8-27-2012

Numerical and Analytical Studies of Electromagnetic Waves: Hermite Methods, Supercontinuum Generation, and Multiple Poles in the SEM

Xi Chen

Recommended Citation

_____

Candidate

_____

Department

This dissertation is approved, and it is acceptable in quality and form for publication:

Approved by the Dissertation Committee:

_____ , Chairperson

_____

_____

_____

_____

_____

_____

_____

Numerical and Analytical Studies of Electromagnetic Waves: Hermite Methods, Supercontinuum Generation, and Multiple Poles in the SEM

by

Xi Chen

B.S., Mathematics, East China Normal University, China, 2001

M.S., Applied Mathematics, University of New Mexico, 2006

DISSERTATION

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

Mathematics

The University of New Mexico

Albuquerque, New Mexico

July, 2012

Dedication

This is dedicated to my family, specially to my wife Yan (Sindy) Qiu, who always believe in me.

To my son, Allan and Evan who come into the world just at the right time.

To my parents, who help me at the time I need it most.

To all of my friends for all kinds of support and great time we spend together.

Acknowledgments

I acknowledge Dr. Daniel Appelö, my advisor and dissertation chair, for his tremendous amount of help, without him, the manuscript can not be done in time.

Special thanks to my advisor, Dr. Thomas Hagstrom, for his continuously support of my work, valuable recommendations and never giving me up.

I thank Dr. Carl E Baum, for helping me complete the electromagnetic scattering part of the manuscript. I also thank Dr. Edl Schamiloglu for supporting me finish the work.

I thank Dr. Alejandro B. Aceves for support me accomplish the supercontinuum generation part of the manuscript. I also thank Michelle Hummel for running the numerical simulations of the supercontinuum generation.

To my family and friends who give me immeasurable support over the years.

Numerical and Analytical Studies of Electromagnetic Waves: Hermite Methods, Supercontinuum Generation, and Multiple Poles in the SEM

by

Xi Chen

B.S., Mathematics, East China Normal University, China, 2001

M.S., Applied Mathematics, University of New Mexico, 2006

PhD., Mathematics, University of New Mexico, 2012

Abstract

The dissertation consists of three parts: Hermite methods, scattering from a lossless sphere, and analysis of supercontinuum generation.

Hermite methods are a new class of arbitrary order algorithms to solve partial differential equations (PDE). In the first chapter, we discuss the fundamentals of Hermite methods in great detail. Hermite interpolation is discussed as well as the different time evolution schemes including Hermite-Taylor and Hermite-Runge-Kutta schemes. Further, an order adaptive Hermite method for initial value problems is described. Analytical studies and numerical simulations in both 1D and 2D are presented. To handle geometry, a hybrid Hermite discontinuous Galerkin methods is introduced. A Discontinuous Galerkin method is used next to the boundaries to handle the geometry and the boundary conditions, while a Hermite method is used in the interior of the computation domain to enhance the performance. Numerical

simulations of 1D wave propagation and the solutions to 2D Maxwell's TM equations are presented along with performance and accuracy data.

In the second chapter, we study the scattering problem concerning the scattering poles from a lossless sphere for both acoustic and electromagnetic waves. We show that in certain cases there exist only first order scattering poles, but in some other cases, arbitrary order scattering poles can be found by imposing certain lossless impedance boundary conditions on the spherical scatterer. A method to construct arbitrary order scattering poles is discussed. The impedance loading function is required to satisfy Foster's theorem so that the scattering problem is lossless.

In the last chapter, we analyse the generation of supercontinua in photonic crystal fibers. We depart from the commonly used approach where a Taylor series expansion of the propagation constant is used to model the dispersive properties in a generalized nonlinear Schrödinger equation (gNLSE). Instead, we develop a mathematical model starting from numerically calculated group velocity dispersion (GVD) curves. Then, we construct a certain function over a broad frequency window and integrate the gNLSE in a way so that the spectral dependence of the propagation constant is preserved. We found that the generation of broadband supercontinua in air-silica microstructured fibers results from a delicate balance of dispersion and nonlinearity. Numerical simulations show that if the nonlinear self-steepening is strong enough, the model produces a shock that is not arrested by dispersion, whereas for weaker nonlinearity the pulse propagates the full extent of the fiber with the generation of a supercontinuum.

Contents

Contents

Contents

List of Figures

List of Tables

Chapter 1

Hermite methods

1.1 Introduction

Hermite methods are a new class of arbitrary order polynomial based spectral methods to solve partial differential equations (PDE). They are different from other nodal based algorithms, as Hermite methods carry and evolve not only the nodal values but the derivatives as well. Hermite methods have a number of unique properties. It converges at the rate $2m + 1$ when degree m polynomials are used in each direction. For constant coefficient equations, the error in time converges at the rate of $2m + 1$ when Taylor time stepping is used. As long as the Courant-Friedrichs-Lewy (CFL) condition holds, i.e. the domain of dependence is contained in a computational cell, the algorithm is stable. The time evolution is completely local which yields a high computation/communication ratio, which is ideal for parallel computing.

The first paper analyzing Hermite methods appeared in 2006 [30]. Since then other experiments using the Hermite methods in different areas and the enhancements of the basic methods have become available. The original Hermite method is ideal for smooth waves, but can be applied to advection of discontinuous or piecewise-

discontinuous waves with favourable results as discussed in [10]. Experiments of simulation of compressible flows are shown in [33], where the built-in dissipation is found to be beneficial. Runge-Kutta time stepping Hermite methods for aeroacoustics are discussed in [11, 9], where large 3D parallel computation results are included. The mesh refinement of Hermite methods is discussed in [11]. Analysis of the Hermite methods for hyperbolic-parabolic systems can be found in [71].

In this work, we introduce two new enhancements of the Hermite methods: An order adaptive Hermite method; and hybrid Hermite discontinuous Galerkin methods. The idea of order adaptive Hermite methods is that when extra accuracy is needed at some location, instead of adaptively refining the mesh, we increase the order. On the computational domain the order adaptive Hermite methods uses polynomials of various degrees. A uniform tolerance is chosen to control the overall accuracy.

Although Hermite methods have many great properties, it is not straightforward to handle curved boundaries. Thus, to handle such boundaries, we develop a hybrid Hermite discontinuous Galerkin method. The discontinuous Galerkin method is used at the curved boundary, while Hermite methods are used in the interior so that the overall performance of the solver will not be compromised. Theoretical results and numerical simulations are presented in both cases.

1.2   Hermite-Taylor methods for periodic problems

To begin with we describe the method for the simple transport equation

$$\frac{\partial u}{\partial t} = a\frac{\partial u}{\partial x}, \quad u(x,0) = f(x),\ x \in [x_{\min}, x_{\max}],\ t > 0, \tag{1.1}$$

where a is a real constant and $f(x)$ is a 2π-periodic smooth function. We use staggered space-time grids for the algorithm; precisely we define the primal and dual

grids as follows:

$$\mathcal{G}^{\mathrm{p}} \;=\; \left\{x_j \mid x_j = x_{min} + jh_x, \; j = 0, \ldots, N_x\right\}, \tag{1.2}$$

$$\mathcal{G}^{\mathrm{d}} \;=\; \left\{x_{j+\frac{1}{2}} \mid x_{j+\frac{1}{2}} = x_{min} + (j + \frac{1}{2})h_x, \; j = 0, \ldots, N_x - 1\right\}. \tag{1.3}$$

Here N_x+1 is the number of grid points we use on the computation domain (interval), with $x_{max} - x_{min} = 2\pi$, and $h_x = \frac{x_{max}-x_{min}}{N_x}$. The Hermite methods evolve not only the function value at the nodes but also the derivatives of the function. Thus, at $t = 0$, we need to compute the derivatives of the initial function $f(x)$ up to m-th order, that is

$$\frac{h_x^i}{i!}\frac{d^i}{dx^i}f(x), \; x \in \mathcal{G}^{\mathrm{p}}, \; i = 0, \ldots, m. \tag{1.4}$$

For complicated initial data, this can be done by deriving a PDE recursion of the function or by numerically mapping the nodal data to the derivatives on the primal grid, details will be provided in the next two sections.
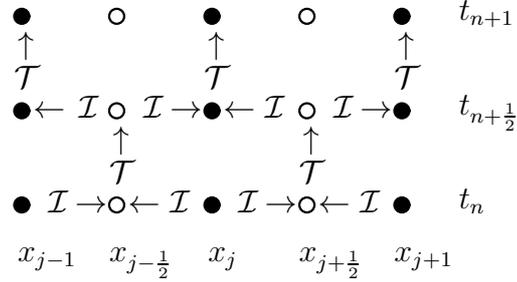
Figure 1.1: Schematic description of the numerical process for a full time step. Solid circles represent the primal grids and open circles represent the dual grids. $\mathcal{I}$ is the Hermite interpolation operator and $\mathcal{T}$ is the time evolution operator.

The algorithmic steps in the solution process can be viewed in Figure 1.1. The method consist of two basic steps: a Hermite interpolation procedure, $\mathcal{I}$, which maps the derivative data (coefficients) at the adjacent points to the center; and an

evolution procedure, $\mathcal{T}$, which uses the PDE to update the solution at the next half time level. Generalizations of the method to more complex PDEs requires a more involved evolution operator $\mathcal{T}$, discussed in [30], while the Hermite interpolation procedure is unchanged.

A detailed description in one dimesion now follows: (I) Starting at t_n, on each cell $\mathsf{D}^j = [x_j, x_{j+1}]$ we construct the Hermite interpolant of the function and derivative data up to order m at the endpoints x_j, x_{j+1}. The interpolant can be expressed as a $2m + 1$ degree polynomial centered around the grid point $x_{j+\frac{1}{2}}$ (in $\mathcal{G}^{\mathrm{d}}$):

$$p_{j+\frac{1}{2}}(x, t_n) = \sum_{l=0}^{2m+1} c_{l0}^{[j+\frac{1}{2}]} \left(\frac{x - x_{j+\frac{1}{2}}}{h_x} \right)^l. \tag{1.5}$$

The polynomial $p_{j+\frac{1}{2}}(x, t_n)$ interpolates the derivative data in the following sense (Figure 1.2): Given

$$p_j(x, t_n) = \sum_{l=0}^{2m+1} c_{l0}^{[j]} \left(\frac{x - x_j}{h_x} \right)^l, \tag{1.6}$$

$$p_{j+1}(x, t_n) = \sum_{l=0}^{2m+1} c_{l0}^{[j+1]} \left(\frac{x - x_{j+1}}{h_x} \right)^l, \tag{1.7}$$

then $p_{j+\frac{1}{2}}(x, t_n)$ satisfies

$$\frac{h_x^i}{i!} \frac{d^i}{dx^i} p_{j+\frac{1}{2}}(x_j, t_n) = c_{l0}^{[j]}, \quad \frac{h_x^i}{i!} \frac{d^i}{dx^i} p_{j+\frac{1}{2}}(x_{j+1}, t_n) = c_{l0}^{[j+1]}, \quad i = 0, \ldots, m. \tag{1.8}$$

Here the coefficients $c_{l0}^{[j]}$, $c_{l0}^{[j+1]}$ are scaled derivatives of p_j, p_{j+1} at x_j, x_{j+1}. The coefficients $c_{l0}^{[j+\frac{1}{2}]}$ can be found by constructing a $(2m + 2) \times (2m + 2)$ Hermite interpolation matrix H such that

$$H \left[c_{00}^{[j]}, \ldots, c_{m0}^{[j]}, c_{00}^{[j+1]}, \ldots, c_{m0}^{[j+1]} \right]^T = \left[c_{00}^{[j+\frac{1}{2}]}, \ldots, c_{2m+1,0}^{[j+\frac{1}{2}]} \right]^T. \tag{1.9}$$

Note that H only depends on the order m and the location where $p_{j+\frac{1}{2}}$ is centered. For example, when $m = 3$ and interpolation is centered at the middle, the interpolation

$$
\begin{array}{cccc}
\text{DOF} & m+1 & 2m+2 & m+1 \\[4pt]
\text{Coeff} & c_{l0}^{[j]} & c_{l0}^{[j+\frac{1}{2}]} & c_{l0}^{[j+1]} \\[4pt]
& \bullet \quad \rightarrow & \circ \quad \leftarrow & \bullet \\[4pt]
& x_j & x_{j+\frac{1}{2}} & x_{j+1}
\end{array}
$$
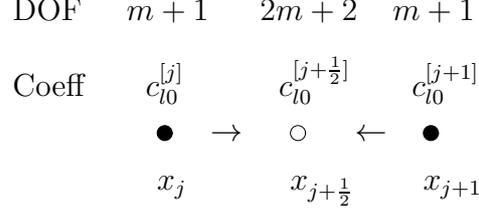
Figure 1.2: Hermite interpolation scheme on a 1D cell, DOF is the degree of freedom. Derivative data (Coefficients) are given at the solid circles to interpolate at the open circle.

matrix is

$$
H = \begin{pmatrix}
0.5 & 0.1719 & 0.04688 & 0.00781 & 0.5 & -0.1719 & 0.04688 & -0.00781 \\
-2.188 & -0.5938 & -0.125 & -0.01563 & 2.188 & -0.5938 & 0.125 & -0.01563 \\
0 & -0.9375 & -0.4375 & -0.09375 & 0 & 0.9375 & -0.4375 & 0.09375 \\
8.75 & 4.375 & 1.25 & 0.1875 & -8.75 & 4.375 & -1.25 & 0.1875 \\
0 & 1.25 & 1.25 & 0.375 & 0 & -1.25 & 1.25 & -0.375 \\
-21 & -10.5 & -4 & -0.75 & 21 & -10.5 & 4 & -0.75 \\
0 & -1 & -1 & -0.5 & 0 & 1 & -1 & 0.5 \\
20 & 10 & 4 & 1 & -20 & 10 & -4 & 1
\end{pmatrix}.
$$

$$(1.10)$$

More details about the Hermite interpolation in 2D will be discussed in next section.

(II) To march from t_n to $t_{n+\frac{1}{2}}$ at each point on the dual grid $x_{j+\frac{1}{2}}$ in $\mathcal{G}^{\mathrm{d}}$, we compute a temporal Taylor series

$$
p_{j+\frac{1}{2}}^n(x,t) = \sum_{l=0}^{2m+1} \sum_{s=0}^{q} c_{ls}^{[j+\frac{1}{2}]} \left(\frac{x - x_{j+\frac{1}{2}}}{h_x} \right)^l \left(\frac{t - t_n}{\Delta t} \right)^s. \tag{1.11}
$$

where $q = 2m + 1 - l$ for the constant coefficient case. For a general case q is chosen to be a little larger than $2m + 1$ if m is large. To compute the coefficients $c_{ls}^{[j+\frac{1}{2}]}$,

we insert equation (1.11) into (1.1), and derive a simple recursion for the coefficients $c_{ls}^{[j+\frac{1}{2}]}$, $s = 0, \ldots, 2m$ (recall we have computed $c_{l0}^{[j+\frac{1}{2}]}$, $l = 0, \ldots, 2m + 1$):

$$c_{l,s+1}^{[j+\frac{1}{2}]} = a \frac{l+1}{s+1} \frac{\Delta t}{h_x} c_{l+1,s}^{[j+\frac{1}{2}]}, \quad l = 0, \ldots, 2m - s, \ s = 0, \ldots, q - 1. \tag{1.12}$$

We note that we have included all nonzero terms in (1.11) as computed by (1.12).

(III) Since we have computed all the coefficients $c_{ls}^{[j+\frac{1}{2}]}$ in (1.11), we can now evaluate (1.11) at $t_{n+\frac{1}{2}} = t_n + \frac{\Delta t}{2}$ to arrive at:

$$\frac{h^i}{i!} \frac{\partial^i}{\partial x^i} u(x_{j+\frac{1}{2}}, t_{n+\frac{1}{2}}) \approx \frac{h^i}{i!} \frac{d^i}{dx^i} p_{j+\frac{1}{2}}^n(x_{j+\frac{1}{2}}, t_n) = \sum_{s=0}^{q} c_{is}^{[j+\frac{1}{2}]} \left(\frac{1}{2}\right)^s, i = 0, \ldots, m \tag{1.13}$$

The Hermite-Taylor method is stable so long as the domain of dependence of the solution at $(x_{j+\frac{1}{2}}, t_{n+\frac{1}{2}})$ is included in the cell, that is,

$$a \frac{\Delta t}{2} < \min(x_{j+1} - x_{j+1/2}, x_{j+1/2} - x_j), \tag{1.14}$$

which is just $a\Delta t < h_x$ in this case. In other words, the method is stable if within the half time step $\Delta t/2$, a wave travels at most from the middle point to the boundary of the cell. Note that the time step restriction is independent of the order m.

(IV) The final step is to repeat the process on the dual grid $\mathcal{G}^{\mathrm{d}}$ starting at $t_{n+\frac{1}{2}}$, this yields a solution on the primal grid $\mathcal{G}^{\mathrm{p}}$ at time t_{n+1}. Periodicity is assumed on $\mathcal{G}^{\mathrm{d}}$ to obtain the desired data on $\mathcal{G}^{\mathrm{p}}$.

1.2.1 General description of the Hermite methods

We consider the numerical solution of hyperbolic (or hyperbolic problems possibly with a small parabolic term) evolution equations

$$u_t = F(u, \mathcal{D}u, \epsilon\mathcal{D}^2 u, x, t) \tag{1.15}$$

in $d + 1$ dimensions. Here $u(x, t) \in \mathbb{R}^M$ and $\mathcal{D}, \mathcal{D}^2$ denotes the arrays of first and second order space derivatives. As the focus of this work is on adaptivity in space,

we assume u is L-periodic in x, that is $u(x + jLe_k, t) = u(x, t)$ for any integer j and standard unit basis vector $e_k \in \mathbb{R}^d$, which we will write as $x \in \mathbb{T}^d(L)$.

A Hermite method in d space dimensions uses staggered computational cells consisting of hypercubes. The standard, fixed-order method is constructed as follows. The degrees-of-freedom are the coefficients of a degree md tensor-product polynomial at each node. That is, at a node $(x_{1,k_1}, \ldots, x_{d,k_d}) \equiv x^{[k]}$ we approximate u by a tensor-product polynomial

$$u \approx \sum_{j_1=0}^{m} \cdots \sum_{j_d=0}^{m} c_{j_1 j_2 \ldots j_d}^{k_1 k_2 \ldots k_d}(t) \left(\frac{(x_1 - x_{1,k_1})}{h_1} \right)^{j_1} \cdots \left(\frac{(x_d - x_{d,k_d})}{h_d} \right)^{j_d}, \tag{1.16}$$

or, using the usual multiindex notation

$$\frac{h^{[j]}}{[j]!} D^{[j]} u(x^{[k]}, t) \approx c_{[j]}^{[k]}(t), \tag{1.17}$$

$$h^{[j]} = h_1^{j_1} \ldots h_d^{j_d}, \ [j]! = j_1! \ldots j_d!, \ D^{[j]} = D^{j_1} \ldots D^{j_d}. \tag{1.18}$$

Here h_i is the grid spacing in the ith coordinate, which we assume to be uniform. The structure of the multidimensional algorithm is the same as in 1D. At a time step, t_n, we construct the degree $(2m + 1)d$ Hermite interpolant centered around the midpoint from the 2^d vertex polynomials. The interpolant is $Q^{[k+1/2]}(t_n)$.

We now consider the evolution problem (1.15) projected onto the degree $(2m+1)d$ tensor-product polynomial with initial data $Q^{[k+1/2]}(t_n)$:

$$\frac{dQ^{[k+1/2]}}{dt} = \mathcal{P}_{2m+1} F(Q^{[k+1/2]}, \mathcal{D}Q^{[k+1/2]}, \epsilon \mathcal{D}^2 Q^{[k+1/2]}, x, t), \tag{1.19}$$

where $\mathcal{P}$ is the projection onto Taylor polynomials

$$\mathcal{P}_{2m+1} w(x, t) = \sum_{j_1=0}^{2m+1} \cdots \sum_{j_d=0}^{2m+1} \frac{D^{[j]} w(x^{[k+1/2]}, t)}{[j]!} \left(\frac{x - x^{[k+1/2]}}{h} \right)^{[j]}. \tag{1.20}$$

Clearly, (1.19) represents a closed system of ordinary differential equations for $(2m + 2)^d$ polynomial coefficients. We approximately evolve it to time $t_{n+1/2}$ using

possibly multiple substeps of some single-step method of Runge-Kutta type or, for linear autonomous problems, temporal Taylor series. At the completion of a time evolution we obtain vertex data on the dual grid by a another projection,

$$c_{[j]}^{[k+1/2]}(t_{n+1/2}) = \frac{h^{[j]}}{[j]!} D^{[j]} Q^{[k+1/2]}(x^{[k+1/2]}, t_{n+1/2}), \quad j_k = 0, \ldots, m. \quad (1.21)$$

The process is repeated on the dual grid to produce $c_{[j]}^{[k]}(t_{n+1})$.

Note that the evolution $\mathcal{T}$ can be accomplished with other ODE solvers. Experiments with Hermite-Runge-Kutta methods for simulating compressible flows can be found in [33]. Using (1.1) as an example, we write the coefficients as functions of time t,

$$p_{j+\frac{1}{2}}^n(x,t) = \sum_{l=0}^{2m+1} c_l^{[j+\frac{1}{2}]}(t) \left(\frac{x - x_{j+\frac{1}{2}}}{h_x} \right)^l. \quad (1.22)$$

Applying (1.1) to (1.22), we derive a system of ODEs for the coefficients $c_l^{[j+\frac{1}{2}]}(t)$,

$$\frac{d}{dt} c_l(t) = (l+1)c_{l+1}(t), \ l = 0, \ldots, 2m. \quad (1.23)$$

Thus, any ODE solver can be used to solve the equation (1.23) to obtain the solution (scaled derivatives $c_l^{[j+\frac{1}{2}]}(t_{n+\frac{1}{2}})$) at the next half time step.

Note that the evolution step is completely local to each cell. Thus no global storage of stage values is needed, nor is any communication of information between cells. If large steps can be taken, which is possible for hyperbolic problems, Hermite methods are essentially optimal from the standpoint of storage and communications.

1.2.2 Stability and convergence

A complete analysis of Hermite methods for linear hyperbolic systems is presented in [30]. The truncation error of the Hermite-Taylor method described above is

$\mathcal{O}(\Delta t^{2m+1} + \Delta h_x^{2m+1})$, i.e. the rate of convergence for a constant coefficient PDE problem is $2m + 1$ in both space and time.

If the system is hyperbolic, the domain of dependence of the solution at $(x^{[k+1/2]}, t_{n+1/2})$ is contained in the region $|x - x^{[k+1/2]}| \leq \frac{a}{2}\Delta t$ where $\Delta t = t_{n+1} - t_n$ is a full time step and a is the maximum wave speed. Assuming a CFL restriction

$$a\Delta t \leq \min_i h_i, \tag{1.24}$$

the solution of (1.19) evaluated at the dual grid node $x^{[k+1/2]}$ is in fact a high-order approximation to the Taylor projection of the exact solution of the evolution equation (1.15) with piecewise polynomial initial data whose restriction to cell $[k + 1/2]$ is $Q^{[k+1/2]}$. If we denote the exact solution of (1.19) restricted to the cell $[k + 1/2]$ by w we have

$$c_{[j]}^{[k+1/2]}(t_{n+1/2}) - \frac{h^{[j]}}{[j]!} D^{[j]} w(x^{[k+1/2]}, t_{n+1/2}) = O(h^{2m+1}\Delta t). \tag{1.25}$$

Stability follows from the fact that the Hermite interpolation process decreases a seminorm of the solution. The basic lemma, proved in [30] via a simple application of integration by parts, reads as follows.

Lemma 1.2.1. *Let f,g be smooth periodic functions, $\mathcal{I}f$ the degree $(2m+1)d$ piecewise Hermite interpolant of the data $D^j f$, $0 \leq j_k \leq m$ on the nodes, $x^{[k]}$, $\mathcal{I}g$ the analogous interpolant of g, and define*

$$(f, g)_{[m+1]} = \left(\prod_{k=1}^{d} D_k^{m+1} f, \prod_{k=1}^{d} D_k^{m+1} g \right)_{L^2}. \tag{1.26}$$

Then

$$(\mathcal{I}f, g - \mathcal{I}g)_{[m+1]} = 0, \tag{1.27}$$

$$\|f\|_{[m+1]}^2 = \|\mathcal{I}f\|_{[m+1]}^2 + \|f - \mathcal{I}f\|_{[m+1]}^2. \tag{1.28}$$

Directly, we conclude that the Hermite interpolation process decreases the semi-norm $\| \cdot \|_{[m+1]}$. This fundamental lemma, combined with standard estimates of interpolation error, can be directly turned into a convergence proof. Precisely, for the method described here, the proofs of Theorems 4.1 and 6.1 in [30] can be adapted to prove that the approximate solution converges at order $2m + \frac{1}{2}$. This can be improved to order $2m + 1$ for linear constant coefficient systems.

A point of emphasis is that, under the assumption that the local evolution problem (1.19) is solved with sufficient accuracy, the outer time step is limited only by the wave speed and the cell size; it is independent of the polynomial degree. In addition, the stability restrictions for the inner time step, that is the local time steps taken within each cell, have a favorable dependence on m. For a standard spectral method, the norm of the differentiation matrix must grow like m^2, and thus the CFL constraint scales like m^{-2} [31]. Although it has been shown how to reduce these norms to $O(m)$ by using dual grid filters for discontinuous Galerkin spectral elements [72], with Hermite methods no special actions are required; the differentiation matrix is always $O(m)$. The reason is that we only differentiate the cell polynomial of degree m at the cell center. Bernstein's inequality then gives

$$|\frac{dp}{dx}| \leq \frac{2m}{h} \max_{\text{cell}} |p|. \tag{1.29}$$

We have not used this fact to formally prove the stated bounds, but we have observed them in practice.

Naturally, these results cannot strictly hold for parabolic equations, as the domain of dependence is the entire spatial domain. However, numerical experiments with the heat equation, the convection-diffusion equation, and the Navier-Stokes equations suggest [22]

$$\Delta t \propto \min \left(\frac{h_i}{c_{\text{max}}}, \frac{h_i^2}{\epsilon m} \right). \tag{1.30}$$

Only in a shock or shear layer, where $\frac{h_i}{m} = O(\epsilon)$, is the additional restriction im-

portant. Again, the m-dependence is significantly better than for standard spectral element methods; see, e.g., [42] for a discussion of discontinuous Galerkin discretizations.

1.2.3 Computational cost of the Hermite methods

The computational costs of the Hermite methods are dominated by two operations, the construction of the interpolant and the evaluation of the right-hand side of (1.19). Using the tensor product structure, the interpolation may be carried out dimension-by-dimension. For example, if $d = 2$, the data consists of $(m + 1)^2$ values at each vertex. We first interpolate along the two edges parallel to the x_1-axis, solving $2m+2$ independent one-dimensional Hermite interpolation problems associated with each power of x_2 (see Figure 1.3). As the cost of the one-dimensional interpolation is that of a matrix-vector multiplication, $O(m^2)$, this step requires $O(m^3)$ flops in 2D. The next step then (see Figure 1.3) involves an additional $2m + 2$ one-dimensional interpolation problems in the y-direction, costing an additional $O(m^3)$ flops. Extending this argument to d-dimensions we conclude

Figure 1.3: Hermite interpolation scheme on a 2D cell. Derivative data (Coefficients) are given at the solid circles to interpolate at the open circle.

$$C_{\text{interpolation}} = O(m^{d+1}) \text{ flops.} \tag{1.31}$$

In another word, the interpolation cost is $O(m)$ per degree of freedom (DOF). The cost of evaluating the right-hand side of (1.19) depends on the structure of F. For a linear, constant-coefficient system the cost is linear in the number of coefficients, $O(m^d)$, or $\mathcal{O}(1)$ flop per DOF. For a system with product nonlinearities, the cost scales with the cost of multiplying degree $(2m + 1)d$ tensor product polynomials. The direct algorithm for accomplishing this exploiting the tensor-product structure costs $O(m^{d+1})$ flops. However, for large m FFTs can be used to reduce the cost to $O(m^d \ln m)$ per DOF.

For general nonlinearities we use a recursive algorithm inspired by automatic differentiation techniques [32]. We illustrate by computing

$$P = \mathcal{P}_{2m+1} e^Q, \tag{1.32}$$

where Q is a degree $(2m + 1)^d$ tensor product polynomial. The starting point is the differential equation

$$D_j P = \mathcal{P}_{2m+1}\left((D_j Q)P\right). \tag{1.33}$$

Taking, for example, $j = d$, (1.33) implies a recursion for the coefficients. Specializing to the node $x_{k,j_k} = 0$ and writing P and Q as

$$P = \sum_{j_d=0}^{2m+1} p_{j_d}(x_1, \ldots, x_{d-1}) \left(\frac{x_d}{h_d}\right)^{j_d}, \quad Q = \sum_{j_d=0}^{2m+1} q_{j_d}(x_1, \ldots, x_{d-1}) \left(\frac{x_d}{h_d}\right)^{j_d} \tag{1.34}$$

we have

$$j_d \cdot p_{j_d} = \sum_{j'=1}^{j_d} j' \cdot q_{j'} \cdot p_{j_d-j'}, \quad j_d = 1, \ldots, 2m+1. \tag{1.35}$$

Directly, given p_0 this allows the computation of all the polynomials p_{j_d} in $O(m^2)$ multiplications of tensor-product polynomials in dimension $d - 1$. To compute p_0

we would apply the analogous recursions in one lower dimension. To begin with we compute the coefficients $p_{j_10\ldots0}$ by

$$j_1 \cdot p_{j_10\ldots0} = \sum_{j'=1}^{j_1} j' \cdot q_{j'0\ldots0} \cdot p_{j_1-j'0\ldots0}, \quad p_{0\ldots0} = e^{q_{0\ldots0}}. \tag{1.36}$$

The cost of the direct implementation of this method is dominated by (1.35) and is $O(m^{d+2})$. If the multiplications are replaced by FFTs this becomes $O(m^{d+1}\ln m)$. Noting that the recursion in (1.35) has a convolutional form, the algorithm of Hairer, Lubich and Schlichte [26] can be adapted to reduce this to $O(m^d \ln^3 m)$, though again the advantage of the fast algorithm is only likely to be felt for m rather large. Griewank [32] has shown that all of the standard transcendental functions can be evaluated in this way. We also note that an alternative pseudospectral approach to evaluating F is possible, but we have not yet tried it.

Letting s be the total number of evaluations of F required to advance (1.19) to the next time level we have a total cost per degree-of-freedom depending on the structure of F and the algorithm employed given by

$$C_{\text{total}} = \begin{cases} O(s) + O(m), & \text{linear constant coeff} \\ O(sm) \leftrightarrow O(s\ln m) + O(m), & \text{product nonlinearity} \\ O(sm^2) \leftrightarrow O(sm\ln m) \leftrightarrow O(s\ln^3 m) + O(m), & \text{general nonlinearity} \end{cases} \tag{1.37}$$

As an example, the degree-of-freedom for problem (1.1) is $O(m)$, the cost of interpolation is $O(m)$ per degree of freedom and cost of evaluating the right hand side is $O(m)$ per degree of freedom (i.e. $s = m$ in (1.37)).

1.3   P-adaptive Hermite methods for initial value problems

For problems exhibiting localized pulses or sharp fronts, adaptivity in space and time is needed. The vast majority of work on adaptive schemes, however, has focused on local mesh refinement and local time-stepping, so-called h-adaptivity. A notable exception to this is the development of hp-adaptive solvers for elliptic boundary value problems by Demkowicz et al [41]. The goal of this work is to exploit the unique features of Hermite discretizations of initial-value problems to develop straightforward and, we believe, efficient purely P-adaptive (order) methods.

We begin with deriving new results on the convergence of the Hermite interpolants in the limit of infinite degree for bandlimited functions. We then develop and test a relatively simple strategy for locally adapting the degree of the interpolants and the time-stepping algorithm. As we can treat the polynomial evolution problem independently in each cell, there is very little overhead required by the proposed technique. Indeed, the basic formulation would allow us to use degrees and time-stepping procedures in each cell limited only by the data available to construct the cell interpolant. The procedure we use constrains the degree of the data used in each cell to guarantee that the interpolation process decreases a certain seminorm. Numerical experiments with the transport equation and Burgers equation are presented to demonstrate the potential of our approach.

1.3.1 P-convergence

The focus of the analysis in [30] is on the h-convergence of the Hermite schemes. Given that our intention here is to improve accuracy by increasing the degree rather than decreasing the cell size, it is of interest to study the convergence of the Hermite

interpolation process in this limit. To this end we take $d = 1$ and assume without loss of generality that $h = 2$. The Peano kernel formula for the interpolation error is

$$f(x) - (\mathcal{I}f)(x) = \int_{-1}^{1} D^{2m+2} f(t) K_m(t, x) dt, \tag{1.38}$$

where we assume $f \in C^{\infty}([-1, 1])$. An interesting characterization of K_m follows from the recognition that $e = f - \mathcal{I}f$ satisfies the boundary value problem

$$D^{2m+2} e = D^{2m+2} f, \quad D^j e(-1) = D^j e(1) = 0, \quad j = 0, \ldots, m. \tag{1.39}$$

Thus K_m is simply the Dirichlet Green's function for D^{2m+2}. As shown in [20] an upper bound for the error is given by

$$\|e\|_{L^{\infty}([-1,1])} \leq \frac{1}{(2m+2)!} \|D^{2m+2} f\|_{L^{\infty}([-1,1])}. \tag{1.40}$$

Specializing to $f = e^{i\omega x}$, we can use the Hermite error formula (e.g. [25]):

$$e^{i\omega x} - (\mathcal{I}e^{i\omega \cdot})(x) = \frac{(x^2 - 1)^{m+1}}{2\pi i} \int_C \frac{e^{i\omega z}}{(z^2 - 1)^{m+1}(z - x)} dz, \tag{1.41}$$

where C is a contour surrounding the real interval $[-1, 1]$. Following, for example, Weideman and Trefethen [73], we consider contours defined by

$$\left| z^2 - 1 \right| = c > 1. \tag{1.42}$$

By direct computation we find that if $z = re^{i\theta}$ these contours are parametrized by

$$r = \sqrt{\cos 2\theta + \sqrt{c^2 - \sin^2 2\theta}}. \tag{1.43}$$

In particular the maximum value of the imaginary part is $\sqrt{c - 1}$. We thus derive the bound for ω large

$$|e^{i\omega x} - (\mathcal{I}e^{i\omega \cdot})(x)| \leq \frac{1}{2\pi\sqrt{c-1}} e^{|\omega|\sqrt{c-1} - (m+1)\ln c}. \tag{1.44}$$

As $|\omega|/\pi$ is the number of wavelengths and $2m+2$ is the number of degrees of freedom we write

$$|\omega| = \frac{2\pi}{\alpha}(m + 1), \tag{1.45}$$

where α is the number of degrees-of-freedom per wavelength. Convergence as $m \to \infty$ is guaranteed if

$$\min_{c>1}\left(\frac{2\pi}{\alpha}\sqrt{c-1} - \ln c\right) < 0. \tag{1.46}$$

The minimum occurs when

$$\frac{\pi}{\alpha\sqrt{c-1}} = \frac{1}{c} \to \frac{\pi}{\alpha} = \frac{\sqrt{c-1}}{c}, \tag{1.47}$$

and thus we require

$$2\frac{c-1}{c} - \ln c < 0 \to c > c_0 = 4.9216\ldots. \tag{1.48}$$

On this interval the formula relating α and c at the minimum, (1.47), implies that α increases with c. Thus our sufficient condition for convergence is

$$\alpha > \frac{c_0}{\sqrt{c_0 - 1}}\pi \approx 2.4853\pi \approx 7.8077\ldots. \tag{1.49}$$

This condition is apparently sharp. In Figure 1.4 we display the maximum interpolation errors for the functions $e^{4\pi i(m+1)x/\alpha}$ on $[0,1]$ for $\alpha = 6 - 11$ and $0 \le m \le 18$, that is for all odd degree interpolants from 1 through 37. We have divergence for $\alpha = 6$ and $\alpha = 7$ and convergence for $\alpha = 8 - 11$. For larger values of m our implementation suffers from poor conditioning. In practice we have limited m to a maximum value of 11 in all of our implementations, that is a maximum degree of 23. (In the experiments here we take $m_{\max} = 8$.)

We conclude that Hermite interpolation requires almost two and one half times as many degrees-of-freedom per wavelength as corresponding methods based on Chebyshev interpolation, which require π points-per-wavelength, and in fact more than the 2 points-per-wavelength required when using equispaced nodes [73]. This is not surprising as we have removed sampling points from the center of the interval. However, we believe that the possibility for reduced the number of timesteps and communications enabled by the Hermite approach compensate for this defect. Preliminary numerical computation of the dispersion relation in 1D indicates that many fewer DOFs/wavelength are needed if the time step is not too small.
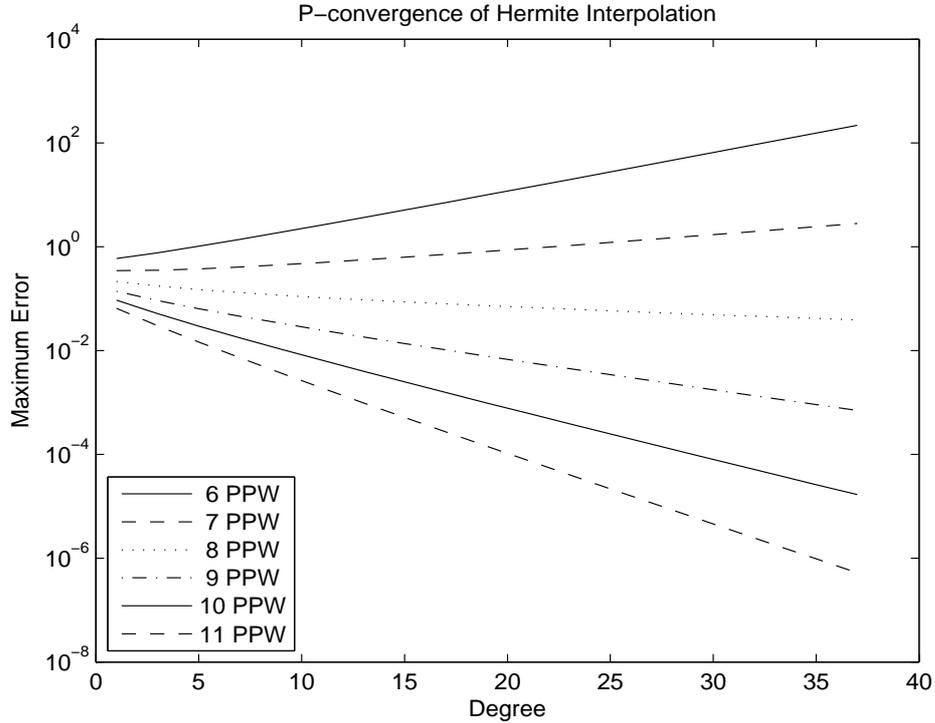
Figure 1.4: Convergence of Hermite interpolation with increasing degree. Here PPW stands for degrees-of-freedom per wavelength, as in all cases we are interpolating on a single cell.

1.3.2   Adaptive implementation in $1 + 1$ dimensions

Instead of using a uniform order m throughout the grids, we can vary the order m_k cell by cell. This requires a slightly different Hermite interpolation scheme, but the time step remains the same, because of the fact that large time steps can be taken in each computational cell, independent of the method order m and independent of data in neighboring cells, once the Hermite interpolant of the vertex data has been computed.

We now set $d = 1$ and consider the implementation of an order adaptive strategy.

Functionally, we simply admit the possibility that at the current time level t_n the coefficient data at the node x_k extends to some previously-determined degree m_k. At $x_{k+1/2}$ we then could compute the Hermite interpolant of the degree m_k polynomial at x_k and the degree m_{k+1} polynomial at x_{k+1} which would have degree $m_k+m_{k+1}+1$. However, we instead set

$$\bar{m}_{k+1/2} = \min\{m_k, m_{k+1}\}, \tag{1.50}$$

and construct the degree $2\bar{m}_{k+1/2}+1$ Hermite interpolant of the function values and derivatives through order $\bar{m}_{k+1/2}$ at each node. Directly this means that we use all of the available data from the node where the polynomial is of lower degree and ignore the highest degree coefficients at the other. The motivation for this choice is Lemma 1.2.1, which can be applied cell-by-cell. It implies that the local cell interpolant satisfies

$$\int_{x_k}^{x_{k+1}} \left(D^{\bar{m}_{k+1/2}+1}Q^{[k+1/2]}(x)\right)^2 dx \leq \int_{x_k}^{x_{k+1}} \left(D^{\bar{m}_{k+1/2}+1}f(x)\right)^2 dx, \tag{1.51}$$

for any function f whose derivatives through order $\bar{m}_{k+1/2}$ agree with the nodal data. We now evolve this polynomial as in the nonadaptive case using (1.19).

To adaptively choose the degree we consider the truncation step (1.21). The full polynomial has degree $2\bar{m}_{k+1/2} + 1$, roughly double what is being carried on the nodes. We simply truncate at whatever order is suggested by the tolerance, τ. That is, find the smallest $m_{k+1/2}$ such that

$$\max_{j>m_{k+1/2}} \frac{h^j}{j!} \left| D^j Q^{[k+1/2]}(x^{[k+1/2]}, t_{n+1/2}) \right| < \tau. \tag{1.52}$$

We also place a limit on the global maximum order m_{max}. We impose the same strategy when marching from $t_{n+1/2}$ to t_{n+1}.

Remark 1. *It is possible to perform an asymmetric Hermite interpolation on each cell (i.e. $m_k \neq m_{k+1}$). However, we have found that this may cause instability of the P-adaptive scheme.*

Application to the transport equation

As a first test of the method we solve the transport equation,

$$u_t = u_x, \quad u(x,0) = e^{-x^2}, \ x \in [-10, \ 10], \ t > 0, \tag{1.53}$$

with periodic boundary conditions for two periods, until final time $T = 40$. We take $h = 0.25$ and $\Delta t = 0.9h$. The derivatives of the initial data are computed utilizing the following recursion

$$\frac{h^i}{i!}\frac{d^i}{dx^i}e^{-ax^2} = g^{(i)}(x)e^{-ax^2}, i = 0, \dots, m, \tag{1.54}$$

$$g^{(i)} = -2ah(xg^{(i-1)} + hg^{(i-2)})/i \ , \ g^{(1)} = -2ahx \ , \ g^{(0)} = 1. \tag{1.55}$$

The time-stepping is performed via Taylor series and is carried out to match the local spatial order. That is, as in [30], we use an order $2\bar{m}_{k+1/2} + 1$ temporal Taylor series. Note that this time-stepping procedure is completely local to each cell, so there is no issue in choosing different temporal orders in different cells. The Taylor method could easily be replaced by multiple substeps of some other Runge-Kutta formula as in [34]. Again, as the local evolution problems are independent, the size of the substeps can be chosen independently in each cell. We set the maximum value of m to be 8 and the minimum to be 1 so that the local method order varies between 3 ($m = 1$) and 17. At $t = 0$, m_k is chosen based on the interpolation of the initial condition.

Varying the tolerance, τ, between 10^{-3} and 10^{-9} we obtain the results summarized in Table 1.1. See also Figure 1.5 for a plot of the solution and m_k when $\tau = 10^{-4}$. We see that the actual error in each case is more than an order of magnitude less than the tolerance, indicating that a less stringent cutoff criterion could be used. For the coarsest tolerance the average density of degrees-of-freedom is almost 2.7 times less than the maximum. Exponential convergence is observed as a decrease in the error of a factor of 10 requires roughly a fixed increase in the number of degrees-of-freedom over the range of tolerances considered.

τ	Maximum Error	$m_{\max}$	m_{ave}
10^{-3}	1.221×10^{-5}	4	1.556
10^{-4}	5.767×10^{-7}	5	1.914
10^{-5}	1.132×10^{-7}	6	2.333
10^{-6}	2.588×10^{-8}	7	2.778
10^{-7}	7.394×10^{-9}	8	3.420
10^{-8}	4.373×10^{-10}	8	3.803
10^{-9}	3.022×10^{-11}	8	4.049

Table 1.1: Errors and orders for various tolerances: transport equation in $1 + 1$ dimensions.
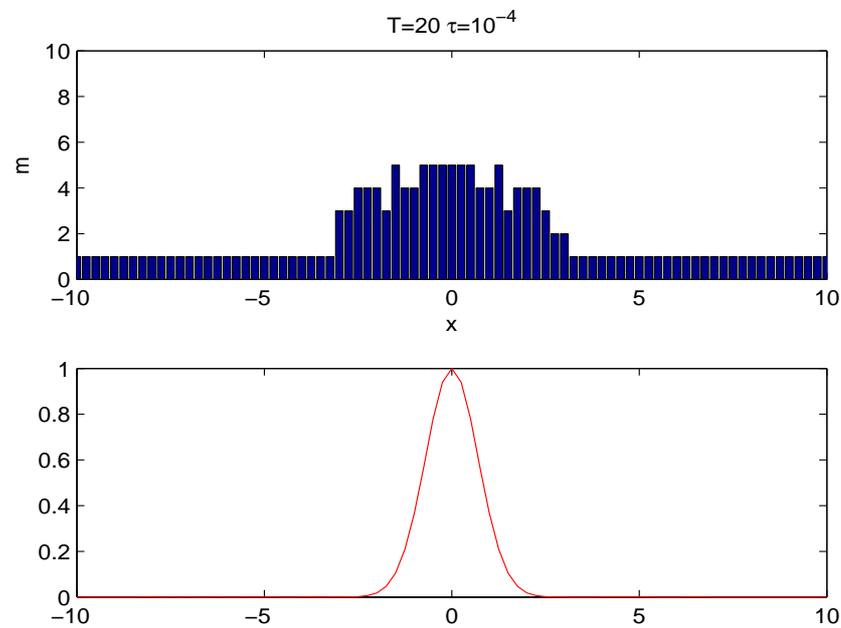
Figure 1.5: Solution and degree for the transport equation in $1 + 1$ dimensions.

Application to Burgers equation

As a second example we solve Burgers equation

$$u_t + uu_x = \epsilon u_{xx}, \quad u(x,0) = \sin x, \ x \in [0, 2\pi], \ 0 \le t \le 2, \tag{1.56}$$

The recursion to compute the coefficients similar to (1.12) with the higher order terms truncated becomes

$$\frac{s+1}{\Delta t}c_{l,s+1} = \tilde{c}_{l,s} + \frac{(l+2)(l+1)}{h^2}c_{l+2,s} \ , \ l = 0,\ldots,2m-1, \ s = 0,\ldots,2m+3, \quad (1.57)$$

$$\sum_{l=0}\sum_{s=0}\tilde{c}_{ls}(\mathsf{X})^l(\mathsf{T})^s = -\left(\sum_{l=0}^{2m+1}\sum_{s=0}^{q}c_{ls}(\mathsf{X})^l(\mathsf{T})^s\right)\left(\sum_{l=0}^{2m}\sum_{s=0}^{q}\frac{l+1}{h}c_{l+1,s}(\mathsf{X})^l(\mathsf{T})^s\right),$$
$$(1.58)$$

$$\mathsf{X} = \frac{x - x_{j+1/2}}{h_x}, \quad \mathsf{T} = \frac{t - t_n}{\Delta t}. \quad (1.59)$$

The cost to evaluate the product nonlinearity is $O(m^2)$ per degree-of-freedom (i.e. $s = m$ in (1.37)).

We choose a uniform grid with $h = \pi/80$, set the maximum value of m to be 8 (maximum order 17) and solve for $\epsilon = 10^{-2}, 10^{-4}$ with $\tau = 10^{-3}, 10^{-6}$. As the time step restriction is now proportional to $h^2/(m\epsilon)$ we chose $\Delta t = h/50$, though for $\epsilon = 10^{-4}$ we could take larger steps. If we define an effective cell Reynolds number based on the finest degree-of-freedom density allowed

$$R_c = u_{\max}\frac{h}{\epsilon \cdot (m_{\max} + 1)}, \quad (1.60)$$

we compute a value of .44 when $\epsilon = 10^{-2}$ and 44 when $\epsilon = 10^{-4}$. Thus the latter case is certainly underresolved and we do not expect to achieve the error tolerances. However we will achieve reasonable accuracy and our solutions will not display spurious oscillations. This demonstrates the robustness of the P-adaptive Hermite methods. Although, based on the results from Section 1.2 we may expect that adaptive implementations of standard spectral element methods could require a little less resolution in the shock layer, the advantage of the proposed scheme is that there is essentially no overhead.

Approximate error data is generated by comparing the solutions to those computed with a Fourier pseudospectral method in space [24] evolved in time with Matlab's ode45 routine. Absolute and relative error tolerances for ode45 were set at 10^{-9}.

With $\epsilon = 10^{-2}$ we used 1280 points for the pseudospectral method and 40960 with $\epsilon = 10^{-4}$. To verify the accuracy of the pseudospectral computations we repeated them with twice as many grid points and the tolerances for the ode solver reduced to 10^{-11}. These indicate that the pseudospectral solutions have errors below 10^{-9} at the times indicated except for $t = 2$ and $\epsilon = 10^{-4}$, when a maximum norm difference of 10^{-4} was recorded. As this is orders-of-magnitude smaller than the Hermite error in that case we deem the error data to be reliable.

Details of the results are presented in Table 1.2 for both the approximate shock formation time, $t = 1.0016$, (note that the shock formation time for the inviscid equations is $t = 1$) and the final time, $t = 2$. Decreasing the tolerance by a factor of 10^{-3} increases the number of degrees of freedom by less than a factor of 3. The error tolerances were approximately achieved in the case of $\epsilon = 10^{-2}$. For $\epsilon = 10^{-4}$ larger errors persist due to the limits placed on R_c. However, with the adaptive strategy we do avoid oscillations at the shock, as is readily apparent in the graphs. (See Figure 1.8 and Figure 1.9.) We also plot the solution and method order at $t = 1$ and $t = 2$ for $\epsilon = 10^{-2}$ and $\epsilon = 10^{-4}$ computed with $\tau = 10^{-3}$.

ϵ	t	τ	$\|u_{\text{adapt}} - u_{PS}\|_\infty$	m_{max}	m_{ave}
10^{-2}	1	10^{-3}	8.32×10^{-5}	7	1.32
10^{-2}	2	10^{-3}	1.80×10^{-3}	8	1.23
10^{-2}	1	10^{-6}	2.99×10^{-10}	8	3.19
10^{-2}	2	10^{-6}	6.09×10^{-6}	8	2.45
10^{-4}	1	10^{-3}	9.0×10^{-3}	8	1.43
10^{-4}	2	10^{-3}	3.0×10^{-2}	8	1.38
10^{-4}	1	10^{-6}	9.2×10^{-3}	8	3.22
10^{-4}	2	10^{-6}	3.2×10^{-2}	8	2.62

Table 1.2: Comparison of Fourier pseudospectral and adaptive order Hermite solutions for various tolerances and viscosities: Burgers equation in $1 + 1$ dimensions.
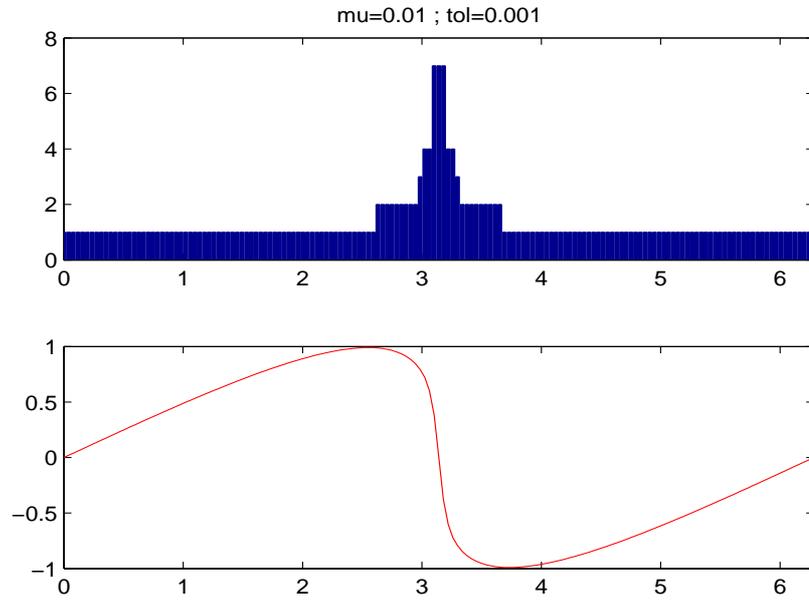
Figure 1.6: Solution and degree at $t = 1$ for Burgers equation, mu=ϵ.
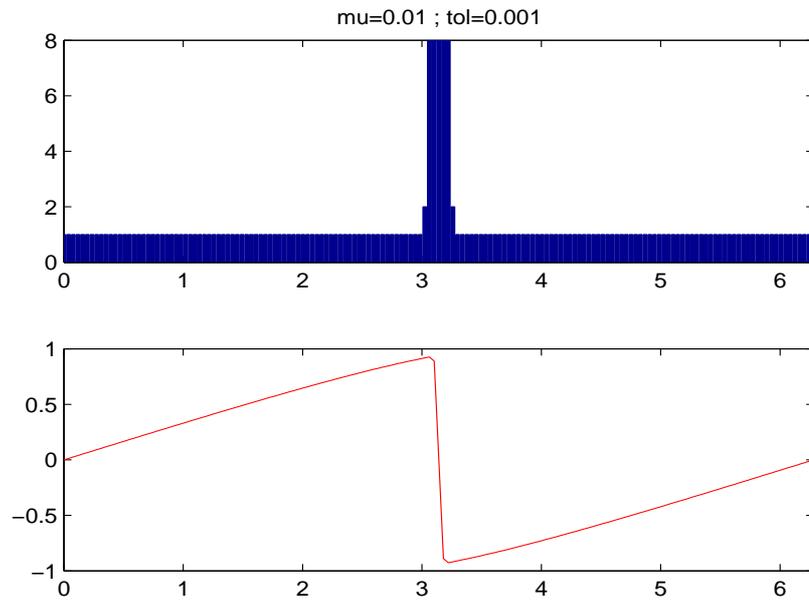
Figure 1.7: Solution and degree at $t = 2$ for Burgers equation, mu=ϵ.
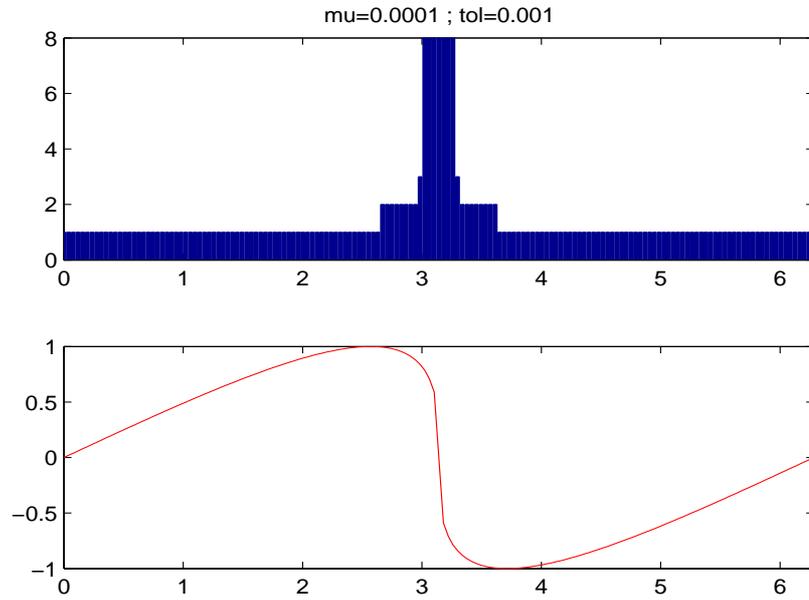
Figure 1.8: Solution and degree at $t = 1$ for Burgers equation, mu=ϵ.
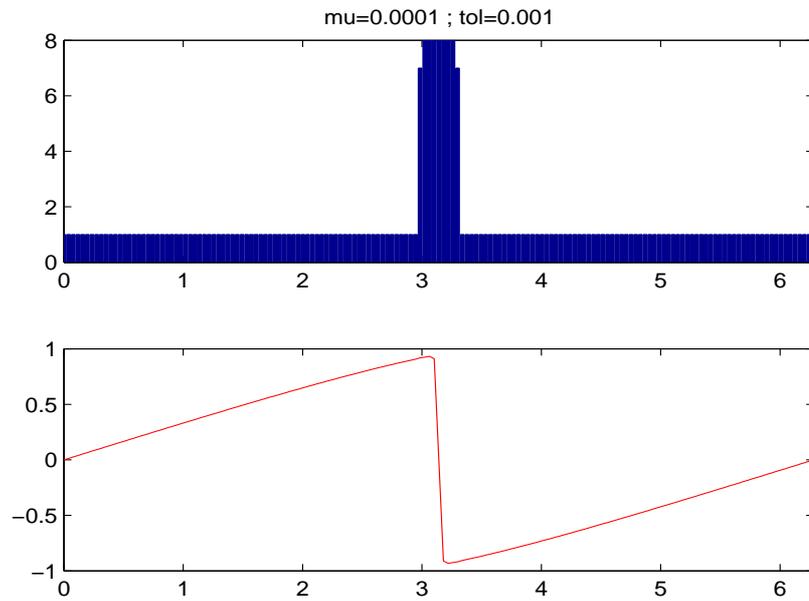
Figure 1.9: Solution and degree at $t = 2$ for Burgers equation, mu=ϵ.

$$y_{j+1} \quad \bullet \quad \rightarrow \quad \odot \quad \leftarrow \quad \bullet$$

$$c^{[i+\frac{1}{2},j+1]}_{l_1,l_2,0} \quad \downarrow$$

$$y_{j+\frac{1}{2}} \quad \bigcirc \; c^{[i+\frac{1}{2},j+\frac{1}{2}]}_{l_1,l_2,0}$$

$$c^{[i,j]}_{l_1,l_2,0} \qquad \uparrow \; c^{[i+\frac{1}{2},j]}_{l_1,l_2,0} \qquad c^{[i+1,j]}_{l_1,l_2,0}$$

$$y_j \quad \bullet \quad \rightarrow \quad \odot \quad \leftarrow \quad \bullet$$

$$x_i \qquad x_{i+\frac{1}{2}} \qquad x_{i+1}$$
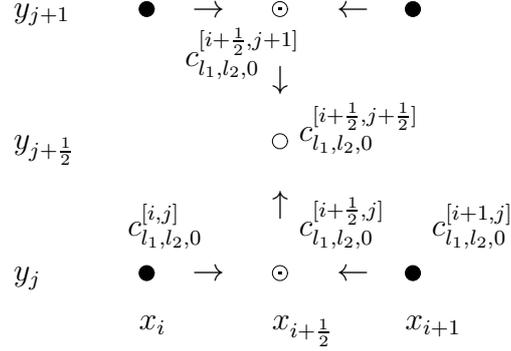
Figure 1.10: Hermite interpolation scheme on a 2D cell. Derivative data (coefficients) are given at the solid circles to interpolate at the open circle.

1.3.3 Extensions to two space dimensions

The interpolation technique used above has a natural extension to any number of space dimensions. The process of the Hermite interpolation in 2D can be done dimension by dimension as shown in Figure 1.10. Denote m_x, m_y as the order in x and y for the Hermite interpolation respectively. At the four vertices (solid circle) we have $(m_x + 1) \times (m_y + 1)$ degrees of freedom. First we interpolate along the x-axis to form a $(2m_x + 1) \times (m_y + 1)$ polynomial at the middle points (circle with dot) Then we interpolate the y-direction from the middle point (circle with dot) to the cell center point (open circle) to form a $(2m_x + 1) \times (2m_y + 1)$ polynomial. We define

$$\bar{m}_{x,[i+\frac{1}{2},j+\frac{1}{2}]} = \min\left(m_{x,[i,j]}, m_{x,[i,j+1]}, m_{x,[i+1,j]}, m_{x,[i+1,j+1]}\right), \tag{1.61}$$

$$\bar{m}_{y,[i+\frac{1}{2},j+\frac{1}{2}]} = \min\left(m_{y,[i,j]}, m_{y,[i,j+1]}, m_{y,[i+1,j]}, m_{y,[i+1,j+1]}\right), \tag{1.62}$$

and compute the tensor-product Hermite interpolant, $Q^{[i+\frac{1}{2},j+\frac{1}{2}]}$ of the vertex data using mixed derivatives up through order $\bar{m}_{x,[i+\frac{1}{2},j+\frac{1}{2}]}$ in x and up through order $\bar{m}_{y,[i+\frac{1}{2},j+\frac{1}{2}]}$ in y. This will result in a cell polynomial of degree $2\bar{m}_{x,[i+\frac{1}{2},j+\frac{1}{2}]} + 1$ in x and $2\bar{m}_{y,[i+\frac{1}{2},j+\frac{1}{2}]} + 1$ in y. The interpolation is a stabilizing step in that Lemma

1.2.1 applied in both variables implies the analogue of (1.51):

$$\int_{x_i}^{x_{i+1}} \int_{y_j}^{y_{j+1}} \left(D_x^{\bar{m}_{x,[i+\frac{1}{2},j+\frac{1}{2}]}+1} D_y^{\bar{m}_{y,[i+\frac{1}{2},j+\frac{1}{2}]}+1} Q^{[i+\frac{1}{2},j+\frac{1}{2}]}(x,y) \right)^2 dx\, dy \ \leq$$

$$\int_{x_i}^{x_{i+1}} \int_{y_j}^{y_{j+1}} \left(D_x^{\bar{m}_x^{[i+\frac{1}{2},j+\frac{1}{2}]}+1} D_y^{\bar{m}_{y,[i+\frac{1}{2},j+\frac{1}{2}]}+1} f(x,y) \right)^2 dx\, dy, \qquad (1.63)$$

for any function f whose mixed derivatives through order $\bar{m}_{x,[i+\frac{1}{2},j+\frac{1}{2}]}$ in x and order $\bar{m}_{y,[i+\frac{1}{2},j+\frac{1}{2}]}$ in y agree with the nodal data. We note that as in the 1D case the restriction (1.61), (1.62) greatly enhances the stability of the order adaptive scheme.

Having computed the cell interpolant, we evolve the data using (1.19). To truncate we seek to satisfy the analogue of (1.52) with the smallest values of $m_{x,[i+\frac{1}{2},j+\frac{1}{2}]}$, $m_{y,[i+\frac{1}{2},j+\frac{1}{2}]}$

$$\max_{j_1 > m_{x,[i+\frac{1}{2},j+\frac{1}{2}]}} \frac{h_1^{j_1} h_2^{j_2}}{j_1!\, j_2!} \left| D_x^{j_1} D_y^{j_2} Q^{[i+\frac{1}{2},j+\frac{1}{2}]}(x_{i+\frac{1}{2}}, y_{j+\frac{1}{2}}, t_{n+\frac{1}{2}}) \right| \ < \ \tau, \qquad (1.64)$$

$$\max_{j_2 > m_{y,[i+\frac{1}{2},j+\frac{1}{2}]}} \frac{h_1^{j_1} h_2^{j_2}}{j_1!\, j_2!} \left| D_x^{j_1} D_y^{j_2} Q^{[i+\frac{1}{2},j+\frac{1}{2}]}(x_{i+\frac{1}{2}}, y_{j+\frac{1}{2}}, t_{n+\frac{1}{2}}) \right| \ < \ \tau. \qquad (1.65)$$

Application to the transport equation in two space dimensions

As a first demonstration of the method in $2+1$ dimensions we have solved the transport equation

$$u_t + \cos\theta \cdot u_x + \sin\theta \cdot u_y = 0, \quad (x,y) \in [-25,25] \times [-25,25], \qquad (1.66)$$

with $\theta = \frac{\pi}{3}$, periodic boundary conditions, and initial data consisting of a Gaussian

$$u(x,y,0) = e^{-(x^2+y^2)}. \qquad (1.67)$$

The recursion similar to (1.12) now becomes

$$c_{l_1,l_2,s+1} = -\cos\theta \frac{l_1+1}{s+1} \frac{\Delta T}{h_x} c_{l_1+1,l_2,s} - \sin\theta \frac{l_2+1}{s+1} \frac{\Delta T}{h_y} c_{l_1,l_2+1,s}. \qquad (1.68)$$
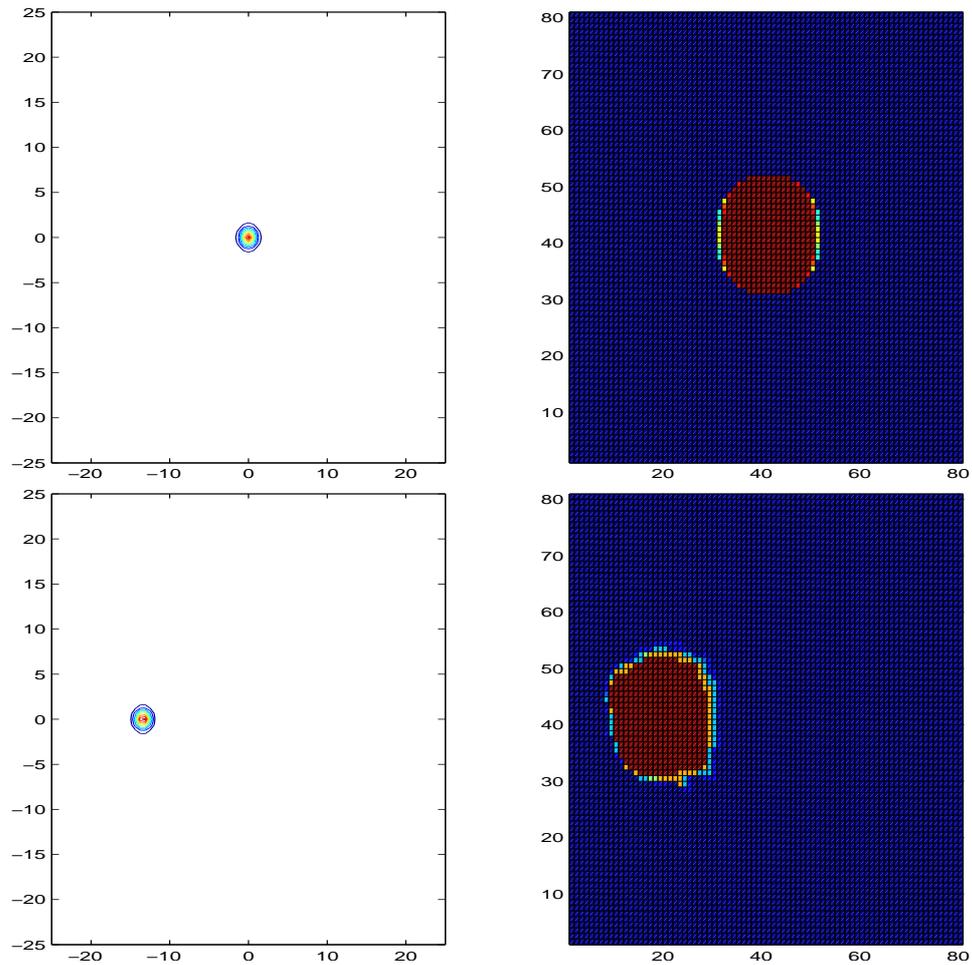
Figure 1.11: Solution to (1.66) at $T = 0$ (top) and $T = 100$ (bottom). On the left is the plot of the solution u, on the right is the adaptive order in x−direction on each grid. $m_{max} = 10$, $m_{min} = 1$, tolerance $\tau = 10^{-15}$, and CFL number equals 0.6. Maximum error at $T = 100$ is 5.56×10^{-11}. Blue is for $m = 1$, brown is for $m = 10$.

We solved up to $t = 100$ using the adaptive method described above with a maximum value of $m_{\max} = 8$ (17th order) and tolerances of $10^{-3}, 10^{-5}$, 10^{-7}, and 10^{-9}. The mesh width was $h = \frac{5}{8}$, $\Delta t \approx .8h$.

The solution to (1.66) along with the plot of the adaptive order on each grid is displayed in Figure 1.11. The region where higher order is needed to maintain

the accuracy doesn't change much after 100 periods of propagation at $T = 100$. The results are summarized in Table 1.3. Generally, the L^2 error is at or below the desired tolerance, though the maximum error is quite a bit larger. That said, the method is seen to be very efficient. Even with a tolerance of 10^{-9}, which produced a maximum error of approximately 10^{-6}, the total number of degrees-of-freedom, as indicated by m_{ave}, exceeds those of the first order method on the same grid by a mere 28%. Timing results of various m_{max} and tolerance τ are listed in Table 1.4, where $\tau = 0$ means using a uniform order m_{max} everywhere in the domain. The total cost of the P-adaptive scheme will be a weight average on all grids for various m as in (1.37). As we can see from the Table 1.4, if we double the mesh, the time cost is approximately 8 times longer for $m_{max} = 2$, and roughly the same accuracy can be obtained when $m_{max} = 3$, $\tau = 10^{-10}$ with only $\frac{1}{7}$ time cost. Since it is a relatively coarse grid, for $\tau = 10^{-10}$, we need at least $m_{max} = 5$ to reach the maximum accuracy for such tolerance. We can also see the maximum error is decreasing at a slightly slower speed than the tolerance when $m_{max} = 8$, and computational cost increases just a little while a lot more accuracy have been achieved.

τ	L^2 Error	Maximum Error	m_{max}	m_{ave}
10^{-3}	3.10×10^{-4}	1.51×10^{-1}	8	0.06
10^{-5}	3.24×10^{-6}	1.80×10^{-3}	8	0.14
10^{-7}	6.07×10^{-8}	1.52×10^{-5}	8	0.20
10^{-9}	1.42×10^{-9}	1.04×10^{-6}	8	0.28

Table 1.3: Errors and orders at $t = 100$ for various tolerances: transport equation in $2 + 1$ dimensions.

m_{max}	N	τ	Maximum Error	L^2 Error	CPU-Time
2	81	1E-010	1.10E-002	1.00E-005	10
2	161	1E-010	4.10E-004	3.60E-007	77
2	81	0	1.10E-002	1.00E-005	16
3	81	1E-010	3.30E-004	2.38E-007	11
3	81	0	2.58E-004	2.22E-007	26
4	81	1E-010	4.38E-006	4.19E-009	12
4	81	0	4.38E-006	4.01E-009	41
5	81	1E-010	2.49E-007	2.52E-010	14
5	81	1E-012	6.33E-008	7.15E-011	14
5	81	0	5.60E-008	6.02E-011	63
6	81	1E-010	2.49E-007	1.91E-010	16
6	81	1E-012	1.07E-008	1.12E-011	17
6	81	0	7.33E-010	7.82E-013	92
7	81	1E-012	1.07E-008	1.03E-011	20
7	81	0	7.64E-012	8.16E-015	130
8	81	1E-011	3.54E-008	5.25E-011	24
8	81	1E-013	1.60E-009	2.28E-012	25
8	81	1E-015	1.58E-010	1.35E-013	26
8	81	1E-017	9.48E-012	6.46E-015	27
8	81	1E-019	7.53E-013	7.70E-016	30
8	81	0	1.87E-013	3.31E-016	179

Table 1.4: Errors and time cost at $t = 20$ for various tolerances τ, number of grid points along each dimension N and maximum adaptive order m_{max}: transport equation in $2 + 1$ dimensions.

1.4 Hybrid Hermite - discontinuous Galerkin method

1.4.1 Introduction

In this section, we will introduce an arbitrary order hybrid Hermite - discontinuous Galerkin (HDG) method. The discontinuous Galerkin (DG) method is used next

to the boundary due to its high flexibility when handling curved geometry and the Hermite method is used in the interior of the region to improve the performance.

High order schemes for curved boundaries can be constructed for example by using overlapping structured grids as done by Henshaw [36] or by using unstructured grids as in the nodal discontinuous Galerkin (DG) method by Hesthaven and Warburton [38]. Unfortunately the efficiency of high order nodal DG methods is limited by the necessity to take very small time steps to maintain stability. The operation count per degree of freedom is also quite high.

Traditional finite element methods (FEM) are useful for handling geometry but they too suffer from high operation count and memory consumption. An efficient alternative to traditional FEM was suggested by Rylander and Bondeson in [70] where they showed how to hybridize FDTD with FEM. The principle idea is to use the more expensive FEM close to the boundaries and the inexpensive FDTD in the interior of the computational domain.

The hybrid Hermite - discontinuous Galerkin method we suggest here uses a description of the geometry similar to that of Rylander and Bondeson but is very different in the discretization strategy of the governing equations. In the interior, on a structured space-time-staggered Cartesian grid, we will use Hermite methods of very high order. Next to the boundary we use the nodal-DG method by Hesthaven and Warburton [38]. The combination of these methods yields high order accurate and geometrically flexible methods.

Hermite methods have exceptional resolving power and excellent explicit time stepping properties. However, enforcing boundary conditions in a Hermite method, in particular in complex geometries, is less straightforward.

Discontinuous Galerkin methods, on the other hand, deal with complicated geometries in a natural way and also have good resolving power. However, they suffer

from severe time-step restrictions due to the evaluation of derivatives throughout the element.

The features of Hermite and DG methods suggest their hybridization, combining the geometrical flexibility of DG with the excellent time stepping properties of the Hermite methods resulting in a class of geometrically flexible and efficient high resolution methods. The hybridization strategy is the same as that of Rylander and Bondeson. It can be summarized as follows: Use DG next to the boundaries to handle the geometry and the more efficient Hermite method in the interior. Of course the time stepping restrictions for the DG scheme will not change, but as the number of DG elements will scale on K for a 2D geometry with $\mathcal{O}(K^2)$ elements while the number of Hermite elements will scale on K^2 the complexity of the method will, to leading order, be the same as for the Hermite method.

1.4.2   Description of the methods in one dimension

This section describes the discretization of the geometry and the two methods and their hybridzation in one dimension.

Discretization of the geometry

The equations at hand will be discretized on a hybrid grid composed of a regular Cartesian grid in the interior and a conforming triangulation or a conforming mesh of quadrilaterals (2D) or hexahedra (3D) outside the Cartesian grid, see Figure 1.12. The Cartesian grid is composed of a primal and a dual grid, denoted $\mathcal{G}^{\mathrm{p}}$ and $\mathcal{G}^{\mathrm{d}}$ and consisting of the points:

$$\mathcal{G}^{\mathrm{p}} = \{x_i\} = x_{\min}^{\mathrm{Her}} + i\, h_x, \quad i = 0, \ldots, N_x, \tag{1.69}$$
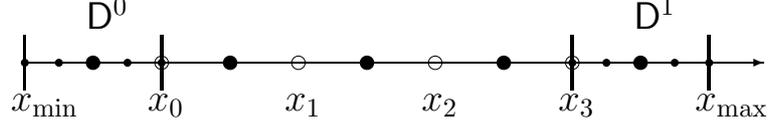
Figure 1.12: Hybrid grid in one dimension. Here the DG-grid $\mathcal{G}^{\mathrm{DG}} = \{\mathsf{D}^0, \mathsf{D}^1\}$, consists of one element on each side of the interior Cartesian grid. The LGL nodes on the elements are denoted by small filled circles. The Cartesian grid is denoted by larger circles, the empty being the primal nodes and the filled being the dual nodes. The communication of Hermite data to the DG solver consists of constructing DG fluxes at x_0 and x_3. The communication of DG data to the Hermite solver consists of evaluating derivatives centered at $x_{-1/2}$ and $x_{3+1/2}$ using the DG solution at the LGL nodes.

and

$$\mathcal{G}^{\mathrm{d}} = \{x_i\} = x_{\min}^{\mathrm{Her}} + i\, h_x, \quad i = \frac{1}{2}, \ldots, N_x - \frac{1}{2}, \tag{1.70}$$

where

$$h_x = \frac{x_{\max}^{\mathrm{Her}} - x_{\min}^{\mathrm{Her}}}{N_x}.$$

The outside DG-grid $\mathcal{G}^{\mathrm{DG}}$ consists of two sets of elements, one on each side of the interior Cartesian grid. The Kth element in $\mathcal{G}^{\mathrm{DG}}$ is denoted D^K and is discretized with N_x^K Legendre-Gauss-Lobatto (LGL) points, $\{x_i^K\}$. In one dimension the LGL points (on the standard element $x \in [-1,1]$) are the zeros of the polynomial $f(x) = (1 - x^2)\tilde{P}'_{N_x^K+1}(x)$ where $\tilde{P}_{N_x^K+1}(x)$, is the classic Legendre polynomials scaled by dividing by $\sqrt{1/(2(N_x^K + 1) - 1)}$, see further p. 47 in Appendix A [37].

We now describe the two methods starting with the Hermite method. We note that the spatial discretization based on Hermite interpolation can be paired with any time stepping algorithm. Here we will either use a Taylor series method (as in the

original Hermite-Taylor algorithm, [30]) or an explicit Runge-Kutta time-stepping method, see [33].

Hermite-Taylor method

We use the same Hermite-Taylor method as discussed before to solve the PDE on the Hermite domain $\mathcal{G}^{\mathrm{p}}$ and $\mathcal{G}^{\mathrm{d}}$, except that, instead of using the periodicity to update the half time solution on $\mathcal{G}^{\mathrm{d}}$, information from the Discontinuous Galerkin (DG) element is used to update the Hermite cells along the boundary of the Hermite domain, which will be detailed in the numerical experiment sections.

Hermite-Runge-Kutta method

The truncation error of the Hermite-Taylor method is $\mathcal{O}(\Delta t^{2m+1} + \Delta h_x^{2m+1})$ and will be the method of choice for linear constant coefficient problems. For problems that have varying coefficients or are nonlinear, the recursion (1.12) becomes prohibitively costly for high time derivatives and large m. Therefore it is better to evolve the approximation of the solution with a one-step method. For such methods it is sufficient to compute the first time derivative, or equivalently, the right hand side of the evolution equations. For a Hermite method evolving (1.11) in time this translates into the need to compute $c_{ls}^{[j+\frac{1}{2}]}$ for $s = 1$, which is the least expensive term of (1.12). Another advantage of using a onestep method is that is straightforward to incorporate an adaptive ODE solver from, for example, the RK-suite library [21].

Remark 2. *Note that when a onestep ODE solver is used, the order of accuracy independent CFL condition* (1.14) *holds in the following sense. The time evolution of* (1.11) *can be performed independently on each cell using many sub-time-steps and communication need only be performed as required by domain of dependence considerations, see also [33].*
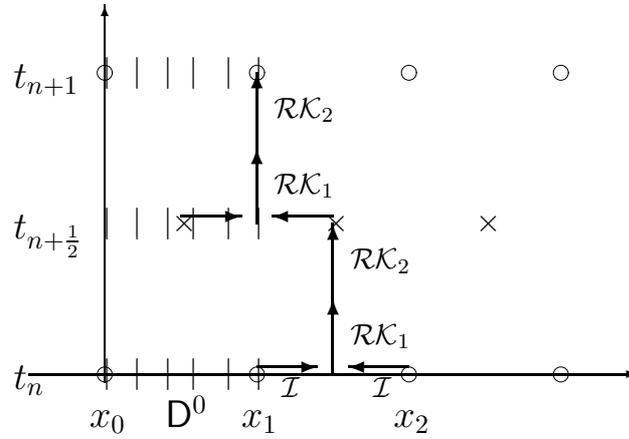
Figure 1.13: Schematic picture of the hybrid Hermite-DG method.

Nodal discontinuous Galerkin method

The Discontinuous Galerkin method used here is that of Hesthaven and Warburton. We outline it here and refer to [37] for more details.

In element K, $x \in [x_l^K, x_r^K] = \mathsf{D}^K$ the solution is approximated by a $N = N_{\mathrm{DG}} - 1$
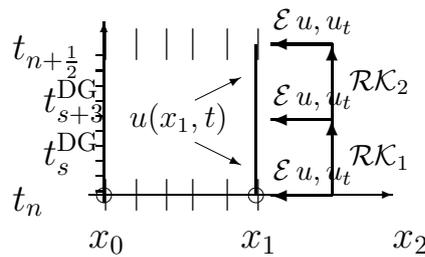
Figure 1.14: The boundary data for the DG-method along x_1 and for the first timestep is obtained by extrapolating u and u_t to x_1 at all Runge-Kutta substeps in $t \in [t_n, t_{n+1/2}]$ and constructing a fourth order accurate interpolant of u.

degree polynomial

$$x \in [x_l^K, x_r^K] = \mathsf{D}^K \quad : \quad u(x,t) \approx u_h(x,t) = \sum_{i=1}^{N_{\mathrm{DG}}} u_h^K(x_i^K,t) l_i^K(x). \tag{1.71}$$

Here x_i are the Legendre-Gauss-Lobatto (LGL) grid points in D^K. The approximation is obtained by requesting the residual

$$R_h(x,t) = \frac{\partial u_h}{\partial t} + \frac{\partial(-u_h)}{\partial x}, \tag{1.72}$$

to be orthogonal to all test functions $l_i^K(x)$ on element K. That is

$$\int_{\mathsf{D}^K} R_h(x,t) l_i^K(x) d\, x = 0, \quad i = 1, \dots, N_{\mathrm{DG}}, \tag{1.73}$$

on element K. Now, integrating by parts two times yields the semi discrete approximation

$$\frac{d}{dt}\mathbf{u}_h^K + (\mathcal{M}^K)^{-1}\mathcal{S}^K(-\mathbf{u}_h^K) = (\mathcal{M}^K)^{-1}\left[\mathbf{l}^K(x)((-u_h^K) - (-u_h^K)^*)\right]_{x_l^K}^{x_r^K}. \tag{1.74}$$

Here $(-u_h^K)^*$ is the numerical flux, $\mathbf{u}_h^K = [u_1^K, \dots, u_{N_{\mathrm{DG}}}^K]^T$ are the degrees of freedom (nodal values of the function) on element K and

$$\mathcal{M}_{ij}^K = \int_{\mathsf{D}^K} l_i^K \, l_j^K dx, \quad \mathcal{S}_{ij}^K = \int_{\mathsf{D}^K} l_i^K \frac{dl_j^K}{dx} dx. \tag{1.75}$$

The differentiation matrix $\mathcal{D}^K = (\mathcal{M}^K)^{-1}\mathcal{S}^K$ can be explicitly constructed from the Vandermonde transformation matrix, $\mathcal{V}$, between the nodal,

$$u_h(r,t) = \sum_{i=1}^{N_{\mathrm{DG}}} u_h(r_i,t) l_i(r),$$

and modal,

$$u_h(x,t) = \sum_{n=1}^{N_{\mathrm{DG}}} \hat{u}_n \tilde{P}_{n-1}(r),$$

representations on the reference element $r \in [-1,1]$. Precisely, $\mathcal{D}^K = \frac{2}{h^K}\mathcal{V}_r\mathcal{V}^{-1}$, with h^K being the length of element K.

At the heart of the DG method is the choice of the numerical flux. To this end we will use an upwind flux, which in this example corresponds to

$$(-u)^* = \{\{-u\}\} + \frac{[\![u]\!]}{2}.$$

Here $\{\{-u\}\} = \frac{u^+ + u^-}{2}$ is the average of the values of the discrete solution ("$-$" is the interior node) at a grid interface and $[\![u]\!] = \hat{n}^- u^- + \hat{n}^+ u^+$ is the jump in u at an interface.

Equation (1.74) can be discretized in time by any ODE method, we choose to use the classic fourth order accurate Runge-Kutta method.

Hybrid Hermite discontinuous Galerkin method

We are now ready to describe the coupling of the Hermite and discontinuous Galerkin method. To make matters concrete we describe the method with an example. Consider the one dimensional system

$$\begin{aligned} u_t - u_x &= 0, \\ v_t + v_x &= 0, \end{aligned} \tag{1.76}$$

for $t \geq 0$, $x_{\min}^{\text{Her}} \leq x \leq x_{\max}^{\text{Her}}$, with initial data

$$u(x, 0) = U_0(x), \quad v(x, 0) = V_0(x),$$

and boundary conditions

$$u(x_{\min}, t) = v(x_{\min}, t), \quad u(x_{\max}, t) = v(x_{\max}, t).$$

Let the primal and dual grids be as in (1.69) - (1.70) with grid spacing $h_x = (x_{\max}^{\text{Her}} - x_{\min}^{\text{Her}})/N_x$ and let

$$x \in (x_0, x_1) = \mathsf{D}^0, \quad x \in (x_{N_x-1}, x_{N_x}) = \mathsf{D}^{N_x-1}, \tag{1.77}$$

be elements, with $N_x^0 = N_x^{N_x-1}$ LGL nodes, at the boundaries of the computational domain.

The computation is initiated at some time level $t = t_n$ by assigning the $(N_x - 1) \times (m + 1)$ coefficients

$$c_{l0}^{\mathrm{p}}[x_i], \quad i = 1, \ldots, N_x - 1, \quad l = 0, \ldots, m, \tag{1.78}$$

in the interior nodes on the primal Hermite grid and the $2 \times N_{\mathrm{DG}}$ nodal values $u^0(x_i, t)$, $u^{N_x-1}(x_i, t)$ on the DG elements.

Evolution of Hermite data   The first step of the evolution of the solution is to construct the Hermite interpolant (1.5) centered at the interior points of the dual grid, i.e. x_i, $i = 3/2, \ldots, N_x - 3/2$. At this stage a recursion relation similar to (1.12) can be used to evaluate the time derivative(s) of the solution needed for Runge-Kutta or Taylor series time stepping. If a Runge-Kutta method is used the interpolant on each cell is advanced independently for one half time step by taking $N_{\mathrm{sub,H}}$ sub-steps of size $\delta_{\mathrm{H}} t = (\Delta t/2)/N_{\mathrm{sub,H}}$. Here we fix $N_{\mathrm{sub,H}}$ at the beginning of the computation but we note that it is straightforward to allow for $\delta_{\mathrm{H}} t$ to change adaptively on cells or in between time levels. If a Taylor series method is used, the solution is advanced by a single time step of size $\Delta t/2$.

Boundary data for the DG method   The next step in the hybrid algorithm is to construct the required boundary data for the DG-method. As the DG-method will be time stepped with a fourth order accurate Runge-Kutta method the boundary data should be constructed with, at least, the same accuracy. Thus, if the Hermite elements are time-stepped with RK4 we construct the DG boundary data by extrapolating $c_{l0}[x_i]$, $c_{l1}[x_i]$, $i = 3/2, N_x - 3/2$ at times $t = t_n + j\delta_{\mathrm{H}} t$, $j = 0, 1, \ldots$, yielding approximations to u and u_t, at the gridpoints x_1 and x_{N_x-1}. The approximations to

u and u_t are then used to construct interpolants along the DG faces:

$$u_\rho(x_q, t) = \sum_{l=0}^{3} \gamma_{l\rho} \left(\frac{t - (t_n + \rho\delta_{\mathrm{H}})}{\delta_{\mathrm{H}}} \right)^l, \quad \rho = 1, \ldots, N_{\mathrm{sub,H}}, \quad q = 1, N_x - 1. \quad (1.79)$$

Here

$$t_{\mathrm{s}} = t_n + (\rho - 1)\delta_{\mathrm{H}} \leq t \leq t_n + \rho\delta_{\mathrm{H}} = t_{\mathrm{e}}, \quad (1.80)$$

and for $q = 1, N_x - 1$

$$\begin{aligned}
\gamma_{0\rho} &= u(x_q, t_{\mathrm{s}}), \\
\gamma_{1\rho} &= u_t(x_q, t_{\mathrm{s}}), \\
\gamma_{2\rho} &= -3u(x_q, t_{\mathrm{s}}) + 3u(x_q, t_{\mathrm{e}}) - 2u_t(x_q, t_{\mathrm{s}}) - u_t(x_q, t_{\mathrm{e}}), \\
\gamma_{2\rho} &= 2u(x_q, t_{\mathrm{s}}) - 2u(x_q, t_{\mathrm{e}}) + u_t(x_q, t_{\mathrm{s}}) + u_t(x_q, t_{\mathrm{e}}).
\end{aligned} \quad (1.81)$$

When the Hermite elements are time stepped with Taylor series the approximate solution (1.11) can be directly evaluated at the boundary of the DG-elements.

Evolution of DG data When the approximate solution has been constructed at x_1 and x_{N_x-1} it can be used to evolve the DG data to time $t_n + \frac{\Delta t}{2}$ using a Runge-Kutta method with time steps $\delta_{\mathrm{DG}} = \Delta t / N_{\mathrm{sub,DG}}$.

Boundary data for the Hermite method When the DG solution has been advanced to time $t_n + \Delta t/2$ we must find the solution and its derivatives at x_i, $i = 1/2, N_x - 1/2$ so that a new Hermite interpolant can be constructed and evolved at the primal grid points x_i, $i = 1, N_x$. Here we take the straightforward approach and use the algorithm suggested by Fornberg [28] to compute the required derivatives using the solution at the LGL points in the DG elements.

With the m derivatives known at the first and last dual grid points (and all the interior) interpolating Hermite polynomials on the form (1.11) can again be advanced

with a Taylor series or Runge-Kutta method. advanced with $N_{\mathrm{sub,H}}$ substeps of Runge-Kutta. The only difference from the first half-step is that the DG boundary data can be constructed without extrapolation, see Figures 1.13 and 1.14.

1.4.3 Experiments in one dimension

This section presents some experiments illustrating the properties of the method in one dimension.

To assess the accuracy and stability we solve (1.76) with $x_{\mathrm{min}}^{\mathrm{Her}} = -1$ and $x_{\mathrm{max}}^{\mathrm{Her}} = 1$ and with initial data

$$U_0(x) = V_0(x) = \sin \omega_0 \pi x. \tag{1.82}$$

The solution is advanced to time $t = 40$ and the error

$$
\begin{aligned}
\mathcal{E} = \max \Big(& \|U_0(\cdot) - u(\cdot, 40)\|_\infty, \|V_0(\cdot) - v(\cdot, 40)\|_\infty, \\
& \|U_0(\cdot) - u^0(\cdot, 40)\|_\infty, \|V_0(\cdot) - v^0(\cdot, 40)\|_\infty, \\
& \|U_0(\cdot) - u^{N_x-1}(\cdot, 40)\|_\infty, \|V_0(\cdot) - v^{N_x-1}(\cdot, 40)\|_\infty \Big),
\end{aligned}
\tag{1.83}
$$

is measured.

Performance comparison with a summation by parts discretization and a Pade method

To study the performance of the Hermite-DG hybrid scheme we solve (1.76), with $\omega_0 = 5$ to time $t = 16$ for matched order Hermite-DG and Taylor time stepping. We keep the CFL number fixed (close to the limit of stability) and decrease the element sizes and plot the time (start to end) it takes to complete the computation against the error at $t = 16$. The number of Runge-Kutta time-steps used in the DG-element are chosen close to the stability limit.
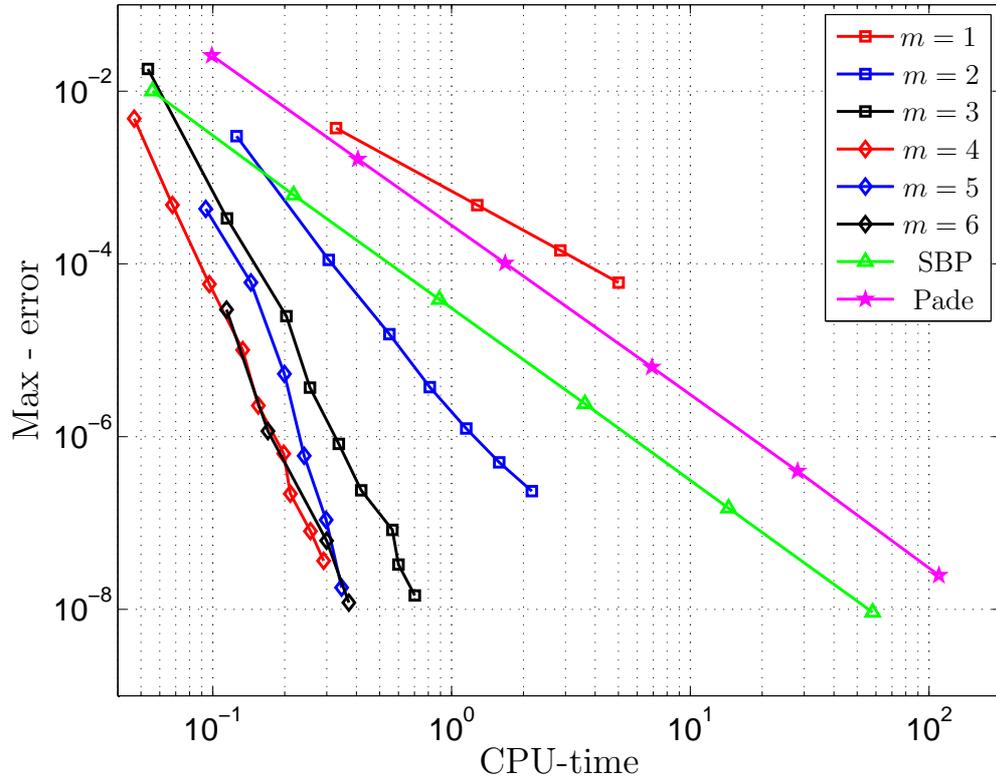
Figure 1.15: Maximum error with CPU-time for $m = 1, \ldots, 7$, using Hermite Runge-Kutta, that is $3 - 15$th order in space and 4th order in time. SBP is summation by parts with 8th in the interior, 4th order exterior. Pade is 4th order.

The same problem is then solved with Runge-Kutta time stepping and with two different spatial discretizations, the classic fourth order accurate compact Pade scheme [23] with fourth order closures at the boundary and the diagonal 4-8 summation by parts operators as given in appendix C.4 in [43].

The results are displayed in Figure 1.15. The maximum error is plotted against the time it takes (start to end) to complete the computation. We observe that for errors smaller than 0.1% the Hermite-DG methods of order greater than 3 are more efficient than the finite difference methods. Obviously the the efficiency of the methods will depend on the computer (in this case a MacBook 2GHz Intel Core 2

Duo), yet the overall trend should be the same. For this hardware it appears that the 9th order method is most efficient over a wide range of error levels.

1.4.4 Description of the hybrid method in two dimensions

This section describes in detail how we couple the Hermite and nodal discontinuous Galerkin methods in 2D for general geometries. We will illustrate the algorithm by applying it to Maxwell's equations. We use Hermite-Taylor here because for a constant coefficient system, it is more efficient to use Taylor series to solve the ODE in time instead of Hermite-Runge-Kutta. In addition, a universal time step $\Delta t_{\mathrm{H}}^{\mathrm{Taylor}}$ can be used on the Hermite domain. As in 1D, the discontinuous Galerkin method is used to solve the equations along the boundary and the Hermite method is used in the interior. The time step size comparison between Hermite-RK and Hermite-Taylor can be seen in Figure 1.16 .
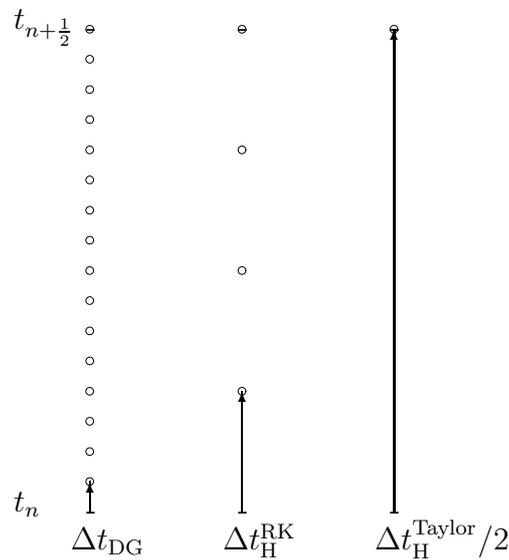
Figure 1.16: Time step size for discontinuous Galerkin t_{DG}, for Hermite-Runge-Kutta $t_{\mathrm{H}}^{\mathrm{RK}}$ where $N_{\mathrm{sub,H}} = 4$, for Hermite Taylor $\Delta t_{\mathrm{H}}^{\mathrm{Taylor}}$, in a half time step.

Consider the two-dimensional Maxwell's equations in transverse magnetic form (TM) and nondimensionalized in vacuum so that the magnetic permeability $\mu = 1$, and the electric permittivity $\epsilon = 1$. That is:

$$\frac{\partial H^x}{\partial t} = -\frac{\partial E^z}{\partial y}, \tag{1.84}$$

$$\frac{\partial H^y}{\partial t} = \frac{\partial E^z}{\partial x}, \tag{1.85}$$

$$\frac{\partial E^z}{\partial t} = \frac{\partial H^y}{\partial x} - \frac{\partial H^x}{\partial y}. \tag{1.86}$$

We assume a perfect electric conductor (PEC) so that the boundary condition is that the tangential component of the electric field, E^z, vanishes on the boundary.
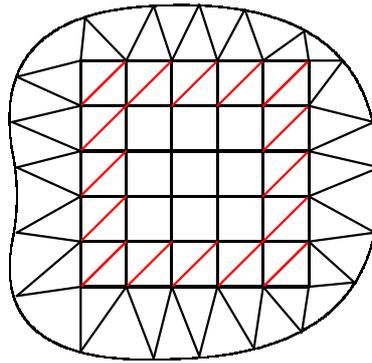
Figure 1.17: A typical hybrid grid used in two dimensions. From the outside is the discontinuous Galerkin grid ($\mathcal{G}^{\mathrm{DG}}$) composed of all black triangles, the coupling grid ($\mathcal{G}^{\mathrm{C}}$), consisting of triangles with one red side and the Hermite grid ($\mathcal{G}^{\mathrm{H}}$) consisting of the squares in the middle.

We describe the algorithm for a simply connected geometry but note that it is trivially extended to domains with holes. For a simply connected geometry the domain can be discretized by three grids denoted: the discontinuous Galerkin grid, $\mathcal{G}^{\mathrm{DG}}$, the Hermite grid, $\mathcal{G}^{\mathrm{H}}$, and the coupling grid, $\mathcal{G}^{\mathrm{C}}$, see Figure 1.17. The grid $\mathcal{G}^{\mathrm{H}}$

consist of a staircased rectilinear grid filling as much as possible of the interior. The coupling grid consist of one layer of rectangular elements immediately outside $\mathcal{G}^{\mathrm{H}}$ that have been split into triangles. The DG grid consists of a triangulation that fills the domain between $\mathcal{G}^{\mathrm{C}}$ and the boundary, see Figure 1.17. In short, we use Hermite solver on $\mathcal{G}^{\mathrm{H}}$, DG solver on $\mathcal{G}^{\mathrm{DG}}$, both of the Hermite solver and the DG solver on $\mathcal{G}^{\mathrm{C}}$.

We use the Hermite solution on $\mathcal{G}^{\mathrm{H}}$ to calculate the numerical flux for the DG solution on $\mathcal{G}^{\mathrm{C}}$, and use the DG solution on $\mathcal{G}^{\mathrm{C}}$ to update the Hermite solution on the dual grid of $\mathcal{G}^{\mathrm{C}}$.

We now describe the individual stages in the algorithm in detail starting with the evolution of the Hermite solution on $\mathcal{G}^{\mathrm{H}}$ for half a Hermite time step, $\Delta t_{\mathrm{H}}/2$.

Hermite Taylor algorithm on a square domain

The dual $\mathcal{G}^{\mathrm{d}}$ and primal $\mathcal{G}^{\mathrm{p}}$ grid now consists of

$$(x_i, y_j) = (x_{\min}^{\mathrm{Her}} + ih_x,\ y_{\min}^{\mathrm{Her}} + jh_y),$$

with $i, j = \frac{1}{2}, \frac{3}{2}, \ldots$ and $i, j = 1, 2, \ldots$ and $\mathcal{G}^{\mathrm{d}}$ is slightly bigger than $\mathcal{G}^{\mathrm{p}}$. As before, at the start of a time-step $t = t_n$, the approximate solution at four adjacent vertices on the primal grid are used to form a local power series (through the 2D Hermite interpolation process) centered at one of the points on the dual grid

$$p_{i+1/2,j+1/2}(x, y, t_n) = \sum_{l_1=0}^{2m+1} \sum_{l_2=0}^{2m+1} c_{l_1,l_2,0} \left(\frac{x - x_{i+1/2}}{h_x}\right)^{l_1} \left(\frac{y - y_{j+1/2}}{h_y}\right)^{l_2}. \qquad (1.87)$$

To advance (1.87) in time it is expanded in a Taylor series

$$p_{i+1/2,j+1/2}(x, y, t) = \sum_{l_1=0}^{2m+1} \sum_{l_2=0}^{2m+1} \sum_{s=0}^{2m+1} c_{l_1,l_2,s} \left(\frac{x - x_{i+1/2}}{h_x}\right)^{l_1} \left(\frac{y - y_{j+1/2}}{h_y}\right)^{l_2} \left(\frac{t - t_n}{\Delta t}\right)^{s},$$

whose coefficients for $s > 0$ are determined by repeated differentiation of the PDE. For example, considering the equation (1.86)

$$\frac{\partial E^z}{\partial t} = \frac{\partial H^y}{\partial x} - \frac{\partial H^x}{\partial y},$$

we may insert the Taylor series into and obtain the recursive relation of the coefficient $c_{l_1,l_2,s}^{E^z}$ as

$$c_{l_1,l_2,s}^{E^z} = \frac{\Delta t}{h_x} \frac{l_1 + 1}{s} c_{l_1+1,l_2,s-1}^{H^y} - \frac{\Delta t}{h_y} \frac{l_2 + 1}{s} c_{l_1,l_2+1,s-1}^{H^x},$$

where

$$l_1, l_2 = 0, \dots, 2m + 1 - s, \ s = 1, \dots, 2m + 1.$$

As in the one dimensional case, evolution in time is performed by evaluating the Taylor series at $t_{n+1/2} = t_n + \Delta t_{\mathrm{H}}/2$. We can relate the mixed derivatives at the half time level to the coefficients

$$\frac{h_x^{l_1} h_y^{l_2}}{l_1! l_2!} \frac{\partial^{l_1+l_2} u(x_{i+1/2}, y_{j+1/2}, t_{n+1/2})}{\partial x^{l_1} \partial y^{l_2}} = \sum_{s=0}^{2m+1} c_{l_1,l_2,s} \left(\frac{\Delta t_{\mathrm{H}}}{2}\right)^s, \ l_1, l_2 = 0, \dots m.$$

Discontinuous Galerkin solver with curved elements

The next step is to call the DG solver to evolve the solution on $\mathcal{G}^{\mathrm{DG}}$ and $\mathcal{G}^{\mathrm{C}}$ until the DG time level reaches $t_{\mathrm{DG}} = t_n + \Delta t_{\mathrm{H}}/2$. Again, the DG method we use is the nodal-DG method by Hesthaven and Warburton [38]. On the Kth element D^K, the solution is approximated by an $N = N_{\mathrm{DG}} - 1$ degree polynomial

$$u(r) \simeq u_h^K(r) = \sum_{i=1}^{N_p} u^K(r_i, t) l_i^K(r),$$

where r_i are the two-dimensional LGL grid points in D^K, $N_p = N(N + 1)/2$ is the number of nodal points in each element and $l_i^K(r)$ is the multidimensional Lagrange polynomial based on the grid points r_i. For the curved elements along the boundary,

the straight sides of the element, as well as the node locations, are changed to fit the exact physical boundary. For details of the implementation of the nodal DG methods on curvilinear we refer to [38, 37].

The local semi-discrete scheme is

$$
\begin{aligned}
\frac{\partial H_h^x}{dt} &= -D_y E_h^z + \frac{1}{2}(JM)^{-1} \int_{\partial D^K} (\hat{n}_y[E_h^z] + \alpha(\hat{n}_x[\![H_h]\!] - [H_h^x]))l(x)dx, \\
\frac{\partial H_h^{yx}}{dt} &= D_x E_h^z + \frac{1}{2}(JM)^{-1} \int_{\partial D^K} (-\hat{n}_x[E_h^z] + \alpha(\hat{n}_y[\![H_h]\!] - [H_h^x]))l(x)dx, \\
\frac{\partial E_h^z}{dt} &= D_x H_h^y - D_y H_h^x + \frac{1}{2}(JM)^{-1} \int_{\partial D^K} (\hat{n}_y[H_h^x] - \hat{n}_x[H_h^y] - \alpha[E_h^z])l(x)dx,
\end{aligned}
$$

where $H = (H^x, H^y)$, M is the mass matrix, J is the Jacobian and D_x, D_y are the derivative matrices. The normal vector is $\hat{n} = (\hat{n}_x, \hat{n}_y)$, $[\![u]\!] = \hat{n}^- \cdot u^- + \hat{n}^+ \cdot u^+$ represent the jump along the normal $\hat{n}$, and $[q] = q^- - q^+ = \hat{n} \cdot [\![q]\!]$.

We set $\alpha = 1$ for the numerical flux, which corresponds to the classic upwind flux scheme. On the straight-sided triangles the differentiation and lift operators are direct generalizations of the 1D operators. On the curvilinear elements differentiation is performed by cubature and the lift integrals are implemented using Gauss quadrature. the time stepping on all DG elements is done by the classic Runge-Kutta method. Finally, a mirror principle is used to enforce the boundary condition on E^z as follows

$$
[E^z] = \hat{n} \cdot [\![E^z]\!] = 2(E^z)^-.
$$

Communication between the discontinuous Galerkin and the Hermite solver

In order to compute the numerical flux at the shared edges between $\mathcal{G}^C$ and $\mathcal{G}^H$, double nodal data values are recorded on that edge of the elements in $\mathcal{G}^C$. This allows for the evolution of the DG solution on $\mathcal{G}^C$. The "exterior" of the double

nodal values are populated by evaluating the local Hermite-Taylor polynomial on the Hermite elements adjacent to the $\mathcal{G}^{\mathrm{C}}$ at the required nodal locations along the shared edges. As the DG solver has to take significantly smaller time steps, this is done at multiple time levels $t_{DG} = t_n + i\Delta t_{\mathrm{DG}}/2 < t_n + \Delta t_{\mathrm{H}}/2$.

Remark 3. *We note that the using the function values from $\mathcal{G}^{\mathrm{H}}$ to construct the numerical fluxes on elements in $\mathcal{G}^{\mathrm{C}}$ is similar to enforcing a time dependent boundary condition, which causes a reduction of the order of convergence when Runge-Kutta methods are used, see [1]. To correct for this we use corrected intermediate boundary conditions as suggested in [1]. These conditions involve the time derivatives of the solution which is easily computed as we know the space-time Hermite-Taylor polynomial.*

When the DG-solution has been evolved a full half step we must communicate the derivative data needed by the Hermite solver. To achieve this we construct a mapping that is used to communicate between two solvers on elements in $\mathcal{G}^{\mathrm{C}}$. Consider a 2D polynomial of order m centered at a rectangle with four vertices (x_i, y_i) around (x_c, y_c)

$$
p(x,y) = \sum_{i=0}^{m} \sum_{j=0}^{m} c_{i,j} \left(\frac{x - x_c}{h_x} \right)^i \left(\frac{y - y_c}{h_y} \right)^j.
$$

Here h_x, h_y are the length of the sides to the rectangle and c_{ij} are scaled derivatives of $p(x,y)$ at (x_c, y_c)

$$
c_{ij} = \frac{h_x^i}{i!} \frac{h_y^j}{j!} p(x,y)|_{(x_c,y_c)}.
$$

Evaluating $p(x,y)$ at a nodal point (x_p, y_p) is equivalent to the scalar-product,

$$
p(x_p, y_p) = M_p C^T,
$$

where M_p, C are row vectors

$$M_p = \left[\left(\frac{x - x_c}{h_x} \right)^0 \left(\frac{y - y_c}{h_y} \right)^j, \left(\frac{x - x_c}{h_x} \right)^1 \left(\frac{y - y_c}{h_y} \right)^j, \dots, \right.$$
$$\left. \left(\frac{x - x_c}{h_x} \right)^m \left(\frac{y - y_c}{h_y} \right)^j, \dots \right] \qquad (1.88)$$

$$C = [c_{0,0}, c_{1,0}, \dots, c_{m,0}, \dots, c_{0,j}, c_{1,j}, \dots, c_{m,j} \dots], \quad j = 0, \dots m. \qquad (1.89)$$

For a set of N_p nodal points this can be formulated as an $N_p \times N_p$ mapping matrix $\mathcal{MP} = [M_p^1, \dots, M_p^{N_p}]^T$ multiplying C^T so that the equation

$$\mathcal{MP} \times C^T = [p(x_p^1, y_p^1), \ p(x_p^1, y_p^1), \dots, \ p(x_p^{N_p}, y_p^{N_p})]^T, \qquad (1.90)$$

can be used to map the nodal data to the scaled derivatives. Once the Hermite data is known the solution process can be repeated in the same way on the dual grid. This concludes a full time step.

As discussed in Section 1.4.4, the derivative data in the communication elements is obtained from $\mathcal{MP}$ which in turn is constructed by evaluating M_p at different nodal points. As there is a double set of nodes at the interface between the two triangles inside each communication element (see Figure 1.18) we can choose to use the points from either of the elements or the average. We have found that the choice of points has little or no effect and to this end we use the values from the upper triangle. The nodes we use to communicate between DG elements and the Hermite cell can be viewed in Figure 1.18, where the mapping is between all the nodal values from two triangular elements and the $x-$, $y-$ derivatives around the center of the square. The mapping is shown in Table 1.5; we map the nodal value to the derivatives and back into nodal values, the error is then calculated from the nodal value.
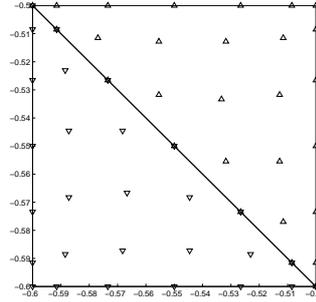
Figure 1.18: Node (LGL) location of two adjacent triangular elements in $\mathcal{G}^{\mathrm{C}}$ for $m = 3$, $N_{DG=7}$.

m	h	Error	ratio
1	0.1	7.67(E-02)	
1	0.08	3.07(E-02)	4.1
1	0.05	6.27(E-03)	3.4
2	0.1	2.93(E-03)	
2	0.08	7.98(E-04)	5.8
2	0.05	5.88(E-05)	5.5
3	0.1	2.99(E-05)	
3	0.08	5.35(E-06)	7.7
3	0.05	1.29(E-07)	7.9
4	0.1	4.08(E-07)	
4	0.08	4.51(E-08)	9.9
4	0.05	4.04(E-10)	10.0

Table 1.5: Maximum mapping error between Hermite derivative data and DG nodal data, where h is the length of the square element, ratio is the rate of convergence, m is the degree of the Hermite-Taylor polynomial, and the degree of polynomial used in DG is $N = 2m$.

1.4.5 Numerical experiments for Maxwell's equations in two dimensions

Hermite discontinuous Galerkin method on a square

We begin our two dimensional experiment without a curved boundary. We solve TM Maxwell's equation (1.84)-(1.86) on the square domain $[-1, 1] \times [-1, 1]$, with

the initial condition:

$$E^z = \sin(2\pi x)\sin(2\pi y) \ , \ H^x = H^y = 0. \tag{1.91}$$

The boundary condition (PEC) is now $H^x = 0$ at the left and right boundary (i.e. at $x = -1$, $x = 1$), and $H^y = 0$ at the top and bottom boundary (i.e. at $y = -1$, $y = 1$). The exact solution to this problem is

$$E^z = \sin(2\pi x)\sin(2\pi y)\cos(\sqrt{8}\pi t), \tag{1.92}$$

$$H^x = -\frac{1}{\sqrt{2}}\sin(2\pi x)\cos(2\pi y)\sin(\sqrt{8}\pi t), \tag{1.93}$$

$$H^y = \frac{1}{\sqrt{2}}\cos(2\pi x)\sin(2\pi y)\sin(\sqrt{8}\pi t). \tag{1.94}$$
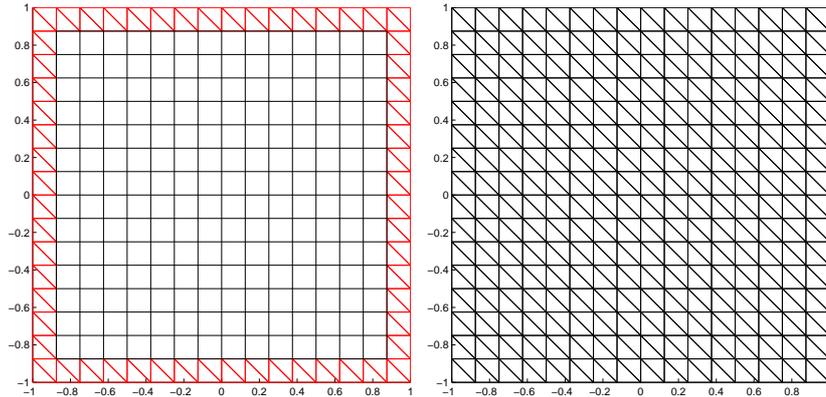
Figure 1.19: Grids for the HDG (left), DG only (right) methods solving TM Maxwell equation on a square $[-1, 1] \times [-1, 1]$ domain, with $h = 0.125$.

We first solve the equations using only the DG method on grids as shown to the right in Figure 1.19 and with different h. Then we solve the same equations using the hybrid method using a grid as shown to the left in Figure 1.19. The results are displayed in Table 1.6. The maximum error is of the same order between HDG

and DG. For DG methods, the computation cost is approximately 8 times larger, when the mesh size h is cut by half. While for the HDG methods, the computation cost only becomes approximately 4 times larger. As a matter of fact, the HDG methods will be one dimensional faster than the DG, because the Hermite methods are so efficient in the interior and the majority of the computation cost is along the boundaries and from the communication between the Hermite cells and the DG elements. For $N_{DG} = 6, 8$ the error in time becomes dominant so that we can not see the actual rate of convergence which suggests a smaller time step is required to keep the error in space and time of the same magnitude.

Hermite discontinuous Galerkin method on a disc

Our experiment with curved boundary in two dimensions is the evolution of a resonant mode of the unit disc. Precisely the mode we consider here is

$$E^z = J_6(\alpha_6 r) \cos(6\theta) \cos(\alpha_6 t), \tag{1.95}$$

$$H^x = \frac{\sin(\alpha_6 t)}{\alpha_6} \left(\frac{6 J_6(\alpha_6 r) \cos(\theta)}{r} \sin(6\theta) - \alpha_6 \sin(\theta) \cos(6\theta) \frac{(J_5(\alpha_6 r) - J_7(\alpha_6 r))}{2} \right), \tag{1.96}$$

$$H^y = \frac{\sin(\alpha_6 t)}{\alpha_6} \left(\frac{6 J_6(\alpha_6 r) \sin(\theta)}{r} \sin(6\theta) - \alpha_6 \cos(\theta) \cos(6\theta) \frac{(J_5(\alpha_6 r) - J_7(\alpha_6 r))}{2} \right), \tag{1.97}$$

which is a solution to the TM Maxwell's equation in a unit radius, cylindrical, metallic cavity. Here $J_l(z)$ is the lth Bessel function of the first kind and $\alpha_6 = 13.589290170541217$. This mode has six periods in the azimuthal direction and one "period" in the r direction. The solution at $t = 1$ of the E^z field is shown in Figure 1.20.

h	m	N_{DG}	Error	Rate
0.25		2	1.86(E-01)	
0.125		2	1.26(E-02)	3.9
0.0625		2	1.73(E-03)	2.9
0.25		4	3.80(E-03)	
0.125		4	1.54(E-04)	4.6
0.0625		4	5.02(E-06)	4.9
0.25		6	6.43(E-05)	
0.125		6	6.60(E-07)	6.6
0.0625		6	2.86(E-08)	4.5
0.25		8	1.81(E-06)	
0.125		8	7.32(E-08)	4.6
0.0625		8	4.43(E-09)	4
0.25	1	2	4.13(E-01)	
0.125	1	2	9.06(E-02)	2.2
0.0625	1	2	1.45(E-02)	2.6
0.25	2	4	1.62(E-02)	
0.125	2	4	5.71(E-04)	4.8
0.0625	2	4	1.62(E-05)	5.1
0.25	3	6	2.53(E-04)	
0.125	3	6	1.45(E-06)	7.4
0.0625	3	6	1.92(E-08)	6.2
0.25	4	8	1.04(E-05)	
0.125	4	8	7.14(E-08)	7.2
0.0625	4	8	1.19(E-08)	2.6

Table 1.6: Results of the HDG (bottom), and DG (top) methods solving TM Maxwell equation on a square $[-1, 1] \times [-1, 1]$ domain, at $T = 5$. m is the order of polynomials used in the Hermite cells, N_{DG} is the order of the polynomials used in the DG elements. Theoretical rate of convergence of the Hermite methods is $2m + 1$, and of the DG method is $N_{DG} + 1$. Rate is the actual rate of convergence on three different grids.

On the DG elements it is straightforward to use the above formulas for the initialization, for the Hermite method we use the mapping $\mathcal{MP}$ to assign the initial derivative data.
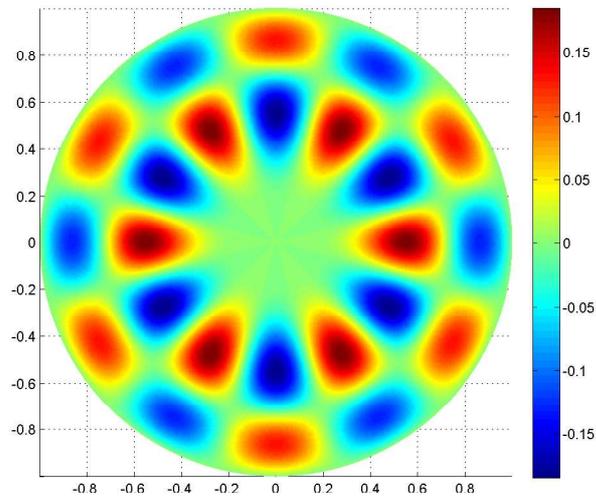
Figure 1.20: The solution at $t = 1$ of the E^z field.
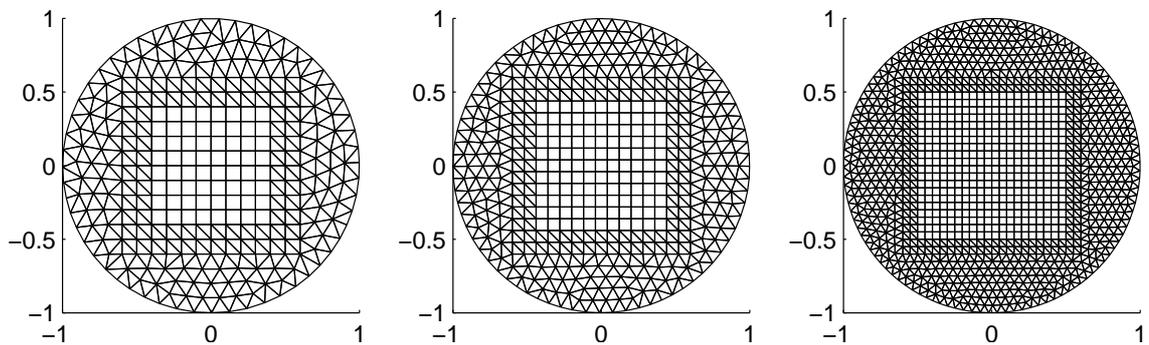
Figure 1.21: Three different HDG grids used in the numerical experiment with $h_x = h_y = 0.1, 0.08, 0.05$, where $h_x = h_y$ is the length of the sides of the square elements. The length of the sides of the triangular elements are roughly the same sides as those of the squares.

The computations are carried out on three grids with grid spacing $h_x = h_y = 0.1, 0.08$, and 0.05 as shown in Figure 1.21.

Grid generation The process of HDG grid generation is shown in Figure 1.22. First we determine the grids of the Hermite method $\mathcal{G}^{\mathrm{H}}$, in this case, a $[-0.5, 0.5] \times [-0.5, 0.5]$ square domain. Then we create the coupling grids $\mathcal{G}^{\mathrm{C}}$ along the outside boundary of the Hermite grids $\mathcal{G}^{\mathrm{H}}$. Finally, we use DISTMESH [69] by Persson and Strang, to generate the triangular DG elements between the coupling grids $\mathcal{G}^{\mathrm{C}}$ and the physical boundary, with the fixed nodes being exact the nodes in the exterior boundary of the coupling grids $\mathcal{G}^{\mathrm{C}}$. We have modified DISTMESH so that there are no nodes located between the nodes of the coupling grid $\mathcal{G}^{\mathrm{C}}$. A more efficient HDG grid, can be constructed by stair casing the Hermite grid so that more Hermite cells are included can be generated in a similar way see Figure 1.23, the thickness of the DG grids $\mathcal{G}^{\mathrm{DG}}$ can be adjusted to make the grids more efficient.
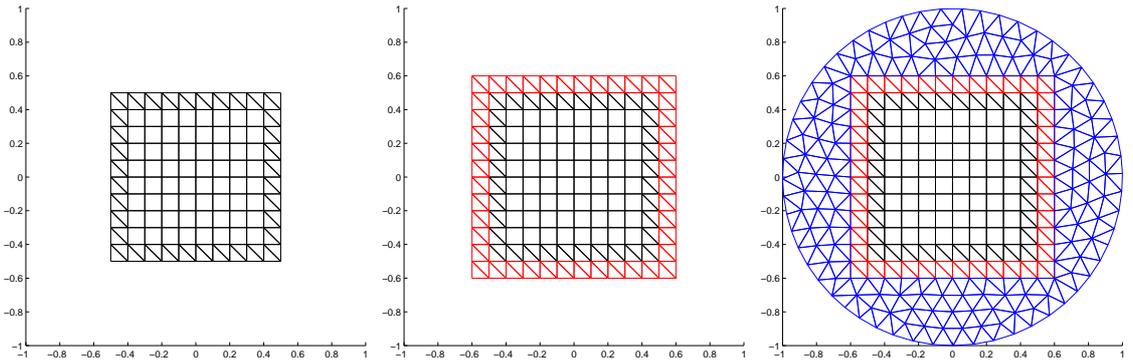
Figure 1.22: The process of HDG grid generation.

As discussed in Section 1.4.4, the derivative data in the communication elements is obtained from $\mathcal{MP}$ which in turn is constructed by evaluating M_p at different nodal points. As there is a double set of nodes at the interface between the two triangles inside each communication element (see Figure 1.18) we can choose to use the points from either of the elements or the average. We have found that the choice of points has little or no effect and to this end we use the values from the upper triangle.
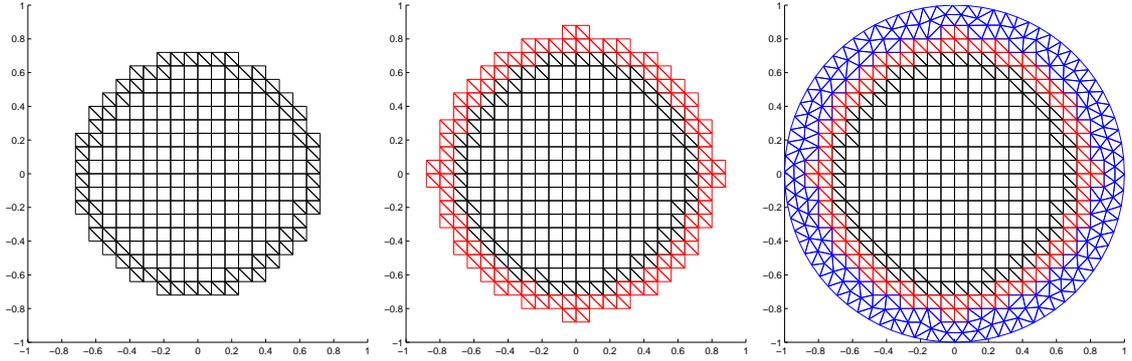
Figure 1.23: The process of HDG grid generation.

In Table we display the maximum error and convergence rates for different m at the final time $T = 5$. The time step in the DG elements has been chosen to match the order of accuracy of the Taylor time stepping used in the Hermite elements.

m	h	K	CFL	Δt_{DG}	Error	rate
1	0.1	447	0.8	1.08(E-02)	3.57(E-02)	
1	0.08	666	0.8	9.03(E-03)	2.26(E-02)	2.1
1	0.05	1525	0.8	5.71(E-03)	6.09(E-03)	2.8
2	0.1	447	0.8	5.11(E-03)	3.20(E-04)	
2	0.08	666	0.8	4.29(E-03)	4.61(E-05)	8.7
2	0.05	1525	0.8	2.71(E-03)	3.92(E-06)	5.2
3	0.1	447	0.7	2.88(E-03)	6.80(E-06)	
3	0.08	666	0.7	2.42(E-03)	1.66(E-06)	6.3
3	0.05	1525	0.7	1.53(E-03)	7.21(E-08)	6.7
4	0.1	447	0.7	1.83(E-03)	1.37(E-08)	
4	0.08	666	0.7	8.56(E-04)	1.42(E-09)	10.1
4	0.05	1525	0.7	4.09(E-04)	1.68(E-11)	9.4

Table 1.7: Maximum error at final time $T = 5$ of approximate solution of Maxwell's equations in a unit disc by the hybrid method. Here h is the length of the sides on the square elements, K is the number of triangular element in $\mathcal{G}^{\mathrm{DG}}$ and $\mathcal{G}^{\mathrm{C}}$, rate is the rate of convergence, m refers to the number of derivatives. The time step in the DG elements, Δt_{DG}, is chosen to match the error in the rest of the method.

m	$error_{l_2}$	$error_\infty$
2	11.52 %	13.96 %
3	0.59 %	0.80 %

Table 1.8: Relative error of the solution E^z as illustrated in Figure 1.24.

Resonant modes in a complex geometry

As a final example we solve Maxwell's equation, (1.84) - (1.86), in a complex geometry consisting of a polygonal perfect electric conductor. The geometry is shown in Figure 1.25. To find the resonant modes we evolve the initial data

$$E^z(x, y, 0) = e^{-200\left((x-0.3)^2 + (y-0.1)^2\right)}, \ H^x(x, y, 0) = 0, \ H^y(x, y, 0) = 0,$$

until time 100 and record E^z in the point $(x, y) = (-0.1, 0.7)$ throughout the computation. The time trace from the recorder is plotted in Figure 1.24 and a plot of E^z at $t = 100$ can be found in Figure 1.25.

As can be seen the solution has many "periods" over the interval $[0, 100]$ which will make it difficult to obtain an accurate solution. To estimate the error in the solution we use $m = 4$, $N = 2m$, $\Delta t_{\text{Her}} = 0.04$ to compute a reference solution. Using the reference solution we repeat the computation with $m = 2$ and 3 and compute the l_∞ and l_2-norm (in time) errors. The errors, given in Table 1.8, are quite big for $m = 2$ but probably small enough for engineering purposes for $m = 3$ and, presumably, even smaller for $m = 4$ (used in the computations below).

To find resonant frequencies we take the Fourier transform, in time, of the time trace. From the peaks in the spectrum, displayed in Figure 1.24, it is easy to identify the resonant frequencies. The shape of any of the resonant modes can be obtained by forcing the solution with a signal with the corresponding resonant frequency. To demonstrate this we pick the first resonant frequency, $\omega_0 \approx 0.359$, and add a forcing
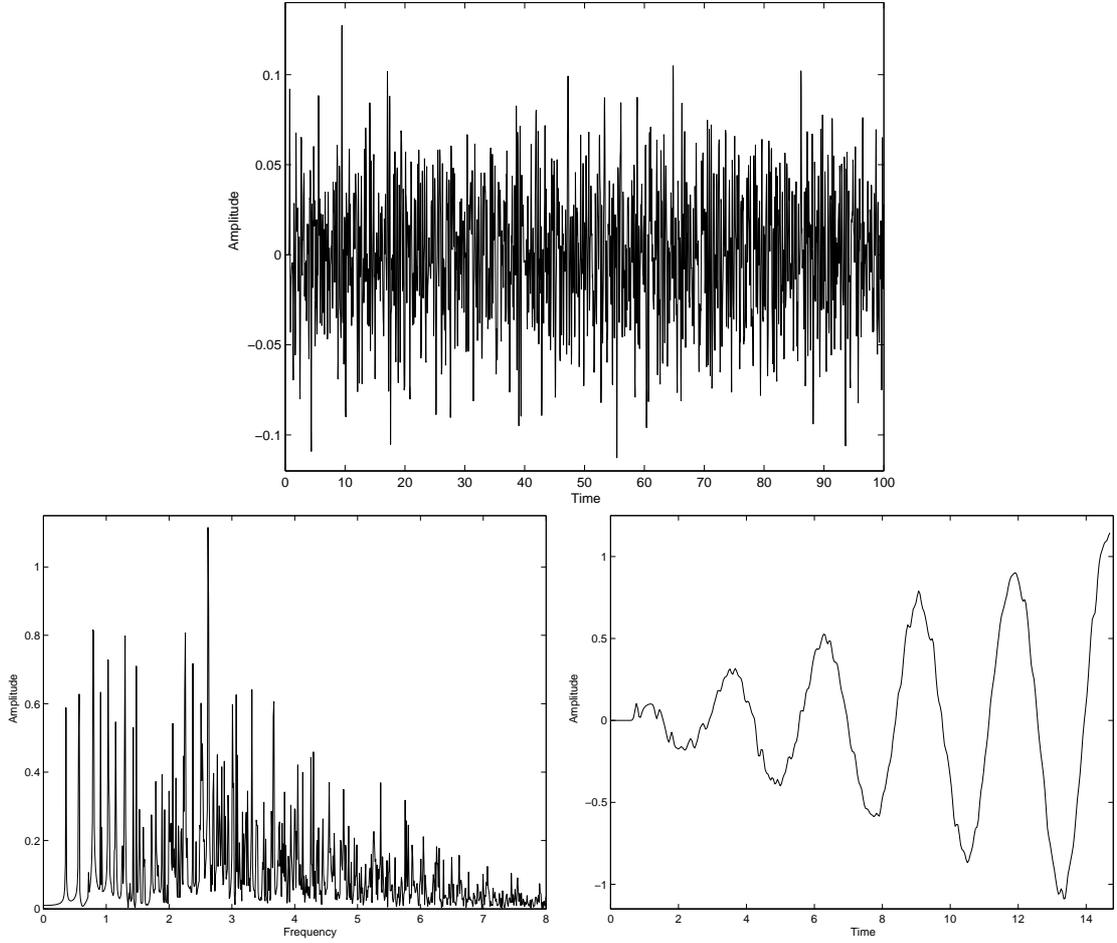
Figure 1.24: Top: The solution at the point $(x, y) = (-0.1, 0.7)$ for time 0 to 100 using $m = 4$, $N = 2m$ and $\Delta t_{Her} = 0.04$. Left: The solution E^z at $t = 100$ in the whole region. Right: Linear growth of the solution forced at a resonant frequency.

term to the equation for the electric field

$$\frac{\partial E^z}{\partial t} = \frac{\partial H^y}{\partial x} - \frac{\partial H^x}{\partial y} + 25\sin(\omega_0 t)e^{-200\left((x-0.3)^2 + (y-0.1)^2\right)}. \tag{1.98}$$

Recording the solution in the same point the expected linear growth is clearly observed, see Figure 1.24. The E^z-field, obtained with this forcing, at time $t = 14.76$ is displayed in Figure 1.25.
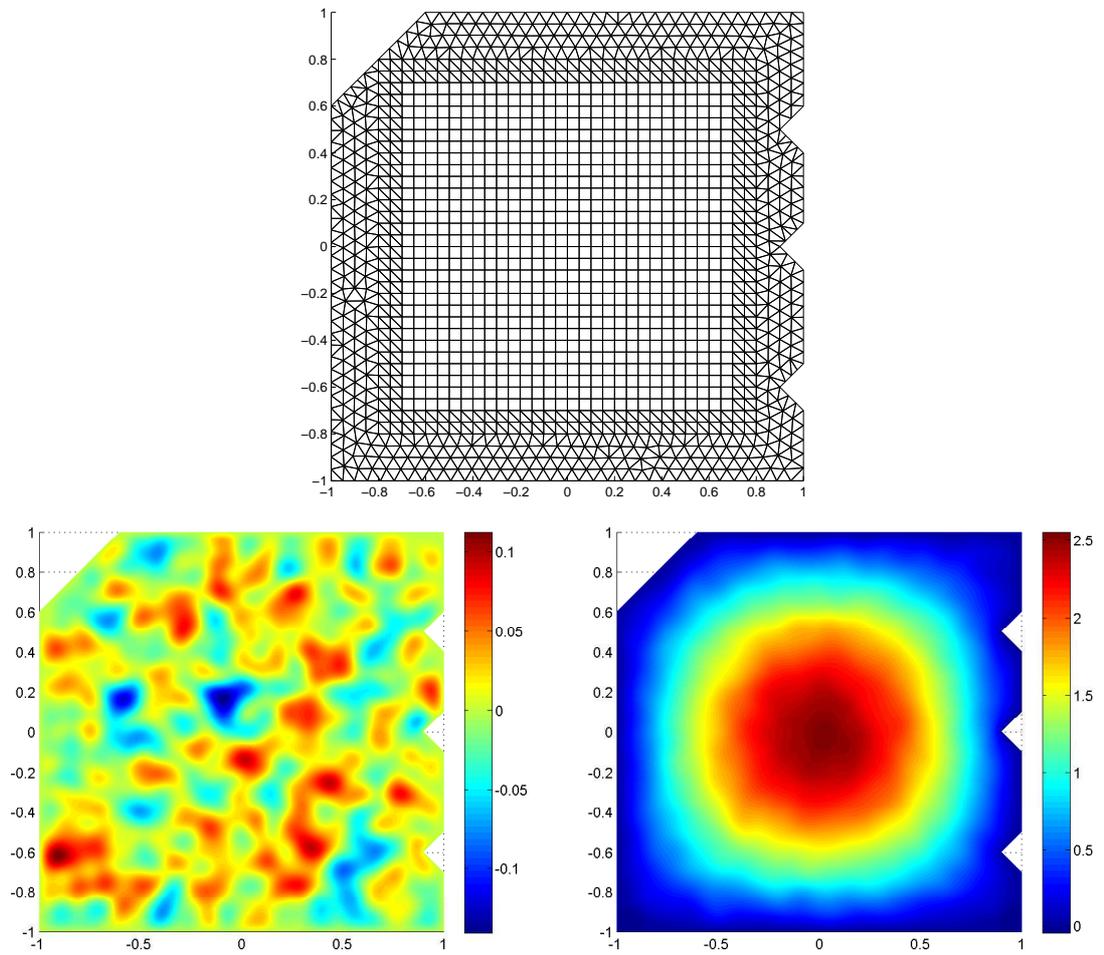
Figure 1.25: Top: The grid used in the numerical experiment for a complex geometry with $h \approx 0.05$. Left: The solution E^z at $t = 100$ in the whole region obtained with no forcing but initial data. Right: E^z at time $t = 14.76$ obtained with the resonant forcing term (1.98).

Chapter 2

Scattering from a lossless sphere

2.1 Introduction

In this chapter, we study a fundamental issue in scattering theory concerning the scattering poles of a lossless sphere. The motivation comes from the Singularity Expansion Method (SEM), where in practice we only deal with first order poles most of the time. However, a theoretically interesting question is whether there exist second or higher order scattering poles in both the acoustic and electromagnetic cases.

Here, Mie theory is used to solve the acoustic and electromagnetic scattering problems for spheres with lossless boundary conditions and an incident plane wave. For certain lossless impedance boundary conditions we construct second order poles in the acoustic and electromagnetic cases. A general procedure to directly construct lossless impedance boundary conditions producing high order poles is discussed.

The Singularity Expansion Method (SEM) [13, 16] was introduced in 1971 as a way to represent the solution of electromagnetic interaction or scattering problems in terms of singularities in the complex-frequency (s of two-sided-Laplace-transform)

plane. Particularly for the pole terms associated with a scatterer (natural frequencies), their factored form separates the dependencies on various parameters of the incident field, observer location, and scatterer characteristics, with an equally simple form in both the frequency (poles) and time (damped sinusoids) domain. An example of the SEM form of the surface current density solution [19, 15] is

$$\tilde{\vec{J}}_s(\vec{r}_s, s) = E_0 \sum_\alpha \tilde{f}(s_\alpha) \eta_\alpha(\vec{1}_1, \vec{1}_p) \vec{j}_{s_\alpha}(\vec{r}_s) [s - s_\alpha]^{-1} e^{-(s - s_\alpha)t_0}$$

$$+ \text{ singularities of } \tilde{f}(s) + \text{possible entire function}, \quad (2.1)$$

where only first-order poles have been included.

The incident field in the scattering problems considered here is taken as a plane wave with electric field

$$\tilde{\vec{E}}^{(inc)}(\vec{r}, s) = E_0 \tilde{f}(s) \vec{1}_p e^{-\gamma \vec{1}_1 \cdot \vec{r}}, \ \vec{E}^{(inc)}(\vec{r}, s) = E_0 f(t - \frac{\vec{1}_1 \cdot \vec{r}}{c}) \vec{1}_p. \quad (2.2)$$

Here $\gamma = s/c$, s is the complex frequency (Laplace-transform variable), c is the speed of light, $\vec{1}_1$ is the direction of incidence, $\vec{1}_p$ is the direction of polarization, $f(t)$ is the wave form, $\sim$ stands for two-sided Laplace transform, $\vec{r}_s$ is the coordinate of the surface of scatterer, s_α is natural frequency, $\gamma_\alpha = s_\alpha/c$, and $\vec{j}_{s_\alpha}$ is natural mode. The coupling coefficient η_α is defined as

$$\eta_\alpha(\vec{1}_1, \vec{1}_p) = \frac{\vec{1}_p \cdot \left\langle e^{-\gamma_\alpha \vec{1}_1 \cdot \vec{r}_s}, \vec{j}_{s_\alpha}(\vec{r}_s{}') \right\rangle}{\left\langle \vec{j}_{s_\alpha}(\vec{r}_s); \frac{\partial}{\partial s} \overleftrightarrow{Z_t}(\vec{r}_s, \vec{r}_s{}'; s)|_{s=s_\alpha}; \vec{j}_{s_\alpha}(\vec{r}_s{}') \right\rangle}, \quad (2.3)$$

where the kernel $\overleftrightarrow{Z_t}(\vec{r}_s, \vec{r}_s{}'; s)$ is symmetric and involves the free-space dyadic Green's function, the symbol $\langle\, ;\, ;\, \rangle$ denotes symmetric product integration over the common coordinates.

Although we typically only encounter first order scattering poles, an interesting question concerning the SEM is the existence of higher order scattering poles. Previously, Baum showed that 2nd order poles can be constructed for a transmission line

problem [18]. As the transmission line problem is finite-dimensional, the scattering matrix can be used directly to find the poles (i.e. the eigenvalues). Here we focus on the infinite dimensional case and consider, as an example, a classical model problem, the scattering of a plane wave from a sphere.

Our main result is that higher order poles can be constructed for certain impedance boundary conditions for the acoustic scattering problem, while for hard and soft spheres, we find that there only exist first order scattering poles. For electromagnetic scattering problems, we find that there only exist simple poles for a perfectly conducting spherical scatterer. For a surface impedance loaded sphere, certain impedance boundary condition can be constructed in a way that is equivalent to the acoustic scattering case, thus arbitrary order scattering poles can be generated. For a sheet impedance loaded sphere, we only find a second order scattering pole for a non-physical case, which leads to the conjecture that there only exist first order scattering poles in this case. Foster's theorem is applied to the impedance boundary condition to ensure that the scatterer is lossless.

2.2   Scattering from a lossless acoustic sphere

Consider an incident plane sound wave propagating in some direction being scattered off a sphere; see Figure 2.1. An explicit formula for the scattered solution can be written in terms of spherical harmonics. By imposing different boundary condition on the scatterer, we are able to construct first, second and higher order poles. Foster's theorem is applied to the impedance boundary condition to ensure that the system is lossless.
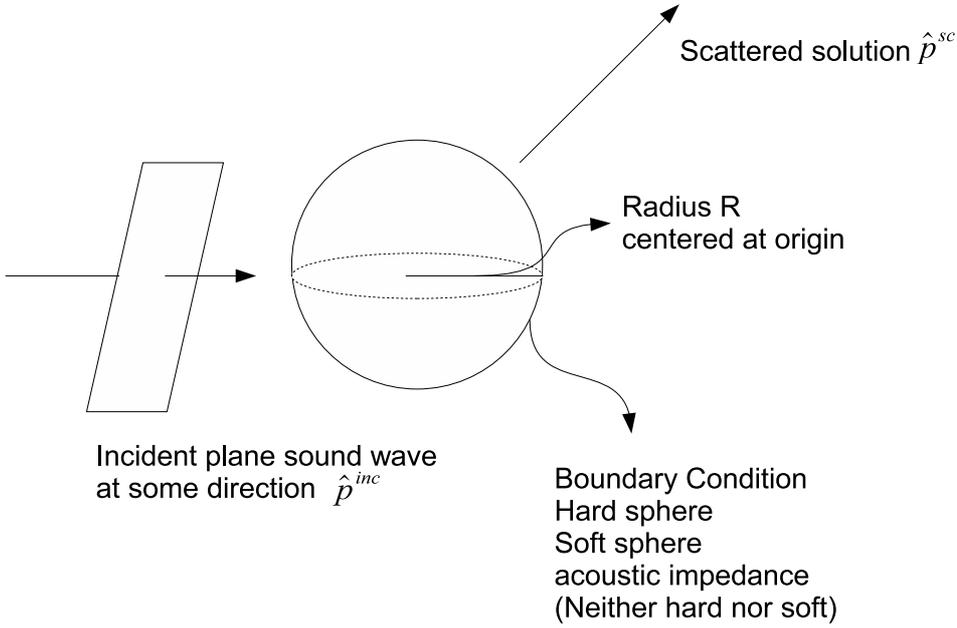
Figure 2.1: Scattering from a lossless sphere

2.2.1　Formulation of the acoustic scattering

The equation of linear acoustics are

$$\frac{\partial \rho}{\partial t} + \rho_0 \nabla \cdot v = 0, \tag{2.4}$$

$$\rho_0 \frac{\partial v}{\partial t} + \nabla p = 0, \tag{2.5}$$

$$p = c^2 \rho, \ c^2 = \left(\frac{\partial p}{\partial \rho}\right)_0, \tag{2.6}$$

where p is the acoustic pressure, ρ is the density, v is the fluid velocity, and c is the speed of sound. Consider a scattering problem for the above equation with an incident plane wave and some different but lossless boundary conditions on a sphere. In this case the SEM is just to use spherical harmonics to solve the equation so that the scattering poles can be located.

We first derive the wave equation from the above equations, so that later on we can treat the problem mathematically, disregarding the physical meaning of the

variables. Inserting $p = c^2 \rho$ into (2.4) gives

$$\frac{\partial p}{\partial t} + c^2 \rho_0 c^2 \nabla \cdot v = 0, \tag{2.7}$$

$$\rho_0 \frac{\partial v}{\partial t} + \nabla p = 0. \tag{2.8}$$

Taking the time derivative of the equation (2.7) and eliminating $\frac{\partial v}{\partial t}$ we find

$$\frac{\partial^2 p}{\partial t^2} + \rho_0 \nabla \cdot (-\frac{1}{\rho_0} \nabla p) = 0,$$

$$\Rightarrow \quad \frac{\partial^2 p}{\partial t^2} - c^2 \nabla^2 p = 0. \tag{2.9}$$

Similarly for v we have

$$\frac{\partial^2 v}{\partial t^2} + \frac{1}{\rho_0} \nabla(-\rho_0 c^2 \nabla \cdot v) = 0 \tag{2.10}$$

$$\Rightarrow \quad \frac{\partial^2 v}{\partial t^2} - c^2 \nabla(\nabla \cdot v) = 0. \tag{2.11}$$

Assuming v is irrotational (i.e. $\nabla \times v = 0$), then $\nabla(\nabla \cdot v) = \nabla^2 v$, we obtain the wave equation

$$\frac{\partial^2 v}{\partial t^2} - c^2 \nabla^2 v = 0. \tag{2.12}$$

Let u be either p or v and perform a Laplace transformation of the wave equation (2.12) to derive

$$(\nabla^2 - \gamma^2)\hat{u} = 0, \quad \text{where} \quad \gamma = \frac{s}{c}. \tag{2.13}$$

Let the scatterer be a sphere with radius 1 and assume that the impedance boundary condition has the following form

$$\frac{\partial \hat{u}}{\partial n} + \alpha(s)\hat{u} = 0, \quad \text{at} \quad r = 1, \tag{2.14}$$

where n is the outward normal. Let the plane wave be

$$\hat{u}^{(inc)} = e^{-\gamma(0,0,1)\cdot(x,y,z)} = e^{-\gamma z}, \tag{2.15}$$

then the scattered solution $\hat{u}^{(sc)}$ can be solved explicitly from (2.13),(2.14),(2.15). We study the behaviour of the scattering problem as a function of the boundary conditions in the following sections.

2.2.2 Hard and soft spherical scatterer

First, let us relate the mathematical impedance function $\alpha(\gamma)$ to the actual acoustic impedance $Z_a(s)$, where the units of acoustic impedance are $Pa \cdot s/m$ or $kg/(m^2 \cdot s)$. By definition

$$Z_a(s) = \frac{\hat{p}(r, s)}{\hat{v}(r, s) \cdot n_{in}} = \rho_0 \hat{v}(s), \tag{2.16}$$

where $\hat{p}, \hat{v}$ are the Laplace transforms of p, v and n_{in} is the inward normal. First perform a Laplace transformation of the equation (2.5), to derive $s\hat{v} + \frac{1}{\rho_0}\nabla\hat{p} = 0$. Then take the inner product with the outward normal direction n, to obtain

$$s\hat{v} \cdot n + \frac{1}{\rho_0}\frac{\partial\hat{p}}{\partial n} = 0$$
$$\Rightarrow \frac{\partial\hat{p}}{\partial n} = -\rho_0 s\hat{v} \cdot n. \tag{2.17}$$

The acoustic impedance can be written as

$$Z(s) = \frac{\hat{p}}{\hat{v} \cdot n_{in}} = \frac{\hat{p}}{\frac{\partial\hat{p}}{\partial n}}\rho_0 s. \tag{2.18}$$

Now inserting (2.17) to (2.14), assuming the impedance boundary condition is on pressure (i.e. $u = p$ here), we can relate the mathematical impedance condition to the acoustic impedance boundary condition in the following way

$$\frac{\hat{p}}{\hat{v} \cdot n_{in}} = -\frac{\rho_0 s}{\alpha(s)} = Z(s). \tag{2.19}$$

The infinite specific-acoustic-impedance limit of $|Z|$ (i.e. $|Z| \to \infty$) corresponds to a hard(rigid) surface, and the limit when Z goes to zero (i.e. $|Z| \to 0$) corresponds to a soft (pressure-release) surface. Thus, for a hard sphere scatterer the mathematical impedance boundary condition become $\frac{\partial\hat{u}}{\partial n} = 0$ and for the soft sphere scatterer $\hat{u} = 0$. To find an expression for the scattered field, expand the incident and scattered

solution in terms of Legendre polynomials [44],

$$\hat{u}^{(inc)} = e^{-\gamma z} = e^{-\gamma r \cos\theta} = \sum_{n=0}^{\infty} (2n+1)(-1)^n i_n(\gamma r) P_n(\cos\theta), \tag{2.20}$$

$$\hat{u}^{(sc)} = \sum_{n=0}^{\infty} \sum_{m=-n}^{m=n} a_{nm}(\gamma) k_n(\gamma r) Y_n^m(\theta, \varphi) \tag{2.21}$$

$$= \sum_{n=0}^{\infty} a_n(\gamma) k_n(\gamma r) P_n(\cos\theta), \tag{2.22}$$

where $k_n(s)$ is the modified Bessel's function, $a_n(\gamma)$ is the coefficient to be determined. Matching the impedance boundary condition to a hard or soft sphere respectively at $r = 1$, we can use the expression for the total field to find $\hat{u} = \hat{u}^{(inc)} + \hat{u}^{(sc)}$

$$\frac{\partial \hat{u}_{hard}^{(sc)}}{\partial n} = -\frac{\partial \hat{u}_{hard}^{(inc)}}{\partial n}, \tag{2.23}$$

$$\hat{u}_{soft}^{(sc)} = -\hat{u}_{soft}^{(inc)}. \tag{2.24}$$

Here the outward normal n is just r. The scattered solutions are

$$\hat{u}_{hard}^{(sc)} = \sum_{n=0}^{\infty} \frac{(-1)^{n+1}(2n+1)[\gamma i_n'(\gamma)]}{\gamma k_n'(\gamma)} k_n(\gamma r) P_n(\cos\theta), \tag{2.25}$$

$$\hat{u}_{soft}^{(sc)} = \sum_{n=0}^{\infty} \frac{(-1)^{n+1}(2n+1)[\gamma i_n(\gamma)]}{k_n(\gamma)} k_n(\gamma r) P_n(\cos\theta). \tag{2.26}$$

Since the modified spherical Bessel function k_n and its derivative only have simple zeros [16], for both the hard and soft sphere scatterer, the scattered solution $\hat{u}^{(sc)}$ only has simple poles.

2.2.3   Lossless impedance loading of a sphere

To ensure that the spherical scatterer is lossless, we consider the total energy defined in the usual way

$$E_D = \int\int\int_D \frac{1}{2}(u_t^2 + |\nabla u|^2) dV,$$

where D is the region of the scatterer. Differentiate the energy in time

$$\frac{dE_D}{dt} = \int\int\int_D (u_t u_{tt} + \nabla u^T \cdot \nabla u_t)dV \tag{2.27}$$

$$= \int\int\int_D u_t(u_{tt} - \nabla^2 u)dV + \int_{\partial D} u_t u_\nu dS \tag{2.28}$$

$$= \int\int_{\partial D} u_t u_\nu dS. \tag{2.29}$$

Then, according to Parseval's Theorem [39] the net energy flux is

$$\int\int_{\partial D} s\hat{u}(s) \cdot \overline{\frac{\partial \hat{u}(y,s)}{\partial r}}dS, \tag{2.30}$$

and in our case D is a sphere. Applying the impedance boundary condition, the net energy flux term becomes

$$-\int\int_{\partial D} s\overline{\alpha(s)}|\hat{u}(y,s)|^2 dS. \tag{2.31}$$

For a lossless scatterer, the total energy change should be zero, which implies

$$\int_{-i\infty}^{i\infty} \int\int_{\partial D} s\hat{u}(s) \cdot \overline{\frac{\partial \hat{u}(y,s)}{\partial r}}dS = 0. \tag{2.32}$$

This suggests that (2.30) should be odd, $\alpha(s)$ should be even and $Z(s)$ be odd. Mathematically, this condition is sufficient to guarantee that energy is conserved. However, in order to extend our results later on to the electromagnetic scattering case, we need to put more constraints on $Z(s)$ or equivalently on $\alpha(s)$, namely, $Z(s)$ should satisfy Foster's theorem to ensure that the scatterer is lossless.

Foster's theorem

A *positive real* function $F(s)$ is an analytic function of the complex variable s, which has the following properties:

1. $F(s)$ is regular(analytic) for $\sigma > 0$,

2. $F(\sigma)$ is real,

3. $\sigma \geq 0$ implies $\Re[F(s)] > 0$.

A *reactance function* is a positive real function that maps the imaginary axis into the imaginary axis. Foster's theorem can be formulated in the following way, as shown in [12].

Theorem 2.2.1. *A positive real rational function of s is a reactance function if and only if all of its poles and zeros are simple, lie on the $j\omega$-axis, and alternate with each other. In other words*

$$\psi(s) = K\frac{s(s^2 + \omega_1^2)(s^2 + \omega_3^2)\cdots(s^2 + \omega_{2n-1}^2)}{(s^2 + \omega_0^2)(s^2 + \omega_2^2)\cdots(s^2 + \omega_{2n}^2)}, \tag{2.33}$$

is a reactance function, where $K > 0$, $0 \leq \omega_0 < \omega_1 < \cdots < \omega_{2n-1} < \omega_{2n} < \infty$.

Second order poles

Inserting (2.20), (2.22) to the boundary condition (2.14), where $\hat{u} = \hat{u}^{(inc)} + \hat{u}^{(sc)}$, the general expansion of the scattered solution for the impedance boundary condition can been written as

$$\hat{u}^{(sc)} = \sum_{n=0}^{\infty} \frac{(-1)^{n+1}(2n + 1)[\alpha(s)i_n(s) + si'_n(s)]}{\alpha(s)k_n(s) + sk'_n(s)}k_n(sr)P_n(\cos\theta). \tag{2.34}$$

For simplicity, assume $c = 1$ here, so $\gamma = s$. The far field pattern solution, following from $k_n(sr) \sim e^{-rs}/s$ [44]

$$\hat{u}_\infty^{(sc)} = \frac{e^{-sr}}{r}\sum_{n=0}^{\infty} \frac{(-1)^{n+1}(2n + 1)[\alpha(s)i_n(s) + si'_n(s)]}{s(\alpha(s)k_n(s) + sk'_n(s))}P_n(\cos\theta). \tag{2.35}$$

In order to construct a second order pole we need both the denominator and the derivative of the denominator in the formula for the scattered solution to be zero at

some specific $s = s_p$. That is, the denominator equals zero,

$$s\alpha(s)k_n(s) + s^2\frac{d}{ds}k_n(s) = 0, \tag{2.36}$$

and the derivative of the denominator equals zero,

$$\alpha(s)k_n(s) + s\left(\frac{d}{ds}\alpha(s)\right)k_n(s) + s\alpha(s)\frac{d}{ds}k_n(s) + 2s\frac{d}{ds}k_n(s) + s^2\frac{d^2}{ds^2}k_n(s) = 0. \tag{2.37}$$

Note that we don't need to solve the above system of ODEs, because equations only need to hold at one specific point $s = s_p$ for some pre-chosen impedance boundary condition $\alpha(s)$. Using Bessel's equation, we can replace the second derivative term with the lower order terms

$$s^2\frac{d^2}{ds^2}k_n(s) = -2s\frac{d}{ds}k_n(s) + \left(s^2 + n(n+1)\right)k_n(s). \tag{2.38}$$

Solving (2.36) and (2.37) by eliminating $\frac{d}{ds}k_n(s)$ we derive

$$k_n(s)\left(s\left(\frac{d}{ds}\alpha(s)\right) + a(s) - \alpha(s)^2 + s^2 + n^2 + n\right) = 0. \tag{2.39}$$

For $k_n(s) \neq 0$ we have

$$s\alpha' + s^2 + \alpha + n(n+1) - \alpha^2 = 0. \tag{2.40}$$

When $n = 1$, $k_1(s) = \frac{\pi}{2}\frac{s+1}{s^2}e^{-s}$, [44]. Solving (2.36) and (2.40) for $\alpha(s)$ and $\alpha'(s)$ respectively, we obtain

$$\alpha(s) = \frac{s^2 + 2s + 2}{s+1}, \quad \alpha'(s) = \frac{s(s+2)}{s^2 + 2s + 1}. \tag{2.41}$$

Note that differentiation of $\alpha(s)$ is not allowed here, since (2.41) holds just at one particular value $s = s_p$. The solution we derived above simply means that, if we want to construct a 2nd order pole at s_p, the impedance function $\alpha(s)$ should satisfy the following conditions:

1. $\alpha(s)$ should be an even function,

2. The value of $\alpha(s)$ at s_p equals $\frac{s_p{}^2 + 2\,s_p + 2}{s_p + 1}$,

3. The value of $\alpha'(s)$ at s_p equals $\frac{s_p(s_p+2)}{s_p{}^2 + 2\,s_p + 1}$.

For example, if we choose $s_p = -2$, then $\alpha(s_p) = -2$, $\alpha'(s_p) = 0$. Assume $\alpha(s)$ has the following form

$$\alpha(s) = \frac{c_1 + c_2 s^2}{1 + c_3 s^2}. \tag{2.42}$$

Then we may solve

$$\left(\frac{c_1 + c_2 s^2}{1 + c_3 s^2} \right)\Big|_{s=-2} = -2, \quad \frac{d}{ds}\left(\frac{c_1 + c_2 s^2}{1 + c_3 s^2} \right)\Big|_{s=-2} = 0, \tag{2.43}$$

to obtain $\alpha(s) \equiv -2$. Also when $n = 1$ the denominator of the scattered solution $\hat{u}^{(sc)}$ is

$$-2sk_1(s) + s^2 k_1'(s) = \frac{\pi}{2}\frac{(s+2)^2}{s}e^{-s}. \tag{2.44}$$

Thus, a 2nd order pole at $s = -2$ is obtained. Note that Foster's theorem is satisfied and the pole location must be in the left half plane (negative if it is real).

Third order poles

The procedure is similar to generating the 2nd order pole except that we need an extra second derivative of the denominator to equal zero at some $s = s_p$. That is, we organize

$$s\alpha(s)k_n(s) + s^2\frac{d}{ds}k_n(s) = 0, \tag{2.45}$$

$$\alpha(s)k_n(s) + s\left(\frac{d}{ds}\alpha(s)\right)k_n(s) + s\alpha(s)\frac{d}{ds}k_n(s)$$
$$+ 2s\frac{d}{ds}k_n(s) + s^2\frac{d^2}{ds^2}k_n(s) = 0, \tag{2.46}$$

$$2 \left(\frac{d}{ds} \alpha \left(s \right) \right) k_n \left(s \right) + 2 \alpha \left(s \right) \frac{d}{ds} k_n \left(s \right) + s \left(\frac{d^2}{ds^2} \alpha \left(s \right) \right) k_n \left(s \right)$$

$$+ 2s \left(\frac{d}{ds} \alpha \left(s \right) \right) \frac{d}{ds} k_n \left(s \right) + s \alpha \left(s \right) \frac{d^2}{ds^2} k_n \left(s \right) + 2 \frac{d}{ds} k_n \left(s \right)$$

$$+ 4s \frac{d^2}{ds^2} k_n \left(s \right) + s^2 \frac{d^3}{ds^3} k_n \left(s \right) = 0. \tag{2.47}$$

Using the following equations derived from Bessel's equation to replace the higher derivative terms with lower order terms,

$$s^2 \frac{d^3}{ds^3} k_n \left(s \right) = - 4 s \frac{d^2}{ds^2} k_n \left(s \right) - 2 \frac{d}{ds} k_n \left(s \right) + 2 s k_n \left(s \right)$$

$$+ \left(s^2 + n \left(n + 1 \right) \right) \frac{d}{ds} k_n \left(s \right), \tag{2.48}$$

$$s^2 \frac{d^2}{ds^2} k_n \left(s \right) = -2 s \frac{d}{ds} k_n \left(s \right) + \left(s^2 + n \left(n + 1 \right) \right) k_n \left(s \right), \tag{2.49}$$

$$\frac{d}{ds} k_n \left(s \right) = - \frac{\alpha \left(s \right) k_n \left(s \right)}{s}, \tag{2.50}$$

equations (2.46), (2.47) can be reduced to

$$k_n \left(s \right) \left(\alpha \left(s \right) + s \frac{d}{ds} \alpha \left(s \right) - \left(\alpha \left(s \right) \right)^2 + s^2 + n^2 + n \right) = 0, \tag{2.51}$$

$$k_n \left(s \right) \left(2 \frac{d}{ds} \alpha \left(s \right) + s \frac{d^2}{ds^2} \alpha \left(s \right) - 2 \left(\frac{d}{ds} \alpha \left(s \right) \right) \alpha \left(s \right) + 2 s \right) = 0. \tag{2.52}$$

Solving (2.45), (2.51), (2.52) for $\alpha, \alpha', \alpha''$ respectively, we find

$$\alpha(s) = - \frac{s \frac{d}{ds} k_n \left(s \right)}{k_n \left(s \right)} \Big|_{s=s_p}, \tag{2.53}$$

$$\alpha'(s) = \frac{1}{s \left(k_n \left(s \right) \right)^2} [s \left(\frac{d}{ds} k_n \left(s \right) \right) k_n \left(s \right) + s^2 \left(\frac{d}{ds} k_n \left(s \right) \right)^2 - s^2 \left(k_n \left(s \right) \right)^2$$

$$- n^2 \left(k_n \left(s \right) \right)^2 - n \left(k_n \left(s \right) \right)^2] |_{s=s_p}, \tag{2.54}$$

$$\alpha''(s) = \frac{2}{s^2 \left(k_n\left(s\right)\right)^3} \{-s \left(\frac{d}{ds} k_n\left(s\right)\right) \left(k_n\left(s\right)\right)^2 - 2k_n\left(s\right) s^2 \left(\frac{d}{ds} k_n\left(s\right)\right)^2$$

$$+ n^2 \left(k_n\left(s\right)\right)^3 + n \left(k_n\left(s\right)\right)^3 - s^3 \left(\frac{d}{ds} k_n\left(s\right)\right)^3 + s^3 \left(\frac{d}{ds} k_n\left(s\right)\right) \left(k_n\left(s\right)\right)^2$$

$$+ s \left(\frac{d}{ds} k_n\left(s\right)\right) n^2 \left(k_n\left(s\right)\right)^2 + s \left(\frac{d}{ds} k_n\left(s\right)\right) n \left(k_n\left(s\right)\right)^2 \}|_{s=s_p}. \qquad (2.55)$$

When $n = 2$, $k_2(s) = \frac{1}{2} \frac{\pi e^{-s}\left(s^2+3\,s+3\right)}{s^3}$, we have

$$\alpha(s_p) = \frac{s_p^3 + 4s_p^2 + 9s_p + 9}{s_p^2 + 3s_p + 3}, \qquad (2.56)$$

$$\alpha'(s_p) = \frac{s_p(s_p^3 + 6s_p^2 + 12s_p + 6)}{(s_p^2 + 3s_p + 3)^2}, \qquad (2.57)$$

$$\alpha''(s_p) = \frac{6(3 + s_p^3 + 6s_p^2 + 9s_p)}{(s_p^2 + 3s_p + 3)^3}. \qquad (2.58)$$

Thus, the impedance function should be constructed to meet the following requirements:

1. $\alpha(s)$ should be an even function,

2. The value of $\alpha(s)$ at s_p equals $\frac{s_p^3+4s_p^2+9s_p+9}{s_p^2+3s_p+3}$,

3. The value of $\alpha'(s)$ at s_p equals $\frac{s_p(s_p^3+6s_p^2+12s_p+6)}{(s_p^2+3s_p+3)^2}$,

4. The value of $\alpha''(s)$ at s_p equals $\frac{6(3+s_p^3+6s_p^2+9s_p)}{(s_p^2+3s_p+3)^3}$.

If we choose $s_p = -4$, then $\alpha(-4) = -\frac{27}{7}$, $\alpha'(-4) = \frac{40}{49}$, and $\alpha''(-4) = -\frac{6}{343}$. Assume $\alpha(s)$ has following form

$$\alpha(s) = \frac{c_1 + c_2 s^2}{1 + c_3 s^2}. \qquad (2.59)$$

Then, after some calculation, we obtain

$$\alpha(s) = \frac{-\frac{47}{27} - \frac{11}{54}s^2}{1 + \frac{1}{54}s^2} = \frac{-11(s^2 + \frac{94}{11})}{s^2 + 54}. \qquad (2.60)$$

The denominator of the scattered solution $\hat{u}^{(sc)}$ when $n = 2$ is

$$\frac{\pi}{2} \frac{e^{-s} \left(s^2 + 3\,s + 12\right) (s + 4)^3}{s^2 \left(54 + s^2\right)}. \tag{2.61}$$

Therefore, a 3rd order pole has been constructed at $s = -4$.

Arbitrary order poles

We can generate high order poles by repeating the above procedure. This can be viewed a the Taylor series around the pole location $s_p < 0$. To begin with, we introduce some properties of the modified Bessel function [35];

$$-s\frac{k_l'(s)}{k_l(s)} = s + 1 + \hat{S}_l(s), \tag{2.62}$$

where, for $l \neq 0$,

$$\hat{S}_l(z) = \frac{P_l(z)}{Q_l(z)}, \tag{2.63}$$

$$P_l(z) = \sum_{k=0}^{l-1} \frac{(2l-k)!}{k!(l-k-1)!}(2z)^k, \tag{2.64}$$

$$Q_l(z) = \sum_{k=0}^{l} \frac{(2l-k)!}{k!(l-k)!}(2z)^k. \tag{2.65}$$

We also have the following continued fraction representation for $\hat{S}_l$

$$\hat{S}_l(z) = \frac{l(l+1)}{2} \frac{1}{z+1+\frac{l(l+1)-1\cdot2}{4(z+2+\frac{l(l+1)-2\cdot3}{4(z+3+\dots)})}}. \tag{2.66}$$

Thus, it is possible for us to rewrite

$$k_l(s) = \frac{k_l^{(1)}(s)}{k_l^{(2)}(s)} \frac{e^{-s}}{s}, \tag{2.67}$$

where, $k_l^{(1)}(s)$, $k_l^{(2)}(s)$ are polynomials in s. Suppose the impedance function $\alpha(s)$ has the following form

$$\alpha(s) = -c_0\frac{(s^2 + a_0)(s^2 + a_2)\cdots(s^2 + a_k)}{(s^2 + a_1)(s^2 + a_3)\cdots(s^2 + a_{2n-1})} = -c_0\frac{\alpha_1(s)}{\alpha_2(s)}, \tag{2.68}$$

where,

$$c_0 > 0, \ 0 \le a_0 < a_1 < a_2 < \cdots < a_{2n-2} < a_{2n-1} < a_{2n} < \infty, \tag{2.69}$$

and $\alpha_1(s)$, $\alpha_2(s)$ are polynomials. For example, for a 2nd order pole we assume 2 free parameters, thus $\alpha(s)$ can been chosen

$$\alpha(s) = -c_0(s^2 + a_0). \tag{2.70}$$

For a 3rd order pole we need 3 free parameters, thus $\alpha(s)$ can been chosen

$$\alpha(s) = -c_0 \frac{(s^2 + a_0)}{(s^2 + a_1)}. \tag{2.71}$$

For a 4th order pole we need 4 free parameters, thus $\alpha(s)$ can been chosen

$$\alpha(s) = -c_0 \frac{(s^2 + a_0)(s^2 + a_2)}{(s^2 + a_1)}, \tag{2.72}$$

and so on. The denominator of the scattered solution $\hat{u}^{(sc)}$ can now be written as

$$-\frac{k_n^{(1)}(s)}{k_n^{(2)}(s)} \frac{e^{-s}}{s} \left(c_0 \frac{\alpha_1(s)}{\alpha_2(s)} + s + 1 + \hat{S}_n(s) \right). \tag{2.73}$$

If we want to construct a j−th order pole at s_p, we need to choose the correct power of $\alpha_1(s)$, $\alpha_2(s)$, set $n = j-1$, and rewrite the numerator of (2.73) using Taylor expansion around s_p,

$$k_n^{(1)}(s) \left(c_0 \alpha_1(s) \hat{Q}_n(s) + \alpha_2(s) \hat{P}_n(s) \right) = \sum_{i=0}^{j-1} \beta_i (s - s_p)^i + O((s - s_p)^j). \tag{2.74}$$

Here $s + 1 + \hat{S}_n(s) = s + 1 + \frac{P_n(s)}{Q_n(s)} = \frac{\hat{P}_n(s)}{\hat{Q}_n(s)}$. We choose $c_0, a_0, a_1, \ldots a_{j-2}$ so that $\beta_i = 0$ for $i = 0, \ldots j - 1$. Note that we want c_0 and a_j to satisfy our assumptions (2.69). Following the process we discussed above, the denominator of the scattered solution $\hat{u}^{(sc)}$ will generate a j−th order zero. Thus, a lossless impedance function $\alpha(s)$ which will produce a j−th order pole at s_p is constructed.

Let's go over the process to generate a 3rd pole again using the Taylor method at $s = s_p$, so $j = 3$, $n = 2$,

$$\alpha(s) = -c\frac{(s^2 + a)}{(s^2 + b)}, \tag{2.75}$$

$$k_2(s) = \frac{e^{-s}}{s}\frac{(s^2 + 3s + 3)}{s^2}, \tag{2.76}$$

$$s + 1 + \hat{S}_n(s) = \frac{\hat{P}_n(s)}{\hat{Q}_n(s)} = \frac{s^3 + 4s^2 + 9s + 9}{s^2 + 3s + 3}. \tag{2.77}$$

According to (2.74), the zeros are determined by

$$(s^2 + 3s + 3)\Big(c(s^2 + a)(s^2 + 3s + 3) + (s^2 + b)(s^3 + 4s^2 + 9s + 9)\Big). \tag{2.78}$$

Expanding (2.78) around $s = s_p$, we obtain the coefficients

$$\beta_0 = s_p{}^7 + 18\,cs_p{}^3 + 7\,bs_p{}^4 + bs_p{}^5 + cs_p{}^6 + 6\,cas_p{}^3 + 9\,cs_p{}^2 + 24\,s_p{}^5 + 7\,s_p{}^6$$
$$+ 15\,cas_p{}^2 + cas_p{}^4 + 48\,s_p{}^4 + 6\,cs_p{}^5 + 54\,s_p{}^3 + 24\,bs_p{}^3 + 15\,cs_p{}^4$$
$$+ 54\,bs_p + 18\,cas_p + 27\,s_p{}^2 + 27\,b + 48\,bs_p{}^2 + 9\,ca, \tag{2.79}$$

$$\beta_1 = 54\,cs_p{}^2 + 7\,s_p{}^6 + 30\,cas_p + 4\,cas_p{}^3 + 6\,cs_p{}^5 + 162\,s_p{}^2 + 18\,cas_p{}^2$$
$$+ 120\,s_p{}^4 + 30\,cs_p{}^4 + 18\,ca + 72\,bs_p{}^2 + 28\,bs_p{}^3 + 18\,cs_p + 42\,s_p{}^5 + 54\,s_p$$
$$+ 96\,bs_p + 192\,s_p{}^3 + 5\,bs_p{}^4 + 54\,b + 60\,cs_p{}^3, \tag{2.80}$$

$$\beta_2 = 72\,bs_p + 6\,cas_p{}^2 + 162\,s_p + 54\,cs_p + 15\,cs_p{}^4 + 18\,cas_p + 21\,s_p{}^5$$
$$+ 105\,s_p{}^4 + 27 + 15\,ca + 60\,cs_p{}^3 + 48\,b + 240\,s_p{}^3 + 9\,c + 10\,bs_p{}^3$$
$$+ 288\,s_p{}^2 + 90\,cs_p{}^2 + 42\,bs_p{}^2. \tag{2.81}$$

Solving $\beta_0 = \beta_1 = \beta_2 = 0$ for a, b, c in terms of s_p we find

$$a = \frac{s_p{}^6 + 12\,s_p{}^5 + 81\,s_p{}^4 + 315\,s_p{}^3 + 648\,s_p{}^2 + 648\,s_p + 216}{3\,s_p{}^4 + 28\,s_p{}^3 + 99\,s_p{}^2 + 153\,s_p + 96}, \tag{2.82}$$

$$b = 3 \, \frac{s_p{}^5 + 9 \, s_p{}^4 + 35 \, s_p{}^3 + 72 \, s_p{}^2 + 72 \, s_p + 24}{s_p{}^3 + 9 \, s_p{}^2 + 27 \, s_p + 24}, \tag{2.83}$$

$$c = -\frac{3 \, s_p{}^4 + 28 \, s_p{}^3 + 99 \, s_p{}^2 + 153 \, s_p + 96}{s_p{}^3 + 9 \, s_p{}^2 + 27 \, s_p + 24}. \tag{2.84}$$

If we choose $s_p = -4$, then $a = \frac{94}{11}$, $b = 54$, $c = 11$. Clearly $c > 0$, and $0 \le a < b < \infty$

$$\alpha(s) = \frac{-11(s^2 + \frac{94}{11})}{s^2 + 54}. \tag{2.85}$$

This is identical to what we derived earlier. Note that sometimes the assumptions of Foster's theorem may hold only in a limited interval. In fact, if we choose $s_p = -3$, then a, b, c will not satisfy our assumptions. In Figure 2.2 plot a, b, c in terms of s_p, and see that in order to satisfy all the assumptions, s_p can only be chosen approximately $s_p < -3.2$. At $s_p = -3$, a will be negative. For a more general case, we may use a similar graph to determine the range of s_p. For example, in this case, the range of s_p will be the interval where the graph of b is above a and all of the graphs a, b, c are above the x-axis.

2.3 Scattering from a lossless electromagnetic sphere

We consider the electromagnetic scattering problem resulting from a plane wave scattered off a perfectly conducting sphere, a lossless surface impedance loaded sphere, and a sheet impedance loaded sphere respectively, as illustrated in Figure 2.3. An E wave has been chosen as an incident electromagnetic plane wave propagating in the $\vec{1}_1$ direction. The scattered solution as well as the surface current density can be written explicitly using vector spherical harmonics (vector wave functions). In the case of a perfectly conducting surface, there exist only first order scattering poles. In the case of a lossless surface impedance loading sphere, we can derive arbitrary order scattering poles like in the acoustic scattering case. In the case of a lossless
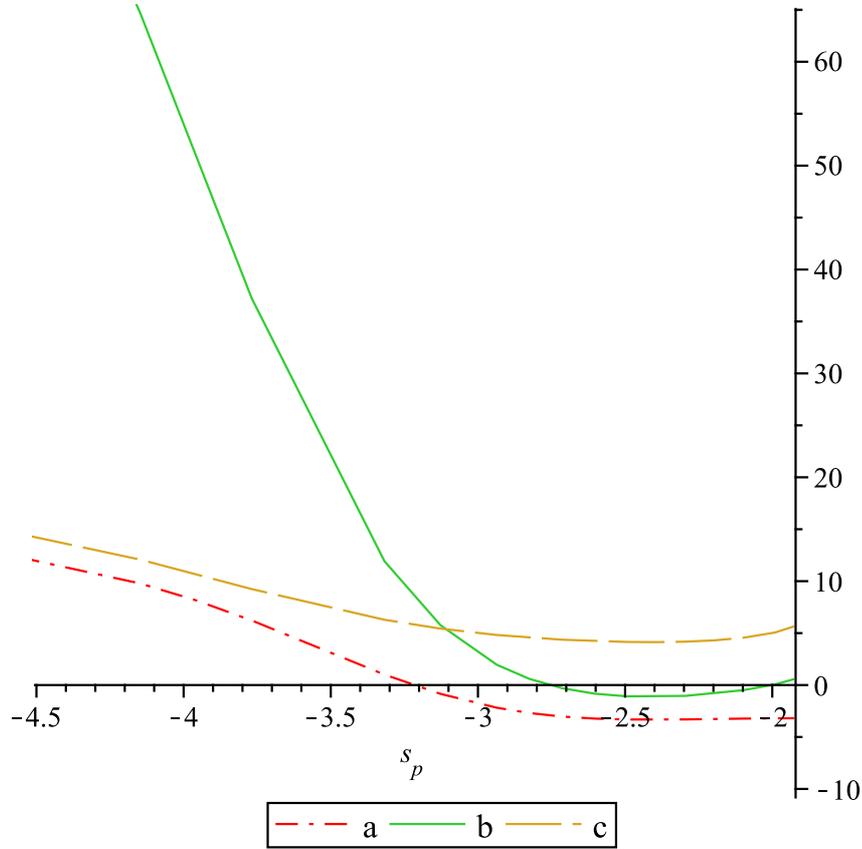
Figure 2.2: plots of a,b,c, region of where the assumption can hold.

sheet impedance loading sphere, we only find second order scattering poles for a non-physical mode, $n = 0$. This leads us to the conjecture that there only exist first order poles for this case. As above Foster's theorem is imposed on the impedance function $\hat{Z}_s(s)$ to guarantee that it is a realizable physical boundary condition.

2.3.1 Formulation of the electromagnetic scattering problem

In this section we formulate the electromagnetic scattering problem in terms of spherical harmonics and derive an expression of the scattered field using the following

setting of the coordinate system.

Define a set of orthogonal (right-handed) unit vectors [16]
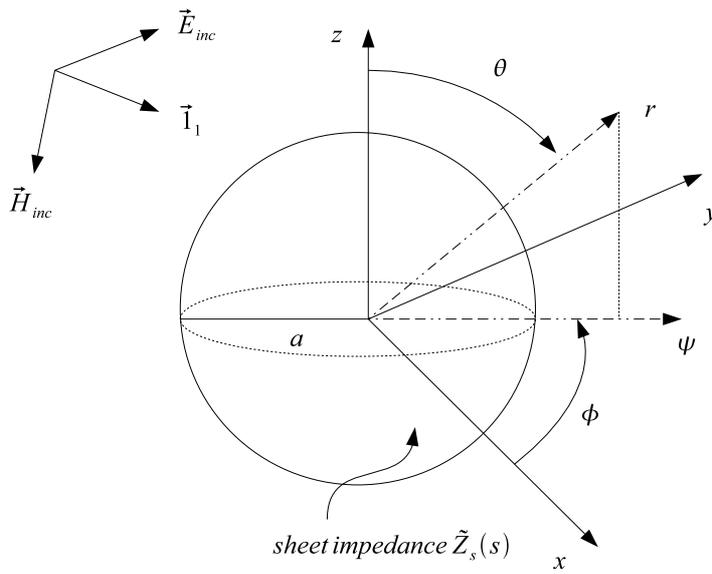
Figure.1

Figure 2.3: Spherical coordinate system with EM incident wave.
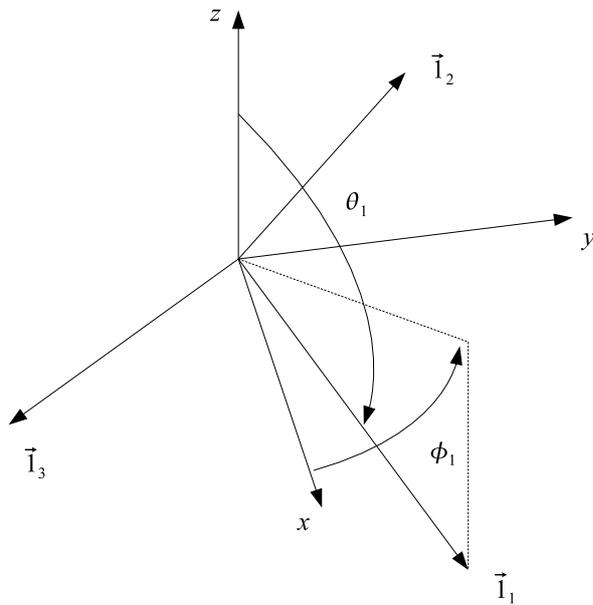
Figure 2.4: Spherical coordinate system with polarization.

$$\vec{1}_1 = \sin(\theta_1)\cos(\phi_1)\vec{1}_x + \sin(\theta_1)\sin(\phi_1)\vec{1}_y + \cos(\theta_1)\vec{1}_z, \tag{2.86}$$

$$\vec{1}_2 = -\cos(\theta_1)\cos(\phi_1)\vec{1}_x - \cos(\theta_1)\sin(\phi_1)\vec{1}_y + \sin(\theta_1)\vec{1}_z, \tag{2.87}$$

$$\vec{1}_3 = \sin(\phi_1)\vec{1}_x - \cos(\phi_1)\vec{1}_y. \tag{2.88}$$

As shown in Figure 2.4, $\vec{1}_1$ is the direction of propagation and $\vec{1}_2$ and $\vec{1}_3$ are mutually orthogonal unit vectors, each orthogonal to $\vec{1}_1$ to indicate the polarization of the electromagnetic fields in the incident plane wave. For convenience $\vec{1}_2$ is chosen in a plane parallel to $\vec{1}_1$ and the z axis (E or TM polarization if the electric field is parallel to $\vec{1}_2$) while $\vec{1}_3$ is parallel to the x, y plane (H or TE polarization if the electric field is parallel to $\vec{1}_3$). In free space, electromagnetic plane waves have both electric and magnetic fields orthogonal to $\vec{1}_1$. Thus only $\vec{1}_2$ and $\vec{1}_3$ are considered. This removes the $\vec{L}$ functions (details are shown later) in the expansion (plane waves have zero-divergence fields).

We can use the relations between Cartesian and spherical coordinates

$$x = r\sin(\theta)\cos(\phi), \tag{2.89}$$

$$y = r\sin(\theta)\sin(\phi), \tag{2.90}$$

$$z = r\cos(\theta), \tag{2.91}$$

$$\vec{1}_x = \sin(\theta)\cos(\phi)\vec{1}_r + \cos(\theta)\cos(\phi)\vec{1}_\theta - \sin(\phi)\vec{1}_\phi, \tag{2.92}$$

$$\vec{1}_y = \sin(\theta)\sin(\phi)\vec{1}_r + \cos(\theta)\sin(\phi)\vec{1}_\theta + \cos(\phi)\vec{1}_\phi, \tag{2.93}$$

$$\vec{1}_z = \cos(\theta)\vec{1}_r - \sin(\theta)\vec{1}_\theta, \tag{2.94}$$

to express the incident-wave unit vectors in terms of (θ_1, ϕ_1) and (θ, ϕ) as

$$\begin{aligned}
\vec{1}_1 = {} & [\cos(\theta_1)\cos(\theta) + \sin(\theta_1)\sin(\theta)\cos(\phi - \phi_1)]\vec{1}_r \\
& + [-\cos(\theta_1)\sin(\theta) + \sin(\theta_1)\cos(\theta)\cos(\phi - \phi_1)]\vec{1}_\theta \\
& + [-\sin(\theta_1)\sin(\phi - \phi_1)]\vec{1}_\phi,
\end{aligned} \tag{2.95}$$

$$\begin{aligned}
\vec{1}_2 = {} & [\sin(\theta_1)\cos(\theta) - \cos(\theta_1)\sin(\theta)\cos(\phi - \phi_1)]\vec{1}_r \\
& - [\sin(\theta_1)\sin(\theta) + \cos(\theta_1)\cos(\theta)\cos(\phi - \phi_1)]\vec{1}_\theta \\
& + [\cos(\theta_1)\sin(\phi - \phi_1)]\vec{1}_\phi,
\end{aligned} \tag{2.96}$$

$$\vec{1}_3 = -\sin(\theta)\sin(\phi - \phi_1)\vec{1}_r$$

$$- \cos(\theta)\cos(\phi - \phi_1)\vec{1}_\theta$$

$$- \cos(\phi - \phi_1)]\vec{1}_\phi. \tag{2.97}$$

Having the direction of incidence and two polarizations expressed in spherical coordinates we can go on to express the response to some delta plane wave functions. For an incident delta function plane wave we need spherical harmonics and vector wave functions to express the expansion in spherical coordinates. In spherical coordinates we have the common differential operators as

$$\nabla F = \vec{1}_r \frac{\partial}{\partial r}F + \vec{1}_\theta \frac{1}{r}\frac{\partial}{\partial \theta}F + \vec{1}_\phi \frac{1}{r\sin(\theta)}\frac{\partial}{\partial \phi}F, \tag{2.98}$$

$$\nabla \cdot \vec{F} = \frac{1}{r^2}\frac{\partial}{\partial r}(r^2 F_r) + \frac{1}{r\sin(\theta)}\frac{\partial}{\partial \theta}(\sin(\theta)F_\theta) + \frac{1}{r\sin(\theta)}\frac{\partial}{\partial \phi}F_\phi, \tag{2.99}$$

$$\nabla \times \vec{F} = \vec{1}_r[\frac{1}{r\sin(\theta)}\frac{\partial}{\partial \theta}(\sin(\theta)F_\phi) - \frac{1}{r\sin(\theta)}\frac{\partial}{\partial \phi}F_\theta] + \vec{1}_\theta[\frac{1}{r\sin(\theta)}\frac{\partial}{\partial \theta}(\sin(\theta)F_r)$$

$$-\frac{1}{r}\frac{\partial}{\partial r}(rF_\phi)] + \vec{1}_\phi[\frac{1}{r}\frac{\partial}{\partial r}(rF_\theta) - \frac{1}{r}\frac{\partial}{\partial \theta}F_r], \tag{2.100}$$

$$\nabla_s F = \vec{1}_\theta \frac{\partial}{\partial \theta}F + \vec{1}_\phi \frac{1}{\sin(\theta)}\frac{\partial}{\partial \phi}F, \tag{2.101}$$

$$\nabla_s \cdot \vec{F} = \frac{1}{\sin(\theta)}\frac{\partial}{\partial \theta}(\sin(\theta)F_\theta) + \frac{1}{\sin(\theta)}\frac{\partial}{\partial \phi}F_\phi, \tag{2.102}$$

$$\nabla_s \times \vec{F} = \vec{1}_r[\frac{1}{\sin(\theta)}\frac{\partial}{\partial \theta}(\sin(\theta)F_\phi) - \frac{1}{\sin(\theta)}\frac{\partial}{\partial \phi}F_\theta] \tag{2.103}$$

$$+\vec{1}_\theta[\frac{1}{\sin(\theta)}\frac{\partial}{\partial \phi}F_r] - \vec{1}_\phi \frac{\partial}{\partial \theta}F_r. \tag{2.104}$$

Spherical Harmonics

The scalar spherical harmonics can be written as [16, 29]

$$Y_{n,m,\substack{e \\ o}}(\theta, \phi) = P_n^{(m)}(cos(\theta)) \left\{ \begin{array}{c} cos(m\phi) \\ sin(m\phi) \end{array} \right\}, \tag{2.105}$$

where $P_n^{(m)}(x)$ is the Legendre functions defined as

$$P_n^m(x) = (-1)^m(1-x^2)^{m/2}\frac{d^m}{dx^m}P_n(x), \quad P_n(x) = P_n^0(x) = \frac{1}{2^n n!}\frac{d^n}{dx^n}(x^2-1)^n. \quad (2.106)$$

The vector spherical harmonics are defined as follows

$$\vec{P}_{n,m,p}(\theta,\phi) = \quad Y_{n,m,p}(\theta,\phi)\vec{1}_r, \tag{2.107}$$

$$\vec{Q}_{n,m,p}(\theta,\phi) = \quad \nabla_s Y_{n,m,p}(\theta,\phi) \quad = \vec{1}_r \times \vec{R}_{n,m,p}, \tag{2.108}$$

$$\vec{R}_{n,m,p}(\theta,\phi) = \nabla_s \times \vec{P}_{n,m,p}(\theta,\phi) \quad = -\vec{1}_r \times \vec{Q}_{n,m,p}. \tag{2.109}$$

They are also mutually orthogonal in an integral sense on the unit sphere. The spherical scalar wave functions are defined as

$$\Xi_{n,m,p}^{(l)}(\gamma\vec{r}) = f_n^{(l)}(\gamma r)P_n^{(m)}(\theta,\phi). \tag{2.110}$$

Here $f_n^{(1)}(\gamma r) = i_n(\gamma r)$, $f_n^{(2)}(\gamma r) = k_n(\gamma r)$ are modified Bessel functions, which satisfy the Wronskian relation

$$W\{si_n(s), sk_n(s)\} = si_n(s)[sk_n(s)]' - [si_n(s)]'sk_n(s) = -1. \tag{2.111}$$

Let $\gamma = [s\mu(\sigma+s\epsilon)]^{1/2}$ with μ, σ, ϵ being permeability, conductivity, and permittivity, respectively (in this work, we only deal with the case $\sigma = 0$). s is the variable of the two-sided Laplace transformation(denote by $\sim$) [16, 14]. The scalar wave function $\Xi_{n,m,p}^{(l)}(\gamma\vec{r})$ satisfies the scalar wave equation which we can write in operator form as [16]

$$[\nabla^2 - \gamma^2]\Xi_{n,m,p}^{(l)}(\gamma\vec{r}) = 0. \tag{2.112}$$

From the solution of the scalar wave equation one constructs solutions of the vector wave equation of three kinds.

$$\hat{L}_{n,m,p}^{(l)}(\gamma\vec{r}) = \quad \frac{1}{\gamma}\nabla\Xi_{n,m,p}^{(l)}(\gamma\vec{r}), \tag{2.113}$$

$$\hat{M}_{n,m,p}^{(l)}(\gamma\vec{r}) = \quad \nabla \times [\vec{r}\Xi_{n,m,p}^{(l)}(\gamma\vec{r})], \tag{2.114}$$

$$\hat{N}_{n,m,p}^{(l)}(\gamma\vec{r}) = \quad \frac{1}{\gamma}\nabla \times \hat{M}_{n,m,p}^{(l)}(\gamma\vec{r}). \tag{2.115}$$

Note that all three kinds of vector wave functions satisfy the vector wave equation, which on the Laplace transform side can be summarized as

$$[\nabla^2 - \gamma^2] \left\{ \begin{array}{c} \hat{L}^{(l)}_{n,m,p} \\ \hat{M}^{(l)}_{n,m,p} \\ \hat{N}^{(l)}_{n,m,p} \end{array} \right\} = 0. \tag{2.116}$$

We can also write a curl curl wave equation for the second and third kind vector wave functions

$$[\nabla \times \nabla \times \gamma^2] \left\{ \begin{array}{c} \hat{M}^{(l)}_{n,m,p} \\ \hat{N}^{(l)}_{n,m,p} \end{array} \right\} = 0. \tag{2.117}$$

The three vector wave functions are related through

$$\hat{M}^{(l)}_{n,m,p}(\gamma\vec{r}) = -\gamma\vec{r} \times \hat{L}^{(l)}_{n,m,p}(\gamma\vec{r}), \tag{2.118}$$

$$\hat{M}^{(l)}_{n,m,p}(\gamma\vec{r}) = -\tfrac{1}{\gamma}\nabla \times \hat{N}^{(l)}_{n,m,p}(\gamma\vec{r}), \tag{2.119}$$

$$\hat{N}^{(l)}_{n,m,p}(\gamma\vec{r}) = \tfrac{1}{\gamma}\nabla \times \hat{M}^{(l)}_{n,m,p}(\gamma\vec{r}). \tag{2.120}$$

It is also useful to write them as

$$\hat{L}^{(l)}_{n,m,p}(\gamma\vec{r}) = [f^{(l)}_n(\gamma r)]'\vec{P}_{n,m,p}(\theta,\phi) + [f^{(l)}_n(\gamma r)]\vec{Q}_{n,m,p}(\theta,\phi)/\gamma r, \tag{2.121}$$

$$\hat{M}^{(l)}_{n,m,p}(\gamma\vec{r}) = [f^{(l)}_n(\gamma r)]\vec{R}_{n,m,p}(\theta,\phi), \tag{2.122}$$

$$\hat{N}^{(l)}_{n,m,p}(\gamma\vec{r}) = \{n(n+1)[f^{(l)}_n(\gamma r)]\vec{P}_{n,m,p}(\theta,\phi) + [\gamma r f^{(l)}_n(\gamma r)]'\vec{Q}_{n,m,p}(\theta,\phi)\}/\gamma r. \tag{2.123}$$

Plane wave in spherical coordinates

The delta function plane waves (transformed) can be written as [2]

$$\vec{\mathrm{I}}_2 e^{-\gamma\vec{\mathrm{I}}_1\cdot\vec{r}} = \sum_{n=1}^{\infty}\sum_{m=0}^{n}\sum_{p=e,o}[a^{(1)}_{n,m,p}\hat{M}^{(1)}_{n,m,p}(\gamma\vec{r}) + b^{(1)}_{n,m,p}\hat{N}^{(1)}_{n,m,p}(\gamma\vec{r})], \tag{2.124}$$

$$\vec{\mathrm{I}}_3 e^{-\gamma\vec{\mathrm{I}}_1\cdot\vec{r}} = \sum_{n=1}^{\infty}\sum_{m=0}^{n}\sum_{p=e,o}[b^{(1)}_{n,m,p}\hat{M}^{(1)}_{n,m,p}(\gamma\vec{r}) - a^{(1)}_{n,m,p}\hat{N}^{(1)}_{n,m,p}(\gamma\vec{r})], \tag{2.125}$$

where

$$a_{n,m,p}^{(1)} = [2 - 1_{0,m}](-1)^{n+1}\frac{2n+1}{n(n+1)}\frac{(n-m)!}{(n+m)!}$$

$$\cdot m\frac{P_n^{(m)}(\cos(\theta_1))}{\sin(\theta_1)}\left\{\begin{array}{c} -\sin(m\phi_1) \\ \cos(m\phi_1) \end{array}\right\}, \quad (2.126)$$

$$b_{n,m,p}^{(1)} = [2 - 1_{0,m}](-1)^n\frac{2n+1}{n(n+1)}\frac{(n-m)!}{(n+m)!}\frac{dP_n^{(m)}(\cos(\theta_1))}{d\theta_1}\left\{\begin{array}{c} \cos(m\phi_1) \\ \sin(m\phi_1) \end{array}\right\}.$$

$$(2.127)$$

Note that we have

$$\frac{1}{\gamma}\nabla \times [\vec{1}_2 e^{-\gamma\vec{1}_1\cdot\vec{r}}] = \vec{1}_3 e^{-\gamma\vec{1}_1\cdot\vec{r}}, \quad (2.128)$$

$$\frac{1}{\gamma}\nabla \times [\vec{1}_3 e^{-\gamma\vec{1}_1\cdot\vec{r}}] = -\vec{1}_2 e^{-\gamma\vec{1}_1\cdot\vec{r}}, \quad (2.129)$$

which is associated with the curl relations between the $\hat{M}_{n,m,p}^{(l)}$ and $\hat{N}_{n,m,p}^{(l)}$ functions. Furthermore any divergence free electric field expansion ($\vec{E}$) can be converted to a magnetic field expansion ($\vec{H}$) by dividing by the wave impedance Z of the medium and changing $\hat{M}_{n,m,p}^{(l)}$ to $-\hat{N}_{n,m,p}^{(l)}$ and $\hat{N}_{n,m,p}^{(l)}$ to $\hat{M}_{n,m,p}^{(l)}$. To go from $\vec{H}$ to $\vec{E}$ multiply by Z and change $\hat{M}_{n,m,p}^{(l)}$ to $\hat{N}_{n,m,p}^{(l)}$ and $\hat{N}_{n,m,p}^{(l)}$ to $-\hat{M}_{n,m,p}^{(l)}$.

Solution for the scattered field

Now define the incident plane wave as an E wave (TM wave) [17]

$$\tilde{\vec{E}}_{inc}(\vec{r}, s) = E_0\vec{1}_2 e^{-\gamma\vec{1}_1\cdot\vec{r}}, \quad (2.130)$$

$$\tilde{\vec{H}}_{inc}(\vec{r}, s) = \frac{E_0}{Z_0}\vec{1}_3 e^{-\gamma\vec{1}_1\cdot\vec{r}}. \quad (2.131)$$

and expand the fields for $r < a$ as

$$\tilde{\vec{E}}_{in}(\vec{r}, s) = E_0 \sum_{n=1}^{\infty} \sum_{m=0}^{n} \sum_{p=e,o} [a_{n,m,p}^{(2)} \hat{M}_{n,m,p}^{(1)}(\gamma\vec{r}) + b_{n,m,p}^{(2)} \hat{N}_{n,m,p}^{(1)}(\gamma\vec{r})], \tag{2.132}$$

$$\tilde{\vec{H}}_{in}(\vec{r}, s) = \frac{E_0}{Z_0} \sum_{n=1}^{\infty} \sum_{m=0}^{n} \sum_{p=e,o} [b_{n,m,p}^{(2)} \hat{M}_{n,m,p}^{(1)}(\gamma\vec{r}) - a_{n,m,p}^{(2)} \hat{N}_{n,m,p}^{(1)}(\gamma\vec{r})]. \tag{2.133}$$

The solution of the scattered fields for $r > a$ can be written as

$$\tilde{\vec{E}}_{sc}(\vec{r}, s) = E_0 \sum_{n=1}^{\infty} \sum_{m=0}^{n} \sum_{p=e,o} [a_{n,m,p}^{(3)} \hat{M}_{n,m,p}^{(2)}(\gamma\vec{r}) + b_{n,m,p}^{(3)} \hat{N}_{n,m,p}^{(2)}(\gamma\vec{r})], \tag{2.134}$$

$$\tilde{\vec{H}}_{sc}(\vec{r}, s) = \frac{E_0}{Z_0} \sum_{n=1}^{\infty} \sum_{m=0}^{n} \sum_{p=e,o} [b_{n,m,p}^{(3)} \hat{M}_{n,m,p}^{(2)}(\gamma\vec{r}) - a_{n,m,p}^{(3)} \hat{N}_{n,m,p}^{(2)}(\gamma\vec{r})]. \tag{2.135}$$

2.3.2 Perfectly conducting sphere

For the scatterer being a perfectly conducting sphere, we show that there only exist first order scattering poles. Here the tangential electric field is constrained to be zero on $r = a$, that is $\vec{1}_r \times [\tilde{\vec{E}}_{inc}(\vec{r}, s) + \tilde{\vec{E}}_{sc}(\vec{r}, s)] = 0$, or

$$\vec{1}_r \times [a_{n,m,p}^{(1)} \hat{M}_{n,m,p}^{(1)}(\gamma a \vec{1}_r) + a_{n,m,p}^{(3)} \hat{M}_{n,m,p}^{(2)}(\gamma a \vec{1}_r)] = \vec{0}, \tag{2.136}$$

$$\vec{1}_r \times [b_{n,m,p}^{(1)} \hat{N}_{n,m,p}^{(1)}(\gamma a \vec{1}_r) + b_{n,m,p}^{(3)} \hat{N}_{n,m,p}^{(2)}(\gamma a \vec{1}_r)] = \vec{0}. \tag{2.137}$$

This yields equations for the coefficients

$$a_{n,m,p}^{(3)} = -\frac{i_n(\gamma a)}{k_n(\gamma a)} a_{n,m,p}^{(1)}, \quad b_{n,m,p}^{(3)} = -\frac{[\gamma a i_n(\gamma a)]'}{[\gamma a k_n(\gamma a)]'} b_{n,m,p}^{(1)}. \tag{2.138}$$

For the poles to be simple we need to show that all the zeros of $k_n(s)$ and $sk_n(s)$ are simple. Since $k_n(s)$ satisfies the spherical Bessel equation we have

$$s^2 \frac{d^2}{ds^2} k_n(s) + 2s \frac{d}{ds} k_n(s) - [s^2 + n(n+1)]k_n(s) = 0. \tag{2.139}$$

Now, suppose there exists higher order zeros, say a 2nd order zero at $s_\alpha \neq 0$. Since both $k_n(s)$ and $k_n'(s)$ have to be zero at s_α, so does $k_n''(s)$, thus, the zero must be

at least a third order zero which is a contradiction. Repeating the process, all the derivatives at s_α will be zero. Therefore the function must be identically zero and there exists only simple poles for $a_{n,m,p}^{(3)}$. The argument for $b_{n,m,p}^{(3)}$ is similar as $[sk_n(s)]$ satisfies the Riccati-Bessel equation,

$$\frac{s^2}{s^2 + n(n+1)}\frac{d^2}{ds^2}[sf_n^{(l)}(s)] - sf_n^{(l)}(s) = 0. \tag{2.140}$$

More details of the perfectly conducting sphere including surface current and charge densities can be found in [16].

2.3.3 Surface-impedance-loaded sphere

With the scatterer being a lossless surface impedance loaded sphere, we show that arbitrary order scattering poles can be generated. Assume we choose the following surface impedance boundary condition

$$\tilde{\vec{E}}_{tan} = \overleftrightarrow{Z}(s) \cdot \tilde{\vec{J}}_s, \quad \overleftrightarrow{Z}(s) = \begin{pmatrix} 0 & \pm\frac{\tilde{Z}_s(s)+2/a}{s} \\ \pm\frac{\tilde{Z}_s(s)+2/a}{s} & 0 \end{pmatrix}, \tag{2.141}$$

where $\tilde{\vec{E}}_{tan} = \vec{1}_r \times \tilde{\vec{E}}$, $\tilde{\vec{J}}_s = \vec{1}_r \times \tilde{\vec{H}}_{tan}$, $\tilde{Z}_s(s)$ is the scalar impedance function, a is the radius of the sphere, the $\pm$ sign is determined by the choice of the coordinate system. Using the standard spherical coordinate system the above surface impedance boundary condition becomes

$$\tilde{E}_\theta = -\frac{\tilde{Z}_s(s) + 2/a}{s}\tilde{H}_\phi, \quad \tilde{E}_\phi = +\frac{\tilde{Z}_s(s) + 2/a}{s}\tilde{H}_\theta. \tag{2.142}$$

Next, $\nabla \cdot \tilde{\vec{E}} = 0$ can be expressed in the spherical coordinate system as

$$\left(\frac{\partial}{\partial r} + \frac{2}{a}\right)\tilde{E}_r = -\frac{1}{a\sin\theta}\frac{\partial}{\partial\theta}(\tilde{E}_\theta\sin\theta) - \frac{1}{a\sin\theta}\frac{\partial\tilde{E}_\phi}{\partial\phi}. \tag{2.143}$$

The $\vec{1}_r$ component of the equation $s\tilde{\vec{E}} = \nabla \times \tilde{\vec{H}}$ can be expressed as

$$s\tilde{E}_r = \frac{1}{a\sin\theta}\left(\frac{\partial}{\partial\theta}(\tilde{H}_\phi\sin\theta) - \frac{\partial\tilde{H}_\theta}{\partial\phi}\right). \tag{2.144}$$

Inserting the boundary conditions (2.142) into the equations (2.143), we derive

$$\left(\frac{\partial}{\partial r} + \frac{2}{a}\right)\tilde{E}_r = \frac{Z + 2/a}{s}\left(\frac{1}{a\sin\theta}\frac{\partial}{\partial\theta}(\tilde{H}_\phi\sin\theta)\right) - \frac{Z + 2/a}{s}\left(\frac{1}{a\sin\theta}\frac{\partial\tilde{H}_\theta}{\partial\phi}\right).$$
(2.145)

Finally, applying (2.144) to the equation (2.145) to find the equivalent surface impedance loaded boundary condition.

$$\frac{\partial}{\partial r}\tilde{E}_r = \tilde{Z}_s(s)\tilde{E}_r.$$
(2.146)

Thus $\tilde{E}_r$ satisfies the same impedance condition that appears in the acoustic scattering problem $n\cdot\nabla u = \frac{\partial u}{\partial n} = \alpha(s)u$. To see that all the results in the acoustic scattering case hold here, we expand the solution for the scattered field using (2.132), (2.133), (2.134), (2.135). From (2.107), (2.108), (2.109), (2.122) and (2.123), the $\vec{1}_r$ component of the electric field $\left(\vec{E}\right)_r = \left(\vec{E}_{in} + \vec{E}_{sc}\right)_r$ can be expanded

$$\left(\vec{E}_{in}\right)_r = E_0\sum_{n=1}^{\infty}\sum_{m=0}^{n}\sum_{p=e,o}b_{n,m,p}^{(2)}n(n+1)i_n(\gamma r)Y_{n,m,p}(\theta,\phi),$$
(2.147)

$$\left(\vec{E}_{sc}\right)_r = E_0\sum_{n=1}^{\infty}\sum_{m=0}^{n}\sum_{p=e,o}b_{n,m,p}^{(3)}n(n+1)k_n(\gamma r)Y_{n,m,p}(\theta,\phi).$$
(2.148)

Applying the surface impedance boundary condition (2.146), we find

$$b_{n,m,p}^{(3)} = \frac{\gamma i_n'(\gamma r) + \tilde{Z}_s(s)i_n(\gamma r)}{\gamma k_n'(\gamma r) + \tilde{Z}_s(s)k_n(\gamma r)}b_{n,m,p}^{(2)},$$
(2.149)

where $\gamma = s/c$. Therefore, one can construct scattering poles of arbitrary order for the surface impedance loaded sphere.

2.3.4   Sheet-impedance-loaded sphere

In this case, we try to find higher order scattering poles by using the same technique we discussed above with the scatterer being a lossless sheet impedance loaded sphere.

However, we only find 2nd order scattering poles for a non-physical mode which leads to the conjecture that there only exist first order scattering poles. Let a sheet impedance be $\tilde{Z}_s(s)$ independent of θ, ϕ (a scalar) located on a spherical surface give by $r = a$. The sheet impedance boundary condition relates tangential electric field and surface current density as

$$\overleftrightarrow{1_t} \cdot \tilde{\vec{E}}(a, \theta, \phi, s) = \tilde{Z}_s(s)\tilde{J}_s(\theta, \phi, s), \tag{2.150}$$

$$\overleftrightarrow{1_t} = \overleftrightarrow{1} - \vec{1}_r\vec{1}_r = \text{transverse dyad,}$$

The surface current density is in turn related to the magnetic field via

$$\vec{1}_r \times [\tilde{\vec{H}}(a+, \theta, \phi, s) - \tilde{\vec{H}}(a-, \theta, \phi, s)] = \tilde{J}_s(\theta, \phi, s). \tag{2.151}$$

The sheet impedance function $\tilde{Z}_s(s)$ also has to satisfy Foster's theorem to guarantee a lossless boundary condition. Matching the boundary condition at $r = a$, with the sheet impedance boundary condition together with the continuity of the tangential electric field gives

$$\overleftrightarrow{1_t} \cdot [\tilde{\vec{E}}_{inc}(a+, \theta, \phi, s) + \tilde{\vec{E}}_{sc}(a+, \theta, \phi, s)] = \overleftrightarrow{1_t} \cdot \tilde{\vec{E}}_{in}(a-, \theta, \phi, s) \tag{2.152}$$

$$= \tilde{Z}_s(s)\tilde{J}_s(\theta, \phi, s) \tag{2.153}$$

$$= \tilde{Z}_s(s) \times [\tilde{\vec{H}}_{inc}(a+, \theta, \phi, s) + \tilde{\vec{H}}_{sc}(a+, \theta, \phi, s) - \tilde{\vec{H}}_{in}(a-, \theta, \phi, s)]. \tag{2.154}$$

Substituting the series (2.124), (2.132), (2.134), (2.125), (2.133), (2.135) for the field and applying the boundary condition (2.152), (2.154), we derive a system of equations involving $a_{n,m,p}^{(2)}, b_{n,m,p}^{(2)}, a_{n,m,p}^{(3)}, b_{n,m,p}^{(3)}$,

$$b_{n,m,p}^{(1)}\frac{[\gamma a i_n(\gamma a)]'}{\gamma a} + b_{n,m,p}^{(3)}\frac{[\gamma a k_n(\gamma a)]'}{\gamma a} = b_{n,m,p}^{(2)}\frac{[\gamma a i_n(\gamma a)]'}{\gamma a}, \tag{2.155}$$

$$b_{n,m,p}^{(2)}\frac{[\gamma a i_n(\gamma a)]'}{\gamma a} = \frac{\tilde{Z}_s(s)}{Z_0}\left(b_{n,m,p}^{(1)}i_n(\gamma a) + b_{n,m,p}^{(3)}k_n(\gamma a) - b_{n,m,p}^{(2)}i_n(\gamma a)\right), \tag{2.156}$$

$$a_{n,m,p}^{(1)}i_n(\gamma a) + a_{n,m,p}^{(3)}k_n(\gamma a) = a_{n,m,p}^{(2)}i_n(\gamma a), \tag{2.157}$$

$$a_{n,m,p}^{(2)} i_n(\gamma a) = \frac{\tilde{Z}_s(s)}{Z_0} \left(a_{n,m,p}^{(1)} \frac{[\gamma a i_n(\gamma a)]'}{\gamma a} + a_{n,m,p}^{(3)} \frac{[\gamma a k_n(\gamma a)]'}{\gamma a} - a_{n,m,p}^{(2)} \frac{[\gamma a i_n(\gamma a)]'}{\gamma a} \right).$$
(2.158)

Eliminating $a_{n,m,p}^{(3)}$, $b_{n,m,p}^{(3)}$, applying the Wronskian relation, and solving for $a_{n,m,p}^{(2)}$ and $b_{n,m,p}^{(2)}$ we get

$$a_{n,m,p}^{(2)} = \frac{a_{n,m,p}^{(1)}}{1 + \frac{Z_0}{\tilde{Z}_s(s)}(\gamma a)^2 i_n(\gamma a) k_n(\gamma a)},$$
(2.159)

$$b_{n,m,p}^{(2)} = \frac{b_{n,m,p}^{(1)}}{1 - \frac{Z_0}{\tilde{Z}_s(s)}[\gamma a i_n(\gamma a)]'[\gamma a k_n(\gamma a)]'}.$$
(2.160)

So the surface current density is given by

$$\tilde{J}_s(\theta, \phi, s) = \frac{E_0}{\tilde{Z}_s(s)} \sum_{n=1}^{\infty} \sum_{m=0}^{n} \sum_{p=e,o} [a_{n,m,p}^{(2)} i_n(\gamma a) \vec{R}_{n,m,p}(\theta, \phi)$$

$$+ b_{n,m,p}^{(2)} \frac{[\gamma a i_n(\gamma a)]'}{\gamma a} \vec{Q}_{n,m,p}(\theta, \phi)]. \quad (2.161)$$

The coefficients relevant to the existence of a second order pole are

$$c_1 = \frac{i_n(\gamma a)}{\tilde{Z}_s(s)} a_{n,m,p}^{(2)},$$
(2.162)

$$c_2 = \frac{[\gamma a i_n(\gamma a)]'}{\tilde{Z}_s(s) \gamma a} b_{n,m,p}^{(2)}.$$
(2.163)

For simplicity, we assume $\gamma a = s$.

Consider the sheet impedance function

$$\tilde{Z}_s(s) = \frac{(\frac{1}{2}e^2 + e + \frac{1}{2})s}{s^2 + \frac{1}{2}e + \frac{1}{4}}.$$
(2.164)

Clearly $\tilde{Z}_s(s)$ satisfies Foster's theorem with $K > 0$ and $\omega_0 > 0$. The expansion of $i_n(s)$ and $k_n(s)$ are

$$k_n(s) = \frac{e^{-s}}{s} \sum_{j=0}^{n} \frac{(n+j)! \, 2^{-j} s^{-j}}{j! \, (n-j)!},$$
(2.165)

$$i_n(s) = \frac{1}{2}[(-1)^{n+1} k_n(s) - k_n(-s)].$$
(2.166)

For $n = 0$, the denominator of c_2 is

$$De(s) = \left(4\,e^{-2\,s} + 4\right)s^2 + \left(4\,e^2 + 8\,e + 4\right)s + 1 + 2\,e^{1-2\,s} + 2\,e + e^{-2\,s}$$

$$= \left(16\,e + 4 + 4\,e^2\right)\left(s + \frac{1}{2}\right)^2$$

$$+ \left(-\frac{56}{3}\,e - \frac{8}{3}\,e^2\right)\left(s + \frac{1}{2}\right)^3 + O\left(s + \frac{1}{2}\right)^4. \tag{2.167}$$

Thus we find a second order pole at $s = -\frac{1}{2}$ in the non-physical case $n = 0$.

In general, to construct a sheet impedance function $\tilde{Z}_s(s) = \frac{Ks}{(s^2+\omega)}$ such that c_2 has a second order pole in the left half plane of s with $K > 0$ and $\omega > 0$, one needs to set the denominator of c_2 to zero and the derivatives of the denominator to zero. The denominator of c_2 has the following form

$$De(s) = Ks + \left(-s^2 i_n\,(s) - s^3 \frac{d}{ds} i_n\,(s) - \omega\,i_n\,(s) - \omega\,s\frac{d}{ds} i_n\,(s)\right) k_n\,(s) \tag{2.168}$$

$$+ \left(-s^3 i_n\,(s) - s^4 \frac{d}{ds} i_n\,(s) - s\omega\,i_n\,(s) - s^2\omega \frac{d}{ds} i_n\,(s)\right) \frac{d}{ds} k_n\,(s).$$

Solving $De(s) = 0$ and $\frac{d}{ds} De(s) = 0$ for K, and ω in terms of s, the solution s_α must satisfy $s_\alpha < 0$, $K(s_\alpha) > 0$ and $\omega(s_\alpha) > 0$. That is, for $n = 0$ we solve

$$2\,Ks + s^2 e^{-2\,s} + s^2 + \omega\,e^{-2\,s} + \omega = 0, \tag{2.169}$$

$$2\,K + 2\,se^{-2\,s} - 2\,s^2 e^{-2\,s} + 2\,s - 2\,\omega\,e^{-2\,s} = 0, \tag{2.170}$$

and find

$$K = -\frac{s\left(e^{-4\,s} + 2\,e^{-2\,s} + 1\right)}{2\,se^{-2\,s} + e^{-2\,s} + 1}, \tag{2.171}$$

$$\omega = -\frac{s^2\left(-e^{-2\,s} + 2\,se^{-2\,s} - 1\right)}{2\,se^{-2\,s} + e^{-2\,s} + 1}. \tag{2.172}$$

We can see from Figure 2.5 that approximately when $-0.64 < s < 0$ both K and ω are positive, thus the assumptions are satisfied. However, for $n = 1$, applying the

same techniques as above K and ω become

$$K = [s^3(10\,s^3\mathrm{e}^{-2\,s} - 3\,s^2 - 5 + 5\,\mathrm{e}^{-2\,s} + 13\,s^2\mathrm{e}^{-2\,s} + 10\,s\mathrm{e}^{-2\,s} + 5\,s^4\mathrm{e}^{-2\,s} + 2\,s^5\mathrm{e}^{-2\,s}$$
$$- s^4)]^{-1}\left(-s^2 + \mathrm{e}^{-2\,s} - 1 - s^4 + s^4\mathrm{e}^{-2\,s} + 3\,s^2\mathrm{e}^{-2\,s} + 2\,s^3\mathrm{e}^{-2\,s} + 2\,s\mathrm{e}^{-2\,s}\right)^2,$$

$$(2.173)$$

$$w = -[s^2(7\,s^2\mathrm{e}^{-2\,s} - 3 + 2\,s^5\mathrm{e}^{-2\,s} + 3\,\mathrm{e}^{-2\,s} + 3\,s^4\mathrm{e}^{-2\,s} + 6\,s\mathrm{e}^{-2\,s} + 6\,s^3\mathrm{e}^{-2\,s} + s^4 - s^2)]^{-1}$$
$$(10\,s^3\mathrm{e}^{-2\,s} - 3\,s^2 - 5 + 5\,\mathrm{e}^{-2\,s} + 13\,s^2\mathrm{e}^{-2\,s} + 10\,s\mathrm{e}^{-2\,s} + 5\,s^4\mathrm{e}^{-2\,s} + 2\,s^5\mathrm{e}^{-2\,s} - s^4).$$

$$(2.174)$$

From Figure 2.6, it can be seen that K and ω can not simultaneously be positive in the left half plane. Similar results hold for c_1. More plots are shown below for both c_1 and c_2 with $n = 0, 1, ..., 5$. Note that it does not help to add extra terms to the expansion of the impedance function $\tilde{Z}_s(s)$, because according to Foster's theorem, additional conditions will be needed to be imposed on the impedance function. To date we have only been able to construct second order scattering poles in the non-physical case $n = 0$ for coefficient c_2, while meeting all the assumptions from the Foster's theorem. Thus, we conjecture that only first order poles exist in the sheet-impedance case.
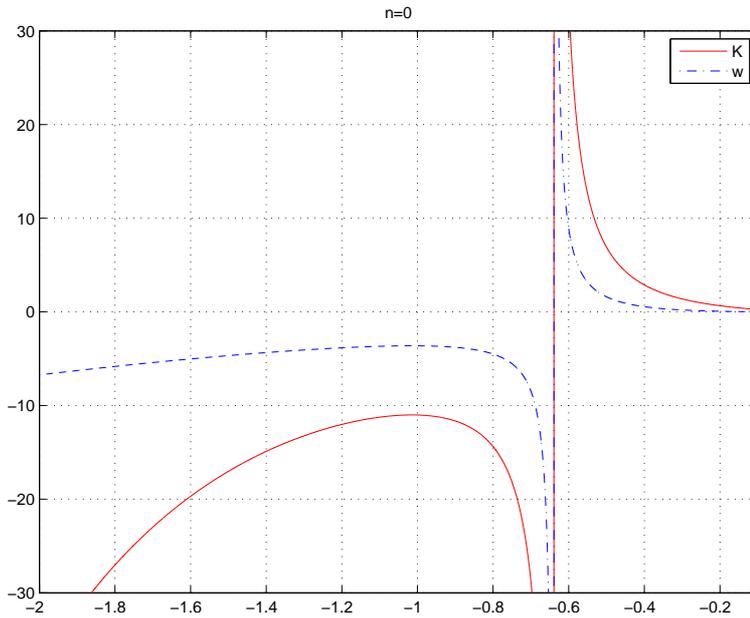
Figure 2.5: Possible region of K and ω on the left half plane of s to construct 2nd order scattering pole associated with coefficient c_2 and $n = 0$.
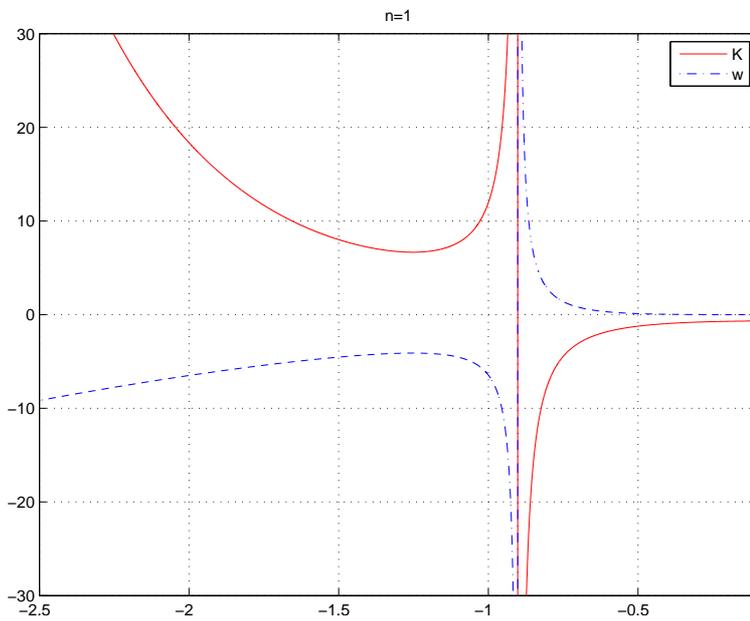

Figure 2.6: Possible region of K and ω on the left half plane of s to construct 2nd order scattering pole associated with coefficient c_2 and $n = 1$.
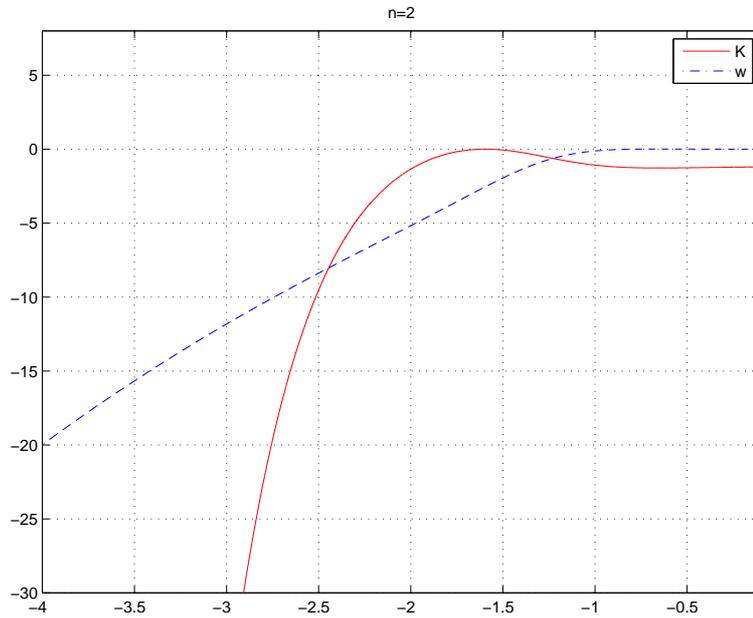
Figure 2.7: Possible region of K and ω on the left half plane of s to construct 2nd order scattering pole associated with coefficient c_2 and $n = 2$.
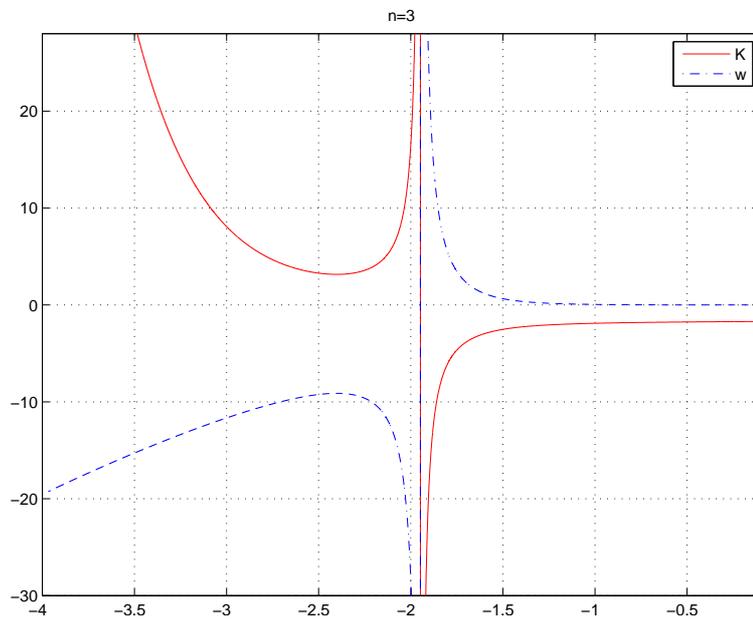

Figure 2.8: Possible region of K and ω on the left half plane of s to construct 2nd order scattering pole associated with coefficient c_2 and $n = 3$.
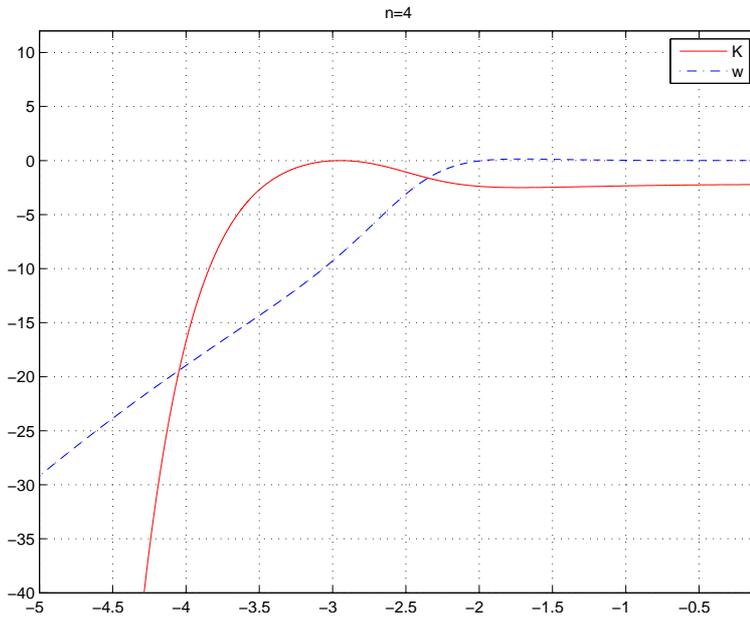
Figure 2.9: Possible region of K and ω on the left half plane of s to construct 2nd order scattering pole associated with coefficient c_2 and $n = 4$.

Figure 2.10: Possible region of K and ω on the left half plane of s to construct 2nd order scattering pole associated with coefficient c_2 and $n = 5$.
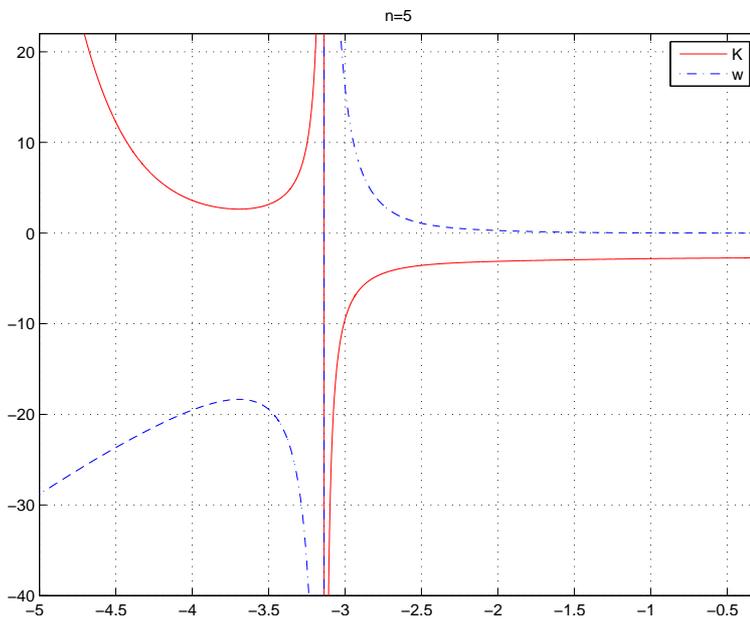
Figure 2.11: Possible region of K and ω on the left half plane of s to construct 2nd order scattering pole associated with coefficient c_1 and $n = 0$.

Figure 2.12: Possible region of K and ω on the left half plane of s to construct 2nd order scattering pole associated with coefficient c_1 and $n = 1$.
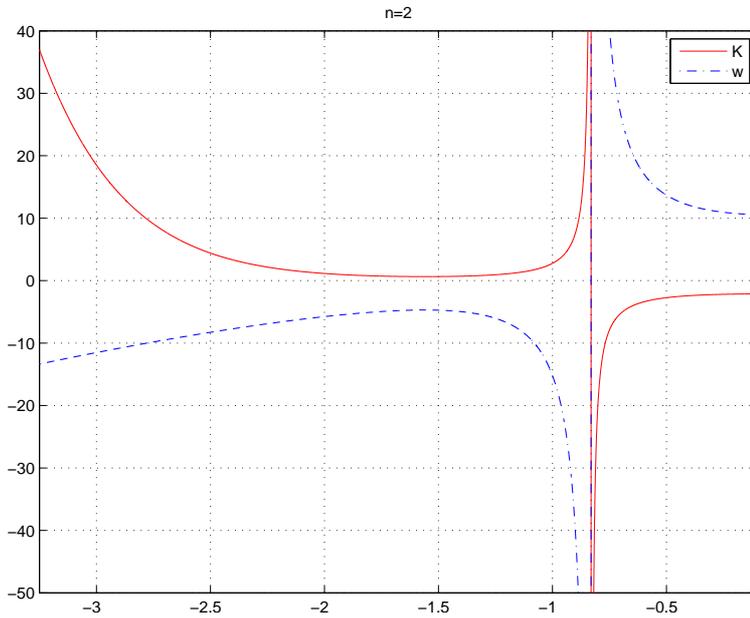
Figure 2.13: Possible region of K and ω on the left half plane of s to construct 2nd order scattering pole associated with coefficient c_1 and $n = 2$.

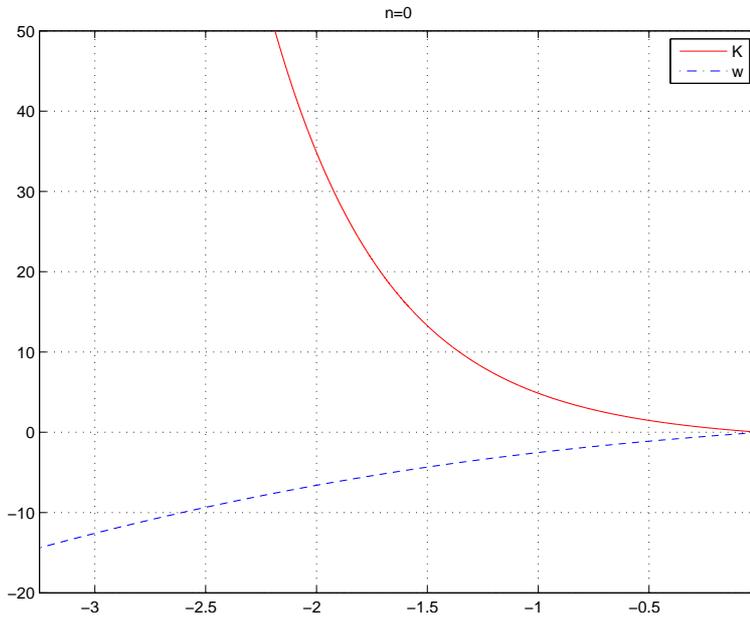Figure 2.14: Possible region of K and ω on the left half plane of s to construct 2nd order scattering pole associated with coefficient c_1 and $n = 3$.

Figure 2.15: Possible region of K and ω on the left half plane of s to construct 2nd order scattering pole associated with coefficient c_1 and $n = 4$.
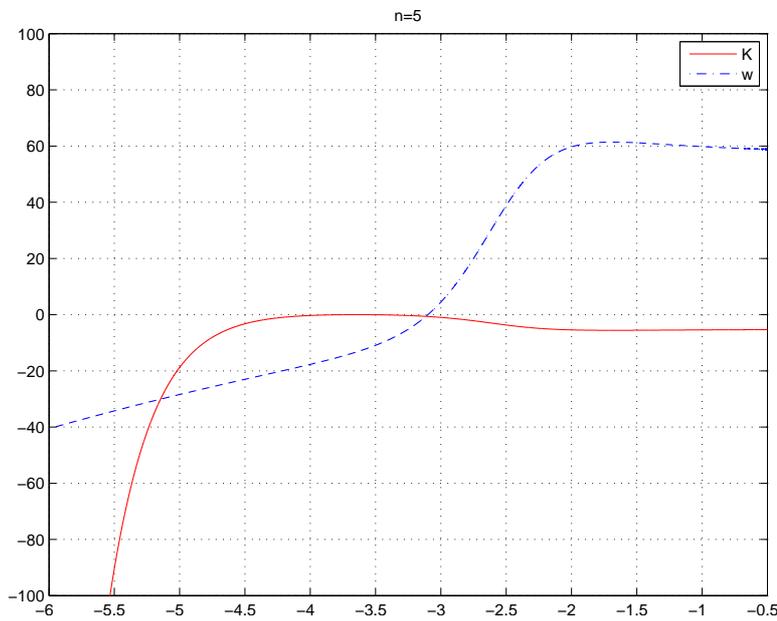
Figure 2.16: Possible region of K and ω on the left half plane of s to construct 2nd order scattering pole associated with coefficient c_1 and $n = 5$.
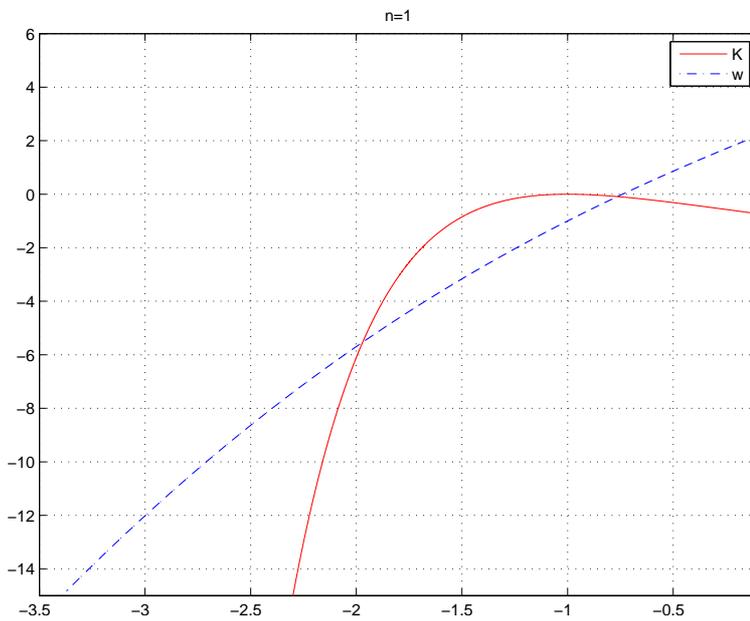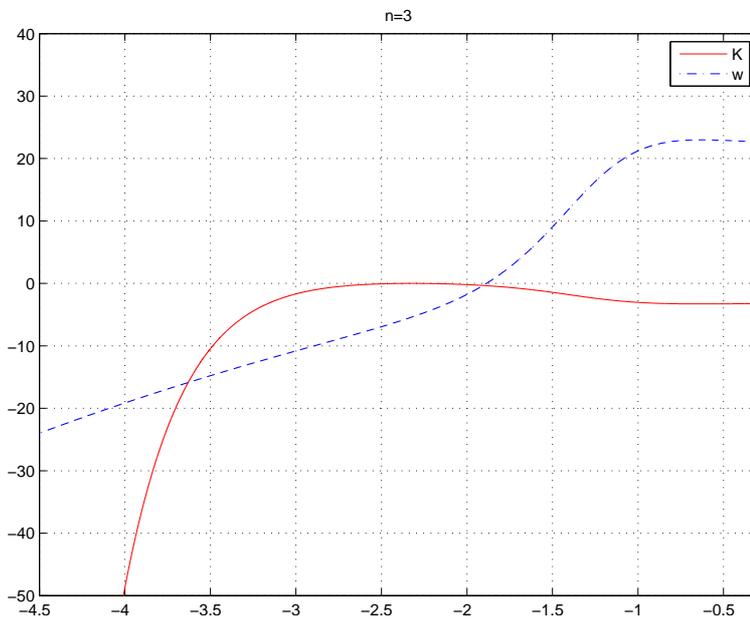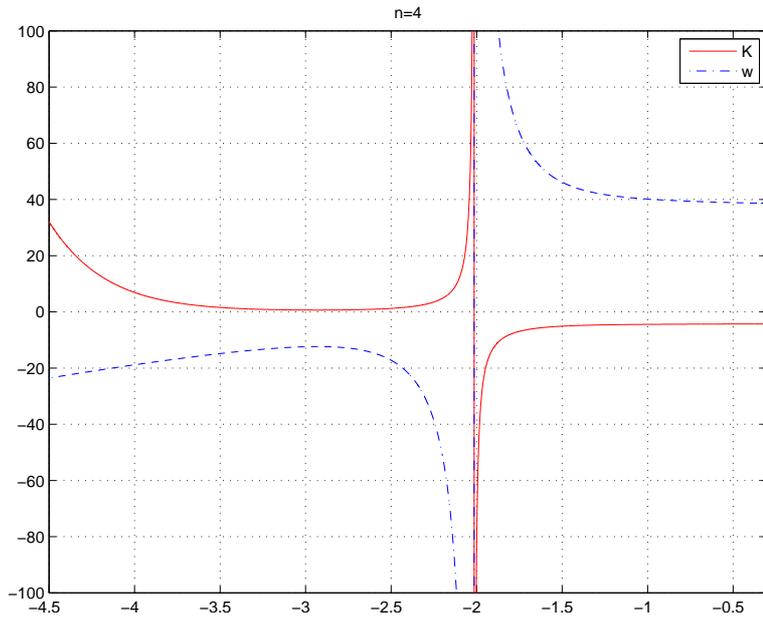
Chapter 3

Analysis of supercontinumm generation

3.1   Introduction

Since the experimental demonstration of optical soliton propagation in single mode fibers some 20 plus years ago, the investigation of pulse dynamics in nonlinear optical fibers has evolved due to the introduction of novel structures with complex properties, such as photonic crystal and holey fibers [52]. In essence these are examples of engineered dielectric structures aimed at tailoring dispersive characteristics and enhancing nonlinear behavior. A direct outcome in terms of the pulse dynamics that has brought much attention from several experimental groups [54, 49, 57, 53, 64, 68, 59, 67, 55, 48] is the ability to generate broadband supercontinuum spectra. Scientifically this is a departure from soliton dynamics that requires careful analytical and numerical modeling in parallel with the experiments. From the applications point of view, it has opened possibilities never seen before in areas such as frequency metrology [65] and medical diagnostics [51, 60, 58].

Supercontinuum generation (SCG) is a process where broad continuous spectra is generated for high power pulses input propagating through a nonlinear media, as first observed in 1970 by Alfano and Shapiro [7, 8]. The term supercontinuum usually refers to extreme pulse broadening. SCG can arise from various physical processes such as self- and cross-phase modulation, and amplitude modulation [6]. Due to the complex interplay of linear and nonlinear phenomena in SCG dynamics, the theoretical formulation of the SCG mechanism imposes considerable challenges, in particular if this process happens in bulk media [61]. The major recent theory that explains the SCG for relatively low intensities in confined waveguides rests on the evolution and fission of higher-order solitons near the zero-dispersion wavelength in [55, 46, 56]. If the input wavelength is close to the zero-dispersion wavelength, then the influence of third-order dispersion is strong, thus a higher-order soliton with number N splits into its constituent solitons with the emission of blueshifted nonsolitonic radiation [62]. Since each soliton and its corresponding radiation has a different central frequency, the width of the generated total spectrum increases with increasing soliton number.

Recent experimental observations of supercontina in soft glass, however, suggest an interesting physical mechanism of SCG that cannot be fully explained by the previously known theories. In these experiments, SCG occurs in a dramatic fashion in the very early states of propagation, in particular at a length scale where solitons start forming. Such a phenomenon can only be explained if, initially, nonlinear effects other than soliton fission dominate the physics. Indeed, the underpinning mechanisms that generate supercontinua as reported in most theoretical and experimental studies are shock generation and its dispersive regularization in combination with multisoliton fission. The shock generation is a well known classical phenomenon in fluids and gas dynamics [63]. It also appears in ultra-short pulse propagation in fibers [3, 47, 50]. For ultrashort pulses, the refractive index depends on the pulse intensity, thus the center of the pulse envelope travels with a different speed than

that of the trailing and leading edges of the pulse; this leads to an asymmetric shape of the pulse, which invokes shock formation. However, in optical propagation, dispersion plays an important role, preventing a sharp discontinuity. On the other hand, multisoliton generation resulting from small dispersion effects is a consequence of the integrability of the nonlinear Schrödinger equation (NLSE) [66]. Its eventual fission is the result of perturbations to the NLSE such as third order dispersion. Altogether, a universal feature of nonlinear dispersive wave phenomena is that the long term dynamics results from the delicate balance between linear and nonlinear effects.

In this work, we recognize the aforementioned outcomes for a more realistic model describing the pulse dynamics in photonic crystal fibers. The extended model accounts for all competing effects including self-steepening, which we believe is as important as the effect from the fully detailed linear dispersion. In order to understand and exploit these phenomena, it is essential to obtain and analyze better these mathematical models. This in addition could explain for each instance in a real experiment what triggered SC generation. To begin with, the accurate broadband modeling of the dispersion relation is required to make sure one does not obtain spurious results, and to do so here we depart from the commonly used approach where a Taylor series expansion of the propagation constant β models the dispersive properties in a generalized nonlinear Schrödinger equation (gNLSE). Instead, we develop a mathematical model starting from calculated group velocity dispersion (GVD) curves. Then, we construct the function $\beta(\omega)$ over a broad frequency window and integrate the gNLSE preserving the spectral dependence of the propagation constant. The generation of broadband supercontinua in air-silica microstructured fibers results from a delicate balance of dispersion and nonlinearity. As an illustration, we present our numerical results based on the calculated GVD for an LP_{01} mode in an air-silica microstructured fiber studied by Dudley *et al.* [57]. Then we carry out a careful numerical analysis. We find that if the nonlinear self-steepening

term is strong enough, the model as it stands produces a shock that is not arrested by dispersion, whereas for weaker nonlinearity the pulse propagates the full extent of the fiber with the generation of a supercontinuum.

3.2   Formulation of the model

To better illustrate this delicate balance we begin by studying a simpler model. The propagation of an electric field wave packet through an optical fiber can be described by gNLSE [3],

$$i\partial_z A + \mathcal{F}^{-1}[(\beta(\omega) - \beta(\omega_0))\hat{A}] + \gamma \left(1 + \frac{i}{\omega_0}\partial_t\right)(A|A|^2) = 0. \tag{3.1}$$

Here, the variables z, T, ω represent propagation distance, time and optical frequency, respectively. The envelope of the wave packet is A, and $c, \lambda, \omega_0, \beta(\omega)$ represent the velocity of light in vacuum, wavelength, central frequency and wave number, respectively. $\mathcal{F}^{-1}$ denotes the inverse Fourier transform, and $\hat{A}$ is the Fourier transform of the pulse envelope. Finally the self-steepening term models the instantaneous nonlinear response function of the medium, which is a good approximation given the temporal lengths of the pulses. The inclusion of a Raman (non-instantaneous) term we believe will only introduce a shift in the peak location in the spectrum (see Figure 3.4 and compare with Figure. 5b in [57]).

The effects of fiber dispersion are accounted for by the propagation constant $\beta(\omega)$ which we calculate based on the dispersion profile presented in [57], without performing a Taylor series expansion around the carrier frequency. Using two high-precision numerical integrations of an accurate rational interpolant of the GVD curve, we obtain the GVD function $D(s)$. Then, the group velocity $\nu_g(\omega)$ is derived from $D(s)$ through the relation,

$$\frac{1}{\nu_g(\omega)} - \frac{1}{\nu_g(\omega_0)} = \int_{\lambda_0}^{\lambda} D(s)ds. \tag{3.2}$$

By setting $F(\lambda) = \int_{\lambda_0}^{\lambda} D(s)ds$, we obtain

$$\nu_g(\omega) = \frac{\nu_g(\omega_0)}{1 + \nu_g(\omega_0)F(\lambda)}. \tag{3.3}$$

Since $\frac{\partial \beta}{\partial \omega} = \frac{1}{\nu_g}$, it follows that

$$\frac{\partial \beta}{\partial \omega} = \frac{1}{\nu_g(\omega_0)} + F(\lambda) \sim \frac{1}{c} + F(\lambda). \tag{3.4}$$

By integrating Eq. (3.4) with respect to ω and using the relation $\lambda = \frac{2\pi c}{\omega}$, we obtain

$$\beta(\omega) - \beta(\omega_0) = \frac{\omega - \omega_0}{\nu_g(\omega_0)} - 2\pi c \int_{\lambda_0}^{\lambda} \frac{F(\lambda)}{\lambda^2} d\lambda. \tag{3.5}$$

We employ a frame of reference moving with the pulse at the group velocity ν_g by making the transformation $t = T - z/\nu_g$. In the end, we obtain

$$i\partial_z A - 2\pi c \mathcal{F}^{-1}\left(\int_{\lambda_0}^{\frac{2\pi c}{\omega}} \frac{F(\lambda)}{\lambda^2} d\lambda \hat{A}(\omega, z)\right) + \gamma \left(1 + \frac{i}{\omega_0}\partial_t\right)(A|A|^2) = 0. \tag{3.6}$$

The resulting equation preserves the complete structure of fiber dispersion which is indeed utilized in experiments. In addition, the equation is valid not only for broad pulses, but also short pulses since the derivation is carried out without the assumption of a pulse centered around a specific carrier frequency (without Taylor series expansion of the propagation constant around a carrier frequency). In the remainder of this section, we present analytical and numerical results obtained from the gNLS (3.6).

3.2.1 Optical shock formation

In order to first pay attention to the nonlinear effects governing the mechanism of shock formation [3, 47], we consider the dispersionless case by setting $F(\lambda) = 0$ in Eq. (3.6). In the absence of dispersion, we first split Eq. (3.6) into an intensity-phase

system by adding and subtracting

$$A^*\frac{\partial A}{\partial z} = -\frac{\gamma A^*}{\omega_0}\frac{\partial}{\partial t}(|A|^2 A) + i\gamma|A|^4, \tag{3.7}$$

$$A\frac{\partial A^*}{\partial z} = -\frac{\gamma A}{\omega_0}\frac{\partial}{\partial t}(|A|^2 A^*) - i\gamma|A|^4, \tag{3.8}$$

where A^* is the complex conjugate of A.

By defining $I = |A|^2$, the addition of Eqs. (3.7), (3.8) gives

$$\frac{\partial I}{\partial z} = -\frac{\gamma}{\omega_0}\left[2|A|^2\frac{\partial|A|^2}{\partial t} + |A|^2\left(A^*\frac{\partial A}{\partial t} + A\frac{\partial A^*}{\partial t}\right)\right]$$

$$= -\frac{3\gamma}{\omega_0}I\frac{\partial I}{\partial t}. \tag{3.9}$$

The general solution of Eq. (3.9) is

$$I(z,t) = f\left(t - \frac{3\gamma}{\omega_0}Iz\right), \tag{3.10}$$

where $f(t)$ is determined by the initial pulse shape, namely, $f(t) = I(0,t)$. The solution form Eq. (3.10) implies that asymmetric distortion of the pulse will occur eventually.

From Eq. (3.10), we also find

$$\frac{\partial I}{\partial t} = \frac{f'}{1 + \frac{3\gamma}{\omega_0}f'z}. \tag{3.11}$$

The resulting equation shows that after a distance $z_s = -\frac{\omega_0}{3\gamma}\frac{1}{f'}$, a singularity in the pulse intensity will be generated, namely, the formation of an optical shock. This shock does play an important role in the spectral broadening once dispersion regularizes it. In other words, the effects of fiber dispersion cannot be ignored. Moreover, the effect of GVD becomes more important as the pulse steepening becomes significant. This phenomenon prevents further steepening of pulse shape, i.e.,an appropriate strength of linear dispersion results in a mechanism that may prevent (or regularize) the shock.

3.2.2 Numerical solutions of the generalized nonlinear Schrödinger equation

We perform our numerical simulations of pulse dynamics based on equation (3.6). In particular, we consider the propagation of $100fs$ pulses at $\omega_0 =780nm$ in a 1-m length air-silica microstructured fiber with $\gamma = 0.1W^{-1}m^{-1}$. We assume that the input pulse has a form of $A_0^2/\cosh^2\frac{t}{t_0}$. As a reference, we find from the previous analysis that the shock length for this input pulse is $z_s = 22.1cm$ which is much shorter than the actual fiber length, thus in this case shock regularization is a likely scenario for SC generation.
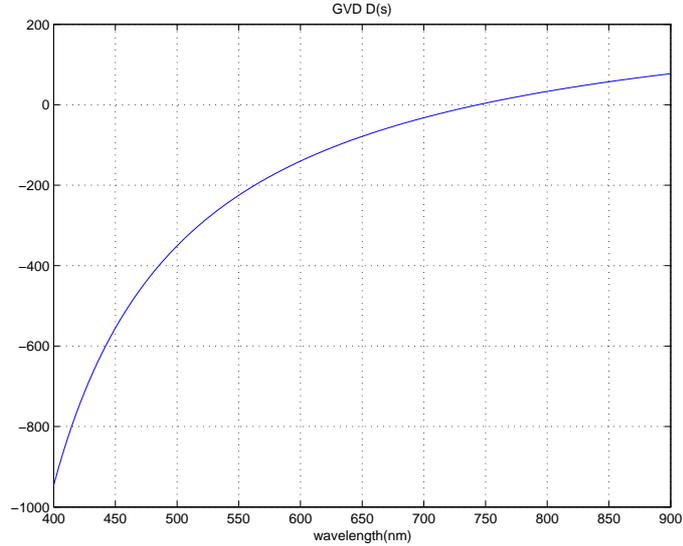

Figure 3.1: Group Velocity Dispersion (GVD) function $D(s)$.

The GVD profile we used here is from Figure 2 in [57]. We used rational interpolation to approximate the GVD function $D(s)$ (3.2) for the LP_{01} mode which is shown in Figure 3.1. The integration spectrally covers the $400 - 900nm$ range and, as stated above, it does not require a Taylor expansion of β. Instead we calculate β shown in Figure 3.5 via two high-precision numerical integrations of the GVD.

The region has been scaled to cover the $400 - 900nm$ range for $D(s)$ so that no extrapolation is used. We use spectral (Fourier) method with $n = 2^{13}$ to calculate the temporal derivatives and the numerical evolution is then completed using a standard adaptive ODE solver.
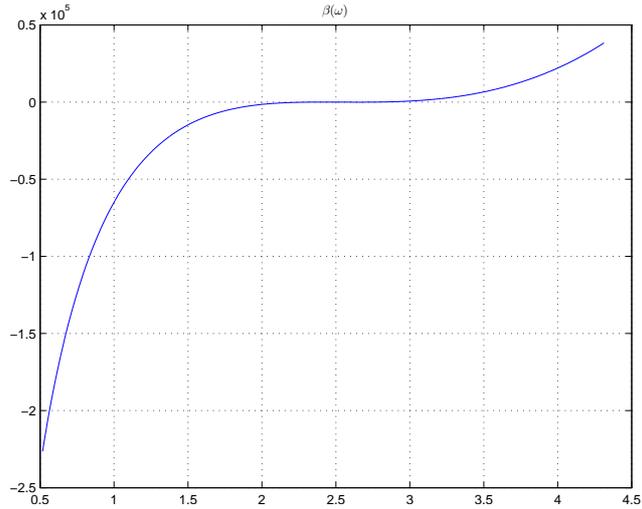

Figure 3.2: Propagation constant $\beta(\omega)$ (1/fs).

The results displayed below (Figures 3.3, 3.4) are in clear qualitative agreement with those in [57]. We should point out two important distinctions between Figure 5 in [57] and Figure 3.3 here: we do not capture the peak in the spectrum at wavelengths close to $1200nm$ in 5a,b of [57] and the corresponding pulse (labeled C) shown in Figure 3.3c. This is because in our approach we computed the dispersion profile based on the calculated GVD curve shown in Figure 2 of [57]. This calculation did not extend to wavelengths beyond $900nm$ and we did not extrapolate such curves. This explains the sharp decay in the spectrum of Figure 3.3 here. Calculations based on a Taylor expansion of β which is commonly used, can in principle be extended to any desired spectral range. On the other hand, our result better reproduces the observed supercontinuum spectrum (Figure 5a of [57]) in the short

(less than $600nm$) wavelength portion and is as good as the Taylor expansion in the intermediate regime.

Finally, Figure 3.3 (right) shows five distinguishable pulses at the output. In Figure 3.4, we spectrally isolate each pulse and find their peaks centered approximately at: $867nm$ (peak 1), $913nm$ (peak 2), $848nm$ (peak 3), $840nm$ (peak 4) $858nm$ (peak 5). As we state below, the spectral separation should be accentuated by the presence of the Raman shift which we did not include in the model. What is most important here is we corroborate spectral broadening and that splitting of pulses occurs, with the spectral shift accounting for differences in soliton velocities.
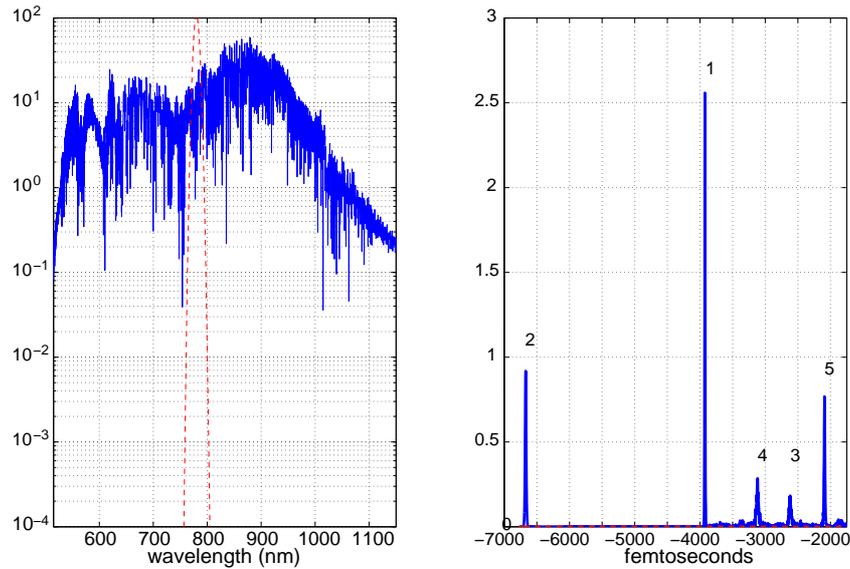

Figure 3.3: Spectral (left) and temporal (right) picture of the output after $1m$ propagation. All relevant parameters are taken from [57]. $100fs$ pulses at $\omega_0=780nm$, $\gamma=0.1W^{-1}m^{-1}$.
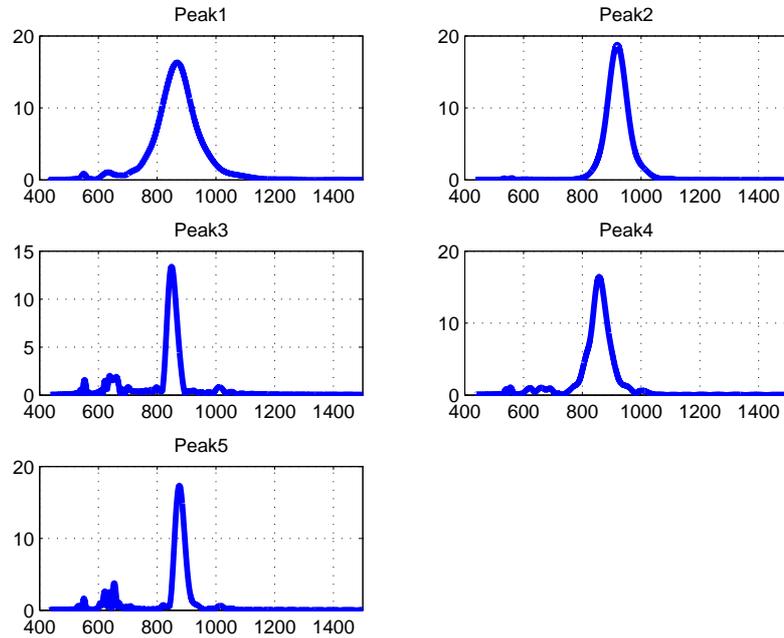
Figure 3.4: Spectral characteristics of the five distinguishable temporal pulses from Figure 3.3.

3.2.3   Numerical balance between the dispersion and nonlinearity

In trying to understand the critical balance between linear and nonlinear effects, we now depart from the concrete example to further illustrate this interplay in a series of simulations of the equation shown below, which is no different than equation (3.6), except that we placed two adjustable constants $c_2(c_3)$ in front of the self-steepening (linear) term to account for their respective strengths.

$$i\partial_z A - c_3 2\pi c \mathcal{F}^{-1}\left(\int_{\lambda_0}^{\frac{2\pi c}{\omega}} \frac{F(\lambda)}{\lambda^2} d\lambda \hat{A}(\omega, z)\right) + \gamma\left(1 + c_2\frac{i}{\omega_0}\partial_t\right)(A|A|^2) = 0. \quad (3.12)$$

By proper re-scaling of the propagation variable z and the pulse peak amplitude, one can eliminate the parameter c_3 but for clarity we analyze our simulations in the 2-parameter space while using the same initial condition. Observe that from (3.11), an increase of c_2 effectively means the shock length is reduced. In practice, this shock length reduction can be induced by having shorter input pulses. By allowing ourselves to modulate dispersion and nonlinearity through these two parameters, we hope to highlight how delicate this balance is.

We begin with the extreme case, (i) $\gamma = 1$, $c_2 = 1$, $c_3 = 0$, only self-steepening occurs in Figure 3.5; (ii) $\gamma = 0$, $c_2 = 0$, $c_3 = 1$, only dispersion occurs in Figure 3.6; (iii) $\gamma = 1$, $c_2 = 0$, $c_3 = 1$, no self-steepening term has been included, soliton fission driven broadening is observed in Figure 3.7 (for simplicity the temporal soliton solution is plotted in a periodic sense); (iv) $\gamma = 0.001$, $c_2 = 2000$, $c_3 = 1$, in Figure 3.9 one can see how numerically the third-order like dispersion regularize the shock formation.

Figure 3.8 which presents a separation between two distinct outcomes was obtained by careful simulations in the (c_2, c_3) parameter space. For the region above the curve (Figure 3.8) the shock is not arrested, a numerical simulation can be seen in Figure 3.10 where the solution blows up; below the curve (Figure 3.8) dispersion regularizes the shock, a numerical simulation can be seen in Figure 3.11. From Figure 3.12, we can see how the shock has been regularized by the dispersion terms. The solution blows up for the top plot; broadening(no blow up) is observed for the middle plot(note that c_2 and c_3 are just on the curve 3.8); solitons are created at an earlier stage for the bottom plot than the middle one. It is important to point out the results shown strongly depend on the dispersion profile. Nonetheless it is intriguing to see from Figure 3.8 that a universal critical value $(\frac{c_2}{c_3})_C \approx 2.1$ emerges. While one could argue that by modifying c_3 one departs from a particular photonic structure, what matters is that for every value c_3 (that is moving vertically in Figure
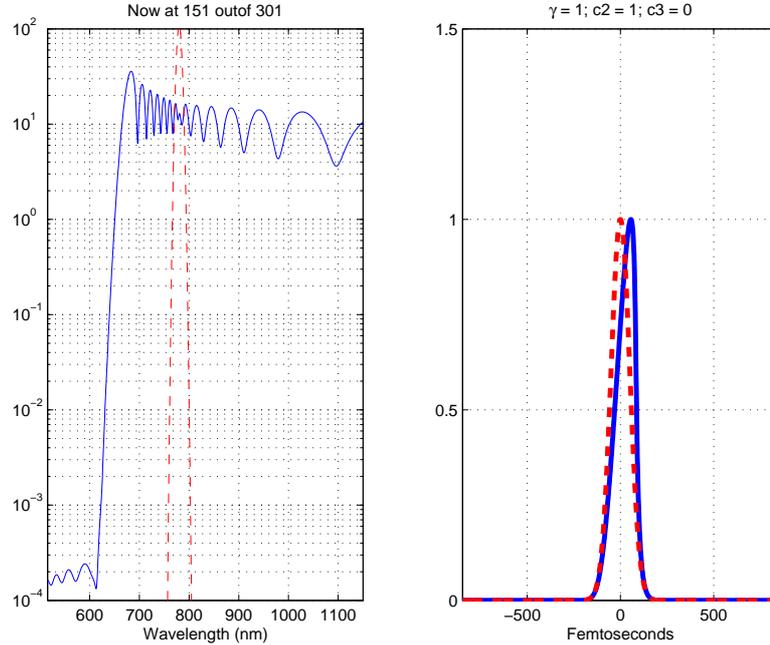
Figure 3.5: Spectral (left) and temporal (right) picture of the output after $0.15m$ propagation, with $\gamma = 1$, $c_2 = 1$, $c_3 = 0$. Other parameters remains the same as in the Figure 3.3. Dashed line stands for the original pulse.

3.8), this transition always occurs. At this time, we do not have an explanation for it. Furthermore, this property should be tested for different dispersion profiles.

To summarize, by performing a series of careful numerical simulations where we look at the relative strengths of the dispersion (measured by a parameter c_3 that multiplies $F(\lambda)$ in Eq. (3.1)) and of the self-steepening term (measured by a parameter c_2) we clearly demonstrate two dynamical regions: one where the singularity due to the shock is not suppressed by dispersion (the region above the curve). In this regime the spectral broadening does not saturate and the numerical solution blows up, clearly suggesting that additional physical mechanisms must be considered. In the second region(the region below the curve) propagation leading to supercontinuum generation. Although we did not show the curve beyond $c_3 = 0.8$, it should be
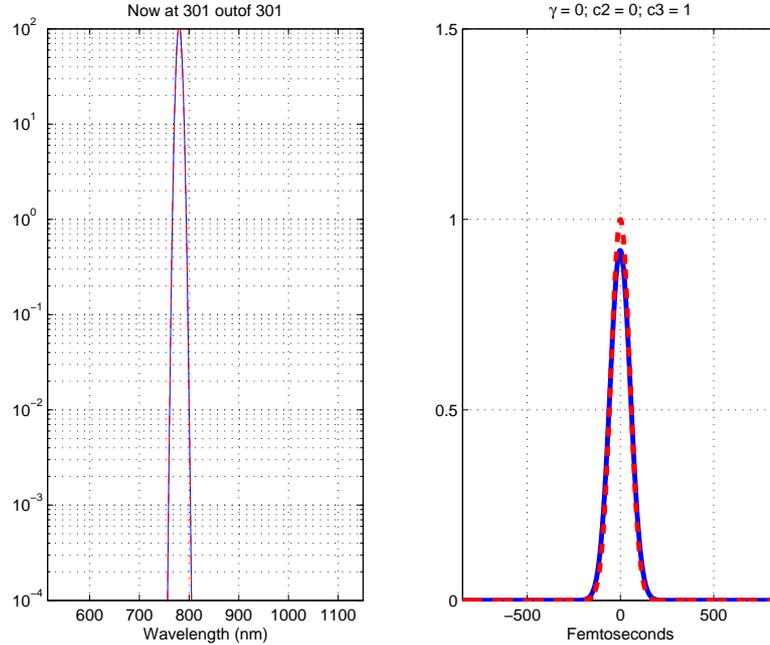
Figure 3.6: Spectral (left) and temporal (right) picture of the output after $0.3m$ propagation, with $\gamma = 0$, $c_2 = 0$, $c_3 = 1$. Other parameters remains the same as in the Figure 3.3.

clear that the point $c_2 = c_3 = 1$ corresponding to the experimental parameters in [57], is as expected below the curve.

3.3   Conclusions

Supercontinuum generation is a fascinating and important phenomenon observed in certain nonlinear wave systems. In this work, we discussed a simple model where we tuned dispersion and nonlinearity so that we could showcase different outcomes. In particular, we showed shock driven SC generation as well as soliton-fission driven SC. Next we moved to a model closely related to an existing photonic crystal fiber and showed both SC generation as well as critical shock formation. It is important
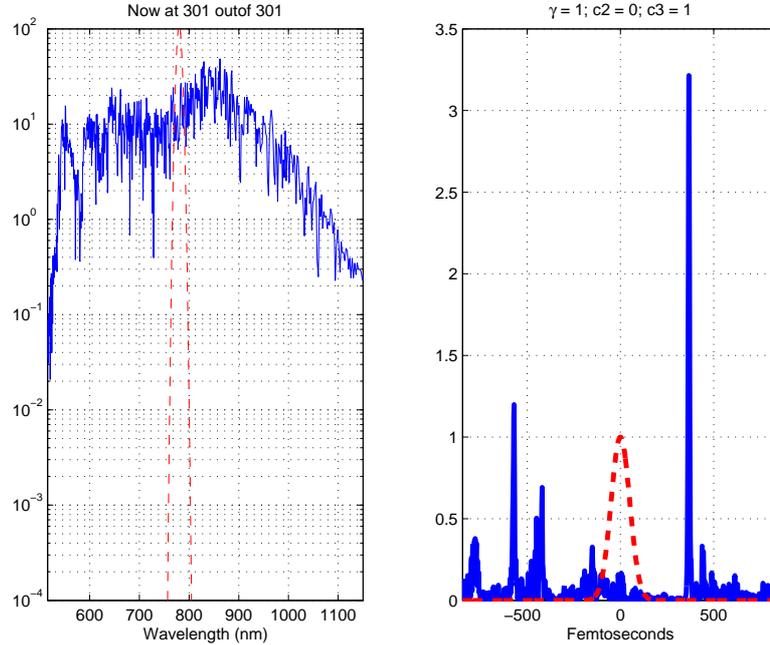
Figure 3.7: Spectral (left) and temporal (right) picture of the output after $0.3m$ propagation, with $\gamma = 1$, $c_2 = 0$, $c_3 = 1$. Other parameters remains the same as in the Figure 3.3.

to emphasize that by properly integrating the dispersive terms for a given photonic microstructured fiber, we capture supercontinuum generation as observed in experiments, likely to greater accuracy than the more common expansion to a finite order of the linear dispersion relation. Our numerical simulations illustrate that for some input conditions, shocks rather than soliton fission appear to be dominant and become the major source of spectral broadening. It is true that soliton fission as seen in many works could be the leading mechanism towards SC generation. Which effect is more dominant and what signatures (if any) of the spectral picture can explain the hierarchy of effects coming into the dynamics remains unclear. Overall, an accurate mathematical model is essential to explain the experimental outcomes. In particular we have a numerical approach at our disposal to study any photonic fiber structure
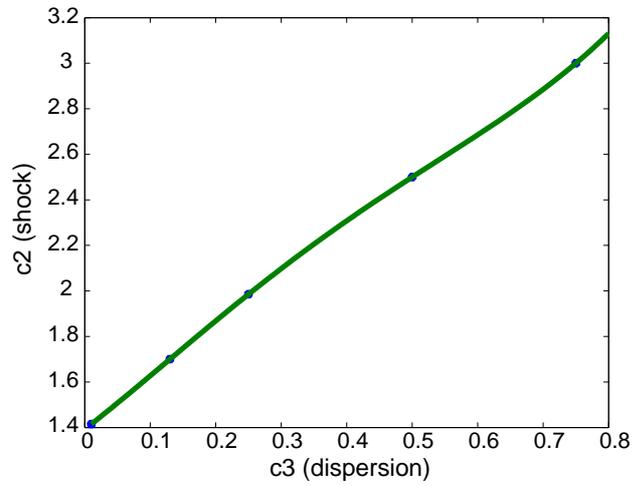
Figure 3.8: c_2 vs c_3 curve that separates regions where the numerical simulation blows up (region above the curve) from the region where supercontinuum is numerically observed (region below the curve).

for which GVD profiles have been or can be computed.

Figure 3.9: Spectral (left) and temporal (right) picture of the output after $0.14m$ (top) $1m$ (bottom) propagation, with $\gamma = 0.001$, $c_2 = 2000$, $c_3 = 1$. Other parameters remains the same as in the Figure 3.3.
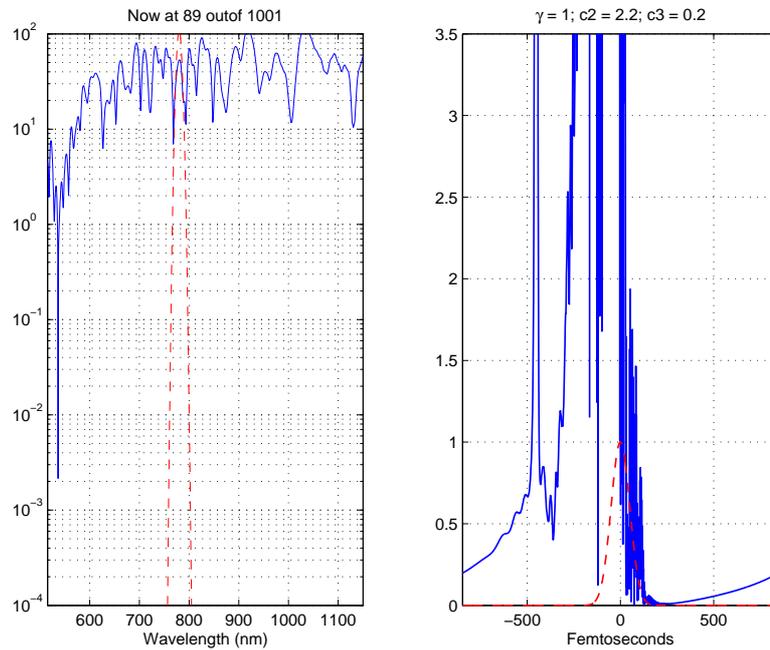
Figure 3.10: Spectral (left) and temporal (right) picture of the output after $8.9cm$ propagation, with $\gamma = 1$, $c_2 = 2.2$, $c_3 = 0.2$. Other parameters remains the same as in the Figure 3.3.
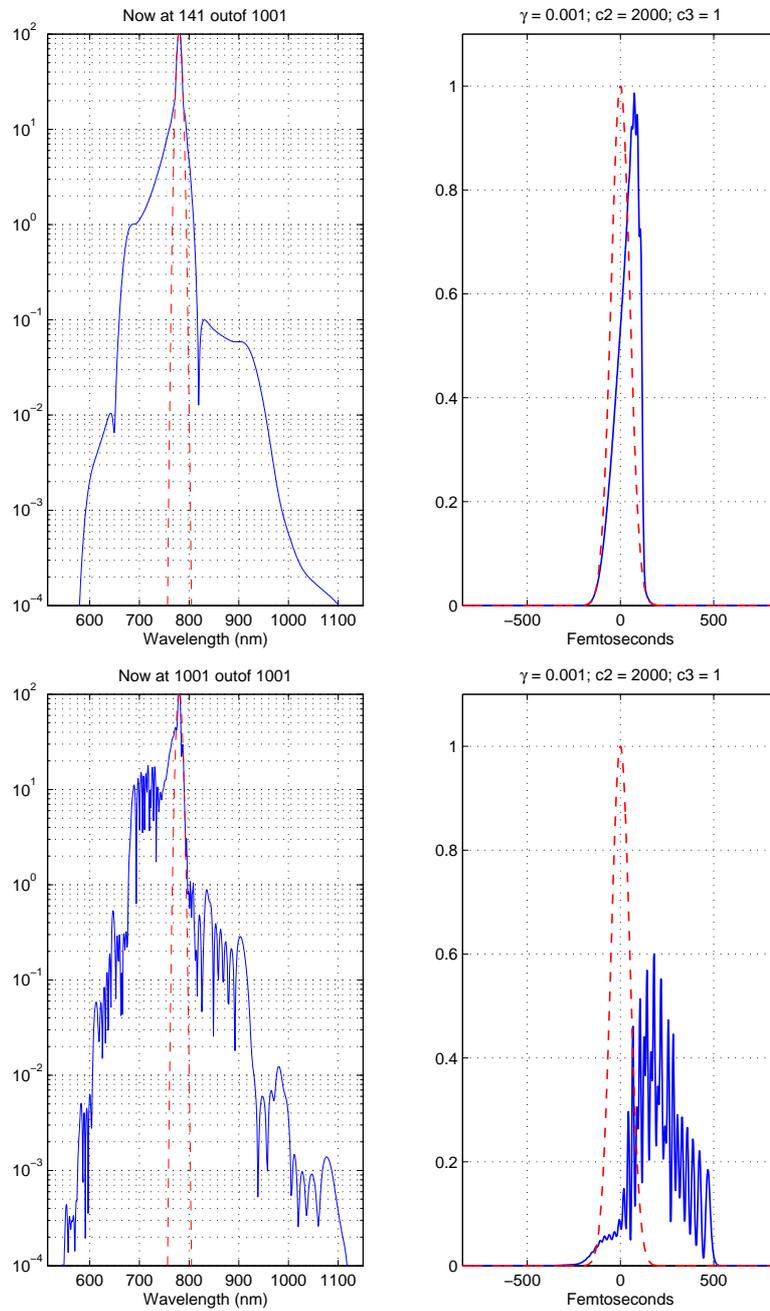

Figure 3.11: Spectral (left) and temporal (right) picture of the output after $1m$ propagation, with $\gamma = 1$, $c_2 = 1.4$, $c_3 = 0.2$. Other parameters remains the same as in the Figure 3.3.

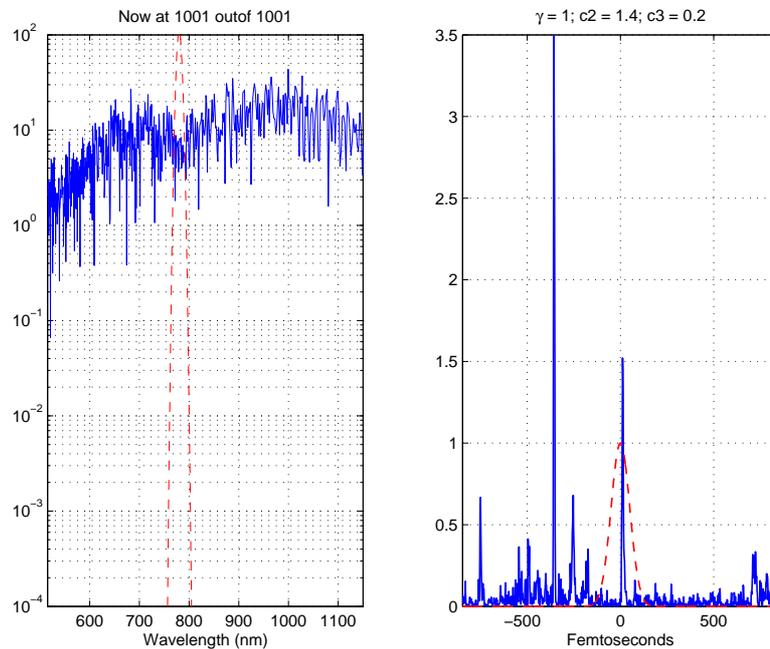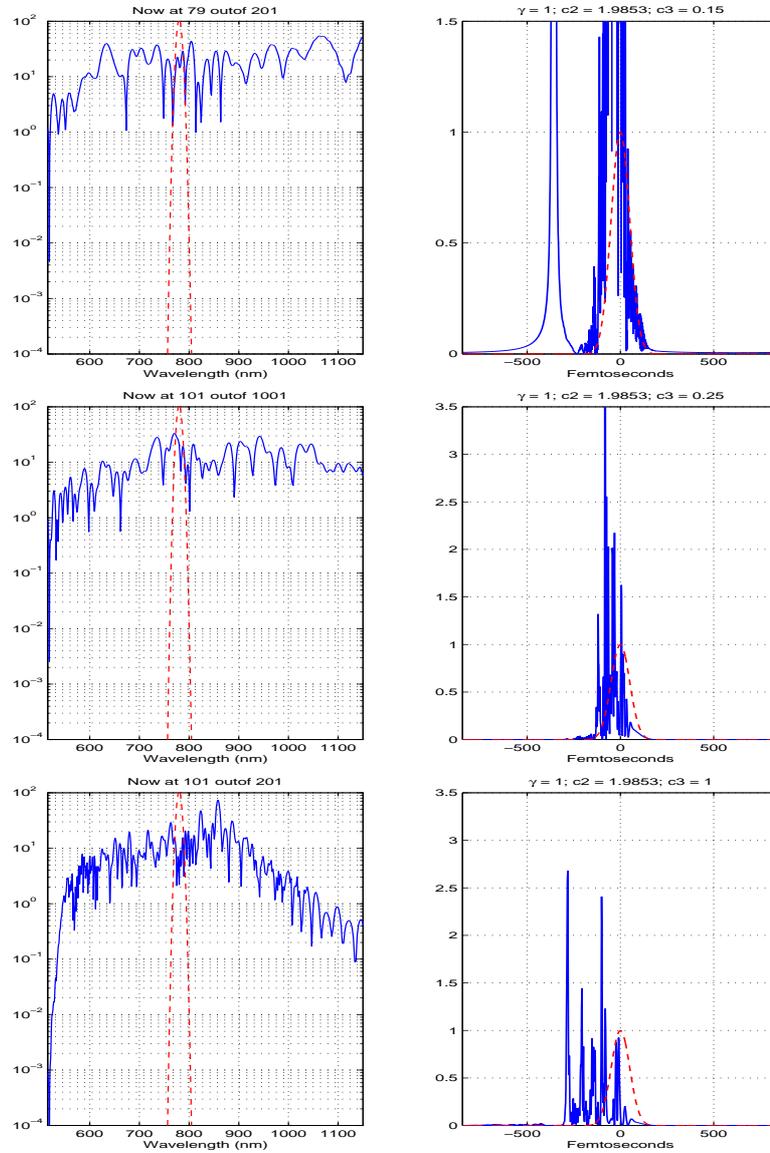Figure 3.12: Spectral (left) and temporal (right) picture of the output after $7.9cm$, $0.1m$, $0.1m$ propagation respectively from top to bottom, with $\gamma = 1$, $c_2 = 1.9853$, $c_3 = 0.15$, $c_3 = 0.25$, $c_3 = 1.0$. Other parameters remains the same as in the Figure 3.3.

References

[1] S Abarbanel, D Gottlieb, and MH Carpenter. On the removal of boundary errors caused by runge-kutta integration of nonlinear partial differential equations. *Siam Journal On Scientific Computing*, 1996:777–782, 1996.

[2] S.J. Adams. *Electromagnetic Theory*. Lightning Source Inc, 2008.

[3] P. Agrawal. Nonlinear fiber optics. 2001. 3rd edition, Academic Press.

[4] M. Ainsworth. Discrete dispersion relation for hp-version finite element approximation at high wave number. *SIAM J. Num. Anal.*, 42:553–575, 2004.

[5] M. Ainsworth. Dispersive and dissipative behavior of high-order discontinuous Galerkin finite element methods. *J. Comput. Phys.*, 198:106–130, 2004.

[6] Alfano. The supercontinuum laser source. 1989. New York, Springer-Verlag.

[7] R. R. Alfano and S. L. Shapiro. Emission in the region 4000 to 7000 via four-photon coupling in glass. *Physical Review Letters 24*, 1970.

[8] R. R. Alfano and S. L. Shapiro. Observation of selfphase modulation and small-scale filaments in crystals and glasses. *Physical Review Letters 24*, 1970.

[9] D. Appelö, T. Colonius, T. Hagstrom, and M. Inkman. Development of arbitrary-order hermite methods for simulation and analysis of turbulent jet noise. *Procedia Engineering*, IUTAM Symposium on Computational Aero-Acoustics for Aircraft Noise Prediction, University of Southampton, 29-31 March, 2010.

[10] D. Appelö and T. Hagstrom. On advection by Hermite methods. *Paciffic Journal Of Applied Mathematics*, 4(2), 2011.

References

[11] D. Appelö, M. Inkman, T. Hagstrom, and T. Colonius. Hermite methods for aeroacoustics: Recent progress. In *17th AIAA/CEAS Aeroacoustics Conference (32nd AIAA Aeroacoustics Conference), Portland, Oregon, June 5-8, 2011*, number AIAA-2011-2757, 2011.

[12] Norman Balabanian and Theodore A. Bickart. *Electrical Network Theory.* John Wiley and Sons, Inc, 1969.

[13] C. Baum. The singularity expansion method. In Leopold Felsen, editor, *Transient Electromagnetic Fields*, volume 10 of *Topics in Applied Physics*, pages 129–179. Springer Berlin / Heidelberg, 1976.

[14] C. Baum. The singularity expansion method: Background and developments. *Antennas and Propagation Society Newsletter, IEEE*, 28(4):14 –23, august 1986.

[15] C. E. Baum. Emerging technology for transient and broad-band analysis and synthesis of antennas and scatterers. *Proceedings of the IEEE*, 64(11):1598–1616, 1976.

[16] Carl E. Baum. On the singularity expansion method for the solution of electromagnetic interaction problems. *EMP Interaction Note*, 88, Dec 1971. Air Force Weapons Laboratory, AD A066905.

[17] Carl E. Baum. Idealized electric- and magnetic-field sensors based on spherical sheet impedances. *Sensor & Simulation Note (SSN) Air Force Weapons Laboratory*, 283, 1983.

[18] Carl E. Baum. Scattering at one-dimensional-transmission-line junctions. *EMP Interaction Note*, 603, 2008.

[19] C.E. Baum, E.J. Rothwell, K.-M. Chen, and D.P. Nyquist. The singularity expansion method and its application to target identification. *Proceedings of the IEEE*, 79(10):1481 –1492, oct 1991.

[20] G. Birkhoff, M. Schultz, and R. Varga. Piecewise hermite interpolation in one and two variables with applications to partial differential equations. *Numerische Mathematik*, 11:232–256, 1968. 10.1007/BF02161845.

[21] R. Brankin, I. Gladwell, and L. Shampine. RKSUITE: a suite of Runge-Kutta codes for the initial value problem for ODEs. Technical Report Softreport 91-1, Math. Dept., Southern Methodist University, Dallas, TX, 1991.

[22] C.Dodson. A high-order hermite compressible navier-stokes solver. Master's thesis, The University of New Mexico, 2003.

References

[23] L. Collatz. *The Numerical Treatment of Differential Equations.* Springer-Verlag, New York, 1960.

[24] Steven A. Orszag David Gottlieb. *Numerical analysis of spectral methods.* SIAM Philadelphia, 1977.

[25] P. Davis. *Interpolation and approximation.* Dover Publications, New York,, 1975.

[26] C.Lubich E.Hairer and M.Schlichte. Fast numerical solution of nonlinear Volterra convolutional equations. *SIAM J. Sci. Stat. Comput*, 6:532–541, 1985.

[27] Bengt Fornberg. On a fourier method for the integration of hyperbolic equations. *SIAM Journal on Numerical Analysis*, 12(4):pp. 509–528, 1975.

[28] Bengt Fornberg. *A practical guide to pseudospectral methods.* Cambridge University Press, 1996.

[29] R. L. Gardner and C. E. Baum. Expansion of a scalar, vector, or dyadic function in terms of the spherical vector wave functions. *Mathematics Notes*, 36, May 1974. Air Force Weapons Laboratory.

[30] J. Goodrich, T. Hagstrom, and J. Lorenz and. Hermite methods for hyperbolic initial-boundary value problems. *Math. Comp.*, 75(254):595–630, 2005.

[31] David Gottlieb and Eitan Tadmor. The CFL condition for spectral approximations to hyperbolic initial-boundary value problems. *Math. Comp.*, 56(194):565–588, 1991.

[32] A. Griewank. *Evaluating Derivatives: Principles and Techniques of Algorithmic Differentiation.* SIAM, Philadelphia, 2000, 2000.

[33] T. Hagstrom and D. Appelö. Experiments with hermite methods for simulating compressible flows: Runge-kutta time-stepping and absorbing layers. In *13th AIAA/CEAS Aeroacoustics Conference*, number 2007-3505, 2007.

[34] T. Hagstrom and D. Appelö. Experiments with hermite methods for simulating compressible flows: Runge-kutta time-stepping and absorbing layers. In *13th AIAA/CEAS Aeroacoustics Conference*, number 2007-3505, 2007.

[35] Thomas Hagstrom. *Radiation boundary conditions for the numerical simulation of waves.* Cambridge University Press, 1999.

[36] W. D. Henshaw. A high-order accurate parallel solver for Maxwell's equations on overlapping grids. *SIAM Journal on Scientific Computing*, 28(5):1730–1765, 2006.

References

[37] Jan S Hesthaven and Tim Warburton. *Nodal discontinuous Galerkin methods: algorithms, analysis, and applications*, volume 54. Springer, New York, 2008.

[38] JS Hesthaven and T Warburton. Nodal high-order methods on unstructured grids - i. time-domain solution of maxwell's equations. *Journal of Computational Physics*, 181(1):186–221, September 2002.

[39] David W. Kammler. *A First Course in Fourier Analysis*. Cambridge University Press, 2008.

[40] H.-O. KREISS and J. OLIGER. Comparison of accurate methods for the integration of hyperbolic equations. *Tellus*, 24:199–215, 1972.

[41] D.Pardo M.Paszynski W.Rachowicz L.Demkowicz, J.Kurtz and A.Zdunek. Computing with *hp*-adaptive finite elements. *Applied Mathematics & Nonlinear Science*, Chapman & Hall/CRC, Boca Raton, 2007.

[42] Frieder Lrcher, Gregor Gassner, and Claus-Dieter Munz. An explicit discontinuous galerkin scheme with local time-stepping for general unsteady diffusion equations. *Journal of Computational Physics*, 227(11):5649 – 5670, 2008.

[43] K Mattsson and J Nordstrom. Summation by parts operators for finite difference approximations of second derivatives. *Journal of Computational Physics*, 199:503–540, 2004.

[44] Irene A. Stegun EDS Milton Abramowitz. *Handbook of Mathematical Functions: with Formulas, Graphs, and Mathematical Tables*. Dover Publications, 1965.

[45] A. Biswas and A. B. Aceves *Dynamics of solitons in optical fibres* Journ. of Modern Optics **48**. 1135-1150. 2001.

[46] A. V. Husakou and Phys. Rev. Lett. **87** J. Herrmann, *Supercontinuum Generation of Higher-Order Solitons by Fission in Photonic Crystal Fibers*. 203901. 2001.

[47] D. Anderson and Phys. Rev. A **27** M. Lisak, *Nonlinear asymmetric self-phase modulation and self-steepening of pulses in long optical waveguides*. 1393–1398. 1983.

[48] M. R. Wehner M. Ross A. Efimov A. J.Taylor V. V. R. K. Kumar A. K. George J. C. Knight N. Y. Joly F. G. Omenetto, N. A. Wolchover and Opt. Exp. **14** P. S. J. Russell, *Spectrally smooth supercontinuum from 350 nm to 3 µm in sub-centimeter lengths of soft-glass photonic crystal fibers*. 4928–4934. 2006.

References

[49] S. Cohen G. Genty and JOSA B **24** J. M. Dudley, *Fiber supercontinuum sources (Invited)*. 1771–1785. 2007.

[50] G. Yang and Opt. Lett. **9** Y. R. Shen, *Spectral broadening of ultrashort pulses in a nonlinear medium*. 510–512. 1984.

[51] C. Chudoba R. K. Ghanta T. H. Ko J. G. Fujimoto J. K. Ranka I. Hartl, X. D. Li and Opt. Lett. **26** R. S. Windeler, *Ultrahigh-resolution optical coherence tomography using continuum generation in an air-silica microstructure optical fiber*. 608–610. 2001.

[52] D. M. Atkin J. C. Knight, T. A. Birks and in Proc. Opt. Fiber Commun. Conf. No. PD3 San Jose California Feb. P. S. J. Russell, *Pure silica single-mode fiber with hexagonal photonic crystal cladding*. 1996. 1996.

[53] J. Dudley and IEEE J. Sel. Top. Quantum Electron **8** S. Coen, *Numerical simulations and coherence properties of supercontinuum generation in photonic crystal and tapered optical fibers*. 651–659. 2002.

[54] G. Genty J. Dudley and Rev. of Mod. Phys. **78** S. Coen, *Supercontinuum generation in photonic crystal fiber*. 1135–1184. 2006.

[55] N. Zhavoronkov A. V. Husakou D. Nickel J. C. Knight W. J. Wadsworth P. S. J. Russell J. Herrmann, U. Griebner and Phys. Rev. Lett. **88** G. Korn, *Experimental evidence for supercontinuum generation by fission of higher-order solitons in photonic fibers*. 173901. 2002.

[56] L. Xu M. Kimmel E. Zeek P. O'Shea R. Trebino S. Coen R. S. Windeler *Cross-correlation frequency resolved optical gating analysis of broadband continuum generation in photonic crystal fiber: simulations and experiments* Opt. Exp. **10** J. M. Dudley, X. Gu. 1215–1221. 2002.

[57] N. Grossard H. Maillotte R. S. Windeler B. J. Eggleton J. M. Dudley, L. Provino and JOSA B **19** S. Coen, *Supercontinuum generation in air-silica microstructured fibers with nanosecond and femtosecond pulse pumping*. 765–771. 2002.

[58] P. Stoller K. Lindfors, T. Kalkbrenner and Phys. Rev. Lett. **93** V. Sandoghdar, *Detection and spectroscopy of gold nanoparticles using supercontinuum white light confocal microscopy*. 037401. 2004.

[59] F. G. Omenetto K. Shi and Opt. Exp. **14** Z. Liu, *Supercontinuum generation in an imaging fiber taper*. 12359–12364. 2006.

[60] S. Yin K. Shi, P. Li and Opt. Exp. **12** Z. Liu, *Chromatic confocal microscopy using supercontinuum light*. 2096–2101. 2004.

References

[61] E. M. Wright M. Kolesik and Opt. Lett. **342** J. V. Moloney, *Supercontinuum and third-harmonic generation accompanying optical filamentation as first order scattering processes.* 2816–2818. 2007.

[62] Y. C. Lee P. K. A. Wai, C. R. Menyuk and Opt. Lett. **11** H. H. Chen, *Nonlinear pulse propagation in the neighborhood of the zero-dispersion wavelength of monomode optical fibers.* 464–466. 1986.

[63] Y. C. Lee P. K. A. Wai, C. R. Menyuk and Opt. Lett. **11** H. H. Chen, *Nonlinear pulse propagation in the neighborhood of the zero-dispersion wavelength of monomode optical fibers.* g. 1986. B. Whitham, "Linear and nonlinear waves," New York, Wiley Ed., 1974.

[64] W. J. Wadsworth T. A. Birks and Opt. Lett. **25** P. S. J. Russell, *Supercontinuum generation in tapered fibers.* 1415-1417. 2000.

[65] R. Holzwarth T. Udem and Nature **416** T. W. Hänsch, *Optical frequency metrology.* 233–237. 2002.

[66] V. E. Zakharov and Sov. Physics JETP **34** A. B. Shabat, *Exact theory of two dimensional self-focusing and one dimensional self-modulation of waves in nonlinear media.* 62–69. 1972.

[67] W. H. Reeves J. C. Knight P. S. J. Russell F. G. Omenetto V. V. R. K. Kumar, A. K. George and Opt. Exp. **10** A. J.Taylor, *Extruded soft glass photonic crystal fiber for ultrabroad supercontinuum generation.* 1520–1525. 2002.

[68] J. C. Knight T. A. Birks T-P. M. Man W. J. Wadsworth, A. Ortigosa-Blanch, JOSA B Special issue on nonlinear optics in photonic crystals P. S. J. Russell, *Supercontinuum generation in photonic crystal fibers and optical fiber tapers: a novel light source*, and **19** fibres. 2148–2155. 2002.

[69] PO Persson and G Strang. A simple mesh generator in matlab. *Siam Review*, 46(2):329–345, June 2004.

[70] T Rylander and A Bondeson. Stable fem-fdtd hybrid method for maxwell's equations. *Computer Physics Communications*, 125(1-3):75–82, March 2000.

[71] D. appelö T. Hagstrom and C. Y. Jang. Hermite methods for hyperbolic-parabolic system. In *extended abstract Waves, Vancouver, Canada*, 2011.

[72] T. Warburton and T. Hagstrom. Taming the cfl number for discontinuous galerkin methods on structured meshes. *SIAM J. on Numerical Analysis / Volume 46 / Issue 6*, 46:3151–3180, 2008.

References

[73] J. A. C. Weideman and L. N. Trefethen. The eigenvalues of second-order spectral differentiation matrices. *SIAM Journal on Numerical Analysis*, 25(6):pp. 1279–1298, 1988.