\Delta t^5\right) \right) \\ &= kc \left(1 - \frac{5}{24}(kc\Delta t)^2 + O\left(\Delta t^4\right)\right),\end{aligned}$$

from which it follows that the phase error is second order.

3.2.2 Semi-discrete Dissipative Schemes

For the dissipative scheme, similarly define $u^n = \hat{u}e^{ikx - i\tilde{\omega}(n\Delta t)}$. Substituting this ansatz into the first order dissipative scheme yields

$$e^{-i\tilde{\omega}(n+1)\Delta t}\hat{u}e^{ikx} = \frac{\alpha}{2} \int_{-\infty}^{\infty} \left(e^{-i\tilde{\omega}n\Delta t}\hat{u}e^{iky} - e^{-i\tilde{\omega}(n-1)\Delta t}\hat{u}e^{iky} \right) e^{-\alpha|x-y|} dy.$$

Once again, cancel the common term $\hat{u}e^{-i\tilde{\omega}n\Delta t}$ to get

$$\begin{aligned} e^{-i\tilde{\omega}\Delta t} &= \frac{\alpha}{2} \int_{-\infty}^{\infty} \left(\hat{u}e^{ik(y-x)} - e^{i\tilde{\omega}\Delta t}\hat{u}e^{ik(y-x)} \right) e^{-\alpha|x-y|} dy \\ &= \frac{\alpha}{2} \int_{-\infty}^{\infty} e^{ik(y-x) - \alpha|x-y|} \left(2 - e^{i\tilde{\omega}\Delta t} \right) dy \\ \implies \frac{e^{-i\tilde{\omega}\Delta t}}{2 - e^{i\tilde{\omega}\Delta t}} &= \frac{\alpha}{2} \int_{-\infty}^{\infty} e^{ik(y-x) - \alpha|x-y|} dy. \end{aligned}$$

Use the change of variables $z = y - x$ to get

$$\frac{e^{-i\tilde{\omega}\Delta t}}{2 - e^{i\tilde{\omega}\Delta t}} = \frac{\alpha}{2} \int_0^{\infty} e^{-\alpha z} \left(e^{ikz} + e^{-ikz} \right) dz.$$

Performing the integration gives

$$\begin{aligned} \frac{e^{-i\tilde{\omega}\Delta t}}{2 - e^{i\tilde{\omega}\Delta t}} &= \frac{\alpha^2}{\alpha^2 + k^2} \\ \implies e^{-i\tilde{\omega}\Delta t} &= \frac{2 - e^{i\tilde{\omega}\Delta t}}{1 + (kc\Delta t)^2}. \end{aligned}$$

From this expression we would like to isolate $\tilde{\omega}$, and so we solve the corresponding

quadratic polynomial for $\lambda = e^{i\tilde{\omega}\Delta t}$,

$$\lambda^2 - 2\lambda + 1 + (kc\Delta t)^2 = 0,$$

which admits complex solutions $\lambda = 1 + \pm ikc\Delta t$. Thus, the temporal frequencies will be complex (dissipative), and are of the form

$$\begin{aligned}\tilde{\omega} &= \frac{1}{i\Delta t} \log(1 + ikc\Delta t) \\ &= kc \left(1 - \frac{ikc\Delta t}{2} + O(\Delta t^2) \right),\end{aligned}$$

which shows that the phase error is first order.

We then repeat this for the final method by putting the ansatz $u^n = \hat{u}e^{ikx - i\tilde{\omega}(n\Delta t)}$ into the second order dissipative scheme. This gives

$$\begin{aligned}e^{-i\tilde{\omega}(n+1)\Delta t}\hat{u}e^{ikx} &= \\ \frac{\alpha}{2} \int_{-\infty}^{\infty} \left(\frac{5}{2}e^{-i\tilde{\omega}n\Delta t}\hat{u}e^{iky} - 2e^{-i\tilde{\omega}(n-1)\Delta t}\hat{u}e^{iky} + \frac{1}{2}e^{-i\tilde{\omega}(n-2)\Delta t}\hat{u}e^{iky} \right) e^{-\alpha|x-y|} dy.\end{aligned}$$

Cancel the common term $\hat{u}e^{-i\tilde{\omega}n\Delta t}$ to get

$$\begin{aligned}e^{-i\tilde{\omega}\Delta t} &= \frac{\alpha}{2} \int_{-\infty}^{\infty} \left(\frac{5}{2}e^{ik(y-x)} - 2e^{i\tilde{\omega}\Delta t}e^{ik(y-x)} + \frac{1}{2}e^{2i\tilde{\omega}\Delta t}e^{ik(y-x)} \right) e^{-\alpha|x-y|} dy \\ &= \frac{\alpha}{2} \int_{-\infty}^{\infty} \left(\frac{5}{2} - 2e^{i\tilde{\omega}\Delta t} + \frac{1}{2}e^{2i\tilde{\omega}\Delta t} \right) e^{ik(y-x) - \alpha|x-y|} dy \\ &\Rightarrow \frac{2e^{-i\tilde{\omega}\Delta t}}{5 - 4e^{i\tilde{\omega}\Delta t} + e^{2i\tilde{\omega}\Delta t}} = \int_{-\infty}^{\infty} e^{ik(y-x) - \alpha|x-y|} dy.\end{aligned}$$

Again, use the change of variables $z = y - x$ to get

$$\frac{2e^{-i\tilde{\omega}\Delta t}}{5 - 4e^{i\tilde{\omega}\Delta t} + e^{2i\tilde{\omega}\Delta t}} = \frac{\alpha}{2} \int_0^\infty e^{-\alpha z} \left(e^{ikz} + e^{-ikz} \right) dz.$$

Performing the integration gives

$$\begin{aligned} \frac{2e^{-i\tilde{\omega}\Delta t}}{5 - 4e^{i\tilde{\omega}\Delta t} + e^{2i\tilde{\omega}\Delta t}} &= \frac{\alpha^2}{\alpha^2 + k^2} \\ \implies 2e^{-i\tilde{\omega}\Delta t} &= \frac{5 - 4e^{i\tilde{\omega}\Delta t} + e^{2i\tilde{\omega}\Delta t}}{1 + \frac{(kc\Delta t)^2}{2}}. \end{aligned}$$

Setting $\lambda = e^{i\tilde{\omega}\Delta t}$ and setting this as a polynomial gives

$$\lambda^3 - 4\lambda^2 + 5\lambda - 2 - (kc\Delta t)^2 = 0.$$

Solving this in Mathematica gives the solutions as

$$\lambda = \left\{ \frac{4}{3} + \frac{\sqrt[3]{2}}{3R} + \frac{R}{3\sqrt[3]{2}}, \frac{4}{3} - \frac{1 - i\sqrt{3}}{3\sqrt[3]{4}R} - \frac{(1 + i\sqrt{3})R}{6\sqrt[3]{2}}, \frac{4}{3} - \frac{1 + i\sqrt{3}}{3\sqrt[3]{4}R} - \frac{(1 - i\sqrt{3})R}{6\sqrt[3]{2}} \right\}, \quad (3.21)$$

where $R = \sqrt[3]{2 + 27(kc\Delta t)^2 + 3\sqrt{3}\sqrt{4(kc\Delta t)^2 + 27(kc\Delta t)^4}}.$

From here, consider the asymptotic limit as $\Delta t \rightarrow 0$, which gives that these roots approach respectively $\{2, 1, 1\}$. Note that the first root, $\frac{4}{3} + \frac{\sqrt[3]{2}}{3R} + \frac{R}{3\sqrt[3]{2}}$ is real and thus corresponds to a parasitic mode instead of a propagating mode and thus it will not be considered any

further. For the two complex roots, consider perturbative roots of the form

$$\lambda = 1 + \delta, \quad \delta = c_1\epsilon + c_2\epsilon^2 + c_3\epsilon^3 + \dots, \quad \epsilon = kc\Delta t. \quad (3.22)$$

We may now substitute $\lambda = 1 + \delta$ into $\lambda^3 - 4\lambda^2 + 5\lambda - 2 - \epsilon^2 = 0$ and this becomes

$$\begin{aligned} (1 + \delta)^3 - 4(1 + \delta)^2 + 5(1 + \delta) - 2 &= \epsilon^2 \\ \implies 1 + 3\delta + 3\delta^2 + \delta^3 - 4 - 8\delta - 4\delta^2 + 5 + 5\delta - 2 &= \epsilon^2 \\ \implies \delta^3 - \delta^2 &= \epsilon^2. \end{aligned} \quad (3.23)$$

From (3.22), we get

$$\begin{aligned} \delta^3 &= c_1^3\epsilon^3 + 3c_1^2c_2\epsilon^4 + O(\epsilon^5) \\ \delta^2 &= c_1^2\epsilon^2 + 2c_1c_2\epsilon^3 + (2c_1c_3 + c_2^2)\epsilon^4 + O(\epsilon^5). \end{aligned}$$

Matching powers of ϵ in (3.23) gives the system of equations

$$\begin{aligned} -c_1^2 &= 1, \\ c_1^3 - 2c_1c_2 &= 0, \\ 3c_1^2c_2 - 2c_1c_3 - c_2^2 &= 0 \end{aligned}$$

which has a solution of

$$c_1 = \pm i, \quad c_2 = \frac{-1}{2}, \quad c_3 = \frac{\mp 5i}{8}$$

and thus, the two complex roots of (3.21) are given by

$$\begin{aligned}
\lambda &= 1 \pm ikc\Delta t - \frac{(kc\Delta t)^2}{2} \mp \frac{5(kc\Delta t)^3 i}{8} + O(\Delta t^4) \\
&= e^{\pm ikc\Delta t} - \frac{11(kc\Delta t)^3 i}{24} + O(\Delta t^4) \\
\Rightarrow e^{i\tilde{\omega}\Delta t} &= e^{\pm ikc\Delta t} - \frac{11(kc\Delta t)^3 i}{24} + O(\Delta t^4).
\end{aligned}$$

Thus, we have that $i\tilde{\omega}\Delta t$ matches $\pm ikc\Delta t$ to the third order which implies that

$$\tilde{\omega}^2 = (kc)^2 + O(\Delta t^2)$$

and the phase error is of second order.

3.2.3 Fully Discrete Non-dissipative Scheme

Consider the non-dissipative scheme employing the midpoint quadrature rule for spatial discretization. The only modification to equation (3.18) will be that the integral is replaced with an infinite sum. Let $u_j^n = \hat{u}e^{i(kj\Delta x - n\tilde{\omega}\Delta t)}$, where now $\tilde{\omega}$ denotes the fully discrete temporal frequency, yielding

$$(e^{-i\tilde{\omega}\Delta t} + e^{i\tilde{\omega}\Delta t}) \hat{u}e^{ikj\Delta x - i\tilde{\omega}(n\Delta t)} = \alpha \sum_{\ell=-\infty}^{\infty} \int_{-\Delta x/2}^{\Delta x/2} e^{-\alpha|x-y|} \hat{u}e^{iky - i\tilde{\omega}(n\Delta t)} dy.$$

After dividing through by $\hat{u}e^{ikj\Delta x - i\tilde{\omega}n\Delta t}$, this becomes

$$\begin{aligned} e^{-i\tilde{\omega}\Delta t} + e^{i\tilde{\omega}\Delta t} &= \sum_{m=-\infty}^{\infty} \alpha e^{ik(m-j)\Delta x} \int_{\Omega_m} e^{-\alpha|x_j-y|} dy \\ \implies 2 \cos(\tilde{\omega}\Delta t) &= \sum_{m=-\infty}^{\infty} \alpha e^{ik(m-j)\Delta x} \int_{\Omega_m} e^{-\alpha|x_j-y|} dy. \end{aligned}$$

The integration produces

$$\int_{\Omega_m} e^{-\alpha|x_j-y|} dy = \begin{cases} \frac{2}{\alpha} \left(1 - e^{(-\nu/2)}\right), & j = m \\ \frac{2}{\alpha} e^{-\alpha|j-m|\Delta x} \sinh\left(\frac{\nu}{2}\right), & j \neq m. \end{cases}$$

This then gives that

$$2 \cos(\tilde{\omega}\Delta t) = 2 \left[1 - e^{(-\nu/2)} + \sum_{m \neq j} e^{ik(m-j)\Delta x} e^{-\alpha|j-m|\Delta x} \sinh\left(\frac{\nu}{2}\right) \right].$$

After dividing by two and centering the sum about $m = j$, this now becomes

$$\begin{aligned} \cos(\tilde{\omega}\Delta t) &= 1 - e^{(-\nu/2)} + \sinh\left(\frac{\nu}{2}\right) \sum_{m \neq 0} e^{ikm\Delta x} e^{-\alpha|m|\Delta x} \\ &= 1 - e^{(-\nu/2)} + \sinh\left(\frac{\nu}{2}\right) \left(\sum_{m=-1}^{-\infty} e^{(ik+\alpha)\Delta x m} + \sum_{m=1}^{\infty} e^{(ik-\alpha)\Delta x m} \right) \\ &= 1 - e^{(-\nu/2)} + \sinh\left(\frac{\nu}{2}\right) \left(\sum_{m=1}^{\infty} e^{(-ik-\alpha)\Delta x m} + \sum_{m=1}^{\infty} e^{(ik-\alpha)\Delta x m} \right) \\ &= 1 - e^{(-\nu/2)} + \sinh\left(\frac{\nu}{2}\right) \left(\frac{e^{(-ik-\alpha)\Delta x}}{1 - e^{(-ik-\alpha)\Delta x}} + \frac{e^{(ik-\alpha)\Delta x}}{1 - e^{(ik-\alpha)\Delta x}} \right). \end{aligned}$$

Finding a common denominator then yields

$$\begin{aligned}\cos(\tilde{\omega}\Delta t) &= 1 - e^{(-\nu/2)} + \sinh\left(\frac{\nu}{2}\right) \left(\frac{e^{(-ik-\alpha)\Delta x} - 2e^{-2\alpha\Delta x} + e^{(ik-\alpha)\Delta x}}{1 - e^{(-ik-\alpha)\Delta x} - e^{(ik-\alpha)\Delta x} + e^{-2\alpha\Delta x}} \right) \\ &= 1 - e^{(-\nu/2)} + \sinh\left(\frac{\nu}{2}\right) \left(\frac{\cos(k\Delta x) - e^{-\nu}}{\cosh(\nu) - \cos(k\Delta x)} \right).\end{aligned}$$

Further algebraic manipulation shows that this becomes

$$\begin{aligned}\cos(\tilde{\omega}\Delta t) &= 1 - e^{(-\nu/2)} + 2\sinh\left(\frac{\nu}{2}\right) \left(\frac{\cos(k\Delta x)e^\nu - 1}{1 + e^{2\nu} - 2\cos(k\Delta x)e^\nu} \right) \\ &= 1 - e^{(\nu/2)} + 2\sinh\left(\frac{\nu}{2}\right) + 2\sinh\left(\frac{\nu}{2}\right) \left(\frac{\cos(k\Delta x)e^\nu - 1}{1 + e^{2\nu} - 2\cos(k\Delta x)e^\nu} \right) \\ &= 1 - e^{(\nu/2)} + 2\sinh\left(\frac{\nu}{2}\right) \left(\frac{1 + e^{2\nu} - 2\cos(k\Delta x)e^\nu + \cos(k\Delta x)e^\nu - 1}{1 + e^{2\nu} - 2\cos(k\Delta x)e^\nu} \right) \\ &= 1 - e^{(\nu/2)} + 2\sinh\left(\frac{\nu}{2}\right) \left(\frac{1 - \cos(k\Delta x)e^{-\nu}}{1 + e^{-2\nu} - 2\cos(k\Delta x)e^{-\nu}} \right).\end{aligned}$$

As in the semi-discrete analysis, we study the phase for a fixed frequency k . Some additional algebraic manipulation, including the double-angle formulas for $\cos$ and $\cosh$, gives

$$\begin{aligned}\cos(\tilde{\omega}\Delta t) &= 1 - e^{(\nu/2)} + 2\sinh\left(\frac{\nu}{2}\right) \left(\frac{1 - \cos(k\Delta x)e^{-\nu}}{1 - 2\cos(k\Delta x)e^{-\nu} + e^{-2\nu}} \right) \\ &= 1 - \frac{e^{(\nu/2)} (1 - 2\cos(k\Delta x)e^{-\nu} + e^{-2\nu}) + 2\sinh(\nu/2) (1 - 2\cos(k\Delta x)e^{-\nu})}{1 - 2\cos(k\Delta x)e^{-\nu} + e^{-2\nu}} \\ &= 1 + \frac{(e^{(-\nu/2)} + e^{(-3\nu/2)}) \cos(k\Delta x) - e^{(-3\nu/2)} - e^{(-\nu/2)}}{1 - 2\cos(k\Delta x)e^{-\nu} + e^{-2\nu}} \\ &= 1 + \frac{(e^{(\nu/2)} + e^{(-\nu/2)}) \cos(k\Delta x) - e^{(-\nu/2)} - e^{(\nu/2)}}{e^\nu - 2\cos(k\Delta x) + e^{-\nu}}.\end{aligned}$$

This then becomes

$$\begin{aligned}
\cos(\tilde{\omega}\Delta t) &= 1 + 2 \cosh\left(\frac{\nu}{2}\right) \left(\frac{\cos(k\Delta x) - 1}{2 \cosh(\nu) - 2 \cos(k\Delta x)} \right) \\
&= 1 + 2 \cosh\left(\frac{\nu}{2}\right) \left(\frac{\cos(k\Delta x) - 1}{4 \sinh^2(\nu/2) + 4 \sin^2(k\Delta x/2)} \right) \\
&= 1 - 2 \sin^2\left(\frac{k\Delta x}{2}\right) \left(\frac{\cosh(\nu/2)}{2 \sinh^2(\nu/2) + 2 \sin^2(k\Delta x/2)} \right) \\
\Rightarrow -\frac{\cos(\tilde{\omega}\Delta t)}{2} &= \frac{-1}{2} + \sin^2\left(\frac{k\Delta x}{2}\right) \left(\frac{\cosh(\nu/2)}{2 \sinh^2(\nu/2) + 2 \sin^2(k\Delta x/2)} \right) \\
\Rightarrow \frac{1}{2} - \frac{\cos(\tilde{\omega}\Delta t)}{2} &= \sin^2\left(\frac{k\Delta x}{2}\right) \left(\frac{\cosh(\nu/2)}{2 \sinh^2(\nu/2) + 2 \sin^2(k\Delta x/2)} \right) \\
\Rightarrow \sin^2\left(\frac{\tilde{\omega}\Delta t}{2}\right) &= \sin^2\left(\frac{k\Delta x}{2}\right) \left(\frac{\cosh(\nu/2)}{2 \sinh^2(\nu/2) + 2 \sin^2(k\Delta x/2)} \right).
\end{aligned}$$

Since we are interested in implicit methods, the relevant inequality is $\Delta x < c\Delta t$, and so consider the limit of the phase error for $\Delta x, \Delta t \rightarrow 0$, while leaving ν as a small, fixed quantity. The error will be better understood if put it in a form similar to equation (3.20).

Therefore, define

$$z = \frac{kc\Delta t}{2}, \quad s_1 = \frac{2}{\nu} \sinh\left(\frac{\nu}{2}\right), \quad s_2 = \frac{2}{k\Delta x} \sin\left(\frac{k\Delta x}{2}\right). \quad (3.24)$$

Then the fully discretized phase error can be given as

$$\begin{aligned}
\sin^2\left(\frac{\tilde{\omega}\Delta t}{2}\right) &= \frac{\frac{1}{2}\sin^2\left(\frac{k\nu c\Delta t}{2\sqrt{2}}\right)\cosh\left(\frac{\nu}{2}\right)}{\sinh^2\left(\frac{\nu}{2}\right) + \sin^2\left(\frac{k\nu c\Delta t}{2\sqrt{2}}\right)} \\
&= \cosh\left(\frac{\nu}{2}\right)\frac{(zs_2)^2}{s_1^2 + 2(zs_2)^2} \\
&= \tilde{z}^2 \cosh\left(\frac{\nu}{2}\right)\frac{1}{1 + 2\tilde{z}^2}
\end{aligned}$$

where

$$\tilde{z} = \frac{s_2}{s_1}z = \frac{\sin(k\Delta x/2)}{2\sinh(\nu/2)} = \frac{1}{2}\sin\left(\frac{k\Delta x}{2}\right)\operatorname{csch}\left(\frac{\nu}{2}\right). \quad (3.25)$$

Notice that since $s_1 = 1 + O(\nu^2)$ and $s_2 = 1 + O((k\Delta x)^2)$, that the semi-discrete limit $\tilde{z} \sim z$ is recovered by letting ν and Δx tend to zero such that Δt is held fixed.

The relative phase error will now depends on ν , Δx and Δt with

$$\Phi(\nu, \Delta x, \Delta t) = \left| \frac{\tilde{\omega}}{kc} - 1 \right|.$$

Following the exact algebra and expansions as given in section 3.2.1, we have

$$\begin{aligned}
\tilde{\omega} &= \cosh(\nu/2)kc \left(1 - \frac{(kc\Delta t)^2}{4}\right) + \frac{\left(\frac{\cosh(\nu/2)kc\Delta t}{2} \left(1 - \frac{(kc\Delta t)^2}{4}\right)\right)^3}{3\Delta t} + O(\Delta t^4) \\
&= \cosh\left(\frac{\nu}{2}\right)kc \left(1 - \frac{(kc\Delta t)^2}{4} + \frac{(\cosh(\nu/2)(kc\Delta t))^3}{24} + O(\Delta t^4)\right).
\end{aligned}$$

Using the expansion

$$\cosh\left(\frac{\nu}{2}\right) = 1 + \frac{\nu^2}{8} + O\left(\nu^4\right),$$

it becomes easy to see that the error is now $O(\Delta t^2 + \nu^2)$. If for a fixed wave number k , we let Δt and $\Delta x \rightarrow 0$ in such a way that ν is held constant, we see that the phase error is non-zero in the static limit

$$\lim_{\Delta x \rightarrow 0} \Phi(\nu, \Delta x, \Delta t) = \left| \frac{\arcsin\left(\tilde{z} \sqrt{\frac{\cosh(\nu/2)}{1 + 2\tilde{z}^2}}\right)}{kc} - 1 \right| = \left| \frac{\sqrt{\cosh(\nu/2)}}{s_1} - 1 \right|,$$

which has been verified numerically for a range of ν . Additionally, letting $\nu \rightarrow 0$ produces zero phase error, but this limit is not realizable, as it implies that either $\Delta x = 0$ or $\Delta t = \infty$.

The fully discrete phase error is shown for the second order method in Figure 3.1, for several values of ν . Clearly the phase error is reduced as ν is reduced, which in the context of the plot corresponds to spatial refinement.

3.2.4 Fully Discrete Dissipative Schemes

We now repeat the same ideas for the dissipative schemes. Begin once again with the ansatz

$u_j^n = \hat{u} e^{i(kj\Delta x - n\tilde{\omega}\Delta t)}$, which when substituted into the first order scheme gives

$$e^{-i\tilde{\omega}\Delta t} \hat{u} e^{ikj\Delta x - i\tilde{\omega}(n\Delta t)} = \alpha \sum_{\ell=-\infty}^{\infty} \int_{-\Delta x/2}^{\Delta x/2} e^{-\alpha|x-y|} \hat{u} e^{iky - i\tilde{\omega}(n\Delta t)} \left(2 - e^{i\tilde{\omega}\Delta t}\right) dy.$$

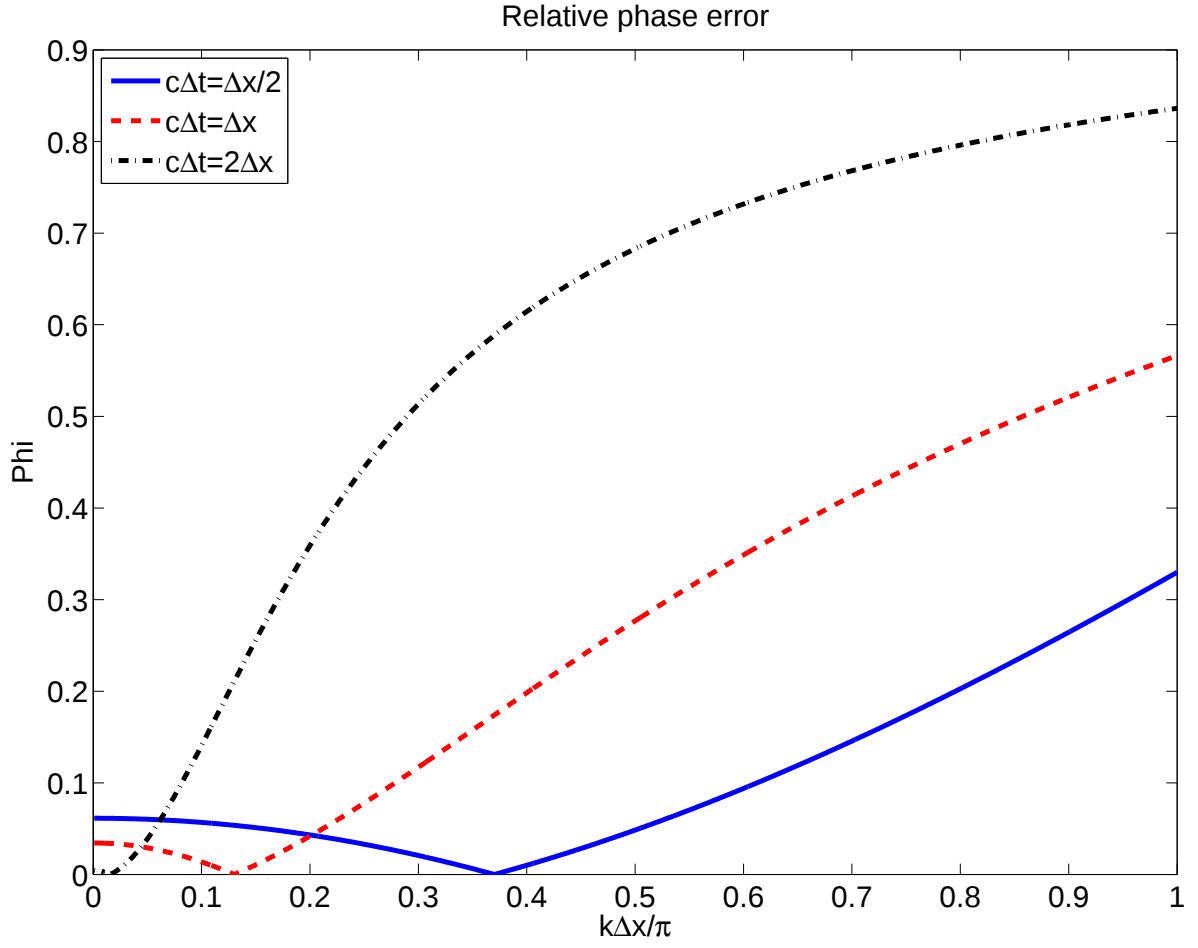


Figure 3.1: Relative Phase Error for the Non-dissipative Scheme

"For interpretation of the references to color in this and all other figures, the reader is referred to the electronic version of this dissertation."

Repeating the integration techniques used in the previous section, we arrive at

$$e^{-i\hat{\omega}\Delta t} = \left(2 - e^{i\hat{\omega}\Delta t}\right) \sigma$$

where

$$\sigma = 1 - e^{(\nu/2)} + 2 \sinh\left(\frac{\nu}{2}\right) \frac{1 - \cos(k\Delta x)e^{-\nu}}{1 - 2 \cos(k\Delta x)e^{-\nu} + e^{-2\nu}}. \quad (3.26)$$

Using the exact algebra as was given above, we have in terms of the expressions (3.24) and (3.25)

$$\begin{aligned} \sigma &= 1 - \cosh\left(\frac{\nu}{2}\right) \frac{\sin^2(k\Delta x/2)}{\sinh^2(\nu/2) + \sin^2(k\Delta x/2)} = 1 - \cosh\left(\frac{\nu}{2}\right) \frac{4\tilde{z}^2}{1 + 4\tilde{z}^2} \\ \Rightarrow e^{-i\hat{\omega}\Delta t} &= \left(2 - e^{i\hat{\omega}\Delta t}\right) \left(1 - \cosh\left(\frac{\nu}{2}\right) \frac{4\tilde{z}^2}{1 + 4\tilde{z}^2}\right). \end{aligned} \quad (3.27)$$

Letting $\lambda = e^{i\hat{\omega}\Delta t}$ and then multiplying (3.27) by $e^{i\hat{\omega}\Delta t}$ yields the quadratic equation

$$\sigma\lambda^2 - 2\sigma\lambda + 1 = 0.$$

This quadratic equation has solution $\lambda = 1 \pm i\sqrt{\frac{1}{\sigma} - 1}$. Here,

$$\begin{aligned} \frac{1}{\sigma} - 1 &= \frac{1 + 4\tilde{z}^2}{1 + 4\tilde{z}^2 - \cosh(\nu/2)4\tilde{z}^2} - 1 \\ &= \frac{4\tilde{z}^2 \cosh(\nu/2)}{1 + 4\tilde{z}^2 - \cosh(\nu/2)4\tilde{z}^2}. \end{aligned}$$

Therefore,

$$\begin{aligned}
\sqrt{\frac{1}{\sigma} - 1} &= \sqrt{\frac{4\tilde{z}^2 \cosh(\nu/2)}{1 + 4\tilde{z}^2 - \cosh(\nu/2)4\tilde{z}^2}} \\
&= 2\tilde{z} \sqrt{\frac{\cosh(\nu/2)}{1 + 4\tilde{z}^2 - \cosh(\nu/2)4\tilde{z}^2}} \\
&= 2\tilde{z} \sqrt{\frac{\cosh(\nu/2)}{1 - 8\tilde{z}^2 \sinh(\nu/4)}}.
\end{aligned}$$

This results in

$$\lambda = e^{i\hat{\omega}\Delta t} = 1 \pm i \sin\left(\frac{k\Delta x}{2}\right) \operatorname{csch}\left(\frac{\nu}{2}\right) \frac{\sqrt{\cosh(\nu/2)}}{\sqrt{1 - 8\sinh^2(\nu/4)\tilde{z}^2}}. \quad (3.28)$$

We now take the following expansions for each of the terms that make up the last term in equation (3.28):

$$\begin{aligned}
\sin\left(\frac{k\Delta x}{2}\right) &= \frac{k\Delta x}{2} - \frac{(k\Delta x)^3}{48} + O\left(\Delta x^5\right), \\
\operatorname{csch}\left(\frac{\nu}{2}\right) &= \frac{2}{\nu} + \frac{\nu}{12} + O\left(\nu^3\right), \\
\sqrt{\cosh\left(\frac{\nu}{2}\right)} &= \sqrt{1 - 1 + \cosh\left(\frac{\nu}{2}\right)} \\
&= 1 + \frac{\cosh(\nu/2) - 1}{2} + O\left(\nu^4\right) = 1 + \frac{\nu^2}{8} + O\left(\nu^4\right), \\
\text{and } \sqrt{1 - 8\sinh^2\left(\frac{\nu}{4}\right)\tilde{z}^2} &= 1 + 4\sinh^2\left(\frac{\nu}{4}\right)\tilde{z}^2 + O\left(\nu^2\right).
\end{aligned}$$

Substituting all of these in to equation (3.28) gives

$$\begin{aligned}\lambda = e^{i\hat{\omega}\Delta t} &= \frac{k\Delta x}{\nu} \left(1 + \frac{\nu^2}{8}\right) \left(1 + 4 \sinh^2\left(\frac{\nu}{4}\right) \tilde{z}^2\right) + O\left(\Delta x^2 + \nu^2\right) \\ &= 1 \pm ikc\Delta t \left(1 + \frac{\nu^2}{8} + \sinh^2\left(\frac{\nu}{4}\right)(kc\Delta t)^2 + O\left(\Delta x^2 + \nu^2\right)\right).\end{aligned}$$

We then take the log of both sides and expand the right hand side using the Taylor expansion for $\log(1+x)$ to get

$$i\hat{\omega}\Delta t = \pm ikc\Delta t \left(1 + \frac{\nu^2}{8} + \sinh^2\left(\frac{\nu}{4}\right)(kc\Delta t)^2 + O\left(\Delta x^2 + \nu^2\right)\right)$$

Dividing through by $i\Delta t$ will give that the error is first order.

Chapter 4

Analysis of the Solution on a Bounded Domain

We now repeat the analysis of these schemes on a bounded domain.

4.1 Consistency

4.1.1 Non-dissipative Scheme

First, consider the purely dispersive solver (2.6) with $\beta = 2$ on the region $[-L, L]$ with either zero Dirichlet or Neumann boundary conditions. The truncation error is now given as

$$\tau^n = \frac{1}{(c\Delta t)^2} \left(v^{n+1} + v^{n-1} - J[v^n] \right),$$

where

$$J[v^n] = \alpha \left(\int_{-L}^L e^{-\alpha|x-y|} v^n(y) dy + \tilde{c}_1(x) \int_{-L}^L e^{-\alpha(L-y)} v^n(y) dy + \tilde{c}_2(x) \int_{-L}^L e^{-\alpha(L+y)} v^n(y) dy \right). \quad (4.1)$$

Zero Dirichlet boundary conditions will give

$$\tilde{c}_1(x) = -\frac{\sinh(\alpha(L+x))}{\sinh(2\alpha L)}, \quad \tilde{c}_2(x) = -\frac{\sinh(\alpha(L-x))}{\sinh(2\alpha L)}$$

while zero Neumann boundary conditions give

$$\tilde{c}_1(x) = \frac{\cosh(\alpha(L+x))}{\sinh(2\alpha L)}, \quad \tilde{c}_2(x) = \frac{\cosh(\alpha(L-x))}{\sinh(2\alpha L)}.$$

This can be written as the integral of a single function $Q(x, y)$ multiplied by v^n with

$$Q(x, y) = \alpha \left[e^{-\alpha|x-y|} + \tilde{c}_1(x) e^{-\alpha(L-y)} + \tilde{c}_2(x) e^{-\alpha(L+y)} \right].$$

and due to the negative exponentials, it is easy to see that

$$|Q(x, y)| < 2, \quad \forall (x, y) \in [-L, L].$$

Observe that a direct consequence of the homogeneous Dirichlet boundary conditions is that at $x = \pm L$, v^n and all of its even spatial derivatives will vanish. That is,

$$\lim_{x \rightarrow \pm L} (\partial_x)^{2m} v(x, t) = \lim_{x \rightarrow \pm L} \frac{1}{c^{2m}} (\partial_t)^{2m} v(x, t) = \frac{1}{c^{2m}} (\partial_t)^{2m} v(\pm L, t) = 0.$$

The order of the limits and the derivatives can be swapped as long as the $2m$ -th derivatives of v^n exist and are continuous.

Similarly, for homogenous Neumann boundary conditions, the odd derivatives will disappear as

$$\lim_{x \rightarrow \pm L} (\partial_x)^{2m+1} v(x, t) = \lim_{x \rightarrow \pm L} \frac{1}{c^{2m}} (\partial_t)^{2m} v_x(x, t) = \frac{1}{c^{2m}} (\partial_t)^{2m} v_x(\pm L, t) = 0.$$

Now, we perform integration by parts on $J[v^n] = J_+[v^n] + J_-[v^n] + \tilde{c}_1(x)K_+[v^n] + \tilde{c}_2(x)K_-[v^n]$, where

$$\begin{aligned} J_{\pm} &= \alpha \int_0^{L \mp x} e^{-\alpha z} v^n(x \pm z) dz \\ &= [-v^n(x \pm z) e^{-\alpha z}]_0^{L \mp x} + \int_0^{L \mp x} (\pm 1) v_x^n(x \pm z) e^{-\alpha z} dz \\ &\vdots \\ &= \left[- \sum_{j=0}^k \left(\frac{\pm 1}{\alpha} \partial_x \right)^j v^n(x \pm z) e^{-\alpha z} \right]_0^{L \mp x} \\ &\quad - \int_0^{L \mp x} \left(\frac{\pm 1}{\alpha} \partial_x \right)^{k+1} v^n(x \pm z) e^{-\alpha z} dz. \end{aligned}$$

Similarly, the boundary terms are

$$\begin{aligned}
K_- &= \alpha \int_{-L}^L e^{-\alpha(L+y)} v^n(y) dy \\
&= \left[-v^n(y) e^{-\alpha(L+y)} \right]_{-L}^L + \int_{-L}^L v_x^n(y) e^{-\alpha(L+y)} dy \\
&\vdots \\
&= \left[-\sum_{j=0}^k \left(\frac{1}{\alpha} \partial_x \right)^j v^n(y) e^{-\alpha(L+y)} \right]_{-L}^L \\
&\quad + \int_{-L}^L \left(\frac{1}{\alpha} \partial_x \right)^{k+1} v^n(y) e^{-\alpha(L+y)} dy
\end{aligned}$$

and

$$\begin{aligned}
K_+ &= \alpha \int_{-L}^L e^{-\alpha(L-y)} v^n(y) dy \\
&= \left[v^n(y) e^{-\alpha(L-y)} \right]_{-L}^L - \int_{-L}^L v_x^n(y) e^{-\alpha(L-y)} dy \\
&\vdots \\
&= \left[\sum_{j=0}^k \left(\frac{-1}{\alpha} \partial_x \right)^j v^n(y) e^{-\alpha(L-y)} \right]_{-L}^L \\
&\quad + \int_{-L}^L \left(\frac{-1}{\alpha} \partial_x \right)^{k+1} v^n(y) e^{-\alpha(L-y)} dy.
\end{aligned}$$

Taking this expansion out to $k = 5$ and after grouping by odd and even derivatives at

the boundary points, we arrive at

$$\begin{aligned}
J[v^n] = & \delta_- \left(v^n(-L) + \frac{1}{\alpha^2} v_{xx}^n(-L) + \frac{1}{\alpha^4} v_{xxxx}^n(-L) \right) \\
& + \delta_+ \left(\frac{1}{\alpha} v_x^n(-L) + \frac{1}{\alpha^3} v_{xxx}^n(-L) + \frac{1}{\alpha^5} v_{xxxxx}^n(-L) \right) \\
& + \gamma_- \left(v^n(L) + \frac{1}{\alpha^2} v_{xx}^n(L) + \frac{1}{\alpha^4} v_{xxxx}^n(L) \right) \\
& + \gamma_+ \left(\frac{1}{\alpha} v_x^n(L) + \frac{1}{\alpha^3} v_{xxx}^n(L) + \frac{1}{\alpha^5} v_{xxxxx}^n(L) \right) \\
& + 2 \left(v^n(x) + \frac{1}{\alpha^2} v_{xx}^n(x) + \frac{1}{\alpha^4} v_{xxxx}^n(x) \right) \\
& + \frac{1}{\alpha^6} \int_0^L Q(x, y) v_{xxxxx}^n(y) dy
\end{aligned}$$

with

$$\begin{aligned}
\delta_{\pm} &= \tilde{c}_2(x) \pm e^{-2\alpha L} \tilde{c}_1(x) \pm e^{-\alpha(L+x)}, \\
\gamma_{\pm} &= \mp \tilde{c}_1(x) - e^{-2\alpha L} \tilde{c}_2(x) - e^{-\alpha(L-x)}.
\end{aligned}$$

For zero Dirichlet boundary conditions,

$$\begin{aligned}
\sinh(2\alpha L) \delta_+ &= -\sinh(\alpha(L-x)) - e^{-2\alpha L} \sinh(\alpha(L+x)) + \sinh(2\alpha L) e^{-\alpha(L+x)} \\
&= \frac{-e^{\alpha(L-x)} + e^{\alpha(-L+x)} - e^{\alpha(-L+x)} + e^{\alpha(-3L-x)} + e^{\alpha(L-x)} - e^{\alpha(-3L-x)}}{2} \\
&= 0
\end{aligned}$$

while

$$\begin{aligned}
\sinh(2\alpha L)\delta_- &= -\sinh(\alpha(L-x)) + e^{-2\alpha L}\sinh(\alpha(L+x)) - \sinh(2\alpha L)e^{-\alpha(L+x)} \\
&= \frac{-e^{\alpha(L-x)} + e^{\alpha(-L+x)} + e^{\alpha(-L+x)} - e^{\alpha(-3L-x)} - e^{\alpha(L-x)} + e^{\alpha(-3L-x)}}{2} \\
&= e^{\alpha(-L+x)} - e^{\alpha(L-x)} = 2\sinh(\alpha(x-L)).
\end{aligned}$$

Additionally,

$$\begin{aligned}
\sinh(2\alpha L)\gamma_+ &= \sinh(\alpha(L+x)) + e^{-2\alpha L}\sinh(\alpha(L-x)) - \sinh(2\alpha L)e^{-\alpha(L-x)} \\
&= \frac{e^{\alpha(L+x)} - e^{\alpha(-L-x)} + e^{\alpha(-L-x)} - e^{\alpha(-3L+x)} - e^{\alpha(L+x)} + e^{\alpha(-3L+x)}}{2} \\
&= 0
\end{aligned}$$

and

$$\begin{aligned}
\sinh(2\alpha L)\gamma_- &= -\sinh(\alpha(L+x)) + e^{-2\alpha L}\sinh(\alpha(L-x)) - \sinh(2\alpha L)e^{-\alpha(L-x)} \\
&= \frac{-e^{\alpha(L+x)} + e^{\alpha(-L-x)} + e^{\alpha(-L-x)} - e^{\alpha(-3L+x)} - e^{\alpha(L+x)} + e^{\alpha(-3L+x)}}{2} \\
&= -e^{\alpha(L+x)} + e^{\alpha(-L-x)} = -2\sinh(\alpha(L+x)).
\end{aligned}$$

However, it was already shown that all even derivatives of v^n are zero at the boundary.

Thus, all of the boundary terms vanish identically, and

$$J[v^n] = 2v^n + \frac{2}{\alpha^2}v_{xx}^n + \frac{2}{\alpha^4}v_{xxxx}^n + \frac{1}{\alpha^6} \int_{-L}^L Q(x,y)v_{xxxxxx}^n(y) dy. \quad (4.2)$$

A comparison of (4.2) and (3.2) shows that $J[v^n]$ and $I[v^n]$ differ by $O(\Delta t^6)$. Thus, for

all cases on a bounded domain with Dirichlet boundary conditions, $J[v^n]$ may be substituted for $I[v^n]$ into the free space consistency proof for the respective solver without having an effect on the order of the method.

For zero Neumann boundary conditions,

$$\begin{aligned}
\sinh(2\alpha L)\delta_+ &= \cosh(\alpha(L-x)) + e^{-2\alpha L}\cosh(\alpha(L+x)) + \sinh(2\alpha L)e^{-\alpha(L+x)} \\
&= \frac{e^{\alpha(L-x)} + e^{\alpha(-L+x)} + e^{\alpha(-L+x)} + e^{\alpha(-3L-x)} + e^{\alpha(L-x)} - e^{\alpha(-3L-x)}}{2} \\
&= e^{\alpha(L-x)} + e^{\alpha(-L+x)} = 2\cosh(\alpha(L-x))
\end{aligned}$$

while

$$\begin{aligned}
\sinh(2\alpha L)\delta_- &= \cosh(\alpha(L-x)) - e^{-2\alpha L}\cosh(\alpha(L+x)) - \sinh(2\alpha L)e^{-\alpha(L+x)} \\
&= \frac{e^{\alpha(L-x)} + e^{\alpha(-L+x)} - e^{\alpha(-L+x)} - e^{\alpha(-3L-x)} - e^{\alpha(L-x)} + e^{\alpha(-3L-x)}}{2} \\
&= 0.
\end{aligned}$$

Additionally,

$$\begin{aligned}
\sinh(2\alpha L)\gamma_+ &= -\cosh(\alpha(L+x)) - e^{-2\alpha L}\cosh(\alpha(L-x)) - \sinh(2\alpha L)e^{-\alpha(L-x)} \\
&= \frac{-e^{\alpha(L+x)} - e^{\alpha(-L-x)} - e^{\alpha(-L-x)} - e^{\alpha(-3L+x)} - e^{\alpha(L+x)} + e^{\alpha(-3L+x)}}{2} \\
&= -e^{\alpha(-L-x)} - e^{\alpha(L+x)} = -2\cosh(\alpha(L+x))
\end{aligned}$$

and

$$\begin{aligned}
\sinh(2\alpha L)\gamma_- &= \cosh(\alpha(L+x)) - e^{-2\alpha L}\cosh(\alpha(L-x)) - \sinh(2\alpha L)e^{-\alpha(L-x)} \\
&= \frac{e^{\alpha(L+x)} + e^{\alpha(-L-x)} - e^{\alpha(-L-x)} - e^{\alpha(-3L+x)} - e^{\alpha(L+x)} + e^{\alpha(-3L+x)}}{2} \\
&= 0.
\end{aligned}$$

However, it was already shown that all odd derivatives of v^n are zero at the boundary.

Thus, all of the boundary terms vanish identically, and

$$J[v^n] = 2v^n + \frac{2}{\alpha^2}v_{xx}^n + \frac{2}{\alpha^4}v_{xxxx}^n + \frac{1}{\alpha^6} \int_{-L}^L Q(x, y)v_{xxxxxx}^n(y) dy.$$

Once again, we may freely substitute $J[v^n]$ for $I[v^n]$ in any of the freespace consistency proofs without changing the order.

4.1.2 Dissipative Schemes

For the dissipative schemes (2.3) and (2.4), we can show consistency based on results obtained in previous sections. From section 3.1.1.2, we have that the first order scheme is given by

$$u^{n+1}(x) = I[u^n] - \frac{1}{2}I[u^{n-1}] \quad (4.3)$$

and the second order scheme is given by

$$u^{n+1}(x) = \frac{5}{4}I[u^n] - I[u^{n-1}] + \frac{1}{4}I[u^{n-2}]. \quad (4.4)$$

We can adapt these situations to the case of a finite mesh as

$$u^{n+1}(x) = J[u^n] - \frac{1}{2}J[u^{n-1}] \quad (4.5)$$

and

$$u^{n+1}(x) = \frac{5}{4}J[u^n] - J[u^{n-1}] + \frac{1}{4}J[u^{n-2}] \quad (4.6)$$

respectively, where J is as defined in (4.1). From section 3.1.1.2, we have that equations (4.3) and (4.4) are first and second order, respectively. From section 4.1.1, we have that $J[u] = I[u] + O(\Delta t^6)$. Thus, (4.5) and (4.6) must similarly be of first and second order.

4.2 Stability

In this section, the focus is on the incorporation of the boundary matrix, such that $A_t = A_p + A_h$ in equation (3.5). We will treat the matrix A_h , given by equation (3.8), perturbatively and look for a bound on the eigenvalues of A_t , based on $\lambda + \delta$, where λ is an eigenvalue of A_p and δ is an eigenvalue of A_h . Stability will follow if it can be shown that $|\mu| \leq 1$, where μ is now given by

$$\begin{aligned} \mu^2 - (\lambda + \delta)\mu + \frac{\lambda + \delta}{2} &= 0 \text{ and} \\ \mu^3 - \frac{5}{4}(\lambda + \delta)\mu^2 + (\lambda + \delta)\mu - \frac{1}{4}(\lambda + \delta)\mu &= 0 \end{aligned}$$

in the respective dissipative schemes and

$$\mu^2 - (\lambda + \delta)\mu + 1 = 0$$

in the non-dissipative scheme. While stability will hold for the non-dissipative scheme if $\lambda + \delta \in [-2, 2]$, the interval is $[-2/3, 2]$ for the first order dissipative scheme and $[-2/5, 2]$ for the second order dissipative scheme. The region for the second order dissipative scheme was found numerically while the other regions may be found by solving a quadratic equation.

Lemma 6. $\lambda + \delta \leq 2$.

Proof. By the Gerschgorin theorem, the maximal eigenvalue of $A_p + A_h$ is bounded above by the maximal row sum of $A_p + A_h$. For zero Dirichlet boundary conditions,

$$\begin{aligned} \sum_j (A_p)_{ij} &= \int_{-L}^L \alpha e^{-\alpha|x_i-y|} dy = 2 - e^{-\alpha(x_i+L)} - e^{-\alpha(L-x_i)} \text{ and} \\ \sum_j (A_h)_{ij} &= -\frac{\sinh(\alpha(L+x_i))}{\sinh(2\alpha L)} \int_{-L}^L \alpha e^{-\alpha|L-y|} dy \\ &\quad - \frac{\sinh(\alpha(L-x_i))}{\sinh(2\alpha L)} \int_{-L}^L \alpha e^{-\alpha|-L-y|} dy \\ &= (1 - e^{-2\alpha L}) \frac{-\sinh(\alpha(L+x_i)) - \sinh(\alpha(L-x_i))}{\sinh(2\alpha L)}. \end{aligned}$$

Thus,

$$\begin{aligned} \sum_j (A_p + A_h)_{ij} &= 2 - e^{-\alpha(x_i+L)} - e^{-\alpha(L-x_i)} \\ &\quad + (1 - e^{-2\alpha L}) \frac{-\sinh(\alpha(L+x_i)) - \sinh(\alpha(L-x_i))}{\sinh(2\alpha L)} < 2. \end{aligned}$$

In a similar manner, zero Neumann boundary conditions will give

$$\begin{aligned}
\sum_j (A_p)_{ij} &= \int_{-L}^L \alpha e^{-\alpha|x_i-y|} dy = 2 - e^{-\alpha(x_i+L)} - e^{-\alpha(L-x_i)} \text{ and} \\
\sum_j (A_h)_{ij} &= \frac{\cosh(\alpha(L+x_i))}{\sinh(2\alpha L)} \int_{-L}^L \alpha e^{-\alpha|L-y|} dy \\
&\quad + \frac{\cosh(\alpha(L-x_i))}{\sinh(2\alpha L)} \int_{-L}^L \alpha e^{-\alpha|-L-y|} dy \\
&= \left(1 - e^{-2\alpha L}\right) \frac{\cosh(\alpha(L+x_i)) + \cosh(\alpha(L-x_i))}{\sinh(2\alpha L)} \\
&= \frac{e^{-\alpha(L+x_i)} + e^{\alpha(L+x_i)} + e^{\alpha(x_i-L)} + e^{\alpha(L-x_i)}}{2\sinh(2\alpha L)} \\
&\quad - \frac{e^{-\alpha(3L+x_i)} - e^{\alpha(x_i-L)} - e^{\alpha(x_i-3L)} - e^{-\alpha(L+x_i)}}{2\sinh(2\alpha L)} \\
&= \frac{e^{-\alpha L} \left(e^{2\alpha L} - e^{-2\alpha L}\right) \left(e^{\alpha x_i} + e^{-\alpha x_i}\right)}{2\sinh(2\alpha L)} \\
&= 2e^{-\alpha L} \cosh(\alpha x_i).
\end{aligned}$$

Thus,

$$\begin{aligned}
\sum_j (A_p + A_h)_{ij} &= 2 - e^{-\alpha(x_i+L)} - e^{-\alpha(L-x_i)} + 2e^{-\alpha L} \cosh(\alpha x_i) \\
&= 2 - e^{-\alpha(x_i+L)} - e^{-\alpha(L-x_i)} + e^{-\alpha(x_i+L)} + e^{-\alpha(L-x_i)} = 2.
\end{aligned}$$

□

Since we already know that $\lambda > 0$, we must ensure that $-2/3 < \delta < 0$ in the first order scheme and $-2/5 < \delta < 0$ in the second order dissipative scheme to ensure stability. Below, we will make use of the Bauer-Fike theorem.

Theorem 7 (Bauer-Fike Theorem [11]). *Let B be a diagonalizable matrix, and Q be the non-singular eigenvector matrix such that $A = Q\Lambda Q^{-1}$, where Λ is the diagonal matrix containing the eigenvalues of B . Then for the matrix $B + E$, with eigenvalues $\lambda + \delta$, and where E is an arbitrary perturbation matrix,*

$$|\delta| \leq \|Q\|_2 \|Q^{-1}\|_2 \|E\|_2.$$

We will apply this theorem with $B = A_p$, which is a real symmetric positive definite matrix, and therefore we have that $\|Q\|_2 \|Q^{-1}\|_2 = 1$ since Q is orthogonal. Also, set $E = A_h$ and thus

$$|\delta| \leq \|A_h\|_2.$$

Therefore, attention will be restricted to the norm of A_h .

Notice that if we define the vectors

$$\begin{aligned} v_d &= (\sinh(\alpha(L - x_1)), \dots, \sinh(\alpha(L - x_N)))^T, \\ v_n &= (\cosh(\alpha(L - x_1)), \dots, \cosh(\alpha(L - x_N)))^T, \\ h &= (e^{-\alpha(L+x_1)}, \dots, e^{-\alpha(L+x_N)}) \end{aligned}$$

then the matrix A_h (3.8) can be expressed as

$$A_h = -2a_0 \frac{\sinh(\nu/2)}{\sinh(2\alpha L)} (v_i \cdot h + (Jv_i) \cdot (Jh)), \quad i \in \{d, n\} \quad (4.7)$$

where a_0 is given by (3.9) and the matrix J is the counter-identity matrix, containing ones

along the anti-diagonal. It has the effect of reversing the elements of a vector that it acts upon, and for an equally spaced symmetric mesh is equivalent to replacing x_j with $-x_j$. Since $v \cdot h$ and $(Jv) \cdot (Jh)$ are both rank one matrices, A_h is a matrix of at most rank two.

Theorem 8. *For the centrosymmetric rank-two matrix with n distinct entries which are negative and of decreasing magnitude less than one, the separation between the eigenvalues is bounded by $2a_n$.*

The proof of this theorem will make use of the following lemmas.

Lemma 9. *For these centrosymmetric matrices, $A_h(1, j)^2 + A_h(1, n - j + 1)^2 = A_h(1, 1) * A_h(j, j) + A_h(1, n) * A_h(j, n - j + 1)$.*

Proof. Letting $\nu = \frac{dx}{c\Delta t}$, $A_h(i, j) = \frac{\sinh((n - i + 0.5)\nu)}{\sinh(n\nu)} e^{-(j-0.5)\nu} + \frac{\sinh((i - 0.5)\nu)}{\sinh(n\nu)} e^{-(n-j+0.5)\nu}$.

This gives then that

$$\begin{aligned}
& \sinh^2(n\nu)(A_h(1, j)^2 + A_h(1, n - j + 1)^2) = \\
& \left(\sinh(-0.5\nu)e^{(j-n-0.5)\nu} + \sinh((0.5 - n)\nu)e^{(0.5-j)\nu} \right)^2 + \\
& \left(\sinh(-0.5\nu)e^{(0.5-j)\nu} + \sinh((0.5 - n)\nu)e^{(j-n-0.5)\nu} \right)^2 \\
& = 4 + e^{(2j-2)\nu} + e^{(2n-2j)\nu} - 4e^{-\nu} + e^{-2j\nu} + e^{(2-2j)\nu} + e^{(2j-2n)\nu} - 4e^{(1-2n)\nu} \\
& + 4e^{-2n\nu} + e^{(2j-4n)\nu} - 4e^{(1-2j)\nu} - 4e^{(2j-2n-1)\nu} + e^{(2-2n-2j)\nu} + e^{(2j-2n-2)\nu}
\end{aligned}$$

which is equal to

$$\begin{aligned}
& \left(\sinh(-0.5\nu)e^{(0.5-n)\nu} + \sinh((0.5-n)\nu)e^{-0.5\nu} \right) \\
& \cdot \left(\sinh((0.5-j)\nu)e^{(j-n-0.5)\nu} + \sinh((j-n-0.5)\nu)e^{(0.5-j)\nu} \right) \\
& + \left(\sinh(-0.5\nu)e^{-0.5\nu} + \sinh((0.5-n)\nu)e^{(0.5-n)\nu} \right) \\
& \cdot \left(\sinh((0.5-j)\nu)e^{(0.5-j)\nu} + \sinh((j-n-0.5)\nu)e^{(j-n-0.5)\nu} \right) \\
& = \sinh^2(n\nu)(A_h(1,1) * A_h(j,j) + A_h(1,n) * A_h(j,n-j+1)).
\end{aligned}$$

For the case of zero Neumann boundary conditions,

$$A(i,j) = \frac{\cosh((n-i+0.5)\nu)}{\sinh(n\nu)} \exp((j-0.5)\nu) + \frac{\cosh((i-0.5)\nu)}{\sinh(n\nu)} \exp((n-j+0.5)\nu),$$

which will consistently change signs in the algebra above and the result will be the same. $\square$

The corollary seen from this is that $\sum_{i=1}^n a_i^2 = \frac{a_1 \text{Trace}(A_n) + a_n \text{SkewTrace}(A_n)}{2}$.

Lemma 10. *For the given matrix, $\text{SkewTrace}(A_n)^2 = \text{Trace}(A_n)^2 - 4M_n$, where M_n consists of the principal 2-minors of A_n .*

Proof. Proof of theorem by induction:

Base Case: Consider the two by two matrix $A_2 = \begin{bmatrix} a_1 & a_2 \\ a_2 & a_1 \end{bmatrix}$.

$$\begin{aligned}
\text{Trace}(A_2)^2 - 4M_n &= (2a_1)^2 - 4(a_1^2 - a_2^2) \\
&= 4a_2^2 = \text{SkewTrace}(A_2)^2.
\end{aligned}$$

Induction Step:

Assume this true for any such matrix of size n . Extend the matrix to A_{n+2} of size $n+2$ by adding on an additional first and last row and column. Due to the nature of the extension from A_n to A_{n+2} , $\text{Trace}(A_{n+2}) = \text{Trace}(A_n) + 2A_{n+2}(1, 1)$. Thus,

$$\text{Trace}(A_{n+2})^2 = \text{Trace}(A_n)^2 + 4A_{n+2}(1, 1)\text{Trace}(A_n) + 4A_{n+2}(1, 1)^2. \quad (4.8)$$

M_n can be extended to M_{n+2} by adding in the principal 2-row minors that have an entry in the first or last row. Thus,

$$\begin{aligned} M_{n+2} = M_n + A_{n+2}(1, 1)^2 - A_{n+2}(1, n+2)^2 \\ + 4A_{n+2}(1, 1)\text{Trace}(A_n) - 2 \sum_{i=2}^n A_{n+2}(1, i)^2. \end{aligned} \quad (4.9)$$

Thus,

$$\begin{aligned} & \text{Trace}(A_{n+2})^2 - 4M_{n+2} \\ &= \text{Trace}(A_n)^2 + 4A_{n+2}(1, 1)^2 + 4A_{n+2}(1, 1)\text{Trace}(A_n) \\ & \quad - 4 \left(M_n + A_{n+2}(1, 1)^2 - A_{n+2}(1, n+2)^2 + 2A_{n+2}(1, 1)\text{Trace}(A_n) - 2 \sum_{i=2}^{n-1} A_{n+2}(1, i)^2 \right) \\ &= \text{Trace}(A_n)^2 - 4M_n + 4A_{n+2}(1, n+2)^2 - 4A_{n+2}(1, 1)\text{Trace}(A_n) + 8 \sum_{i=2}^n A_{n+2}(1, i)^2 \end{aligned}$$

From the base case, this is now

$$\begin{aligned}
& \text{SkewTrace}(A_n)^2 + 4A_{n+2}(1, n+2)^2 - 4A_{n+2}(1, 1)\text{Trace}(A_n) + 8 \sum_{i=2}^n A_{n+2}(1, i)^2 \\
&= \text{SkewTrace}(A_n)^2 + 4A_{n+2}(1, n+2)^2 + 4A_{n+2}(1, n+2)\text{SkewTrace}(A_n) \\
&= (\text{SkewTrace}(A_n) + 2A_{n+2}(1, n+2))^2 \\
&= \text{SkewTrace}(A_{n+2})^2.
\end{aligned}$$

□

Proof. Proof of theorem 8 by induction:

Base case 1:

Consider the case of the one by one centrosymmetric matrix $[a]$. This obviously has one eigenvalue a and there is no separation to consider.

Base case 2:

Consider the case of the two by two centrosymmetric rank-two Hankel matrix $\begin{bmatrix} a_1 & a_2 \\ a_2 & a_1 \end{bmatrix}$ which obviously has eigenvalues $\lambda = a_1 \pm a_2$, which have a difference of $2a_2 = 2a_n$.

Induction Step:

Assume this is true for any matrix A_n of size n . Extend the matrix to A_{n+2} of size $n+2$ by adding on an additional first and last row and column. The characteristic polynomial is given as

$$P(\lambda) = \sum_{i=0}^{2n} c_i \lambda^i. \quad (4.10)$$

Since this is the characteristic polynomial, $c_{2n} = 1$. Further, because the matrix is of rank-two, $c_i = 0$ for $i < 2n - 2$. Also, $c_{2n-1} = \text{Trace}(A_{n+2})$ and c_{2n-2} is the sum of the principal 2-row minors. Solving the resultant quadratic equation yields that the eigenvalues of the matrix are

$$\frac{1}{2} \left(\text{Trace}(A_{n+2}) + \sqrt{\text{Trace}(A_{n+2})^2 - 4M_{n+2}} \right) \quad (4.11)$$

where M_{n+2} denotes the sum of the principal 2-row minors. Since A_n and A_{n+2} are real and symmetric, all the eigenvalues must be real, which implies that $\text{Trace}(A_{n+2})^2 \geq 4M_{n+2}$.

By lemma 10, the eigenvalues for A_{n+2} are $\text{Trace}(A_{n+2}) \pm \text{SkewTrace}(A_{n+2})$.

□

From the above theorem, the eigenvalues of A_h are given by $\frac{\text{Trace}(A_h) \pm \text{Trace}(JA_h)}{2}$.

Furthermore, it is clear from equation (3.8) that $A_{ij} = A_{ji}$, giving that A_h is symmetric.

Therefore, the 2-norm of the matrix is equal to its spectral radius

$$\begin{aligned} \|A_h\|_2 &= -\frac{\text{Trace}(A_h) + \text{Trace}(JA_h)}{2} \\ &= \frac{\sinh(\nu/2)}{\sinh(2\alpha L)} (v \cdot h + (Jv) \cdot (Jh) + (Jv) \cdot h + v \cdot (Jh)). \end{aligned}$$

These dot products are each given by a geometric series determined by the boundary

conditions. For Dirichlet boundary conditions,

$$\begin{aligned}
v \cdot h &= (Jv) \cdot (Jh) = \sum_{j=1}^N e^{-\alpha(L+x_j)} \sinh(\alpha(L-x_j)) \\
&= \frac{1}{2} \sum_{j=1}^N e^{-2\alpha x_j} - e^{-2\alpha L} \\
&= \frac{1}{2} \left(\frac{\sinh(N\nu)}{\sinh(\nu)} - N e^{-N\nu} \right)
\end{aligned}$$

and

$$\begin{aligned}
(Jv) \cdot h &= v \cdot (Jh) = \sum_{j=1}^N e^{-\alpha(L+x_j)} \sinh(\alpha(L+x_j)) \\
&= \frac{1}{2} \sum_{j=1}^N 1 - e^{-2\alpha(L+x_j)} \\
&= \frac{1}{2} \left(N - e^{-N\nu} \frac{\sinh(N\nu)}{\sinh(\nu)} \right),
\end{aligned}$$

where we recall that $2\alpha L = N\nu$. Thus,

$$\begin{aligned}
|\delta| \leq \|A_h\| &= a_0 \left(1 - e^{-N\nu} \right) \left(\frac{\sinh(\nu/2)}{\sinh(N\nu)} \right) \left(\frac{\sinh(N\nu)}{\sinh(\nu)} + N \right) \\
&= a_0 \left(1 - e^{-N\nu} \right) \left(\frac{\sinh(\nu/2)}{\sinh(\nu)} + N \frac{\sinh(\nu/2)}{\sinh(N\nu)} \right).
\end{aligned}$$

Neumann boundary conditions give

$$\begin{aligned}
v \cdot h &= (Jv) \cdot (Jh) = \sum_{j=1}^N e^{-\alpha(L+x_j)} \cosh(\alpha(L-x_j)) \\
&= \frac{1}{2} \sum_{j=1}^N e^{-2\alpha x_j} + e^{-2\alpha L} \\
&= \frac{1}{2} \left(\frac{\sinh(N\nu)}{\sinh(\nu)} + N e^{-N\nu} \right)
\end{aligned}$$

and

$$\begin{aligned}
(Jv) \cdot h &= v \cdot (Jh) = \sum_{j=1}^N e^{-\alpha(L+x_j)} \cosh(\alpha(L+x_j)) \\
&= \frac{1}{2} \sum_{j=1}^N 1 + e^{-2\alpha(L+x_j)} \\
&= \frac{1}{2} \left(N + e^{-N\nu} \frac{\sinh(N\nu)}{\sinh(\nu)} \right).
\end{aligned}$$

Thus,

$$\begin{aligned}
|\delta| \leq \|A_h\| &= a_0 \left(1 + e^{-N\nu} \right) \left(\frac{\sinh(\nu/2)}{\sinh(N\nu)} \right) \left(\frac{\sinh(N\nu)}{\sinh(\nu)} + N \right) \\
&= a_0 \left(1 + e^{-N\nu} \right) \left(\frac{\sinh(\nu/2)}{\sinh(\nu)} + N \frac{\sinh(\nu/2)}{\sinh(N\nu)} \right).
\end{aligned}$$

Figures 4.1 and 4.2 give numerical plots of these eigenvalue bounds versus $\nu = \Delta x/(c\Delta t)$ and N . The figure encompasses $1 \leq N \leq 75$ and $0.01 \leq \nu \leq 2$. The white strips indicate a region where the bound $\delta > -2/3$ is not satisfied, and therefore indicates a potential for instabilities to arise. The blue area indicates a region where $-2/3 < \delta < -2/5$ indicating stability of the first order scheme but not of the second order scheme. However, no such

instabilities have been observed in my numerical experiments thus far, and it is likely that if a tighter bound were taken, this region of instability would not be present. It should be emphasized that this stability region only applies to the dissipative schemes and has no bearing on the second order purely dispersive scheme.

4.3 Numerical Results

In this section we present preliminary results for the convergence of the outlined numerical schemes for a test case. The wave equation is computed on the unit interval, with wave speed $c = 2$, up to a time $t = T = 1$, and zero Dirichlet conditions are imposed. For the initial condition, we prescribe a Gaussian function

$$f(x) = e^{-\left(\frac{12(x-0.5)}{L}\right)^2}, \quad g(x) = 0$$

which will be initially of negligible amplitude at the boundaries, and will have the exact solution

$$v(x, t) = \frac{e^{-\left(\frac{12(x-0.5-ct)}{L}\right)^2} + e^{-\left(\frac{12(x-0.5+ct)}{L}\right)^2}}{2}.$$

The discretization is made so that the ratio $c\Delta t/\Delta x$ is held fixed at 10 in the first order method, and 20 in the second order methods. Note that in both cases, this quantity is well beyond the CFL stability limit imposed on standard finite difference schemes, and so this method is an order of magnitude faster. The spatial step $\Delta x = 1/N$ is varied in the range

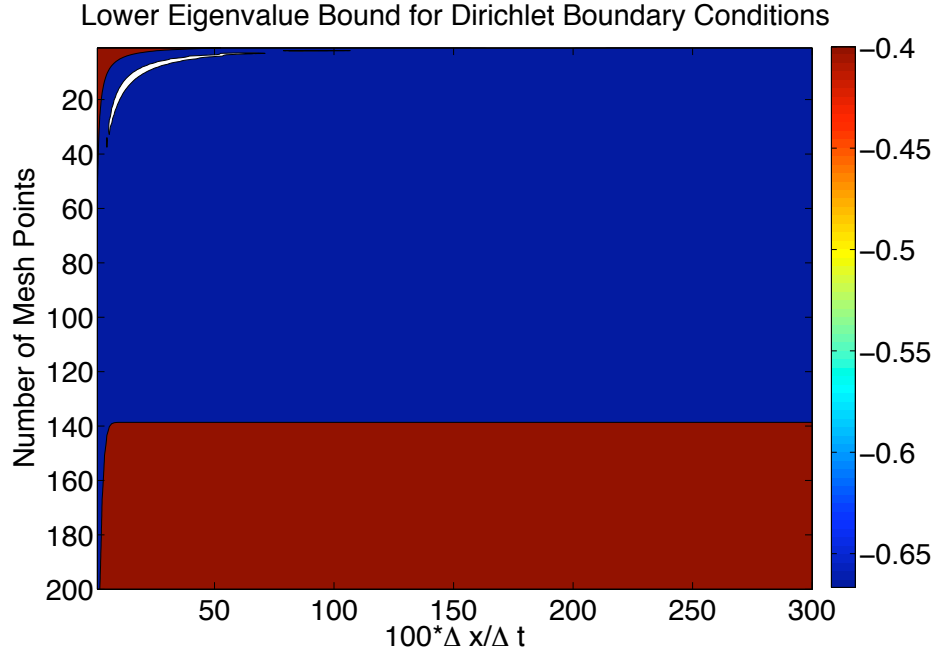


Figure 4.1: Contour Plot of the Minimum Eigenvalue of the Dissipative Schemes for Zero Dirichlet Boundary Conditions with Midpoint Quadrature

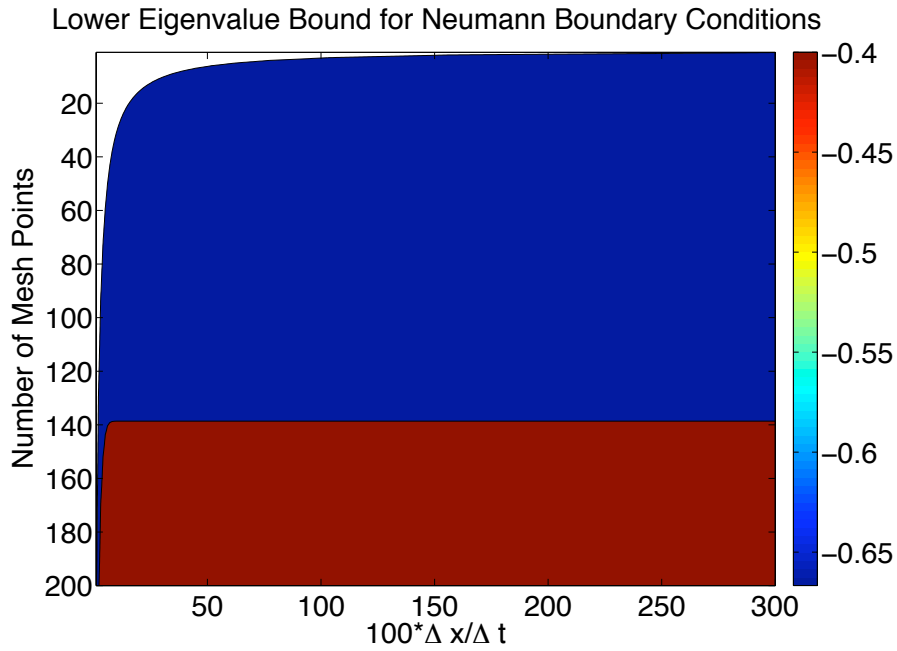


Figure 4.2: Contour Plot of the Minimum Eigenvalue of the Dissipative Schemes for Zero Neumann Boundary Conditions with Midpoint Quadrature

corresponding to $N = 50 * 2^k, k = \{1, 2, \dots, 8\}$, and Δt is scaled accordingly. The error is computed as

$$\|v(x, t) - u(x, t)\|_{L^2[0,1]}.$$

The error for the first order dissipative scheme is shown in Figure 4.3. The convergence is as expected, with the exception of the gradual decrease for larger values of Δt . Similar trends are observed in Figures 4.4 and 4.5, the second order schemes that are dissipative and non-dissipative, respectively. Since our interest is in taking time steps larger than the CFL condition, it is reassuring to see the error drop with increasing time step over some nontrivial interval.

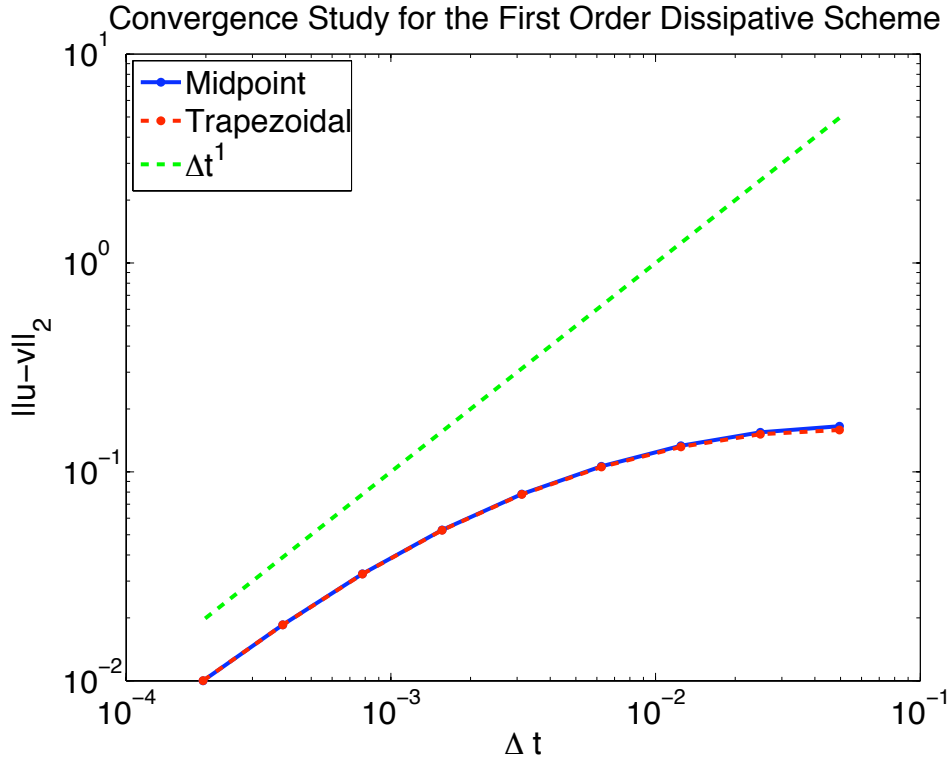


Figure 4.3: Convergence of the First Order Dissipative Scheme with $\nu = \Delta x / (c\Delta t) = 0.1$.

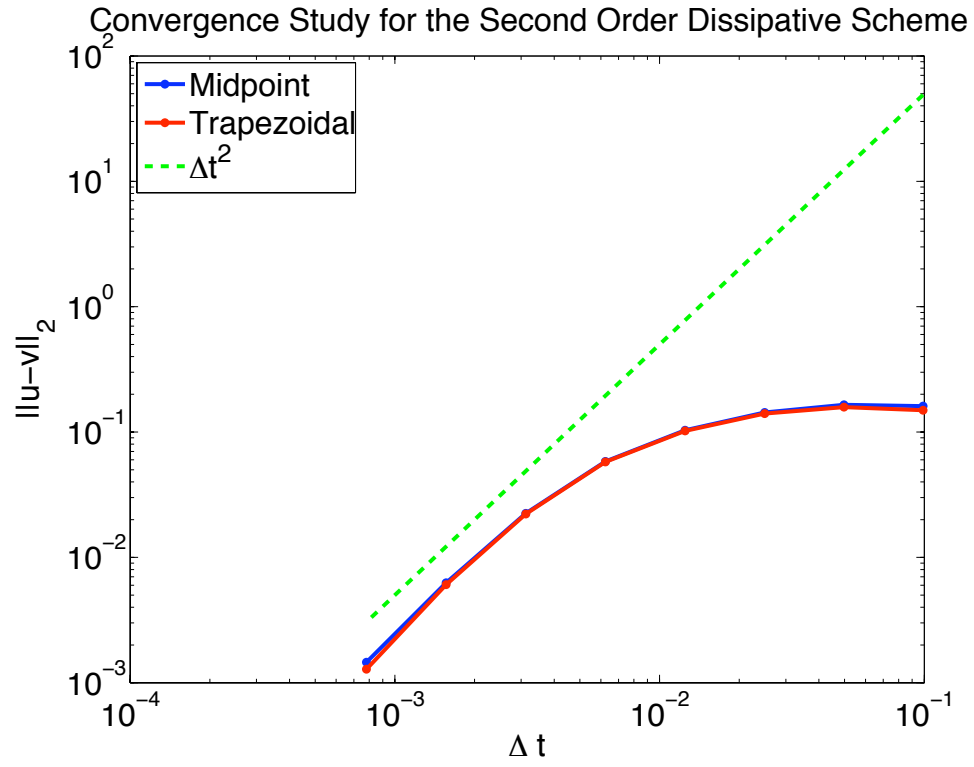


Figure 4.4: Convergence of the Second Order Dissipative Scheme with $\nu/\sqrt{2} = \Delta x/(c\Delta t) = 0.05$.

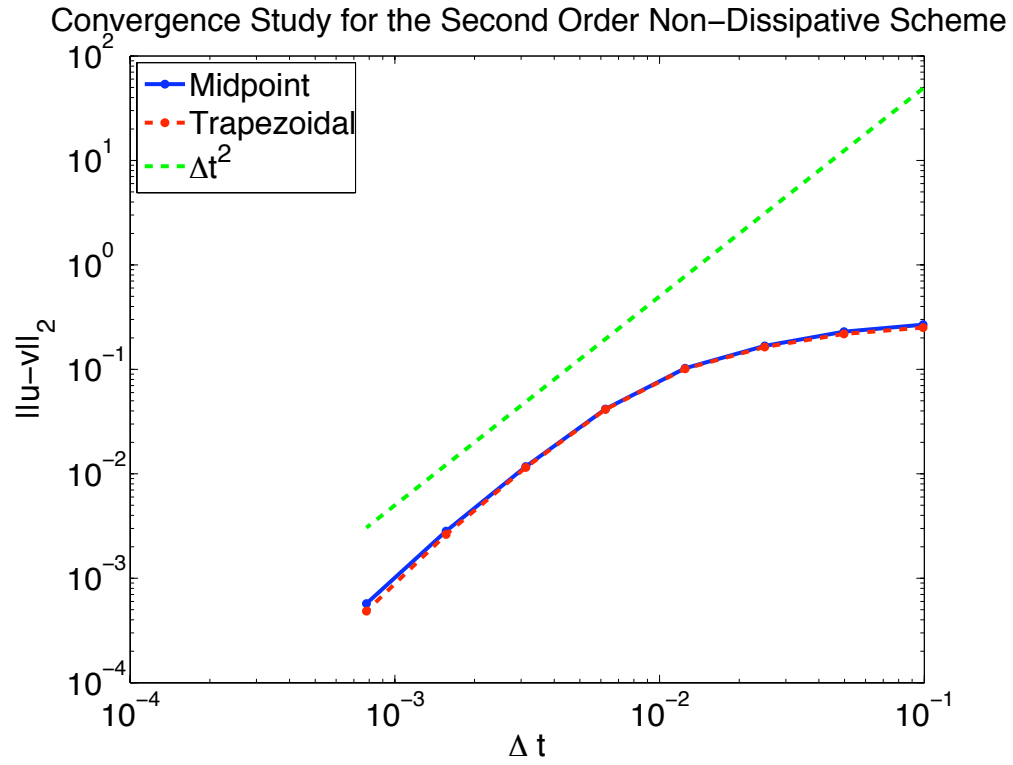


Figure 4.5: Convergence of the Second Order Non-dissipative Scheme with $\nu/\sqrt{2} = \Delta x/(c\Delta t) = 0.05$.

4.4 Fast Implementation

All of the structure underlying both the matrices A_p and A_h creates opportunities for either fast implementation or minimized storage. The matrix A_p (3.10), which corresponds to the free space solution, can be written as

$$A_p = \tilde{b}_0 I + \tilde{b}_1 (L + U),$$

where I is the standard identity matrix and L and U are strictly lower and upper triangular matrices respectively. The entries of L and U have been defined as

$$L_{ij} = U_{ji} = \begin{cases} 0, & j \geq i \\ e^{-|i-j|\nu}, & j < i. \end{cases}$$

Consider the matrix vector multiplication $A_p \cdot \bar{x}$ with $\bar{x} = [x_1, x_2, \dots, x_N]^T$. Then

$$A_p \cdot \bar{x} = \tilde{b}_1 L\bar{x} + \tilde{b}_1 U\bar{x} + \tilde{b}_0 I\bar{x}.$$

$\tilde{b}_0 I\bar{x}$ requires only that $\tilde{b}_0$ be stored and multiplied by each entry in $\bar{x}$, requiring precisely N operations.

Both $L\bar{x}$ and $U\bar{x}$ can be quickly solved in a manner akin to Horner's method. The structure of L will yield that

$$L\bar{x} = [0, e^{-\nu}x_1, e^{-\nu}x_2 + e^{-2\nu}x_1, e^{-\nu}x_3 + e^{-2\nu}x_2 + e^{-3\nu}x_1, \dots].$$

Starting with the second entry, each entry can be found recursively based solely on the previous entry using

$$L(x_1) = 0, \quad L(x_k) = e^{-\nu} * (L(x_{k-1}) + x_{k-1}), \quad k = \{2, \dots, N\}.$$

Similarly,

$$U(x_N) = 0, \quad U(x_{N-k}) = e^{-\nu} * (U(x_{N-k+1}) + x_{N-k+1}), \quad k = \{2, \dots, N\}.$$

Each of these matrix multiplications will require N multiplications and N additions. Therefore, the matrix vector multiplication $A_p \cdot \bar{x}$ requires storing only the quantities a_0, a_1 , and $e^{-\nu}$ as well as $O(N)$ computations.

The matrix A_h , corresponding to the boundary correction can also be stored as two $O(N)$ vectors as can easily be seen by equation (4.7). Additionally, since A_h may be defined as a constant multiplied by the summand of two outer products, $v \cdot h$ and $Jv \cdot Jh$, it may be implemented fast as well. Note that multiplying the outer product by another vector $\bar{x}$ gives, for example,

$$(v \cdot h)\bar{x} = v \cdot h \cdot \bar{x} = v \cdot (h \cdot \bar{x}) = (h \cdot \bar{x})v.$$

Here, $h \cdot \bar{x}$ is simply a dot product which requires only $2N - 1$ operations and that scalar is simply multiplied by v . Thus, the matrix vector multiplication for A_h will require this to be done twice, yielding a total of $O(N)$ operations.

Chapter 5

Non-Reflecting Boundary Conditions

A common problem that must be resolved when solving the wave equation is choosing how to truncate the domain for numerical computation. For a wave propagating on a finite domain, boundary conditions should be applied. While Dirichlet or Neumann conditions may be used, a more practical approach may be to create an artificial non-reflecting boundary which will allow the wave to flow out of the domain without re-entering [37]. One example of such an approach is to use a radiation condition at infinity, which states that the waves are outgoing. For $\mathcal{R}^d, d \in \{1, 2, 3\}$, the most well-known such condition is the Sommerfeld radiation condition [64], given by

$$\lim_{r \rightarrow \infty} r^{(d-1)/2} (u_r - iku) = 0, \quad (5.1)$$

where r is the radial coordinate, u is the scattered field and k is the wavenumber. For the case of one dimensional time harmonic waves, $u(x, t) = \tilde{u}(x)e^{-i\omega t}$, the Sommerfeld radiation

condition (5.1) will exactly reduce to

$$u_t \pm cu_x = 0, \quad (5.2)$$

where the sign depends on the direction the wave is traveling. Both the accuracy [35] and the stability [43] of solvers using various finite difference approximations implementing (5.2) have been studied. In one dimension still, applying either of (5.2) is analogous to implementing a one-way wave, that is, a wave that is allowed to propagate in only one direction [44, 68].

While a viable option in one dimension, radiation conditions may experience problems in higher dimensions, however. The implementation of such conditions in two or three dimensions has typically suffered from significant spurious reflections at the boundary [37]. Several attempts have been made to correct this, most notably [32, 33].

An alternative approach introduced by Berenger [15] is the idea of Perfectly Matched Layers (PML), specifically designed for the context of solving Maxwell's equation in two dimensions, wherein the computational domain is surrounded by a layer of absorbing material. Berenger split the magnetic field H_z into directional components H_{zx} and H_{zy} and then showed that if the electric conductivity σ and the magnetic conductivity σ^* satisfied

$$\frac{\sigma}{\epsilon_0} = \frac{\sigma^*}{\mu_0},$$

then the interface between the computational domain and the absorbing material would be perfectly matched allowing the wave to pass through into the absorbing medium without reflection. The wave would then decay exponentially in the absorbing layer. Berenger would later go on to adapt this method for the three dimensional case of Maxwell's equations [16].

This idea of a perfectly matched layer has since been extended to wave problems in other fields such as the linearized Euler equations [47], advective acoustics [1], shallow water waves [58], etc.

The PML method proved numerically advantageous over other such methods for truncating domains, such as radiating boundary conditions and absorbing boundary conditions. However, Abarbanel and Gottlieb [3] showed that Berenger’s initial PML implementation with field splitting was only weakly well-posed and thus when slightly perturbed, this numerical solution was subject to becoming divergent due to explosive modes. This splitting was later shown to be unnecessary by Zhao and Cangellaris [72]. Many other studies, including [2, 4, 7, 13, 12, 14, 30], have investigated the key properties of well-posedness and stability of PML in its various schemes.

Another key tenet of PML that has been frequently studied is the method’s efficiency. These implementations include further component splitting, coordinate stretching, the use of absorbing boundaries in the absorbing medium [17], reductions in number of auxiliary variables [42] and optimization of both the absorbing function and the absorbing material grid [9, 18, 26]. While significant time has been invested into finding an optimal absorbing function, the problem is still regarded as open.

5.1 Perfectly Matched Layers

Let $v(x, t)$ be a true solution to the 1D wave equation (1.1). We wish to compute a numerical solution $u(x, t)$ to this problem on the truncated computational domain $[-L, L]$ by introducing perfectly matched layers of length d . Boundary conditions at the end of the

perfectly matched layer may be enforced, but are unnecessary. Thus, we wish to compute

$$u_{tt} = c^2 u_{\tilde{x}\tilde{x}}, \quad t > 0, -L - d < \tilde{x} < L + d \quad (5.3)$$

$$u(x, 0) = f(x), \quad u_t(x, 0) = g(x)$$

$$u(\pm(L + d), t) = 0.$$

Here, we have chosen to enforce zero Dirichlet conditions at the end of the PML region for convenience.

The new variable $\tilde{x}$ will be defined using an analytic continuation of the wave equation and the Laplace transform. Let

$$\hat{u}(x, s) = \int_0^\infty e^{-st} u(x, t) dt$$

be the Laplace transform of the solution to equation (1.1). Now, define $\tilde{x}$ by the mapping

$$\frac{d\tilde{x}}{dx} = 1 + \frac{\sigma(x)}{s},$$

where $\sigma(x)$ is some spatially varying damping rate, which is non-decreasing into the PML, and nonzero only in the PML regions. Next, since

$$\frac{\partial u}{\partial \tilde{x}} = \frac{dx}{d\tilde{x}} \frac{\partial u}{\partial x} = \frac{s}{s + \sigma(x)} \frac{\partial u}{\partial x},$$

we can take the Laplace transform of equation (5.3) (with $f(x) = g(x) = 0$), and get

$$\begin{aligned}
s^2 \hat{u} &= c^2 \left(\frac{s}{s + \sigma} \partial_x \left(\frac{s}{s + \sigma} \hat{u}_x \right) \right) \\
\implies s^2 \hat{u} &= \frac{(sc)^2}{(s + \sigma)^2} \left(\hat{u}_{xx} - \frac{\sigma'}{s + \sigma} \hat{u}_x \right) \\
\implies \left(\frac{s + \sigma}{c} \right)^2 \hat{u} &= \hat{u}_{xx} - \hat{v}
\end{aligned}$$

where

$$\hat{v} = \frac{\sigma'}{s + \sigma} \hat{u}_x \implies (s + \sigma) \hat{v} = \sigma' \hat{u}_x.$$

Now, using the inverse Laplace transform on the expressions for u and v gives the following system of equations:

$$\frac{1}{c^2} \left(u_{tt} + 2\sigma u_t + \sigma^2 u \right) = u_{xx} - v, \quad (5.4)$$

$$v_t + \sigma v = \sigma' u_x. \quad (5.5)$$

Alternatively, we can write a second order integro-differential equation involving only u , by solving for v using an integrating factor. Elimination then gives

$$\frac{1}{c^2} \left(u_{tt} + 2\sigma u_t + \sigma^2 u \right) = u_{xx} - \sigma' \int_0^t u_x(x, s) e^{-\sigma(t-s)} ds. \quad (5.6)$$

5.1.1 Constant Damping Function

As a first attempt to understand this approach, consider the case where $\sigma(x)$ is a piecewise positive constant, with the wave equation satisfied for $x < L$, and the PML region for $x > L$; then

$$\sigma(x) = \begin{cases} 0, & x < L \\ \sigma_0 > 0, & x > L \end{cases}.$$

Then, $\sigma(x) = \sigma_0 H(x - L)$, where $H(x)$ is the Heaviside step function, and therefore

$$\sigma'(x) = \sigma_0 \delta(x - L). \quad (5.7)$$

5.1.1.1 Implementation

Making use of (5.7) in equation (5.6) produces

$$\frac{1}{c^2} \left(u_{tt} + 2\sigma u_t + \sigma^2 u \right) = u_{xx} - \sigma_0 \delta(x - L) \int_0^t u_x(x, s) e^{-\sigma(t-s)} ds.$$

The delta function will not prove problematic for my method of solution since we intend to integrate over space with the Green's function.

We now discretize equation (5.4) in time, using the standard first order discretization to get

$$\frac{1}{(c\Delta t)^2} \left(u^{n+1} - 2u^n + u^{n-1} + 2\sigma\Delta t \left(u^{n+1} - u^n \right) + (\sigma\Delta t)^2 u^{n+1} \right) = u_{xx}^{n+1} - v^{n+1}$$

or

$$\begin{aligned} & \frac{1}{(c\Delta t)^2} \left(2u^{n+1} - 5u^n + 4u^{n-1} - u^{n-2} \right. \\ & \quad \left. + 2\sigma\Delta t \left(\frac{3}{2}u^{n+1} - 2u^n + \frac{1}{2}u^{n-1} \right) + (\sigma\Delta t)^2 u^{n+1} \right) = u_{xx}^{n+1} - v^{n+1} \end{aligned}$$

by using the standard second order discretization.

The term v^{n+1} can be discretized in a number of ways. If we apply the integrating factor to equation (5.5) and integrate over $[t_n, t_{n+1}]$ using the trapezoidal rule, we get

$$v^{n+1} = \left(e^{-\sigma\Delta t} v^n + \delta(x-L) \frac{\sigma_0 \Delta t}{2} \left(u_x^{n+1} + e^{-\sigma\Delta t} u_x^n \right) \right).$$

Now, we can apply the Helmholtz operator, which gives

$$\mathcal{L} \left[u^{n+1} \right] (x) = f \left(u^{n+1}, u^n, u^{n-1}, v^{n+1} \right)$$

where

$$\mathcal{L}[u] := \left(\frac{\partial^2}{\partial x^2} - \alpha^2 \right) u$$

and α is dependent upon the discretization used.

The right hand side, depending on the desired order, is then given either by

$$\begin{aligned} f(u^{n+1}, u^n, u^{n-1}, v^{n+1}) &= -\alpha^2 (2u^n - u^{n-1} - 2zu^{n+1} + 2zu^n - z^2u^{n+1}) + v^{n+1} \text{ or} \\ f(u^{n+1}, u^n, u^{n-1}, u^{n-2}, v^{n+1}) &= \\ & -\frac{\alpha^2}{2} (5u^n - 4u^{n-1} + u^{n-2} - 3zu^{n+1} + 4zu^n - zu^{n-1} - z^2u^{n+1}) + v^{n+1} \end{aligned}$$

with

$$z = \sigma \Delta t.$$

The Green's function is given by

$$G(x|y) = \frac{-1}{2\alpha} e^{-\alpha|x-y|}$$

and upon inverting the Helmholtz operator, we have either

$$\begin{aligned} u^{n+1} &= \alpha \int_{-\infty}^{\infty} e^{-\alpha|x-y|} \left(u^n(y) - \frac{u^{n-1}(y)}{2} \right) dy + \alpha \int_L^{\infty} ze^{-\alpha|x-y|} u^n(y) dy \\ & - \alpha \int_L^{\infty} \left(\frac{z^2}{2} + z \right) e^{-\alpha|x-y|} u^{n+1}(y) dy + \tilde{v}^{n+1}(L) \end{aligned} \quad (5.8)$$

or

$$\begin{aligned} u^{n+1} &= \alpha \int_{-\infty}^{\infty} e^{-\alpha|x-y|} \left(\frac{5}{4}u^n(y) - u^{n-1}(y) + \frac{1}{4}u^{n-2}(y) \right) dy \\ & + \alpha \int_L^{\infty} e^{-\alpha|x-y|} \left(\left(-\frac{z^2}{4} - \frac{3z}{4} \right) u^{n+1}(y) + zu^n(y) - \frac{z}{4}u^{n-1}(y) \right) dy + \tilde{v}^{n+1}(L). \end{aligned} \quad (5.9)$$

Here

$$\tilde{v}^{n+1} = \frac{-1}{2\alpha} \int \delta(x-L) e^{-\alpha|x-y|} v^{n+1}(y) dy = \frac{-1}{2\alpha} e^{-\alpha|x-L|} v^{n+1}(L).$$

Hence, due to the delta function, the term v is localized to the space-PML interface.

Consider the discretization of equation (5.8) or (5.9) using the midpoint rule. Let u_j^n denote this fully discretized solution at $(x_{j-1/2}, t_n)$, with $x_{j-1/2} = -L + (j-1/2)\Delta x$. The gridpoints corresponding to $j = 1, 2, \dots, N$ are in the computational region and those corresponding to $j = N+1, N+2, \dots, N+M$ are in the PML region. The solution can then be written as

$$\begin{aligned} u_j^{n+1} = & \sum_{i=1}^{N+M} A_{ji} \left(u_i^n - \frac{u_i^{n-1}}{2} \right) - e^{-\alpha|x_{j-1/2}-x_N|} v^{n+1} \\ & + \sum_{i=N+1}^{N+m} A_{ji} \left(z u_i^n - \left(\frac{z^2}{2} + z \right) u_i^{n+1} \right) \end{aligned} \quad (5.10)$$

or

$$\begin{aligned} u_j^{n+1} = & \sum_{i=1}^{N+M} A_{ji} \left(\frac{5}{4} u_i^n - u_i^{n-1} + \frac{1}{4} u_i^{n-2} \right) - e^{-\alpha|x_{j-1/2}-x_N|} v^{n+1} \\ & + \sum_{i=N+1}^{N+m} A_{ji} \left(\left(-\frac{z^2}{4} - \frac{3z}{4} \right) u_i^{n+1} + z u_i^n - \frac{z}{4} u_i^{n-1} \right) \end{aligned} \quad (5.11)$$

with

$$v^{n+1} = e^{-z} v^n + \lambda \left(u_{N+1+p}^{n+1} - u_{N-p}^{n+1} + e^{-z} \left(u_{N+1+p}^n - u_{N-p}^n \right) \right).$$

The coefficient λ can be simplified using $\nu = \alpha\Delta x$, so that

$$\lambda = \frac{\sigma_0}{2\alpha} \frac{\Delta t}{2} \frac{1}{(2p+1)\Delta x} = \frac{z}{4(2p+1)\nu},$$

and p is any non-negative integer of my choice to represent the derivative at $x = L$. The matrix A is as given in section 2.3,

$$A_{ji} = 2 \left(1 - e^{-\nu/2}\right) \delta_{ji} + 2 \sinh\left(\frac{\nu}{2}\right) e^{-|i-j|\nu} (1 - \delta_{ji}), \quad 1 \leq j, i \leq N + M$$

in terms of the Kronecker delta δ_{ji} .

The important observation to be made here is that while the solution is given implicitly, these implicit terms arise solely in the PML region. This will allow the continued use of the fast summation methods used in the non-outflow algorithms, after further analysis of the implicit equations which consitute the subsystem, $N + 1 \leq j \leq N + M$.

5.1.1.2 Implicit Update

The matrix A decomposes as

$$A = (b_0 - b_1) I + b_1 T$$

where

$$b_0 = 2 \left(1 - e^{-\nu/2}\right), \quad b_1 = 2 \sinh\left(\frac{\nu}{2}\right)$$

and T is the Toeplitz matrix given in terms of $d = e^{-\nu}$ by

$$T = \begin{pmatrix} 1 & d & \dots & d^{N+M-1} \\ d & 1 & \dots & d^{N+M-2} \\ \vdots & \vdots & \ddots & \vdots \\ d^{N+M-1} & d^{N+M-2} & \dots & 1 \end{pmatrix},$$

whose inverse is given by the tridiagonal matrix with perturbed corners

$$T^{-1} = \frac{1}{1-d^2} \begin{pmatrix} 1 & -d & \dots & 0 \\ -d & 1+d^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \vdots & \ddots & 1 \end{pmatrix}.$$

Now, define

$$\bar{u} = (u_{N+1}, \dots, u_{N+M})^T,$$

then the subsystem corresponding to (5.10) or (5.11) can be rearranged and expressed by

$$\left(I + \left(\frac{z^2}{2} + z \right) A \right) \bar{u}^{n+1} = A \left((z+1) \bar{u}^n - \frac{1}{2} \bar{u}^{n-1} \right) - \bar{v}^{n+1} \quad (5.12)$$

or

$$\left(I + \left(\frac{z^2}{4} + \frac{3z}{4} \right) A \right) \bar{u}^{n+1} = A \left(\left(z + \frac{5}{4} \right) \bar{u}^n - \left(\frac{z}{4} + 1 \right) \bar{u}^{n-1} + \frac{1}{4} \bar{u}^{n-2} \right) - \bar{v}^{n+1}$$

with

$$\bar{v}_i = e^{-\alpha|x_{i-1/2}-x_N|} v^{n+1}, \quad N+1 \leq i \leq N+M.$$

We can rewrite the first order system as

$$\begin{aligned} \left(I + \left(\frac{z^2}{2} + z \right) A \right) \bar{u}^{n+1} &= A \left((z+1)\bar{u}^n - \frac{1}{2}\bar{u}^{n-1} \right) - \bar{v}^{n+1} \\ &= ((1+c_0)I + c_1T) \bar{u}^{n+1} = (c_0I + c_1T) \left((z+1)\bar{u}^n - \frac{1}{2}\bar{u}^{n-1} \right) - \bar{v}^{n+1} \end{aligned}$$

with

$$c_0 = (b_0 - b_1) \left(\frac{1}{2}z^2 + z \right), \quad c_1 = b_1 \left(\frac{1}{2}z^2 + z \right).$$

Multiply by T^{-1} to get

$$\left((1+c_0)T^{-1} + c_1I \right) \bar{u}^{n+1} = \left(c_0T^{-1} + c_1I \right) \left((z+1)\bar{u}^n - \frac{1}{2}\bar{u}^{n-1} \right) - T^{-1}\bar{v}^{n+1}.$$

From here, note that the matrix on the left hand side, $(1+c_0)T^{-1} + c_1I$ is tri-diagonal and thus this equation can be solved for $\bar{u}^{n+1}$ using a tri-diagonal solver.

An alternative approach is to bring all of the implicit terms to the left hand side of the equation. We begin with the breakdown of the temporal derivative as

$$\bar{v}^{n+1} = e^{-\alpha|x_{i-1/2}-x_N|} \left(e^{-z}v^n + \lambda \left(u_{N+1+p}^{n+1} - u_{N-p}^{n+1} + e^{-z} \left(u_{N+1+p}^n - u_{N-p}^n \right) \right) \right).$$

Rewrite this as

$$\bar{v}^{n+1} = \lambda ED \bar{u}^{n+1} + e^{-\alpha|x_{i-1/2}-x_N|} \left(e^{-z} v^n + \lambda \left(e^{-z} \left(u_{N+1+p}^n - u_{N-p}^n \right) \right) \right),$$

where E is the matrix with $e^{-\alpha|x_{i-1/2}-x_N|}$ on the i th diagonal entry and D is the matrix with -1 in the $(N-p)$ th column and 1 in the $(N+p+1)$ th column. Note that the matrix multiplication ED results in a matrix with only two non-zero columns.

We can then rewrite (5.12) as

$$\begin{aligned} & \left(I + \left(\frac{z^2}{2} + z \right) A + \lambda ED \right) \bar{u}^{n+1} = \\ & A \left((z+1) \bar{u}^n - \frac{1}{2} \bar{u}^{n-1} \right) - e^{-\alpha|x_{i-1/2}-x_N|} \left(e^{-z} v^n + \lambda \left(e^{-z} \left(u_{N+1+p}^n - u_{N-p}^n \right) \right) \right). \end{aligned}$$

Once again, reformulate A in terms of I and T as

$$\begin{aligned} ((1+c_0)I + c_1T + \lambda ED) \bar{u}^{n+1} &= (c_0I + c_1T) \left((z+1) \bar{u}^n - \frac{1}{2} \bar{u}^{n-1} \right) \\ &- e^{-\alpha|x_{i-1/2}-x_N|} \left(e^{-z} v^n + \lambda \left(e^{-z} \left(u_{N+1+p}^n - u_{N-p}^n \right) \right) \right) \end{aligned}$$

Multiply by T^{-1} as before to get

$$\begin{aligned} \left((1+c_0)T^{-1} + c_1I + \lambda T^{-1}ED \right) \bar{u}^{n+1} &= (c_0T^{-1} + c_1I) \left((z+1) \bar{u}^n - \frac{1}{2} \bar{u}^{n-1} \right) \\ &- T^{-1} e^{-\alpha|x_{i-1/2}-x_N|} \left(e^{-z} v^n + \lambda \left(e^{-z} \left(u_{N+1+p}^n - u_{N-p}^n \right) \right) \right) \end{aligned}$$

It must be noted here that, due to the $\lambda T^{-1}ED$ term, the left hand side is no longer tri-diagonal. However, since T^{-1} is tri-diagonal and ED is a matrix with only two non-zero

columns, $\lambda T^{-1}ED$ will also be a tridiagonal matrix with two additional non-zero columns. Thus, the left hand side will still be sparse. Furthermore, this has the benefit of creating a right hand side that is truly explicit as it depends solely upon values from previous time steps.

The second order system may be solved in a very similar way, resulting in either the tri-diagonal system

$$\begin{aligned}
& \left(I + \left(\frac{z^2}{4} + \frac{3z}{4} \right) A \right) \bar{u}^{n+1} = A \left(\left(z + \frac{5}{4} \right) \bar{u}^n - \left(\frac{z}{4} + 1 \right) \bar{u}^{n-1} + \frac{1}{4} \bar{u}^{n-2} \right) - \bar{v}^{n+1} \\
\Rightarrow & ((1 + c_0)I + c_1 T) \bar{u}^{n+1} = \\
& (c_0 I + c_1 T) \left(\left(z + \frac{5}{4} \right) \bar{u}^n - \left(\frac{z}{4} + 1 \right) \bar{u}^{n-1} + \frac{1}{4} \bar{u}^{n-2} \right) - \bar{v}^{n+1} \\
\Rightarrow & ((1 + c_0)T^{-1} + c_1 I) \bar{u}^{n+1} = \\
& (c_0 T^{-1} + c_1 I) \left(\left(z + \frac{5}{4} \right) \bar{u}^n - \left(\frac{z}{4} + 1 \right) \bar{u}^{n-1} + \frac{1}{4} \bar{u}^{n-2} \right) - T^{-1} \bar{v}^{n+1}.
\end{aligned}$$

Again, we have the option to remove all the implicit terms from the right hand side with

$$\begin{aligned}
& \left((1 + c_0)T^{-1} + c_1 I + \lambda T^{-1}ED \right) \bar{u}^{n+1} = \\
& (c_0 T^{-1} + c_1 I) \left(\left(z + \frac{5}{4} \right) \bar{u}^n - \left(\frac{z}{4} + 1 \right) \bar{u}^{n-1} + \frac{1}{4} \bar{u}^{n-2} \right) \\
& - T^{-1} e^{-\alpha |x_{i-1/2} - x_N|} \left(e^{-z} v^n + \lambda \left(e^{-z} \left(u_{N+1+p}^n - u_{N-p}^n \right) \right) \right).
\end{aligned}$$

Once again, the left hand side is now a tri-diagonal matrix with two additional non-zero columns for every change in σ_0 .

For the second order non-dissipative MOL^T formulation, (5.6) becomes

$$\begin{aligned} & \frac{1}{(c\Delta t)^2} \left(u^{n+1} - 2u^n + u^{n-1} + \sigma\Delta t \left(u^{n+1} - u^{n-1} \right) + (\sigma\Delta t)^2 \left(\frac{u^{n+1} + u^{n-1}}{2} \right) \right) \\ &= \frac{u_{xx}^{n+1} + u_{xx}^{n-1}}{2} - v^{n+1}. \end{aligned}$$

We treat v in exactly the same fashion as above in the dissipative solvers. After reformulating this as a Helmholtz equation and inverting the operator, this leads to

$$\begin{aligned} u^{n+1} + u^{n-1} &= \alpha \int_{-\infty}^{\infty} e^{-\alpha|x-y|} u^n(y) dy + \alpha \int_L^{\infty} 2ze^{-\alpha|x-y|} u^{n-1}(y) dy \\ &\quad - \alpha \int_L^{\infty} (z^2 + z) e^{-\alpha|x-y|} \left(u^{n+1}(y) + u^{n-1}(y) \right) dy + \tilde{v}^{n+1}(L), \end{aligned}$$

with

$$z = \frac{\sigma\Delta t}{2}.$$

This gives

$$\begin{aligned} u_j^{n+1} + u_j^{n-1} &= \sum_{i=1}^{N+M} A_{ji} u_i^n - e^{-\alpha|x_{j-1/2}-x_N|} v^{n+1} \\ &\quad + \sum_{i=N+1}^{N+m} A_{ji} \left(2zu_i^n - (z^2 + z) \left(u_i^{n+1} + u_i^{n-1} \right) \right), \\ v^{n+1} &= e^{-2z} v^n + \lambda \left(u_{N+1+p}^{n+1} - u_{N-p}^{n+1} + e^{-2z} \left(u_{N+1+p}^n - u_{N-p}^n \right) \right) \end{aligned}$$

where

$$\lambda = \frac{\sigma_0}{\alpha} \frac{\Delta t}{2(2p+1)\Delta x} = \frac{z}{(2p+1)\nu}.$$

The subsection corresponding to the PML region can then be expressed as

$$\begin{aligned} \left(I + \left(z^2 + z \right) A \right) \left(\bar{u}^{n+1} + \bar{u}^{n-1} \right) &= A \left(\bar{u}^n + 2z\bar{u}^{n-1} \right) - \bar{v}, \\ \bar{v}_i &= e^{-\alpha |x_{i-1/2} - x_N|} v^{n+1}, \quad N+1 \leq i \leq N+M. \end{aligned}$$

Letting $c_0 = (b_0 - b_1)(z^2 + z)$ and $c_1 = b_1(z^2 + z)$, this now becomes

$$((1 + c_0)I + c_1T) \left(\bar{u}^{n+1} + \bar{u}^{n-1} \right) = (c_0I + c_1T) \left(\bar{u}^n + 2z\bar{u}^{n-1} \right) - \bar{v}.$$

Multiply both sides by T^{-1} to get the tri-diagonal system

$$\left((1 + c_0)T^{-1} + c_1I \right) \left(\bar{u}^{n+1} + \bar{u}^{n-1} \right) = \left(c_0T^{-1} + c_1I \right) \left(\bar{u}^n + 2z\bar{u}^{n-1} \right) - T^{-1}\bar{v}.$$

Furthermore, we can obtain the system with only explicit terms on the right hand side

as

$$\begin{aligned} \left(I + \left(z^2 + z \right) A + \lambda ED \right) \left(\bar{u}^{n+1} + \bar{u}^{n-1} \right) - \lambda ED \bar{u}^{n-1} = \\ A \left(\bar{u}^n + 2z\bar{u}^{n-1} \right) + e^{-\alpha |x_{i-1/2} - x_N|} \left(e^{-2z} v^n + \lambda \left(e^{-2z} \left(u_{N+1+p}^n - u_{N-p}^n \right) \right) \right). \end{aligned}$$

This implies that

$$\begin{aligned} ((1 + c_0) I + c_1 T + \lambda ED) \left(\bar{u}^{n+1} + \bar{u}^{n-1} \right) &= \lambda ED \bar{u}^{n-1} + (c_0 I + c_1 T) \left(\bar{u}^n + 2z \bar{u}^{n-1} \right) \\ &+ e^{-\alpha |x_{i-1/2} - x_N|} \left(e^{-2z} v^n + \lambda \left(e^{-2z} \left(u_{N+1+p}^n - u_{N-p}^n \right) \right) \right). \end{aligned}$$

Multiplication by T^{-1} then gives

$$\begin{aligned} \left((1 + c_0) T^{-1} + c_1 I + \lambda T^{-1} ED \right) \left(\bar{u}^{n+1} + \bar{u}^{n-1} \right) &= \\ \lambda T^{-1} ED \bar{u}^{n-1} + \left(c_0 T^{-1} + c_1 I \right) \left(\bar{u}^n + 2z \bar{u}^{n-1} \right) \\ &+ T^{-1} e^{-\alpha |x_{i-1/2} - x_N|} \left(e^{-2z} v^n + \lambda \left(e^{-2z} \left(u_{N+1+p}^n - u_{N-p}^n \right) \right) \right). \end{aligned}$$

5.1.1.3 Performance

While using a constant damping factor $\sigma(x) = \sigma_0 H(x - L)$ may be fairly easy to implement, it does not necessarily provide a good method for truncating the domain. The biggest issue is that, due to the Heaviside function, $\sigma(x)$ is not smooth when switching from the interior of the domain to the PML region and thus, the regions are no longer ‘perfectly’ matched. This break in continuity may cause reflections of the wave as it passes the boundary into the absorbing layer. To check for these reflections and to quantify the damping potential of a piece-wise constant damping function, we start with an initial Gaussian pulse and let it propagate for a full pass through the domain with a perfectly matched region on the right side and see what percentage of the initial function still remains. The parameters were:

- Initial conditions are $f(x) = e^{-100(x-0.5)^2}$, $g(x) = 0$.
- The mesh sizes are $\Delta x = \Delta t = 0.1$ with a wave speed of $c = 1$.

- The interior domain is given by $[-1, 1]$.
- The damping layer is $[1, 5]$.
- The initial damping constant was given by one of $\{0.1, 0.01, 0.001, 0.0001\}$.
- The damping constant was doubled halfway through damping region.

At time $t = 1$, a check is performed on both $\|u(x)\|_2$ and $\|u(x)\|_\infty$ in the range $[0, 1]$. At this point, it is safe to assume that the right bound wave has entered the PML region while the left bound wave has not yet reached the matched boundary. This will allow us to compare the spurious numerical reflection coming off of the boundary from the right bound wave caused by the discontinuity in the damping function. Finally, at a time of $t = 12$, which corresponds to the time necessary for the wave to make one full pass through the domain, we once again check both $\|u(x)\|_2$ and $\|u(x)\|_\infty$ to see how much of the initial pulse remains. The results are shown in tables [5.1](#) and [5.2](#).

As can easily be seen from the tables, the reflection caused by the discontinuity in the damping function may eliminate much of the benefit that comes from the actual damping. In addition, the actual damping itself does not perform well, removing no more than 70% of the wave despite the damping region taking up well over half of the domain.

Table 5.1: Reflection and Damping for the Dissipative 1st Order Scheme with a Piecewise Constant Damping Function

σ_0	∞ -norm at $t = 1$	2-norm at $t = 1$	∞ -norm at $t = 12$	2-norm at $t = 12$
0	0.000014176	0.000022378	0.17580	1.48488
0.0001	0.000016310	0.000029892	0.17559	1.48313
0.001	0.000035517	0.00018632	0.17374	1.46747
0.01	0.00022812	0.0018358	0.15625	1.31976
0.1	0.0022091	0.018325	0.054052	0.45717

Table 5.2: Reflection and Damping for the Dissipative 2nd Order Scheme with a Piecewise Constant Damping Function

σ_0	∞ -norm at $t = 1$	2-norm at $t = 1$	∞ -norm at $t = 12$	2-norm at $t = 12$
0	0.00000098983	0.0000013516	0.35042	2.44360
0.0001	0.0000039184	0.000020798	0.35001	2.44077
0.001	0.000030283	0.00020579	0.34638	2.41540
0.01	0.00029472	0.0020593	0.31205	2.17588
0.1	0.0030214	0.020930	0.11020	0.76839

5.1.2 Polynomial Damping Functions

To correct the matching deficiencies of the piece-wise constant damping functions from the previous section, consider damping functions of the form

$$\sigma(x) = \sigma_0 \left(\frac{x - L}{d} \right)^m, \quad m \geq 2.$$

We make use of a Heaviside function to extend this to the entire domain as

$$\sigma(x) = \sigma_0 \left(\frac{x - L}{d} \right)^m H(x - L), \quad m \geq 2. \quad (5.13)$$

These polynomial damping functions will ensure that $\sigma(x) \in \mathbb{C}^m$.

5.1.2.1 Implementation

We now wish to create a fast implementation of equations (5.4) and (5.5) incorporating (5.13). We first note that by use of an integrating factor, the solution to (5.5) is given by

$$\begin{aligned}
v(x, t) &= \sigma'(x) \int_0^t e^{-\sigma(x)(t-s)} u_x(x, s) ds \\
\Rightarrow v(x, t^{n+1}) &= \sigma'(x) \int_0^{t^{n+1}} e^{-\sigma(x)(t^{n+1}-s)} u_x(x, s) ds \\
\Rightarrow v(x, t^{n+1}) &= \sum_{m=0}^n \sigma'(x) \int_{m\Delta t}^{(m+1)\Delta t} e^{-\sigma(x)((n+1)\Delta t-s)} u_x(x, s) ds.
\end{aligned}$$

We then approximate each integral using trapezoidal quadrature and arrive at

$$\begin{aligned}
v^{n+1} &= \frac{\sigma'(x)\Delta t}{2} \left(\sum_{m=0}^n u_x(x, m\Delta t) e^{\sigma(x)(m-n-1)\Delta t} \right. \\
&\quad \left. + u_x(x, (m+1)\Delta t) e^{\sigma(x)(m-n)\Delta t} \right).
\end{aligned} \tag{5.14}$$

Applying the standard first order discretization to equation (5.4) gives

$$\begin{aligned}
&\frac{(1 + 2\sigma(x)\Delta t + \sigma(x)^2\Delta t^2) u^{n+1}(x)}{\Delta t^2} - \frac{2(1 + \Delta t\sigma(x))u^n}{\Delta t^2} + \frac{u^{n-1}(x)}{\Delta t^2} \\
&= c^2 u_{xx}^{n+1}(x) - \sigma'(x) \int_0^{t^{n+1}} u_x(x) e^{-\sigma(x)(t-s)}.
\end{aligned} \tag{5.15}$$

Combining equations (5.15) and (5.14) produces the following scheme which we would

like to solve quickly:

$$\begin{aligned} & \frac{(1 + 2\sigma(x)\Delta t + \sigma(x)^2\Delta t^2) u^{n+1}(x)}{\Delta t^2} - \frac{2(1 + \Delta t\sigma(x))u^n}{\Delta t^2} + \frac{u^{n-1}(x)}{\Delta t^2} = c^2 u_{xx}^{n+1}(x) \\ & - \frac{\sigma'(x)\Delta t}{2} \sum_{m=0}^n u_x(x, m\Delta t) e^{\sigma(x)(m-n-1)\Delta t} + u_x(x, (m+1)\Delta t) e^{\sigma(x)(m-n)\Delta t}. \end{aligned}$$

This then may be re-arranged into a Helmholtz equation as

$$(\partial_{xx} - \alpha^2) u^{n+1} = \alpha^2 \left((2\sigma\Delta t + \sigma^2\Delta t^2) u^{n+1} - 2(1 + \sigma\Delta t)u^n + u^{n-1} \right) + v^{n+1}.$$

An inversion of the Helmholtz operation then yields

$$\begin{aligned} u^{n+1} = & -\frac{\alpha}{2} \int e^{-\alpha|x-y|} \left((2\sigma(y)\Delta t + \sigma^2(y)\Delta t^2) u^{n+1}(y) \right. \\ & \left. - 2(1 + \sigma(y)\Delta t)u^n(y) + u^{n-1}(y) \right) dy + \tilde{v}^{n+1} \end{aligned} \quad (5.16)$$

with

$$\begin{aligned} \tilde{v}^{n+1} = & -\frac{1}{2\alpha} \int e^{-\alpha|x-y|} \frac{\sigma'(y)\Delta t}{2} \\ & \cdot \sum_{m=0}^n \left(u_y(y, m\Delta t) e^{\sigma(y)(m-n-1)\Delta t} + u_y(y, (m+1)\Delta t) e^{\sigma(y)(m-n)\Delta t} \right) dy. \end{aligned} \quad (5.17)$$

Define

$$\omega_0(y) = \frac{\sigma'(y)\Delta t}{4\alpha} e^{-\alpha|x-y|}, \quad \omega_{k+1} = e^{-\sigma(y)\Delta t} \omega_k(y)$$

such that integration by parts on (5.17) will yield

$$\begin{aligned}
\tilde{v}^{n+1} &= -\frac{1}{4\alpha} \sum_{m=0}^n \int_L^\infty \sigma'(y) \Delta t e^{-\alpha|x-y|-(n-m)\Delta t\sigma(y)} \left(u_y^{m+1}(y) + e^{-\Delta t\sigma(y)} u_y^m(y) \right) dy \\
&= -\sum_{m=0}^n \int_L^\infty \omega_{n-m}(y) u_y^{m+1}(y) + \omega_{n+1-m}(y) u_y^m(y) dy \\
&= -\sum_{m=0}^n \left[\omega_{n-m}(y) u^{m+1}(y) + \omega_{n+1-m}(y) u^m(y) \right]_L^\infty \\
&\quad + \sum_{m=0}^n \int_L^\infty \omega'_{n-m}(y) u^{m+1}(y) + \omega'_{n+1-m}(y) u^m(y) dy.
\end{aligned}$$

The left boundary term will vanish due to the stipulation that the minimum order of our polynomial was at least 2, which guarantees $\omega_k(L) = 0, k \in \{0, 1, \dots, n\}$. Additionally, the right boundary term will vanish due to the negative exponential. This leaves that

$$\tilde{v}^{n+1} = \sum_{m=0}^n \int_L^\infty \omega'_{n-m}(y) u^{m+1}(y) + \omega'_{n+1-m}(y) u^m(y) dy.$$

Furthermore, for a truncated domain, we can ensure the right boundary term disappears by enforcing zero Dirichlet boundary conditions at the end of the damping region.

Using the same techniques as in section 5.1.1.1, we now consider the fully discretized form of equation (5.16), given by

$$\begin{aligned}
u^{n+1}(x_j) &= \sum_{i=1}^{N+M} A_{ji} \left(u_i^n - \frac{1}{2} u_i^{n-1} \right) \\
&\quad - \sum_{i=N+1}^{N+m} \left(\Sigma_{ji}^{(1)} u_i^{n+1} + \Sigma_{ji}^{(2)} \frac{1}{2} u^{n+1} - \Sigma_{ji}^{(1)} u_i^n \right) + \tilde{v}_j^{n+1},
\end{aligned} \tag{5.18}$$

where

$$\begin{aligned}\Sigma_{ji}^{(1)} &= \alpha \int_{x_{i-1}}^{x_i} \sigma(y) \Delta t e^{-\alpha |x_{j-1/2} - y|} dy \text{ and} \\ \Sigma_{ji}^{(2)} &= \alpha \int_{x_{i-1}}^{x_i} (\sigma(y) \Delta t)^2 e^{-\alpha |x_{j-1/2} - y|} dy,\end{aligned}\tag{5.19}$$

which are defined for $1 \leq i, j \leq N + M$, but noting that $\sigma(y) = 0$ outside of the PML region.

Further, we can simplify the term $\tilde{v}_j^{n+1}$. After application of the midpoint rule on the u terms, we have

$$\begin{aligned}\tilde{v}_j^{n+1} &= \sum_{m=0}^n \sum_{i=N+1}^{N+M} \int_{L+(i-1)\Delta x}^{L+i\Delta x} \omega'_{n-m}(y) u_i^{m+1} + \omega'_{n+1-m}(y) u_i^m dy \\ &= \sum_{m=0}^n \sum_{i=N+1}^{N+M} [\omega_{n-m}(y)]_{x_{i-1}}^{x_i} u_i^{m+1} + [\omega_{n+1-m}(y)]_{x_{i-1}}^{x_i} u_i^m.\end{aligned}$$

This implies that

$$\begin{aligned}\tilde{v}_j^{n+1} &= \sum_{i=N+1}^{N+M} \sum_{m=0}^n \left[\frac{\sigma'(y) \Delta t}{4\alpha} e^{-\alpha |x_{j-1/2} - y|} e^{-(n-m)\sigma(y)\Delta t} \right]_{x_{i-1}}^{x_i} u_i^{m+1} \\ &\quad + \sum_{i=N+1}^{N+M} \sum_{m=0}^n \left[\frac{\sigma'(y) \Delta t}{4\alpha} e^{-\alpha |x_{j-1/2} - y|} e^{-(n+1-m)\sigma(y)\Delta t} \right]_{x_{i-1}}^{x_i} u_i^m \\ &= \sum_{i=N+1}^{N+M} \frac{\sigma'(x_i) \Delta t}{4\alpha} e^{-\alpha |x_{j-1/2} - x_i|} \\ &\quad \cdot \sum_{m=0}^n e^{-(n-m)\sigma(x_i)\Delta t} \left(u^{m+1}(x_i) + e^{-\sigma(x_i)\Delta t} u^m(x_i) \right) \\ &\quad - \frac{\sigma'(x_{i-1}) \Delta t}{4\alpha} e^{-\alpha |x_{j-1/2} - x_{i-1}|} \\ &\quad \cdot \sum_{m=0}^n e^{-(n-m)\sigma(x_{i-1})\Delta t} \left(u^{m+1}(x_i) + e^{-\sigma(x_{i-1})\Delta t} u^m(x_i) \right).\end{aligned}$$

Therefore,

$$\begin{aligned}\tilde{v}_j^{n+1} &= \sum_{i=N+1}^{N+M} e^{-\alpha|x_{j-1/2}-x_i|} v_{1,i}^{n+1} - e^{-\alpha|x_{j-1/2}-x_{i-1}|} v_{2,i}^{n+1} \\ &= \sum_{i=N+1}^{N+M} B_{ji}^{(1)} v_{1,i}^{n+1} - B_{ji}^{(2)} v_{2,i}^{n+1}\end{aligned}$$

where

$$B_{ji}^{(1)} = e^{-\nu|j-i-1/2|}, \quad B_{ji}^{(2)} = e^{-\nu|j-i+1/2|},$$

and the sums are defined by the auxiliary variables

$$\begin{aligned}v_{1,i}^{n+1} &= \frac{\sigma'(x_i)\Delta t}{4\alpha} \sum_{m=0}^n \left(u^{m+1}(x_i) + e^{-\sigma(x_i)\Delta t} u^m(x_i) \right), \\ v_{2,i}^{n+1} &= \frac{\sigma'(x_{i-1})\Delta t}{4\alpha} \sum_{m=0}^n \left(u^{m+1}(x_i) + e^{-\sigma(x_{i-1})\Delta t} u^m(x_i) \right).\end{aligned}$$

We can simplify this further by using a recurrence relation to remove the sum. Thus, the auxiliary variables can be updated locally in time by

$$v_{1,i}^{n+1} = E_i v_{1,i}^n + F_i \left(u_i^{n+1} + E_i u_i^n \right), \quad v_{2,i}^{n+1} = E_{i-1} v_{2,i}^n + F_{i-1} \left(u_i^{n+1} + E_{i-1} u_i^n \right) \quad (5.20)$$

where

$$E_i = e^{-\sigma(x_i)\Delta t}, \quad F_i = \frac{\sigma'(x_i)\Delta t}{4\alpha}, \quad N \leq i \leq N+M. \quad (5.21)$$

Thus, for each time step, we update the equations (5.20) as well as solving for equation

(5.18) in the following form

$$\begin{aligned}
u^{n+1}(x_j) = & \sum_{i=1}^{N+M} A_{ji} \left(u_i^n - \frac{1}{2} u_i^{n-1} \right) - \sum_{i=N+1}^{N+m} \left(\Sigma_{ji}^{(1)} u_i^{n+1} + \Sigma_{ji}^{(2)} \frac{1}{2} u^{n+1} - \Sigma_{ji}^{(1)} u_i^n \right) \quad (5.22) \\
& + \sum_{i=N+1}^{N+m} B_{ji}^{(1)} \left(E_i v_{1,i}^n + F_i \left(u_i^{n+1} + E_i u_i^n \right) \right) \\
& - \sum_{i=N+1}^{N+m} B_{ji}^{(2)} \left(E_{i-1} v_{2,i}^n + F_{i-1} \left(u_i^{n+1} + E_{i-1} u_i^n \right) \right).
\end{aligned}$$

We now specifically consider the damping function

$$\sigma(x) = \sigma_0 \left(\frac{x - L}{d} \right)^2$$

for the purpose of fast implementation of this scheme. The techniques that will follow can be adapted in a very straight-forward manner for polynomials of higher order. However, we have chosen to use a polynomial of degree 2 as it will be the easiest to implement as well as performing the best as a damping function as can be seen in figure 5.1.

Since we have assumed the damping region to be M grid points wide, we have $d = M\Delta x$ and can write

$$s_i = \sigma(x_i)\Delta t = \sigma_0\Delta t \left(\frac{i - N}{M} \right)^2, \quad N \leq i \leq N + M.$$

Similarly,

$$s'_i = \sigma'(x_i)\Delta t = 2\sigma_0\Delta t \left(\frac{i - N}{M} \right), \quad N \leq i \leq N + M.$$

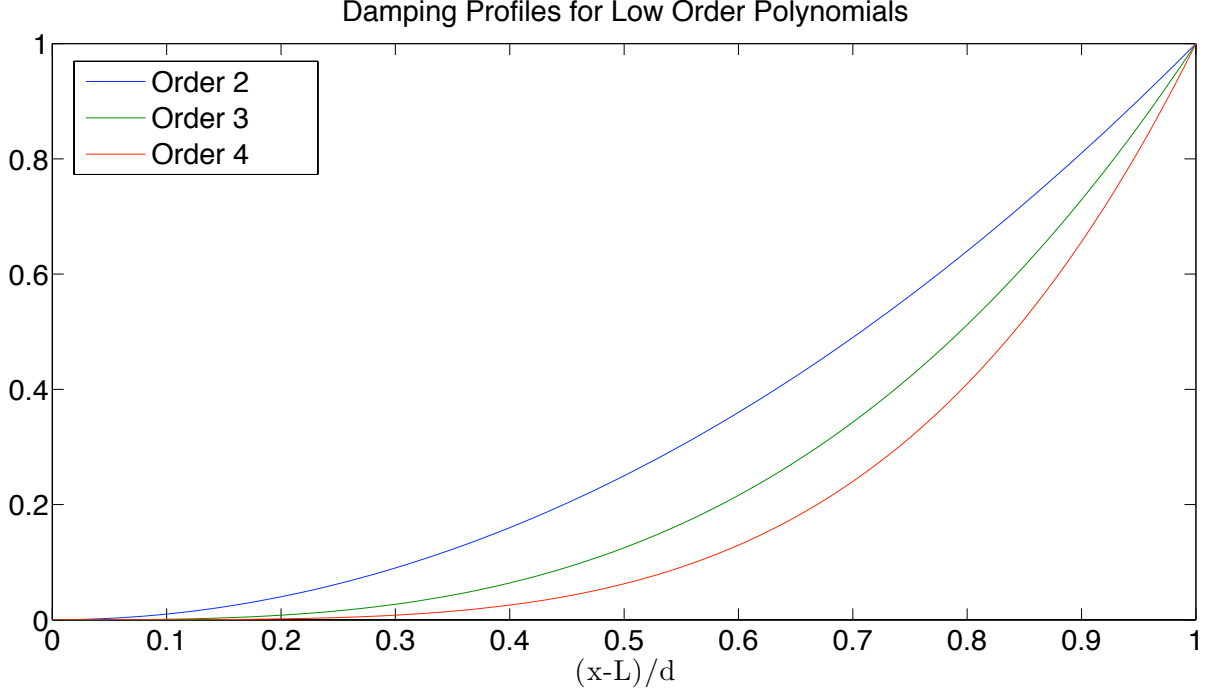


Figure 5.1: Damping Profiles for Low Order Polynomials

From these expressions, we can now evaluate (5.19)

$$\begin{aligned}\Sigma_{ji}^{(1)} &= \alpha \sigma_0 \Delta t \int_{x_{i-1}}^{x_i} \left(\frac{y-L}{d} \right)^2 e^{-\alpha |x_{j-1/2}-y|} dy \\ &= \nu \sigma_0 \Delta t \int_{-1/2}^{1/2} \left(\frac{i-1/2-N+z}{M} \right)^2 e^{-\nu |j-i-z|} dz\end{aligned}$$

using the change of variables $z = (y - x_{i-1/2})/dx$. This then becomes

$$\Sigma_{ji}^{(1)} = \nu \frac{\sigma_0 \Delta t}{M^2} \int_{-1/2}^{1/2} \left(z^2 + (i-N-1/2)^2 + (2i-2N-1)z \right) e^{-\nu |j-i-z|} dz.$$

Note that $\Sigma^{(1)}$ can be decomposed into terms of the form

$$G_{ji}^{(k)} = \nu \int_{-1/2}^{1/2} z^k e^{-\nu |j-i-z|} dz.$$

We now evaluate these expressions for each value of k and the cases $i = j$ and $i \neq j$.

When $k = 0$, we recover

$$G_{ji}^{(0)} = A_{ji} = 2 \left(1 - e^{-\nu/2} \right) \delta_{ji} + 2 \sinh \left(\frac{\nu}{2} \right) e^{-|i-j|\nu} (1 - \delta_{ji}).$$

More generally, for the diagonal terms and arbitrary k ,

$$\begin{aligned} G_{ii}^{(k)} &= \nu \int_{-1/2}^{1/2} z^k e^{-\nu|z|} dz \\ &= \begin{cases} 0, & k \text{ odd} \\ \frac{2}{\nu^k} \int_0^{\nu/2} w^k e^{-w} dw, & k \text{ even} \end{cases} \\ &= \begin{cases} 0, & k \text{ odd} \\ \frac{2k!}{\nu^k} \left(1 - b_k \left(\frac{\nu}{2} \right) e^{-\nu/2} \right) & k \text{ even,} \end{cases} \end{aligned}$$

where

$$b_k(z) = \sum_{\ell=0}^k \frac{z^\ell}{\ell!}.$$

Note that $b_k \left(\frac{\nu}{2} \right)$ is exactly the Taylor polynomial for $e^{\nu/2}$ of order k , so that $G_{ii}^{(k)}$ is well defined as ν tends to 0.

For $j \neq i$, we similarly have

$$\begin{aligned}
G_{ji}^{(k)} &= \nu e^{-|j-i|\nu} \int_{-1/2}^{1/2} z^k e^{-\nu z} dz \\
&= e^{-|j-i|\nu} \frac{1}{\nu^k} \int_{-\nu/2}^{\nu/2} w^k e^{-\operatorname{sgn}(j-i)w} dw \\
&= e^{-|j-i|\nu} \frac{k!}{\nu^k} \left(\operatorname{sgn}(j-i)^k b_k \left(-\frac{\nu}{2} \right) e^{\nu/2} - b_k \left(\frac{\nu}{2} \right) e^{-\nu/2} \right) \\
&= e^{-|j-i|\nu} (\operatorname{sgn}(j-i))^k \frac{2k!}{\nu^k} \left(c_k \left(\frac{\nu}{2} \right) \sinh \left(\frac{\nu}{2} \right) - s_k \left(\frac{\nu}{2} \right) \cosh \left(\frac{\nu}{2} \right) \right),
\end{aligned}$$

where $c_k(z)$ and $s_k(z)$ are the even and odd terms, respectively, of $b_k(z)$. Because of the sign change due to odd values of k , it will be useful to define the upper and lower triangular Toeplitz matrices U and $L = U^T$, where

$$U = \begin{pmatrix} 0 & d & d^2 & \dots & d^{N+M-1} \\ & 0 & d & \ddots & \vdots \\ & & \ddots & \ddots & \vdots \\ & & & 0 & d \\ & & & & 0 \end{pmatrix}$$

with $d = e^{-\nu}$. Further, we have the Toeplitz matrix

$$T^{-1} = I + U + L = \begin{pmatrix} 1 & d & \dots & \dots & d^{N+M-1} \\ d & 1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 & d \\ d^{N+M-1} & \dots & \dots & d & 1 \end{pmatrix}$$

which is the inverse of the tridiagonal matrix

$$T = \frac{1}{1-d^2} \begin{pmatrix} 1 & -d & & & \\ -d & 1+d^2 & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & 1+d^2 & -d \\ & & & -d & 1 \end{pmatrix}.$$

Assembling all of this, we can now write

$$G^{(k)} = \begin{cases} g_{k0}I + g_{k1}(U + L), & k \text{ even} \\ g_{k1}(U - L), & k \text{ odd} \end{cases}$$

where

$$g_{k0} = \frac{2k!}{\nu^k} \left(1 - b_k \left(\frac{\nu}{2} \right) e^{-\nu/2} \right), \quad g_{k1} = \frac{2k!}{\nu^k} \left(c_k \left(\frac{\nu}{2} \right) \sinh \left(\frac{\nu}{2} \right) - s_k \left(\frac{\nu}{2} \right) \cosh \left(\frac{\nu}{2} \right) \right).$$

Having fully determined expressions for $G^{(k)}$, we now express

$$\begin{aligned} \Sigma^{(1)} &= \frac{\sigma_0 \Delta t}{M^2} \left(G^{(2)} + 2G^{(1)}D + G^{(0)}D^2 \right) \\ &= \frac{\sigma_0 \Delta t}{M^2} \left(g_{20}I + g_{00}D^2 + U \left(g_{21}I + 2g_{11}D + g_{01}D^2 \right) + L \left(g_{21}I - 2g_{11}D + g_{01}D^2 \right) \right) \\ &= \left(T^{-1}\Sigma_0^1 + U \left(\Sigma_1^1 - \Sigma_0^1 \right) + L \left(\Sigma_2^1 - \Sigma_0^1 \right) \right) \end{aligned} \tag{5.23}$$

where we have used the fact that $I = T^{-1} - U - L$ and D is the diagonal matrix with

elements $\left(0, \dots, 0, \frac{1}{2}, \frac{3}{2}, \dots, M - \frac{1}{2}\right)$. Additionally, we have defined

$$\begin{aligned}\Sigma_0^1 &= \frac{\sigma_0 \Delta t}{M^2} \left(g_{20} I + g_{00} D^2 \right), \\ \Sigma_1^1 &= \frac{\sigma_0 \Delta t}{M^2} \left(g_{21} I + 2g_{11} D + g_{01} D^2 \right), \\ \Sigma_2^1 &= \frac{\sigma_0 \Delta t}{M^2} \left(g_{21} I - 2g_{11} D + g_{01} D^2 \right).\end{aligned}$$

Likewise,

$$\begin{aligned}\Sigma^{(2)} &= \left(\frac{\sigma_0 \Delta t}{M^2} \right)^2 \left(G^{(4)} + 4G^{(3)} D + 6G^{(2)} D^2 + 4G^{(1)} D^3 + G^{(0)} D^4 \right) \\ &= \left(\frac{\sigma_0 \Delta t}{M^2} \right)^2 \left(g_{40} I + 6g_{20} D^2 + g_{00} D^4 \right. \\ &\quad \left. + U \left(g_{41} I + 4g_{31} D + 6g_{21} D^2 + 4g_{11} D^3 + g_{01} D^4 \right) \right. \\ &\quad \left. + L \left(g_{41} I - 4g_{31} D + 6g_{21} D^2 - 4g_{11} D^3 + g_{01} D^4 \right) \right) \\ &= \left(T^{-1} \Sigma_0^2 + U \left(\Sigma_1^2 - \Sigma_0^2 \right) + L \left(\Sigma_2^2 - \Sigma_0^2 \right) \right),\end{aligned}\tag{5.24}$$

with

$$\begin{aligned}\Sigma_0^2 &= \left(\frac{\sigma_0 \Delta t}{M^2} \right)^2 \left(g_{40} I + 6g_{20} D^2 + g_{00} D^4 \right), \\ \Sigma_1^2 &= \left(\frac{\sigma_0 \Delta t}{M^2} \right)^2 \left(g_{41} I + 4g_{31} D + 6g_{21} D^2 + 4g_{11} D^3 + g_{01} D^4 \right), \\ \Sigma_2^2 &= \left(\frac{\sigma_0 \Delta t}{M^2} \right)^2 \left(g_{41} I - 4g_{31} D + 6g_{21} D^2 - 4g_{11} D^3 + g_{01} D^4 \right).\end{aligned}$$

We can also write the matrices $B^{(1)}$ and $B^{(2)}$ using combinations of T^{-1} , U and L as

$$\begin{aligned} B^{(1)} &= e^{-\nu/2}(I + U) + e^{\nu/2}L = e^{-\nu/2}T^{-1} + 2 \sinh\left(\frac{\nu}{2}\right)L, \\ B^{(2)} &= e^{-\nu/2}(I + L) + e^{\nu/2}U = e^{-\nu/2}T^{-1} + 2 \sinh\left(\frac{\nu}{2}\right)U. \end{aligned} \quad (5.25)$$

Due to the symmetry and structure of T ,

$$TU = \frac{1}{1-d^2} \begin{pmatrix} 0 & d & & & \\ 0 & -d^2 & \ddots & & \\ & \ddots & \ddots & & \\ & & & -d^2 & d \\ & & & 0 & -d^2 \end{pmatrix},$$

and similarly

$$TL = \frac{1}{1-d^2} \begin{pmatrix} -d^2 & 0 & & & \\ d & -d^2 & \ddots & & \\ & \ddots & \ddots & & \\ & & & -d^2 & 0 \\ & & & d & 0 \end{pmatrix}.$$

This makes it possible to write expressions for $TB^{(1)}$, etc. as a combination of diagonal matrices and products with TU and TL . From this, we will show that equation (5.22) can be written as a tridiagonal system.

Upon collecting common terms in equation (5.22), and considering only $N+1 \leq j \leq$

$N + M$, we have

$$\begin{aligned}
& \left(I + \Sigma^{(1)} + \frac{1}{2}\Sigma^{(2)} - B^{(1)}F^+ + B^{(2)}F^- \right) u^{n+1} \\
&= \left(A + \Sigma^{(1)} + B^{(1)}F^+E^+ - B^{(2)}F^-E^- \right) u^n - \frac{1}{2}Au^{n-1} \\
&+ B^{(1)}E^+v_1^n - B^{(2)}E^-v_2^n,
\end{aligned} \tag{5.26}$$

where the plus and minus superscripts of E and F indicate whether all but the first or last entry of the $M + 1$ diagonal entries given by equations (5.21) is taken. Multiply by T and make use of the decompositions given by (5.23), (5.24) and (5.25), so that

$$T_2u^{n+1} = T_1u^n - T_0u^{n-1} + T_+v_1^n - T_-v_2^n,$$

where

$$\begin{aligned}
T_2 &= T \left(I + \Sigma^{(1)} + \frac{1}{2}\Sigma^{(2)} - B^{(1)}F^+ + B^{(2)}F^- \right) \\
&= \left[T + \Sigma_0^1 + \frac{1}{2}\Sigma_0^2 - e^{-\nu/2}(F^+ - F^-) \right] \\
&- TU \left[\Sigma_0^1 - \Sigma_1^1 + \frac{1}{2}\Sigma_0^2 - \frac{1}{2}\Sigma_1^2 - 2 \sinh \left(\frac{\nu}{2} \right) F^- \right] \\
&- TL \left[\Sigma_0^1 - \Sigma_2^1 + \frac{1}{2}\Sigma_0^2 - \frac{1}{2}\Sigma_2^2 - 2 \sinh \left(\frac{\nu}{2} \right) F^+ \right],
\end{aligned}$$

$$\begin{aligned}
T_1 &= T \left(A + \Sigma^{(1)} + B^{(1)} F^+ E^+ - B^{(2)} F^- E^- \right) \\
&= \left[TA + \Sigma_0^1 + e^{-\nu/2} (E^+ F^+ - E^- F^-) \right] \\
&\quad - TU \left[\Sigma_0^1 - \Sigma_1^1 - 2 \sinh \left(\frac{\nu}{2} \right) F^+ E^+ \right] \\
&\quad - TL \left[\Sigma_0^1 - \Sigma_2^1 - 2 \sinh \left(\frac{\nu}{2} \right) F^- E^- \right],
\end{aligned}$$

$$T_0 = \frac{TA}{2},$$

$$T_+ = \left[e^{-\nu/2} E^+ \right] + 2 \sinh \left(\frac{\nu}{2} \right) TL [E^+],$$

and

$$T_- = \left[e^{-\nu/2} E^- \right] + 2 \sinh \left(\frac{\nu}{2} \right) TU [E^-].$$

The system is solved with a basic tridiagonal algorithm for u_j^{n+1} in the PML region. It is then used to update v_1^{n+1} and v_2^{n+1} .

In the free space region, note that equation (5.22) gives the solution for u_j^{n+1} explicitly for $1 \leq j \leq N$, since the sums involving PML terms will not be carried over this range. Thus, all terms arising from the PML will only appear as a constant vector in the $N \times N$ subsystem. Furthermore, this vector has a simple form due again to the structure of T , U and L .

If we instead choose to use a second order dissipative discretization for equation (5.4),

we now recover

$$\begin{aligned} & \frac{(2 + 3\sigma(x)\Delta t + \sigma(x)^2\Delta t^2) u^{n+1}(x)}{\Delta t^2} - \frac{(5 + 4\sigma(x)\Delta t) u^n}{\Delta t^2} \\ & + \frac{(4 + \sigma(x)\Delta t) u^{n-1}(x)}{\Delta t^2} - \frac{u^{n-2}(x)}{\Delta t^2} = c^2 u_{xx}^{n+1}(x) - \sigma'(x) \int_0^{t^{n+1}} u_x(x) e^{-\sigma(x)(t-s)} \end{aligned}$$

instead, but still treat

$$v^{n+1} = \sigma'(x) \int_0^{t^{n+1}} u_x(x) e^{-\sigma(x)(t^{n+1}-s)}$$

in exactly the same manner. Once again, rearrange this into a Helmholtz equation as

$$\begin{aligned} \left(\partial_{xx} - \alpha^2 \right) u^{n+1} &= \frac{\alpha^2}{2} \left(\left(3\sigma\Delta t + \sigma^2\Delta t^2 \right) u^{n+1} - (5 + 4\sigma\Delta t) u^n \right. \\ &\quad \left. + (4 + \sigma(x)\Delta t) u^{n-1} - u^{n-2} \right) + v^{n+1}, \end{aligned}$$

with $\alpha = \frac{\sqrt{2}}{c\Delta t}$ here. Invert the Helmholtz operator to get

$$\begin{aligned} u^{n+1}(x) &= -\frac{\alpha}{4} \int e^{-\alpha|x-y|} \left(\left(3\sigma(y)\Delta t + \sigma^2(y)\Delta t^2 \right) u^{n+1}(y) \right. \\ &\quad \left. - (5 + 4\sigma(y)\Delta t) u^n(y) + (4 + \sigma(y)\Delta t) u^{n-1}(y) - u^{n-2}(y) \right) dy + \tilde{v}^{n+1}. \end{aligned}$$

with $\tilde{v}^{n+1}$ defined as before.

Discretization using the midpoint approximation will then result in

$$u^{n+1}(x_j) = \sum_{i=1}^{N+M} A_{ji} \left(\frac{5}{4}u_i^n - u_i^{n-1} + \frac{1}{4}u_i^{n-2} \right) - \sum_{i=N+1}^{N+m} \left(\frac{3}{4}\Sigma_{ji}^{(1)} u_i^{n+1} + \frac{1}{4}\Sigma_{ji}^{(2)} u_i^{n+1} - \Sigma_{ji}^{(1)} u_i^n + \frac{1}{4}\Sigma_{ji}^{(1)} u_i^{n-1} \right) + \tilde{v}_j^{n+1},$$

with $\Sigma^{(1)}$ and $\Sigma^{(2)}$ defined as before.

Following the exact same procedure as for the 1st order dissipative scheme will yield that equation (5.26) now becomes

$$\begin{aligned} & \left(I + \frac{3}{4}\Sigma^{(1)} + \frac{1}{4}\Sigma^{(2)} - B^{(1)}F^+ + B^{(2)}F^- \right) u^{n+1} \\ &= \left(\frac{5}{4}A + \Sigma^{(1)} + B^{(1)}F^+E^+ - B^{(2)}F^-E^- \right) u^n \\ &- \left(A + \frac{1}{4}\Sigma^{(1)} \right) u^{n-1} + \frac{1}{4}Au^{n-2} + B^{(1)}E^+v_1^n - B^{(2)}E^-v_2^n. \end{aligned}$$

Multiply by T to give that all the matrices on the left hand side become tri-diagonal at worst, which gives

$$T_3u^{n+1} = T_2u^n - T_1u^{n-1} + T_0u^{n-2} + T_+v_1^n - T_-v_2^n,$$

where

$$\begin{aligned}
T_3 &= T \left(I + \frac{3}{4} \Sigma^{(1)} + \frac{1}{4} \Sigma^{(2)} - B^{(1)} F^+ + B^{(2)} F^- \right) \\
&= \left[T + \frac{3}{4} \Sigma_0^1 + \frac{1}{4} \Sigma_0^2 - e^{-\nu/2} (F^+ - F^-) \right] \\
&\quad - TU \left[\frac{3}{4} \Sigma_0^1 - \frac{3}{4} \Sigma_1^1 + \frac{1}{4} \Sigma_0^2 - \frac{1}{4} \Sigma_1^2 - 2 \sinh \left(\frac{\nu}{2} \right) F^- \right] \\
&\quad - TL \left[\frac{3}{4} \Sigma_0^1 - \frac{3}{4} \Sigma_2^1 + \frac{1}{4} \Sigma_0^2 - \frac{1}{4} \Sigma_2^2 - 2 \sinh \left(\frac{\nu}{2} \right) F^+ \right],
\end{aligned}$$

$$\begin{aligned}
T_2 &= T \left(\frac{5}{4} A + \Sigma^{(1)} + B^{(1)} F^+ E^+ - B^{(2)} F^- E^- \right) \\
&= \left[\frac{5}{4} T A + \Sigma_0^1 + e^{-\nu/2} (E^+ F^+ - E^- F^-) \right] \\
&\quad - TU \left[\Sigma_0^1 - \Sigma_1^1 - 2 \sinh \left(\frac{\nu}{2} \right) F^+ E^+ \right] \\
&\quad - TL \left[\Sigma_0^1 - \Sigma_2^1 - 2 \sinh \left(\frac{\nu}{2} \right) F^- E^- \right],
\end{aligned}$$

$$\begin{aligned}
T_1 &= T \left(A + \frac{1}{4} \Sigma^{(1)} \right) \\
&= T A + \frac{1}{4} \left(\Sigma_0^1 - TU \left[\Sigma_0^1 - \Sigma_1^1 \right] - TL \left[\Sigma_0^1 - \Sigma_1^1 \right] \right),
\end{aligned}$$

$$T_0 = \frac{1}{4} T A,$$

$$T_+ = \left[e^{-\nu/2} E^+ \right] + 2 \sinh \left(\frac{\nu}{2} \right) TL \left[E^+ \right],$$

and

$$T_- = \left[e^{-\nu/2} E^- \right] + 2 \sinh \left(\frac{\nu}{2} \right) TU \left[E^- \right].$$

We can then find u^{n+1} in the PML region using a tri-diagonal solver before explicitly updating for u^{n+1} in the entire domain.

For the second order non-dissipative scheme, after the standard discretization, equations (5.4) and (5.5) become

$$\frac{(1 + 2s + 2s^2) w - 2u^n - 4su^{n-1}}{(c\Delta t)^2} = \frac{1}{2}w_{xx} - v^{n+1},$$

$$v^{n+1} = s' \sum_{m=0}^n e^{-2s(n-m)} \left(u_x^{m+1} + e^{-2s} u_x^m \right),$$

where

$$w = u^{n+1} + u^{n-1}, \quad s(x) = \frac{\sigma(x)\Delta t}{2}.$$

Arranging this in a Helmholtz type equation gives

$$\left(\partial_{xx} - \alpha^2 \right) w = -2\alpha^2 \left(u^n + 2su^{n-1} - \left(s + s^2 \right) w \right) + v^{n+1}$$

with $\alpha = \frac{\sqrt{2}}{c\Delta t}$. After inverting the operator, this now yields

$$w(x) = \alpha \int e^{-\alpha|x-y|} \left(u^n(y) + 2s(y)u^{n-1}(y) - \left(s(y) + s^2(y) \right) w(y) \right) dy + \tilde{v}^{n+1}$$

with

$$\tilde{v}^{n+1}(x) = -\frac{1}{2\alpha} \int e^{-\alpha|x-y|} s' \sum_{m=0}^n e^{-2(n-m)s(y)} \left(u_y^{m+1}(y) + e^{-2s(y)} u_y^m(y) \right) dy.$$

After a full discretization of the domain, we obtain

$$w_j = \sum_{i=1}^{N+M} A_{ji} u_i^n + \sum_{i=N+1}^{N+m} \left(2\Sigma_{ji}^{(1)} u_i^{n-1} - \left(\Sigma_{ji}^{(1)} + \Sigma_{ji}^{(2)} \right) w_i \right) + \tilde{v}_j^{n+1}.$$

Handling the temporal integral $\tilde{v}_j^{n+1}$ in the same way as for the dissipative cases then gives

$$\begin{aligned} w_j = & \sum_{i=1}^{N+M} A_{ji} u_i^n + \sum_{i=N+1}^{N+m} 2\Sigma_{ji}^{(1)} u_i^n - \left(\Sigma_{ji}^{(2)} + \Sigma_{ji}^{(1)} \right) w_i \\ & + \sum_{i=N+1}^{N+m} B_{ji}^{(1)} \left(E_i v_{1,i}^n + F_i \left(E_i u_i^n + w_i - u_i^{n-1} \right) \right) \\ & - \sum_{i=N+1}^{N+m} B_{ji}^{(2)} \left(E_{i-1} v_{2,i}^n + F_{i-1} \left(E_{i-1} u_i^n + w_i - u_i^{n-1} \right) \right). \end{aligned}$$

This then becomes

$$\begin{aligned} & \left(I + \Sigma^{(1)} + \Sigma^{(2)} - B^{(1)} F^+ + B^{(2)} F^- \right) w \\ & = \left(A + B^{(1)} E^+ F^+ - B^{(2)} E^- F^- \right) u^n + \left(2\Sigma^{(1)} - B^{(1)} F^+ + B^{(2)} F^- \right) u^{n-1} \\ & + B^{(1)} E^+ v_1^n - B^{(2)} E^- v_2^n. \end{aligned}$$

Multiplication by T then produces

$$T_w w = T_1 u^n + T_0 u^{n-1} + T_+ v_1^n - T_- v_2^n + S, \quad (5.27)$$

where

$$\begin{aligned} T_w = & \left[I + \Sigma_0^1 + \Sigma_0^2 - e^{-\nu/2}(F^+ - F^-) \right] \\ & - TU \left[I + \Sigma_0^1 - \Sigma_1^1 + \Sigma_0^2 - \Sigma_1^2 - a_1 F^- \right] \\ & - TL \left[I + \Sigma_0^1 - \Sigma_2^1 + \Sigma_0^2 - \Sigma_2^2 + a_1 F^+ \right], \end{aligned}$$

$$\begin{aligned} T_1 = & \left[a_0 I + e^{-\nu/2}(E^+ F^+ - E^- F^-) \right] \\ & - TU \left[(a_0 - a_1) I + a_1 E^- F^- \right] \\ & - TL \left[(a_0 - a_1) I - a_1 E^+ F^+ \right], \end{aligned}$$

$$\begin{aligned} T_0 = & \left[2\Sigma_0^1 + 2\Sigma_0^2 - e^{-\nu/2}(F^+ - F^-) \right] \\ & - TU \left[2\Sigma_0^1 - 2\Sigma_1^1 + 2\Sigma_0^2 - 2\Sigma_1^2 - a_1 F^- \right] \\ & - TL \left[2\Sigma_0^1 - 2\Sigma_2^1 + 2\Sigma_0^2 - 2\Sigma_2^2 + a_1 F^+ \right], \end{aligned}$$

$$T_+ = \left[e^{-\nu/2} E^+ \right] + a_1 TL \left[E^+ \right],$$

$$T_- = \left[e^{-\nu/2} E^- \right] + a_1 TU \left[E^- \right],$$

and finally the vector S is nonzero only in the first entry, and is comprised of the matrix vector product over $j = 1, 2, \dots N$

$$S_j = \delta_{j,N+1} \sum_{i=1}^N (TL)_{N+1,i} u_i^n, \quad N+1 \leq j \leq N+M.$$

Equation (5.27) may then be solved in the PML region using a tri-diagonal solver. This solution may then be used to calculate an explicit solution for u^{n+1} over the entire domain.

5.1.2.2 Optimization

Optimizing the damping of a wave in a PML layer has been a subject of significant study, but most authors believe that the result is very much problem specific. Thus, we will conduct a numerical study for optimizing the absorbance of a perfectly matched layer for a quadratic damping function. The variables will consist of both σ_0 and the length of the PML region. The other parameters of the study are as such:

- The initial conditions are $f(x) = e^{-100x^2}$ and $g(x) = 0$.
- The interior domain is $[-1, 1]$.
- The PML region has a length of $\{0.5, 2, 3\}$.
- $\Delta x = \Delta t = 0.01$ with a wave speed of $c = 1$.
- The chosen final time is $tf = 20$, such that all waves would end on a recombination.
- $\sigma_0 = \{1, 2, 3, 4, 5, 10, 15, 20, 25\}$.

Table 5.3: Optimization of the Dissipative 1st Order Quadratic Damping Function

	$d = 0.5$		$d = 2$		$d = 3$	
σ_0	∞ -norm	2-norm	∞ -norm	2-norm	∞ -norm	2-norm
0	0.14901	1.35786	0.13877	1.30731	0.13422	1.28520
1	0.046120	0.40325	0.021976	0.19768	0.016763	0.16238
2	0.014676	0.12161	0.0047455	0.045587	0.0029356	0.035231
3	0.0046670	0.036095	0.0014279	0.018349	0.00087016	0.011461
4	0.0014033	0.010184	0.00086525	0.010557	0.0010083	0.011400
5	0.00035072	0.0031107	0.00045273	0.0058206	0.0010008	0.011423
10	0.000045637	0.00032578	0.00044261	0.0056118	0.00065305	0.0076492
15	0.000044945	0.00032420	0.00047148	0.0044793	0.00050443	0.0053598
20	0.000038350	0.00027405	0.00053009	0.0050003	0.00050236	0.0066065
25	0.000030196	0.00023973	0.00038082	0.0032459	0.00036747	0.0044241

Both the 2-norm and ∞ -norm of u was taken at the end of the run. A baseline with $\sigma_0 = 0$ was run for each PML length to determine how much damping had actually occurred. The results can be seen in table 5.3. From the table, while it appears that the first few increases of σ_0 are quite helpful, the results very much level out shortly thereafter. This could be due to the larger co-efficient exacerbating the error that is caused by the numerical approximation of the integral. The best scenario appears to be taking a shorter PML region accompanied with a relatively larger σ_0 , especially in consideration of the fact that a shorter PML region will constitute less work being done by the tri-diagonal solver when solving the implicit step which appears only in the PML region.

5.2 One-Way Waves

An alternative approach to the idea of a perfectly matched damping layer is to construct a boundary that appears invisible that will simply let the outward propagating wave pass through without creating any reflection. While this may be rather difficult to do in the general case, it is quite feasible in the one dimensional case under a few basic assumptions.

Assume that the initials conditions $f(x)$ and $g(x)$ in equation (1.1) are of compact support and that this support does not extend beyond Ω_0 , the desired domain of computation. For the sake of simplicity, let $\Omega_0 = [-L, L]$.

The first order dissipative MOL^T solution to the 1D wave equation is then given by

$$u^{n+1}(x) = \int_{\Omega_{t_n}} \alpha \left(u^n(y) - \frac{1}{2}u^{n-1}(y) \right) e^{-\alpha|x-y|} dy, \quad (5.28)$$

where Ω_{t_n} is the domain of dependence for $u(x)$ at time t_n .

Since the wave speed propagation is limited by c , it is easy to see that Ω_{t_n} is fully contained in $[-L - ct_n, L + ct_n]$, which gives that (5.28) now becomes

$$u^{n+1}(x) = \int_{-L-ct_n}^{L+ct_n} \alpha \left(u^n(y) - \frac{1}{2}u^{n-1}(y) \right) e^{-\alpha|x-y|} dy. \quad (5.29)$$

This is equivalent to the free space solution as $u^{n+1}(x)$ will be identically zero outside of this region.

We now wish to compute the solution only over the initial region Ω_0 while still factoring in the contribution from the additional domain of $\Omega_{t_{n+1}}/\Omega_0$. This is accomplished by writing (5.29) as

$$\begin{aligned} u^{n+1}(x) &= \int_{-L}^L \alpha \left(u^n(y) - \frac{1}{2}u^{n-1}(y) \right) e^{-\alpha|x-y|} dy + S_1^{n+1} + S_2^{n+1}, \\ S_1^{n+1} &= \int_{-L-ct_n}^{-L} \alpha \left(u^n(y) - \frac{1}{2}u^{n-1}(y) \right) e^{-\alpha(x-y)} dy, \\ S_2^{n+1} &= \int_L^{L+ct_n} \alpha \left(u^n(y) - \frac{1}{2}u^{n-1}(y) \right) e^{-\alpha(y-x)} dy. \end{aligned} \quad (5.30)$$

Since we are considering the case of outflow boundary conditions, the key step here is to

restrict only to outgoing waves that will not reflect back into the domain. This means that we must impose non-reflecting boundary conditions at $x = \pm L$. For this one dimensional case, this condition simplifies to the respective transport equations

$$u_t + cu_x = 0, \quad x \geq L$$

$$u_t - cu_x = 0, \quad x \leq -L.$$

We now consider the case where $x \geq L$ and the case where $x \leq -L$ will be similar.

Tracing along the characteristics gives that

$$u(L + y, t) = u\left(L, t - \frac{y}{c}\right) \text{ for } y > 0.$$

Thus, for $x < L$, we can write

$$\begin{aligned} S_2^{n+1} &= \int_L^{L+ct_n} \alpha \left(u^n(y) - \frac{1}{2}u^{n-1}(y) \right) e^{-\alpha(y-x)} dy \\ &= \int_L^{L+ct_n} \alpha \left(u(y, t_n) - \frac{1}{2}u(y, t_{n-1}) \right) e^{-\alpha(y-x)} dy \\ &= \alpha e^{-\alpha(L-x)} \int_0^{ct_n} \left(u(L + y, t_n) - \frac{1}{2}u(L + y, t_{n-1}) \right) e^{-\alpha y} dy \\ &= \alpha c e^{-\alpha(L-x)} \int_0^{t_n} \left(u(L, t_n - s) - \frac{1}{2}u(L, t_{n-1} - s) \right) e^{-\alpha cs} ds. \end{aligned}$$

This, in effect, has transformed the spatial integral into a temporal integral and therefore, we must only track the time history at the boundary points $x = \pm L$. Additionally, we may proceed to track this temporal integral using a recurrence relation which limits storage to

the previous two time steps. Note that

$$\begin{aligned} S_2^{n+1} &= e^{-\alpha(L-x)} \alpha c \int_0^{t_n} \left(u(L, t_n - s) - \frac{1}{2} u(L, t_{n-1} - s) \right) e^{-\alpha c s} ds \\ &= e^{-\alpha(L-x)} \tilde{S}_2^{n+1}. \end{aligned}$$

Making use of the change of variables $s = s + \Delta t$ gives

$$\begin{aligned} \tilde{S}_2^{n+1} &= \alpha c \int_0^{t_n} \left(u(L, t_n - s) - \frac{1}{2} u(L, t_{n-1} - s) \right) e^{-\alpha c s} ds \\ &= e^{-\alpha c \Delta t} \tilde{S}_2^n + \alpha c \int_0^{\Delta t} \left(u(L, t_n - s) - \frac{1}{2} u(L, t_{n-1} - s) \right) e^{-\alpha c s} ds \\ &= e^{-1} \tilde{S}_2^n + \alpha c \int_0^{\Delta t} \left(u(L, t_n - s) - \frac{1}{2} u(L, t_{n-1} - s) \right) e^{-\alpha c s} ds. \end{aligned}$$

We now choose to integrate $\int_0^{\Delta t} u(L, t_n - s) e^{-\alpha c s} ds$ by approximating $u(L, t_n - s)$ with a linear function through points at $u(L, t_n)$ and $u(L, t_{n-1})$. This linear approximation is given by

$$u(L, t_n - s) \approx \frac{u(L, t_{n-1}) - u(L, t_n)}{\Delta t} s + u(L, t_n) = z s + b. \quad (5.31)$$

Thus,

$$\begin{aligned}
\alpha c \int_0^{\Delta t} u(L, t_n - s) e^{-\alpha c s} ds &= \frac{1}{\Delta t} \int_0^{\Delta t} (zs + b) e^{-\frac{s}{\Delta t}} ds \\
&= \frac{1}{\Delta t} \left(-(zs + b) \Delta t e^{-\frac{s}{\Delta t}} \Big|_0^{\Delta t} + \int_0^{\Delta t} z \Delta t e^{-\frac{s}{\Delta t}} ds \right) \\
&= \frac{1}{\Delta t} \left(-(zs + b) \Delta t e^{-\frac{s}{\Delta t}} \Big|_0^{\Delta t} - \Delta t^2 z e^{-\frac{s}{\Delta t}} \Big|_0^{\Delta t} \right) \\
&= \frac{1}{\Delta t} \left(-(z\Delta t + b) \Delta t e^{-1} - \Delta t^2 z e^{-1} + b\Delta t + z\Delta t^2 \right) \\
&= -(z\Delta t + b) e^{-1} - \Delta t z e^{-1} + b + z\Delta t \\
&= - \left(u^{n-1}(L) - u^n(L) + u^n(L) \right) e^{-1} - \left(u^{n-1}(L) - u^n(L) \right) e^{-1} \\
&\quad + u^n(L) + u^{n-1}(L) - u^n(L) \\
&= e^{-1} u^n(L) + (1 - 2e^{-1}) u^{n-1}(L).
\end{aligned}$$

This also gives that

$$\alpha c \int_0^{\Delta t} u(L, t_{n-1} - s) e^{-\alpha c s} ds = e^{-1} u^{n-1}(L) + (1 - 2e^{-1}) u^{n-2}(L).$$

Thus, estimating this integral using the trapezoidal rule results in

$$\begin{aligned}
& \alpha c \int_0^{\Delta t} \left(u(L, t_n - s) - \frac{1}{2} u(L, t_{n-1} - s) \right) e^{-\alpha c s} ds \\
&= e^{-1} u^n(L) + \left(1 - 2e^{-1} \right) u^{n-1}(L) - \frac{1}{2} \left(e^{-1} u^{n-1}(L) + \left(1 - 2e^{-1} \right) u^{n-2}(L) \right) \\
&= e^{-1} u^n(L) + \left(1 - \frac{5}{2} e^{-1} \right) u^{n-1}(L) - \frac{1 - 2e^{-1}}{2} u^{n-2}(L) \\
&= \frac{\alpha c \Delta t}{2} \left(e^{-\alpha c \Delta t} \left(u(L, t_{n-1}) - \frac{1}{2} u(L, t_{n-2}) \right) + u(L, t_n) - \frac{1}{2} u(L, t_{n-1}) \right) \\
&= \frac{1}{2} \left(u(L, t_n) + \left(e^{-1} - \frac{1}{2} \right) u(L, t_{n-1}) - \frac{e^{-1}}{2} u(L, t_{n-2}) \right).
\end{aligned}$$

This implies that

$$\tilde{S}_{n+1}^2 = e^{-1} \tilde{S}_n^2 + e^{-1} u^n(L) + \left(1 - \frac{5}{2} e^{-1} \right) u^{n-1}(L) - \frac{1 - 2e^{-1}}{2} u^{n-2}(L).$$

Similarly,

$$\begin{aligned}
S_1^{n+1}(x) &= e^{-\alpha(L+x)} \tilde{S}_1^{n+1} \\
\tilde{S}_1^{n+1} &= e^{-1} \tilde{S}_1^n + e^{-1} u^n(-L) + \left(1 - \frac{5}{2} e^{-1} \right) u^{n-1}(-L) \\
&\quad - \frac{1 - 2e^{-1}}{2} u^{n-2}(-L) \frac{1}{2} \left(u(-L, t_n) + \left(e^{-1} - \frac{1}{2} \right) u(-L, t_{n-1}) - \frac{e^{-1}}{2} u(-L, t_{n-2}) \right).
\end{aligned}$$

Thus, the fully discretized solution over Ω_0 will be given by

$$\begin{aligned}
u_j^{n+1} &= \sum_{i=1}^m A_{ji} \left(u_i^n - \frac{1}{2} u_i^{n-1} \right) + S_1^{n+1} + S_2^{n+1} \\
&= \sum_{i=1}^m A_{ji} \left(u_i^n - \frac{1}{2} u_i^{n-1} \right) + e^{-\alpha(L+x_j)} \tilde{S}_1^{n+1} + e^{-\alpha(L-x_j)} \tilde{S}_2^{n+1} \\
&= \sum_{i=1}^m A_{ji} \left(u_i^n - \frac{1}{2} u_i^{n-1} \right) \\
&\quad + e^{-\alpha(L+x_j)} \left(e^{-1} \tilde{S}_1^n + e^{-1} u^n(-L) + \left(1 - \frac{5}{2} e^{-1} \right) u^{n-1}(-L) - \frac{1-2e^{-1}}{2} u^{n-2}(-L) \right) \\
&\quad + e^{-\alpha(L-x_j)} \left(e^{-1} \tilde{S}_2^n + e^{-1} u^n(L) + \left(1 - \frac{5}{2} e^{-1} \right) u^{n-1}(L) - \frac{1-2e^{-1}}{2} u^{n-2}(L) \right).
\end{aligned}$$

For the second order dissipative scheme, equation (5.29) now becomes

$$u^{n+1}(x) = \int_{-L-ct_n}^{L+ct_n} \alpha \left(\frac{5}{4} u^n(y) - u^{n-1}(y) + \frac{1}{4} u^{n-2}(y) \right) e^{-\alpha|x-y|} dy.$$

which in turn changes equations (5.30) to

$$\begin{aligned}
u^{n+1}(x) &= \int_{-L}^L \alpha \left(\frac{5}{4} u^n(y) - u^{n-1}(y) + \frac{1}{4} u^{n-2}(y) \right) e^{-\alpha|x-y|} dy + S_1^{n+1} + S_2^{n+1}, \\
S_1^{n+1} &= \int_{-L-ct_n}^{-L} \alpha \left(\frac{5}{4} u^n(y) - u^{n-1}(y) + \frac{1}{4} u^{n-2}(y) \right) e^{-\alpha(x-y)} dy, \\
S_2^{n+1} &= \int_L^{L+ct_n} \alpha \left(\frac{5}{4} u^n(y) - u^{n-1}(y) + \frac{1}{4} u^{n-2}(y) \right) e^{-\alpha(y-x)} dy.
\end{aligned}$$

Repeating the same change of variables will give

$$\begin{aligned}
\tilde{S}_2^{n+1} &= \alpha c \int_0^{t_n} \left(\frac{5}{4} u(L, t_n - s) - u(L, t_{n-1} - s) + \frac{1}{4} u(L, t_{n-2} - s) \right) e^{-\alpha c s} ds \\
&= e^{-\alpha c \Delta t} \tilde{S}_2^n + \alpha c \int_0^{\Delta t} \left(\frac{5}{4} u(L, t_n - s) - u(L, t_{n-1} - s) + \frac{1}{4} u(L, t_{n-2} - s) \right) e^{-\alpha c s} ds \\
&= e^{-\sqrt{2}} \tilde{S}_2^n + \alpha c \int_0^{\Delta t} \left(\frac{5}{4} u(L, t_n - s) - u(L, t_{n-1} - s) + \frac{1}{4} u(L, t_{n-2} - s) \right) e^{-\alpha c s} ds.
\end{aligned}$$

Repeating the estimation of the linearized integral (5.32) now gives

$$\begin{aligned}
\alpha c \int_0^{\Delta t} u(L, t_n - s) e^{-\alpha c s} ds &= \frac{\sqrt{2}}{\Delta t} \int_0^{\Delta t} (zs + b) e^{-\frac{\sqrt{2}s}{\Delta t}} ds \\
&= \frac{\sqrt{2}}{\Delta t} \left(-(zs + b) \frac{\Delta t}{\sqrt{2}} e^{-\frac{\sqrt{2}s}{\Delta t}} \Big|_0^{\Delta t} + \int_0^{\Delta t} z \frac{\Delta t}{\sqrt{2}} e^{-\frac{\sqrt{2}s}{\Delta t}} ds \right) \\
&= \frac{\sqrt{2}}{\Delta t} \left(-(zs + b) \frac{\Delta t}{\sqrt{2}} e^{-\frac{\sqrt{2}s}{\Delta t}} \Big|_0^{\Delta t} - \frac{\Delta t^2}{2} z e^{-\frac{s}{\Delta t}} \Big|_0^{\Delta t} \right) \\
&= \frac{\sqrt{2}}{\Delta t} \left(-(z\Delta t + b) \frac{\Delta t}{\sqrt{2}} e^{-\sqrt{2}} - \frac{\Delta t^2}{2} z e^{-\sqrt{2}} + b \frac{\Delta t}{\sqrt{2}} + z \frac{\Delta t^2}{2} \right) \\
&= -(z\Delta t + b) e^{-\sqrt{2}} - \frac{\Delta t}{\sqrt{2}} z e^{-\sqrt{2}} + b + z \frac{\Delta t}{\sqrt{2}}.
\end{aligned}$$

Substituting back in for z and b from (5.31) produces

$$\begin{aligned}
&= -u^{n-1}(L) e^{-\sqrt{2}} - \frac{(u^{n-1}(L) - u^n(L)) e^{-\sqrt{2}}}{\sqrt{2}} + u^n(L) + \frac{u^{n-1}(L) - u^n(L)}{\sqrt{2}} \\
&= \frac{\sqrt{2} - 1 + e^{-\sqrt{2}}}{\sqrt{2}} u^n(L) + \frac{1 - \sqrt{2} e^{-\sqrt{2}} - e^{-\sqrt{2}}}{\sqrt{2}} u^{n-1}(L) \\
&= \epsilon_1 u^n(L) + \epsilon_2 u^{n-1}(L)
\end{aligned}$$

with

$$\begin{aligned}\epsilon_1 &= \frac{\sqrt{2} - 1 + e^{-\sqrt{2}}}{\sqrt{2}} \text{ and} \\ \epsilon_2 &= \frac{1 - \sqrt{2}e^{-\sqrt{2}} - e^{-\sqrt{2}}}{\sqrt{2}}.\end{aligned}$$

Thus

$$\tilde{S}_2^{n+1} = e^{-\sqrt{2}}\tilde{S}_2^n + \frac{5}{4}\epsilon_1 u^n(L) + \left(\frac{5}{4}\epsilon_2 - \epsilon_1\right) u^{n-1}(L) - \left(\epsilon_2 + \frac{1}{4}\epsilon_1\right) u^{n-2}(L) + \frac{1}{4}\epsilon_2 u^{n-3}(L).$$

$\tilde{S}_1^{n+1}$ will be similar.

Therefore, the fully discretized scheme will be given by

$$u_j^{n+1} = \sum_{i=1}^m A_{ji} \left(\frac{5}{4}u_i^n - u_i^{n-1} + \frac{1}{4}u_i^{n-2} \right) + e^{-\alpha(L+x_j)} \tilde{S}_1^{n+1} + e^{-\alpha(L-x_j)} \tilde{S}_2^{n+1}.$$

Similarly, for the second order non-dissipative scheme,

$$\begin{aligned}u^{n+1}(x) + u^{n-1}(x) &= \alpha \int_{-L-ct_n}^{L+ct_n} e^{-\alpha|x-y|} u^n(y) dy \\ &= \alpha \int_{-L}^L e^{-\alpha|x-y|} u^n(y) dy + S_1^n + S_2^n.\end{aligned}$$

We then follow exactly the same procedure outlined for the second order dissipative scheme to get

$$\tilde{S}_2^{n+1} = e^{-\sqrt{2}}\tilde{S}_2^n + \epsilon_1 u^n(L) + \epsilon_1 u^{n-1}(L)$$

with a similar result for $\tilde{S}_1^{n+1}$.

Therefore,

$$u_j^{n+1} = -u_j^{n-1} + \sum_{i=1}^N A_{ji} u_i^n + e^{-\alpha(L+x_j)} \tilde{S}_1^n + e^{-\alpha(L-x_j)} \tilde{S}_2^n. \quad (5.32)$$

5.2.1 Performance

Instituting this non-reflecting boundary condition definitely produces an increase in performance as can be seen in table 5.4, based on the following parameters:

- The initial conditions are $f(x) = e^{-100x^2}$ and $g(x) = 0$.
- The domain is $[-1, 1]$.
- $\Delta x = \Delta t = 0.01$ with a wave speed of $c = 1$.
- The 2-norm and ∞ -norm were taken at time $t = 2k, k \in \{0, 1, 2, 3, 4, 5\}$ as to measure based on recombinations.

Implementing a non-reflecting boundary condition produces significantly better results than an absorbing layer, especially when considering that the non-reflecting boundary needs neither additional mesh points or an implicit solve. However, it does have certain drawbacks that will force continued consideration of PML as a viable option for further work. First, the non-reflecting boundary condition established here is strictly one dimensional. The idea translates well to higher dimensions, but demands a considerably greater amount of work and based on the literature, will produce lesser results. PML, on the other hand, was initially created for a two dimensional setting and has been well established in three dimensions as well.

Further, the non-reflecting boundary may not be stable under certain perturbations of the problem. When first implementing this code, the mesh points x_i were located at $-L + i * \Delta x$, meaning the domain truly covered $[-L - \Delta x/2, L + \Delta x/2]$. This actually caused the code to diverge as exponential growth was seen at the boundary. When the domain was reduced to $[-L, L]$ and the boundary points were estimated by using the end mesh points $\pm(L - \Delta x/2)$, the code converged. Currently, a Taylor approximation is used for the value at the boundary based upon the two mesh points nearest the boundary.

Table 5.4: Performance of the Dissipative 1st Order Scheme with a Non-reflecting Boundary

<i>Time</i>	2-norm	∞ -norm
0	3.54022	1.00000
2	0.072160	0.013608
4	0.0017082	0.00024755
6	0.000043195	0.00000047568
8	0.0000011799	0.00000012592
10	0.00000039215	0.0000000051711

Chapter 6

MOL^T in Higher Dimensions Utilizing an Alternating Directly Implicit Scheme

The adaptation of the scheme to higher dimensions creates a more difficult, yet more interesting challenge. Further, the MOL^T scheme will be more valuable in higher dimensions as the true known solution for the wave equation is considerably more complex in multiple dimensions. Some of the difficulties that will arise due to higher dimensionality include an exponential growth in the number of mesh points as well as Green's functions integrals that can no longer be evaluated analytically. Additionally, an increase in geometrical complexity is to be expected as higher dimensions are more inclined to non-Cartesian coordinates, such as spherical coordinates for three dimensions.

While there is currently no way to avoid the increase in the number of mesh points, we can avoid the latter difficulties by implementing an Alternating Direct Implicit (ADI) scheme, which will split the problem dimensionally and treat it as a series of one dimensional problems.

6.1 Freespace Solution

Consider the wave equation (1.1) specifically in two dimensions with no applied boundary conditions. That is,

$$\begin{aligned}\frac{\partial^2 v}{\partial t^2} &= c^2 \nabla^2 v, \quad x \in \mathbb{R}^2, \quad t \geq 0, \\ v(x, 0) &= f(x), \quad v_t(x, 0) = g(x).\end{aligned}$$

Once again, the Method of Lines Transpose begins by discretizing the temporal derivative, yielding

$$\frac{v^{n+1} - 2v^n + v^{n-1}}{\Delta t^2} = c^2 \nabla^2 v.$$

This can then be re-arranged into a Helmholtz type equation as

$$\left(\nabla^2 - \alpha^2 \right) v = \frac{-2v^n + v^{n-1}}{c^2 \Delta t^2}, \quad \alpha = \frac{1}{c \Delta t}. \quad (6.1)$$

From here, note that this Helmholtz operator may be split dimensionally as

$$\begin{aligned}\nabla^2 - \alpha^2 &= \frac{\partial}{\partial x} \frac{\partial}{\partial x} + \frac{\partial}{\partial y} \frac{\partial}{\partial y} - \alpha^2 \\ &= \frac{-1}{\alpha^2} \left(\frac{\partial^2}{\partial x^2} - \alpha^2 \right) \left(\frac{\partial^2}{\partial y^2} - \alpha^2 \right) + \frac{1}{\alpha^2} \frac{\partial^2}{\partial x^2} \frac{\partial^2}{\partial y^2} \\ &= \frac{-1}{\alpha^2} \left(\frac{\partial^2}{\partial x^2} - \alpha^2 \right) \left(\frac{\partial^2}{\partial y^2} - \alpha^2 \right) + O\left(\Delta t^2\right).\end{aligned} \quad (6.2)$$

While this dimensional split will introduce an $O\left(\Delta t^2\right)$ error if we drop the $\frac{1}{\alpha^2} \frac{\partial^2}{\partial x^2} \frac{\partial^2}{\partial y^2}$ term,

this is acceptable as none of the schemes we are implementing are of order greater than 2.

While this may affect the accuracy of the scheme, it will not affect the order of the scheme.

We then take the split operator (6.2) and substitute it into equation (6.1) to obtain

$$\frac{-1}{\alpha^2} \left(\frac{\partial^2}{\partial_x^2} - \alpha^2 \right) \left(\frac{\partial^2}{\partial_y^2} - \alpha^2 \right) v^{n+1} = \frac{-2v^n + v^{n-1}}{c^2 \Delta t^2}.$$

This may then be arranged as the two step system of Helmholtz equations

$$\mathcal{L}_x \mathcal{L}_y v^{n+1} = -\alpha^2 \frac{-2v^n + v^{n-1}}{c^2 \Delta t^2}, \quad (6.3)$$

where

$$\mathcal{L}_x = \frac{\partial^2}{\partial_x^2} - \alpha^2 \text{ and } \mathcal{L}_y = \frac{\partial^2}{\partial_y^2} - \alpha^2.$$

Inverting the operators in order in equation (6.3) using our MOL^T solver gives

$$\begin{aligned} \mathcal{L}_y v^{n+1}(x, y) &= \int -\alpha^2 \frac{-2v^n(z) + v^{n-1}(z)}{c^2 \Delta t^2} e^{-\alpha|x-z|} dz \\ \Rightarrow v^{n+1}(x, y) &= \int \left(\int -\alpha^2 \frac{-2v^n(z) + v^{n-1}(z)}{c^2 \Delta t^2} e^{-\alpha|x-z|} dz \right) e^{-\alpha|y-z|} dz \\ &= \alpha^4 \int \left(\int \left(2v^n(z) - v^{n-1}(z) \right) e^{-\alpha|x-z|} dz \right) e^{-\alpha|y-z|} dz. \end{aligned}$$

This is the method for which we will implement the first order dissipative method. In

exactly the same fashion, we may calculate the second order dissipative scheme, given by

$$v^{n+1}(x, y) = \alpha^4 \int \left(\int \left(\frac{5}{4}v^n(z) - v^{n-1}(z) + \frac{1}{4}v^{n-2}(z) \right) e^{-\alpha|x-z|} dz \right) e^{-\alpha|y-z|} dz$$

and the second order non-dissipative scheme, given by

$$v^{n+1}(x, y) = -v^{n-1}(x, y) + \alpha^4 \int \left(\int 2v^n(z) e^{-\alpha|x-z|} dz \right) e^{-\alpha|y-z|} dz.$$

To minimize any error that may arise from always propagating first in one direction, we will solve the problem twice, changing only the order of the dimensions. That is, find numerical solutions to both

$$\begin{aligned} \mathcal{L}_x \mathcal{L}_y v^{n+1} &= f(v^n, v^{n-1}, v^{n-2}) \quad \text{and} \\ \mathcal{L}_y \mathcal{L}_x v^{n+1} &= f(v^n, v^{n-1}, v^{n-2}) \end{aligned}$$

and take $v^{n+1}(x, y)$ to be the average of these two solutions. While this is helpful for two dimensions, solving for all possible orders of dimensions may become cost prohibitive as the number of dimensions increases and $n!$ solutions are needed for a domain in $\mathbb{R}^n$.

6.2 Finite Domain

Thanks to the splitting of the operator in (6.2), enforcing any of the boundary conditions from chapters 4 and 5 is as simple as applying the boundary condition in each step corresponding to a dimension.

For example, to apply one-way wave boundary conditions to the 1st order dissipative

method, apply the temporal integral from 5.2 to each of the dimensions in equation (6.3), which can be rewritten as

$$\tilde{v} = \int \alpha^2 \left(2v^n(z) - v^{n-1}(z) \right) e^{-\alpha|x-z|} dz, \quad \tilde{v} = \left(\frac{\partial^2}{\partial_y^2} - \alpha^2 \right) v^{n+1}, \quad (6.4)$$

$$v^{n+1} = \alpha^2 \int \tilde{v}(z) e^{-\alpha|y-z|} dz. \quad (6.5)$$

Thus, the fully discretized solution for (6.4) for each row $\check{u}$, upon which y is held constant, will be given by

$$\begin{aligned} \check{u} = & \sum_{i=1}^m A_{ji} \left(\check{u}_i^n - \frac{1}{2} \check{u}_i^{n-1} \right) \\ & + e^{-\alpha(L+x_j)} \left(e^{-1} \tilde{S}_{x1}^n + e^{-1} \check{u}^n(-L, y) + \left(1 - \frac{5}{2} e^{-1} \right) \check{u}^{n-1}(-L, y) - \frac{1-2e^{-1}}{2} \check{u}^{n-2}(-L, y) \right) \\ & + e^{-\alpha(L-x_j)} \left(e^{-1} \tilde{S}_{x2}^n + e^{-1} \check{u}^n(L, y) + \left(1 - \frac{5}{2} e^{-1} \right) \check{u}^{n-1}(L, y) - \frac{1-2e^{-1}}{2} \check{u}^{n-2}(L, y) \right), \end{aligned}$$

where $\tilde{S}_{x1}$ and $\tilde{S}_{x2}$ are the natural equivalents to the original definition of $\tilde{S}_1$ and $\tilde{S}_2$ taken at the boundaries $x = -L$ and $x = L$, respectively.

This will then give that the fully discretized solution for (6.5) is

$$\begin{aligned} u^{n+1}(x, y) = & \sum_{i=1}^m A_{ji} \left(\check{u}_i \right) + e^{-\alpha(L+y_j)} \left(e^{-1} \tilde{S}_{y1}^n + e^{-1} u^n(x, -L) \right. \\ & + \left(1 - \frac{5}{2} e^{-1} \right) u^{n-1}(x, -L) - \frac{1-2e^{-1}}{2} u^{n-2}(x, -L) \Big) \\ & + e^{-\alpha(L-y_j)} \left(e^{-1} \tilde{S}_{y2}^n + e^{-1} u^n(x, L) + \left(1 - \frac{5}{2} e^{-1} \right) u^{n-1}(x, L) - \frac{1-2e^{-1}}{2} u^{n-2}(x, L) \right). \end{aligned}$$

Any of the previous boundary conditions may be applied to any of the three schemes already presented in a similar manner.

An example of the first order dissipative method conjoined with ADI with the one way wave boundary condition of section 5.2 is shown in figures 6.1 - 6.3. The initial pulse was given by $f(x, y) = e^{-5(x^2+y^2)}$ and allowed to propagate outward at a speed of $c = 1$. The wave was then tracked at times $t = \{0, 0.5, 1.25\}$.

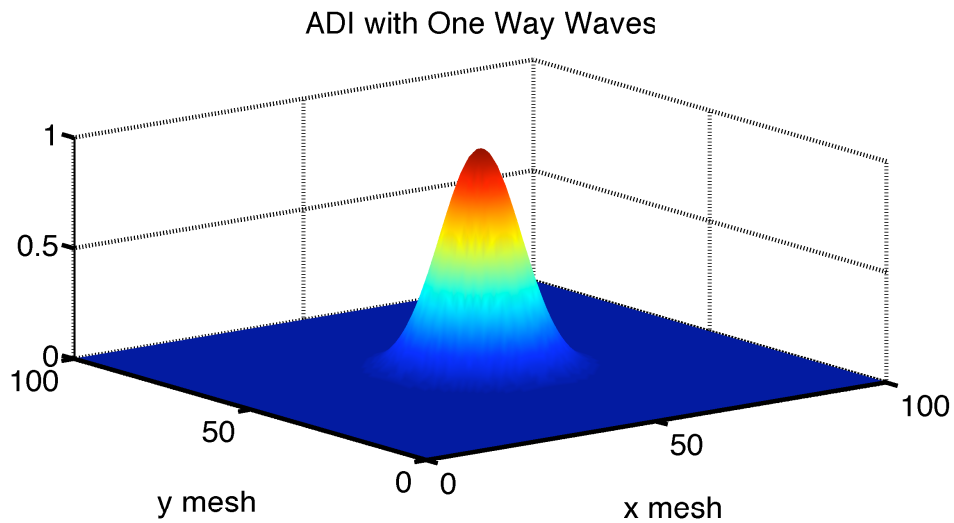


Figure 6.1: Propagation of the First Order Dissipative ADI Scheme at $t = 0$

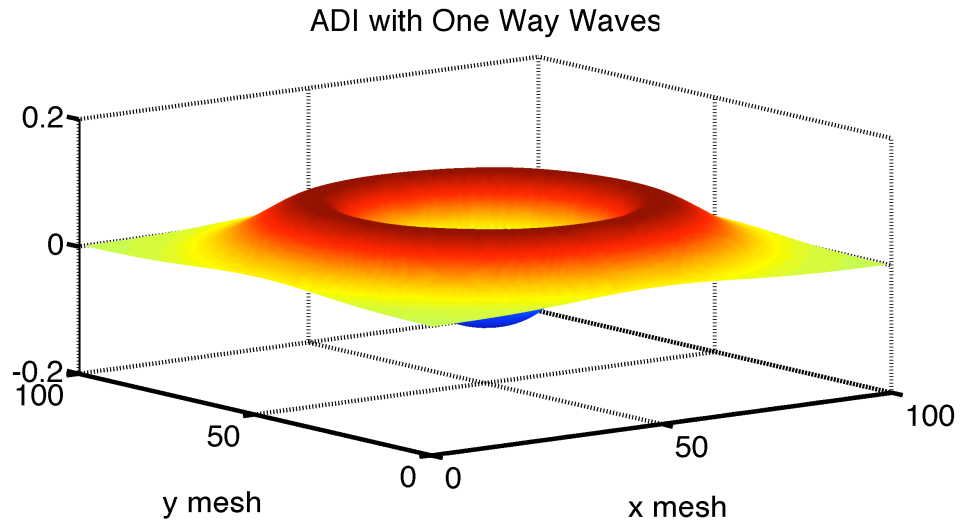


Figure 6.2: Propagation of the First Order Dissipative ADI Scheme at $t = 0.5$

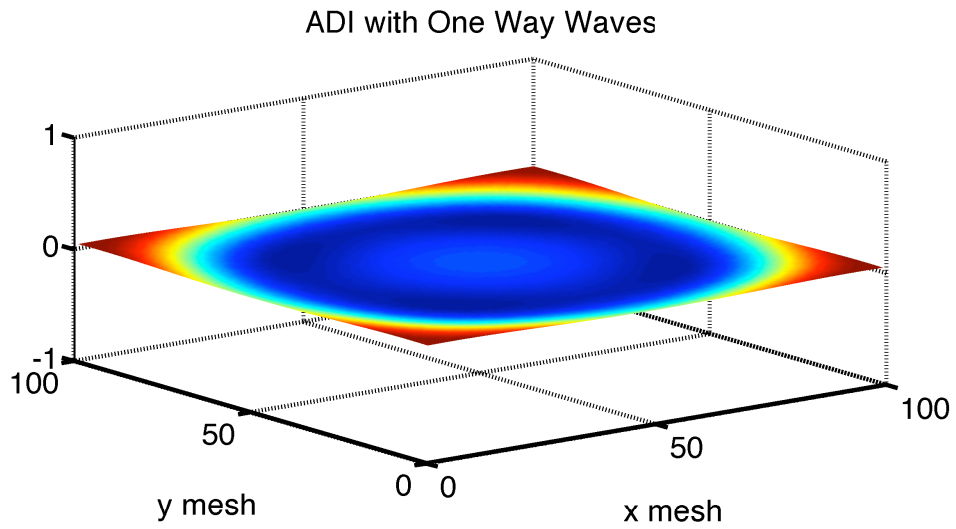


Figure 6.3: Propagation of the First Order Dissipative ADI Scheme at $t = 1.25$

Chapter 7

Soft Sources

When computing the wave equation, a soft source is a prescribed source which is time varying and introduces propagating waves into the medium at a given spatial location. The source is said to be soft since, while it will emit waves at that given location, the source does not interact with incoming waves and thus creates no reflection. To implement a soft source for the one dimensional wave equation, begin with the equation (1.1) with zero initial conditions which satisfies the soft source condition

$$v(x_s, t) = S(t). \quad (7.1)$$

This soft source condition is not a boundary condition as the domain is free to extend to the left and right of x_s and waves may freely travel in either direction. Although we are interested in time harmonic sources, this formulation will be for a general temporal function $S(t)$. We wish to understand how this problem might be reformulated by

$$v_{tt} = c^2 v_{xx} + Q(x, t), \quad t > 0, \quad -\infty < x < \infty \quad (7.2)$$

with zero initial conditions, without explicitly enforcing condition (7.1). That is, we would like to find a source term $Q(x, t)$ that ensures $v(x_s, t) = S(t)$ will hold. We proceed by constructing the Laplace solution to equation (1.1) with the soft source condition (7.1) and equation (7.2) independently and equating the results. After applying the Laplace transform to the first equation, it is easily seen that

$$\left(\frac{s}{c}\right)^2 V = V_{xx}, \quad -\infty < x < \infty,$$

and after imposing $V(x_s, s) = \hat{S}(s)$, along with proper decay at ∞ , this produces

$$V(x, s) = \hat{S}(s) e^{-\frac{s}{c}|x-x_s|}. \quad (7.3)$$

Likewise, after applying the Laplace transform to the second equation,

$$\left(\frac{s}{c}\right)^2 V = V_{xx} + \hat{Q}(x, s), \quad -\infty < x < \infty,$$

which can be solved using the Green's function for the Helmholtz operator. Thus

$$V(x, s) = \frac{c}{2s} \int_{-\infty}^{\infty} \hat{Q}(y, s) e^{-\frac{s}{c}|x-y|} dy. \quad (7.4)$$

We now impose upon $\hat{Q}(y, s)$ that equations (7.3) and (7.4) be equal, yielding

$$\frac{c}{2s} \int_{-\infty}^{\infty} \hat{Q}(y, s) e^{-\frac{s}{c}|x-y|} dy = \hat{S}(s) e^{-\frac{s}{c}|x-x_s|}.$$

This may be solved by separating the variables in $\hat{Q}(y, s)$ as $\frac{2s}{c}\hat{S}(s)g(y)$. This will require that the undetermined function $g(y)$ satisfies

$$\int_{-\infty}^{\infty} g(y) e^{-\frac{s}{c}|x-y|} dy = e^{-\frac{s}{c}|x-x_s|}.$$

While the function $g(y)$ may not be unique, there is the natural choice of

$$g(y) = \delta(y - x_s),$$

which ensures that the solutions to equations (7.3) and (7.4) are equal. This gives that

$$\hat{Q}(x, s) = \frac{2s}{c}\hat{S}(s)\delta(x - x_s)$$

or, after taking the inverse Laplace transform,

$$Q(x, t) = \frac{2}{c}S'(t)\delta(x - x_s).$$

Therefore, solving the wave equation (1.1) with the soft source condition (7.1) is equivalent to solving the problem

$$u_{tt} = c^2 u_{xx} + \frac{2}{c}S'(t)\delta(x - x_s), \quad t > 0, -\infty < x < \infty. \quad (7.5)$$

7.1 Implementation

The implementation of a soft source based on equation (7.5) follows in a manner similar to the methods already presented. For the first order dissipative scheme, first discretize the temporal derivative and then rearrange to form the Helmholtz equation

$$\left(\partial_{xx} - \frac{1}{(c\Delta t)^2} \right) u^{n+1} = \frac{-2u^n(x) + u^{n-1}(x)}{(c\Delta t)^2} + \frac{2}{c^3} S'(t) \delta(x - x_s).$$

Upon inversion of the Helmholtz operator, we have

$$\begin{aligned} u^{n+1} &= \alpha \int \left(u^n(y) - \frac{1}{2} u^{n-1}(y) \right) e^{-\alpha|x-y|} dy - \int \frac{1}{\alpha c^3} S'(t) \delta(y - x_s) e^{-\alpha|x-y|} dy \quad (7.6) \\ &= \alpha \int \left(u^n(y) - \frac{1}{2} u^{n-1}(y) \right) e^{-\alpha|x-y|} dy - \frac{1}{\alpha c^3} S'(t) e^{-\alpha|x-x_s|}. \end{aligned}$$

Some care should be taken in choosing the source function $S(t)$. Ideally, we would prefer a source term that creates a wave and thus should be some combination of sin and cos terms. Additionally, we would like the source to initially be off, i.e., $S(0) = 0$ and thus, we will consider source terms of the form $S(t) = \sin(\omega(t - (x - x_s)))$. Further, we would like to gradually introduce the source term in order to limit numerical error in the initial introduction of the term; therefore, the source term will look like

$$S(t) = E(\sigma t) \sin(\omega(t - (x - x_s))),$$

where $E(t)$ satisfies $E(0) = 0, E(\infty) = 1$ and $E \in \mathbb{C}^\infty$. A good example of this is to choose $E(t) = \tanh(t)$. The constants σ and ω are chosen to ensure that $E(\sigma t)$ and $\sin(\omega(t - (x - x_s)))$ have different frequencies. Finally, if we wish to turn the signal off,

we can multiply $S(t)$ by $E_2(t)$, where $E_2(t)$ satisfies $E_2(0) = 1, E_2(\infty) = 0$ and $E_2 \in \mathbb{C}^\infty$.

On a finite domain, we may include boundary conditions by simply adding in the boundary correction term, as was done in section 6.2. However, if we choose to utilize boundary conditions, it makes more sense to include outflow boundary conditions as to be able to avoid the continuous buildup of reflected waves that would occur with non-outflow boundary conditions.

The implementation of a soft source may be seen in figure 7.1. It simulates the first order dissipative scheme with the one way wave boundary conditions from section 5.2. The wave is created in the center of the domain using the source term $S(t) = \sin(\pi t) \tanh(\pi/4)$. The waves shown were taken at times $t = \{1, 3, 5\}$ and show the general introduction of the wave. The long term behavior, showing the waves flowing out of the domain is displayed in figure 7.2.

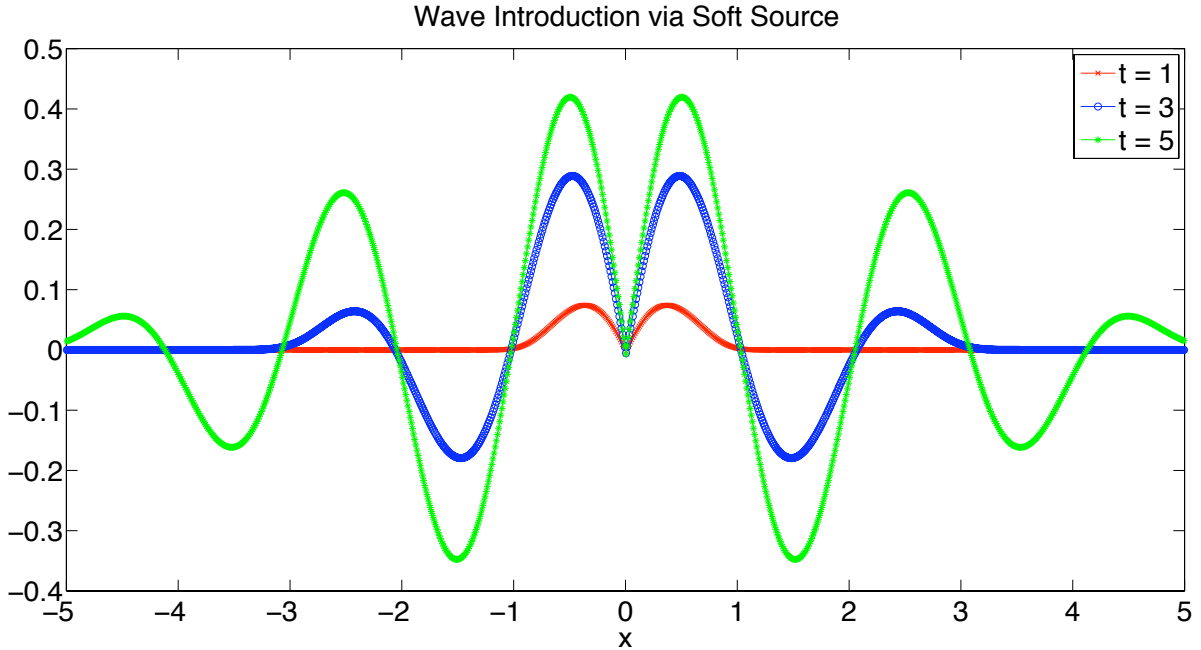


Figure 7.1: Wave Introduction via Soft Source

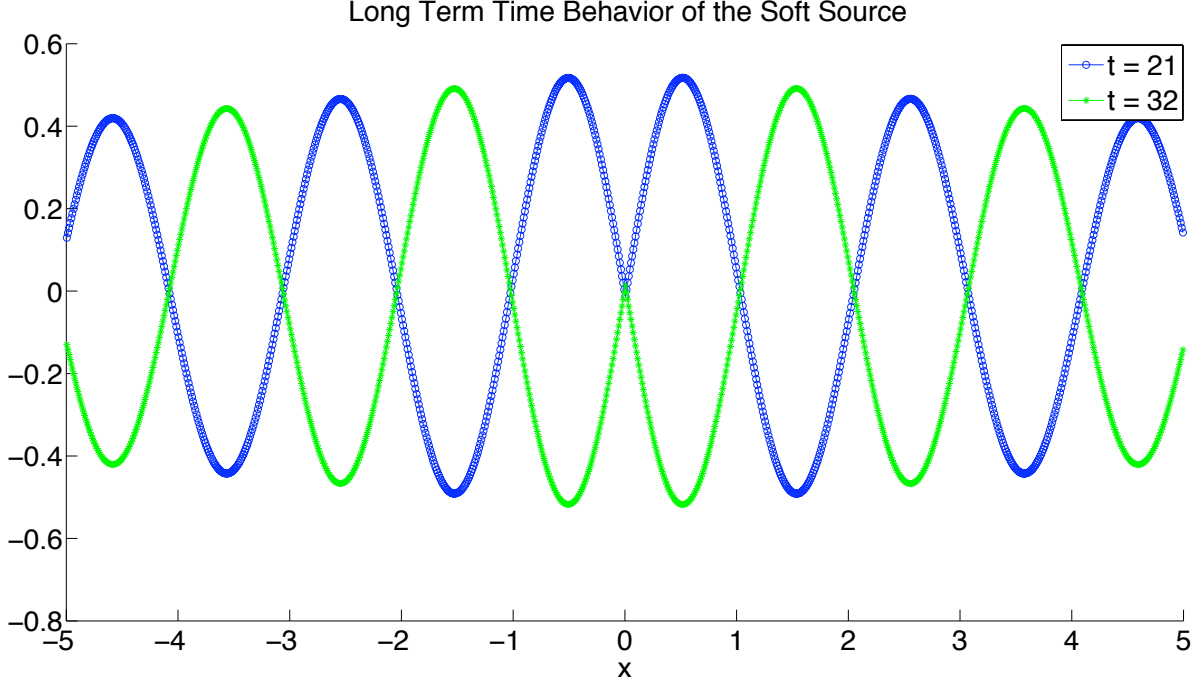


Figure 7.2: Long Term Time Behavior of the Soft Source

7.2 Higher Dimensions

When we moved the MOL^T solver to a higher dimension, ADI was incorporated to split the problem dimensionally. Thus, the natural choice for creating a soft source in a domain encapsulating more than one dimension would be to split the source dimensionally as well. The latter term from equation (7.6) will be applied to each dimension as it is split in the ADI framework.

The simplest case of the implementation of a two-dimensional soft source in a dimensionally split environment is given as such. Begin with the free-space split problem, given by equation (6.3) as

$$\mathcal{L}_x \mathcal{L}_y v^{n+1} = -\alpha^2 \frac{-2v^n + v^{n-1}}{c^2 \Delta t^2}.$$

In the first step, invert the operator in the x -dimension, $\mathcal{L}_x$ as well as propagating the wave emanating from my source, as given by equation (7.6). This results in

$$w(x, y) = \int_{\Omega} \alpha^3 \frac{-2v^n(z, y) + v^{n-1}(z, y)}{2} e^{-\alpha|x-z|} dz - \frac{1}{\alpha c^3} S'(t) e^{-\alpha|x-x_s|},$$

$$w(x, y) = \mathcal{L}_y v^{n+1}(x, y).$$

Repeating the process in the y -dimension will then give that

$$v^{n+1}(x, y) = \int_{\Omega} -\frac{w(x, z)}{2} e^{-\alpha|y-z|} dz - \frac{1}{\alpha c^3} S'(t) e^{-\alpha|y-y_s|}.$$

As in section 6.1, it may be useful to solve the problem twice, varying the dimension order and averaging the solution to minimize any dimensional order error.

Chapter 8

Future Directions

8.1 Further Problems

As the implementation of d'Alembert's solution to the one-dimensional wave equation is fairly straight-forward, more interesting problems lie in higher dimensions. The results that have been proved here are useful as a basis for extending the work already done in one dimension to higher dimensions. We have used an ADI scheme to dimensionally split the problem and lift the numerical solution from one dimension to two dimensions. However, this approach may not be practical for interesting problems in plasma physics, which will likely require the six dimensions of space and velocity-space. The usefulness of the MOL^T method may be determined by how well it can be scaled for use in a higher dimensional setting.

We would like to apply this work and determine its usefulness in solving other PDEs.

The most natural extension would be to the heat equation,

$$v_t = c\nabla^2 v, \quad v(x, 0) = f(x).$$

While the Method of Lines Transpose idea has been used to numerically solve the heat equation [51], the convergence of the method was never proven. The extension of the proofs of consistency and stability should be fairly straight forward to extend from the wave equation to the heat equation.

One of the factors impeding the practicality of the Method of Lines Transpose scheme is its limitation to second order. The natural extensions to a third order dissipative scheme and a third and fourth order purely dispersive scheme have also been investigated. Alas, these methods were not feasible due to stability issues. For the third order dissipative scheme, one of the eigenvalues of the matrix A_t was slightly larger than one, leading to a long-term instability. Due to its symmetric temporal structure, the fourth order purely dispersive scheme required that all the eigenvalues of A_t multiply to one. Numerical testing showed two of the eigenvalues to be complex with norm equal to one, but unfortunately, the other two were shown to be real and of different magnitude, causing the method to be completely unstable. The issue with the third order purely dispersive scheme is slightly different. The third order approximation to the temporal derivative v_{tt} about the point $t^{n-\frac{1}{2}}$ is given by

$$v_{tt} = \frac{v^{n+1} - v^n - v^{n-1} + v^{n-2}}{2}.$$

A centered third order equivalence for v_{xx} is given by

$$v_{xx} = \frac{-v_{xx}^{n+1} + 9v_{xx}^n + 9v_{xx}^{n-1} - v_{xx}^{n-2}}{16}.$$

Thus, rearranging the wave equation (1.1) into a Helmholtz type equation would give

$$\begin{aligned}\mathcal{L}\left(v^{n+1} - v^n - v^{n-1} + v^{n-2}\right) &= 8v_{xx}^n + 8v_{xx}^{n-1}, \\ \mathcal{L}(v) &= \left(\partial_{xx} + \frac{8}{(c\Delta t)^2}\right)v.\end{aligned}$$

Due to the Helmholtz operator having a positive sign, the local nature of the inverted operator is destroyed as the corresponding Green's function now becomes complex and oscillatory instead of a negative exponential, yielding this as an unfeasible scheme.

Defect correction schemes [20, 23, 24, 31, 57] may provide a way to overcome these order constraints. In a defect correction setting, the differential equation is originally solved with a low order based scheme. An integral equation for the error is then determined, yielding a correction term which is then applied to the solution, which in turn will raise the order of the numerical solution. These methods would have to be adapted to fit this second order partial differential equation, but may provide an avenue to obtaining higher order numerical solutions for the wave equation using the MOL^T as its base scheme.

8.2 Numerical Work

In this dissertation, we have discussed the advantages of implementing the MOL^T scheme for the wave equation. As we have the ability to take a large CFL, we expect to be able

to outperform standard explicit methods, however, a comparison on the accuracy and computational complexity of this method against standard finite difference methods and finite element methods should be done.

Additionally, some of the proofs of the numerical properties could be more completely solved. Namely, in section 3.2, we showed the dispersion relationship for the semi-discrete methods. Attempts have been made to show the dispersion relation for the fully discrete methods, but to this point, have been unsuccessful in showing that the second order dissipative scheme actually achieves second order. This is something that we hope to be able to show in the future. Also, the stability bounds from section 4.2 for the finite domain schemes can be tightened down. All of the numerical tests that we have run for the eigenvalues of the matrix $A_t = A_p + A_h$, defined in section 4.2, have shown the eigenvalues to be positive and within the predetermined bounds. The expectation is that the method is fully stable and we would like to find a way to prove this, getting rid of the dependence on the Bauer-Fike theorem to determine a region on which stability is guaranteed.

BIBLIOGRAPHY

Bibliography

- [1] S. Abarbanel, D. Gottlieb, and J.S. Hesthaven. Well-posed perfectly matched layers for advective acoustics. *J. Comput. Phys.*, 154(2):266–283, September 1999.
- [2] S. Abarbanel, D. Gottlieb, and J.S. Hesthaven. Long time behavior of the perfectly matched layer equations in computational electromagnetics. *J. Sci. Comput.*, 17(1-4):405–422, 2002.
- [3] Saul Abarbanel and David Gottlieb. A mathematical analysis of the pml method. *J. Comput. Phys.*, 134:357–363, 1997.
- [4] Saul Abarbanel and David Gottlieb. On the construction and analysis of absorbing layers in cem. *Appl. Num. Math.*, 27:331–340, 1998.
- [5] Bradley Alpert, Leslie Greengard, and Thomas Hagstrom. An integral evolution formula for the wave equation. *J. Comput. Phys.*, 162:536–543, 2000.
- [6] Bradley Alpert, Leslie Greengard, and Thomas Hagstrom. Nonreflecting boundary conditions for the time-dependent wave equation. *J. Comput. Phys.*, 180:270–296, 2002.
- [7] Daniel Appelo, Thomas Hagstrom, and Gunilla Kreiss. Perfectly matched layers for hyperbolic systems: general formulation, well-posedness and stability. *J. Appl. Math.*, 67(1):1–23, 2006.
- [8] Uri M. Ascher, Robert M. M. Mattheij, and Robert D. Russell. *Numerical solution of boundary value problems for ordinary differential equations*, volume 13 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1995. Corrected reprint of the 1988 original.
- [9] Sergey Asvadurov, Vladimir Druskin, Murthy Guddati, and Leonid Knizhnerman. On optimal finite-difference approximation of pml. *J. Numer. Anal.*, 41(1):287–305, 2003.
- [10] Peter Bartello and Stephen Thomas. The cost-effectiveness of semi-lagrangian advection. *Mon. Weather Rev.*, 124:2883–2924, 1996.
- [11] F.L. Bauer and E.T. Fike. Norms and exclusion theorems. *Numer. Math*, 2:137–144, 1960.
- [12] E. Becache, S. Fauqueux, and P. Joly. Stability of perfectly matched layers, group velocities and anisotropic waves. *J. Comput. Phys.*, 188:399–433, 2003.

- [13] Eliane Becache and Patrick Joly. On the analysis of berenger’s perfectly matched layers for maxwell’s equations. *Mathematical Modelling and Numerical Analysis*, 36(1):87–119, 2002.
- [14] Eliane Becache, Peter Petropoulos, and Stephen Gedney. On the long-time behavior of unsplit perfectly matched layers. *IEEE Trans. Antennas and Propagation*, 52(5):1335–1342, 2004.
- [15] Jean-Pierre Berenger. A perfectly matched layer for the absorption of electromagnetic waves. *J. Comput. Phys.*, 114:185–200, 1994.
- [16] Jean-Pierre Berenger. Three-dimensional perfectly matched layer for the absorption of electromagnetic waves. *J. Comput. Phys.*, 127:363–379, 1996.
- [17] Jean-Pierre Berenger. Improved pml for the fdtd solution of wave-structure interaction problems. *IEEE Trans. Antennas and Propagation*, 45(3):466–473, March 1997.
- [18] A. Bermudez, L. Hervella-Nieto, A. Prieto, and R. Rodriguez. An exact bounded perfectly matched layer for time-harmonic scattering problems. *J. Sci. Comput.*, 30, 2007.
- [19] L. Bonaventura. An introduction to semi - lagrangian methods for geophysical scale flows. Lecture Notes, ERCOFTAC Leonhard Euler Lectures, SAM-ETH Zurich, 2004.
- [20] Anne Bourlioux, Anita Layton, and Michael Minion. High-order multi-implicit spectral deferred correction methods for problems of reactive flow. *J. Comput. Phys.*, 189(2):651–675, 2003.
- [21] Roman Chapko and Rainer Kress. Rothe’s method for the heat equation and boundary integral equations. *J. Integral Equations and Appl.*, 9:47–69, 1997.
- [22] H. Cheng, L. Greengard, and V. Rokhlin. A fast adaptive multipole algorithm in three dimensions. *J. Comput. Phys.*, 155(2):468–498, 1999.
- [23] Andrew Christlieb, Benjamin Ong, and Jing-Mei Qiu. Comments on high order integrators embedded within integral deferred correction methods. *Comm. Appl. Math and Comp. Sci.*, 4(1):27–56, 2009.
- [24] Andrew Christlieb, Benjamin Ong, and Jing-Mei Qiu. Integral deferred correction methods constructed with high order runge-kutta integrators. *Math. Comp.*, 79(270):761–783, 2010.
- [25] R. Coifman, V. Rokhlin, and S. Wandzura. The fast multipole method for the wave equation: A pedestrian prescription. *IEEE Trans. Antennas and Propagation*, 35(3):7–12, 1993.
- [26] F. Collino and P.B. Monk. Optimizing the perfectly matched layer. *Comput. Methods Appl. Mech. Engrg.*, 164:157–171, 1998.

- [27] Richard Courant, Eugene Isaacson, and Mina Rees. On the solution of nonlinear hyperbolic differential equations by finite differences. *Communications on Pure and Applied Mathematics*, 5:243–255, 1952.
- [28] Richard Courant and Peter Lax. On nonlinear partial differential equations with two independent variables. *Communications on Pure and Applied Mathematics*, 2:255–273, 1949.
- [29] Stephen Davis. A rotationally biased upwind difference scheme for the euler equations. *J. Comput. Phys.*, 56:65–92, 1984.
- [30] J. Diaz and P. Joly. A time domain analysis of pml models in acoustics. *Comput. Methods Appl. Mech. Engrg.*, 195:3820–3852, 2006.
- [31] Alok Dutt, Leslie Greengard, and Vladimir Rokhlin. Spectral deferred correction methods for ordinary differential equations. *BIT*, 40(2):241–266, 2000.
- [32] Bjorn Engquist and Andrew Majda. Absorbing boundary conditions for the numerical simulation of waves. *Math. Comp.*, 31(139):629–651, 1977.
- [33] Bjorn Engquist and Andrew Majda. Radiation boundary conditions for acoustic and elastic wave calculations. *Commun. Pure Appl. Math*, 32:313–358, 1979.
- [34] Maruizio Falcone and Roberto Ferretti. Convergence analysis for a class of high-order semi-lagrangian advection schemes. *SIAM J. Numer. Anal.*, 35(3):909–940, June 1998.
- [35] M.G.G. Foreman. An accuracy analysis of boundary conditions for the forced shallow water equations. *J. Comput. Phys.*, 64:334–367, 1986.
- [36] Zydrunas Gimbutas and Vladimir Rokhlin. A generalized fast multipole method for nonoscillatory kernels. *SIAM J. Sci. Comput.*, 24(3):796–817 (electronic), 2002.
- [37] Dan Givoli. Non-reflecting boundary conditions. *J. Comput. Phys.*, 94(1):1–29, 1991.
- [38] S.K. Godunov. A difference scheme for numerical computation of discontinuous solutions of equations of fluid dynamics. *Math. Sbornik*, 47:271–306, 1959.
- [39] A. Greenbaum, L. Greengard, and G.B. McFadden. Laplace’s equation and the dirichlet-neumann map in multiply connected domains. *J. Comput. Phys.*, 105:267–278, 1993.
- [40] L. Greengard, M.C. Kropinski, and A. Mayo. Integral equation methods for stokes flow and isotropic elasticity in the plane. *J. Comput. Phys.*, 125(2):403–414, 1996.
- [41] L. Greengard and V. Rokhlin. A fast algorithm for particle simulations. *J. Comput. Phys.*, 73(2):325–348, 1987.
- [42] Marcus Grote and Imbo Sim. Efficient pml for the wave equation. *Preprint*, 2010.
- [43] Laurence Halpern. Absorbing boundary conditions for the discretization schemes of the one-dimensional wave equation. *Math. Comp.*, 38(158):415–429, 1982.

- [44] Laurence Halpern and Lloyd Trefethen. Wide-angle one-way wave equations. *J. Acoust. Soc. Amer.*, 84:1397–1404, 1988.
- [45] Amiram Harten, Peter Lax, and Bram Van Leer. On upstream differencing and godunov-type schemes for hyperbolic conservation laws. *SIAM Review*, 25(1):34–61, 1983.
- [46] U. Hornung. A parabolic-elliptic variational inequality. *Manuscripta Math.*, 39(2):155–172, 1982.
- [47] Fang Hu. On absorbing boundary conditions for linearized euler equations by a perfectly matched layer. *J. Comput. Phys.*, 129:201–219, 1996.
- [48] J. Jia and J. Huang. Krylov deferred correction accelerated method of lines transpose for parabolic problems. *J. Comput. Phys.*, 227(3):1739–1753, 2008.
- [49] D.J Jones, J.C South Jr., and E.B Klunker. On the numerical solution of elliptic partial differential equations by the method of lines. *J. Comput. Phys.*, 9(3):496–527, 1972.
- [50] J. Kacur. *Method of Rothe in evolution equations*, volume 1192. Springer Berlin/Heidelberg, 1986.
- [51] Mary Catherine Kropinski and Bryan Quaife. Fast integral equation methods applied for rothe’s method applied to the isotropic heat equation. *Comp. Math. Appl.*
- [52] Peter Lax. Nonlinear hyperbolic equations. *Communications on Pure and Applied Mathematics*, 6:231–258, 1953.
- [53] M. Lentine, J.T. Grotzinger, and R. Fedkiw. An unconditionally stable fully conservative semi-lagrangian method. *J. Comput. Phys.*, 230:2857–2895, 2011.
- [54] Randall J. LeVeque. *Finite Volume Methods for Hyperbolic Problems*. Cambridge University Press, 2002.
- [55] Peijun Li, Hans Johnston, and Robert Krasny. A Cartesian treecode for screened Coulomb interactions. *J. Comput. Phys.*, 228(10):3858–3868, 2009.
- [56] A. Mazzia and F. Mazzia. High-order transverse schemes for the numerical solution of PDEs. *J. Comput. Appl. Math.*, 82(1-2):299–311, 1997.
- [57] Michael Minion. Semi-implicit spectral deferred correction methods for ordinary differential equations. *Commun. Math Sci.*, 1(3):471–500, 2003.
- [58] I.M. Navon, B. Neta, and M.Y. Hussaini. A perfectly matched layer formulation for the linearized shallow water equation models: The split equation approach. *Monthly Weather Review*, 132(6):1369–1378, 2004.
- [59] Haim Nessyahu, Eitan Tadmor, and Tamir Tassa. The convergence rate of godunov type schemes. *SIAM J. Numer. Anal.*, 31(1):1–16, 1994.

- [60] Satish Reddy and Lloyd Trefethen. Stability of the method of lines. *Numer. Math.*, 62:235–267, 1992.
- [61] V. Rokhlin. Rapid solution of integral equations of scattering theory in two dimensions. *J. Comput. Phys.*, 86:414–439, 1990.
- [62] J.S. Sawyer. A semi-lagrangian method of solving the vorticity advection equation. *Tellus*, 15(4):336–342, 1963.
- [63] Andrew Selle, Ronald Fedkiw, ByungMoon Kim, Yingjie Liu, and Jarek Rossignac. An unconditionally stable maccormack method. *J. Sci. Comput.*, 35:350–371, 2008.
- [64] A. Sommerfeld. *Lectures on Theoretical Physics*. Academic Press, New York, 1964.
- [65] Eric Sonnendrucker, Jean Roche, Pierre Bertrand, and Alain Ghizzo. The semi-lagrangian method for the numerical resolution of the vlasov equation. *J. Comput. Phys.*, 149:201–220, 1999.
- [66] A. Staniforth and J. Cote. Semi-lagrangian integration schemes for atmospheric models - a review. *Mon. Weather Rev.*, 119:2206–2223, 1991.
- [67] John Strikwerda. *Finite Difference Schemes and Partial Differential Equations*, volume 2. Springer, 2004.
- [68] Lloyd Trefethen and Laurence Halpern. Well-posedness of one-way wave equations and absorbing boundary conditions. *Math. Comp.*, 47(176):421–435, 1986.
- [69] J. Verwer and J.M. Sanz-Serna. Convergence of method of lines approximations to partial differential equations. *Computing*, 33(3-4):297–313, 1984.
- [70] Wen-Chyuan Yueh and Sui Sun Cheng. Explicit eigenvalues and inverses of tri-diagonal toeplitz matrices with four perturbed corners. *ANZIAM*, pages 361–387, 2008.
- [71] A. Zafarullah. Application of the method of lines to parabolic partial differential equations with error estimates. *J. ACM*, 17:294–302, 1970.
- [72] Li Zhao and Andreas Cangellaris. A general approach for the development of unsplit-field time-domain implementations of perfectly matched layers for fdtd grid truncation. *IEEE Microwave and Guided Wave Letters*, 6(5):209–211, 1996.