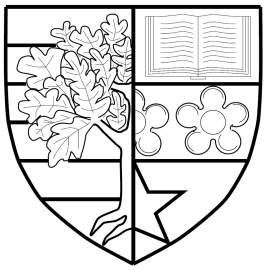
ENRICHED DISCRETE SPACES FOR TIME DOMAIN WAVE EQUATIONS

by

Oluwaseun Francis Lijoka



Submitted for the degree of
Doctor of Philosophy

Department of Mathematics
School of Mathematical and Computer Sciences
Heriot-Watt University

13th January 2017

Abstract

The second order linear wave equation is simple in representation but its numerical approximation is challenging, especially when the system contains waves of high frequencies. While 10 grid points per wavelength is regarded as the rule of thumb to achieve tolerable approximation with the standard numerical approach, high resolution or high grid density is often required at high frequency which is often computationally demanding.

As a contribution to tackling this problem, we consider in this thesis the discretization of the problem in the framework of the space-time discontinuous Galerkin (DG) method while investigating the solution in a finite dimensional space whose building blocks are waves themselves. The motivation for this approach is to reduce the number of degrees of freedom per wavelength as well as to introduce some analytical features of the problem into its numerical approximation.

The developed space-time DG method is able to accommodate any polynomial bases. However, the Trefftz based space-time method proves to be efficient even for a system operating at high frequency. Comparison with polynomial spaces of total degree shows that equivalent orders of convergence are obtainable with fewer degrees of freedom. Moreover, the implementation of the Trefftz based method is cheaper as integration is restricted to the space-time mesh skeleton.

We also extend our technique to a more complicated wave problem called the telegraph equation or the damped wave equation. The construction of the Trefftz space for this problem is not trivial. However, the flexibility of the DG method enables us to use a special technique of propagating polynomial initial data using a wave-like solution (analytical) formula which gives us the required wave-like local solutions for the construction of the space.

This thesis contains important a priori analysis as well as the convergence analysis for the developed space-time method, and extensive numerical experiments.

Acknowledgements

I would like to acknowledge the contribution of individual that has made this work a success in one way or the other. The greatest share of my gratitude goes to my supervisor Professor Lehel Banjai for being there always as a constant source of support, idea, courage and encouragement. I also appreciate him for his patience and for being a good teacher whom every student would dream to have. I can't thank him enough for teaching me how to think outside the box, analyse a problem and eventually provide a solution to the problem. I appreciate you a lot !.

My next big appreciation goes to Dr. Emmanuil Georgoulis for his support and his contribution to the success of this work. I appreciate him for being part of the brain stormy process as well as his contribution in the analysis segment.

Next, I would like to appreciate Ian McCrone for always being there whenever I found trouble with my computer. I appreciate him for his time, support and help. I must say he is unique and I will never forget him. I also appreciate Professor Dugald Duncan, Dr. Heiko Gimperlein, Dr. Markus Schmuck, and Dr. Lyonel Boulton for their helpful and fruitful conversation during the course of the program. My "Thanks" also goes to Dr. Anastasia Doikou for reading my first draft and for her corrections. I will never forget to appreciate late Professor Jack Carr for his encouragement, friendship, and his support since I met him and before he passed away.

I am also grateful to Christine McBride for her support and her constant supply of milk in the refrigerator. Thanks for saving not only me from midnight hunger. I extend my thanks to Morag Jones for her kind gestures and for making my conference trips easy and splendid.

I gratuitously acknowledged the presence of some individuals that add fun to my life during the program. In particular, I would like to appreciate Hola Adrakey for his assistance in LaTeX and for being funny even at difficult times, Akin Ojagbemi, Ihechukwu Chinyere, Ariana Bianchi, Getachew Alemu, Benard Bainson, Dina Abla, Doaa El-Sakaout, Laila Gandour for their support and friendship at every moment, Hovey Assionvi not only for organising football games, Nadia Solime for being patience and for her excellent soup, Kehinde Olowookere for being a good flatmate, Gwendolyn Barnes and Hannah Jones for their help in English, David Bolea for fruitful chat on MATLAB and numerical simulations, Temidayo Ayoola for filling my life with happiness and even more, all the people I didn't mention here. Finally, I appreciate God Almighty who is the source of all wisdom.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Time-dependent wave models such as the acoustic and the elastic wave equations are very important in the scientific and engineering fields. Many industries such as the mining industry, the aviation industry, and other engineering industries have found these models indispensable in their applications. Areas of applications include medical ultrasonic, seismology, electromagnetism and non-destructive testing.

Despite their numerous applications, it has remained challenging to develop efficient numerical methods capable of approximating and predicting the propagation as well as the behavior of these models, especially when the system contains waves of high frequency (equivalently short wavelength). This problem is due to the oscillatory nature of the solutions of the models, and standard discrete spaces with standard numerical methods only approximate the solutions for exceedingly fine mesh which often results in high computational cost.

A similar challenge is encountered in the time-harmonic wave problems as approximations of the Helmholtz equation with large wave number (see [66] and [132]).

The development of better and improved numerical schemes capable of representing the oscillatory phenomena with reduced computational time has been an important and active area of research in the field of computational acoustics, and we are motivated to show in this thesis the development of such a method supported with mathematical analysis by investigating special discrete spaces whose elements are local solutions of the transient wave equations.

We focus on developing and analyzing a method in the framework of finite ele-

ment methods which utilizes the wave-like functions of the targeted partial differential equation (PDE) in its approximation space. This new method is in the class of the Trefftz methods for wave equations in the time-domain. The space-time quasi-optimality result proven in this thesis does not depend on the use of Gronwall's inequality in contrast to the standard approach for time-dependent a priori bounds.

In order to reduce computational complexity and to increase the speed of approximation, the proposed Trefftz spaces give us the advantage of reducing the number of degrees of freedom per wavelength and also the advantage of restricting integration only to the space-time skeleton of the mesh. This idea has been proven successful in the time-harmonic regime and we show in this thesis that it is equally successful also in the time-domain.

In the rest of this chapter, we briefly describe the initial-boundary value problems (IBVPs) that will be considered in this thesis. In Section 1.2, we discuss the standard discretization method. We conclude the chapter with an outline of the thesis.

1.1 Wave problems in time domain

We are interested in two cases: The general acoustic wave equation and the damped wave equation which is also known as the telegraph equation. We consider the relevant IBVPs with different classical boundary conditions in the bounded spatial domain Ω subset of $\mathbb{R}^d$, $d = \{1, 2, 3\}$. We also present variational or weak formulations for each case as solutions are sought for in the subspace of Sobolev space $H^1(\Omega)$.

Although the proposed method should be applicable to other wave equations in the time-domain such as Maxwell's equations which govern the realm of electromagnetism and the Schrödinger equation which describes the evolution of quantum particles, we do not consider them in this thesis. We refer readers to [16, 63, 97] and [42] for extensive description, applications and analytical motivations.

1.1.1 The acoustic wave equation

The propagation of acoustic waves in heterogeneous isotropic media with small amplitude can be represented by the wave equation

$$\ddot{u} - \nabla \cdot (a \nabla u) = 0, \tag{1.1.1}$$

which is often coupled with initial data

$$u(x,0) = u_0 \quad \dot{u}(x,0) = v_0. \tag{1.1.2}$$

For a derivation, see Chapter 3 of [16]. The solution u is the perturbed pressure which depends on the position vector x and on time variable t; $a \equiv a(x)$ could be a matrix or a scalar function of x while ∇ and $\nabla\cdot$ denote the usual gradient and divergence operators. If $a \equiv 1$, then $\nabla \cdot \nabla = \Delta$ which is the usual Laplace operator. This model can be used to represent mechanical vibrations with small amplitude. In this case the scalar solution u could represent the vertical displacement of an elastic membrane in 2 spatial dimensions (2d) (or a string in 1 spatial dimension (1d)).

The acoustic wave model is usually equipped with different boundary conditions (BCs). If the value of u is given on the boundary, we have a Dirichlet boundary condition; if the normal derivative $\mathbf{n}\cdot\nabla u$ is prescribed ($\mathbf{n}$ is the outward unit normal to the boundary $\partial\Omega$), we have a Neumann boundary condition. There is also an impedance boundary condition $\dot{u}t + \gamma \frac{\partial u}{\partial n} = 0$ which models semi-reflecting boundaries and often appears when approximating problems of unbounded domains by absorbing boundary conditions. The combination of the Dirichlet and the Neumann data is called the Robin boundary condition. Mixture of these boundary conditions on different parts of the domain is also possible.

Another boundary condition (may be regarded as interface condition) different from the above classical BCs is the so called transmission condition that is usually prescribed at the interface of two subdomains of Ω. This condition is important when dealing with wave problems consisting of piecewise homogeneous subdomains or materials, such as piecewise change in speed or density at different parts of Ω (see

Chapter 2 of [42]). For instance let Ω_1 and Ω_2 be two subdomains of Ω and let Γ_{12} be the boundary between them. Let $u_1 = u|_{\Omega_1}$ and $u_2 = u|_{\Omega_2}$ with $a \equiv a_1$ in Ω_1 and $a \equiv a_2$ in Ω_2 be the restricted values of u in Ω_1 and Ω_2 respectively where a_1 and a_2 are constant in the respective subdomains, then we have the following transmission conditions defined on the interface

$$u_j = u_k, \quad a_j \partial_{\mathbf{n}} u_j = a_k \partial_{\mathbf{n}} u_k \quad \text{on } \Gamma_{jk}, \tag{1.1.3}$$

where $\mathbf{n}$ is the exterior normal to Ω_j (or Ω_k).

Before we present different weak formulations for the acoustic wave equation, we give some notation. Let u be a scalar function of $\mathbb{R}^d$, $d = \{1, 2, 3\}$, we define the general partial differential operator

$$D^\alpha = \frac{\partial^p}{\partial x_1^{\alpha_1} \dots \partial x_d^{\alpha_d}}, \tag{1.1.4}$$

where $\alpha = \{(\alpha_1 \dots \alpha_d) \in \mathbb{N}^d : |\alpha| = \sum_{j=1}^d \alpha_j = p, p \in \mathbb{N}\}$. We denote by $H^m(\Omega)$ the Sobolev space [37]

$$H^m(\Omega) = \left\{ u \in L^2(\Omega) : D^\alpha u \in L^2(\Omega), \forall \alpha, |\alpha| \le m \right\}, \tag{1.1.5}$$

where, when $m = 1$, we have

$$H^1(\Omega) = \left\{ u \in L^2(\Omega) : \frac{\partial u}{\partial x_j} \in L^2(\Omega) \ \forall j = 1, \dots, d \right\}, \tag{1.1.6}$$

and $L^2(\Omega)$ represents the space of square integrable functions over Ω. The subspace of functions in $H^1(\Omega)$ whose traces vanish at the boundary $\partial\Omega$ is given by

$$H_0^1(\Omega) = \left\{ u \in H^1(\Omega) : u = 0 \text{ on } \partial\Omega \right\}. \tag{1.1.7}$$

Now if we set up the homogeneous wave equation with zero Dirichlet boundary condition, with the assumption that $u_0 \in H_0^1(\Omega)$ and $v_0 \in L^2(\Omega)$, then the following

variational formulation defined holds: Find $u(\cdot,t) \in H_0^1(\Omega)$ such that for $t \in \mathbb{R}^+$

$$\int_\Omega \ddot{u}v \,\mathrm{d}\mathbf{x} + \int_\Omega a\nabla u \cdot \nabla v \,\mathrm{d}\mathbf{x} = 0 \quad \forall v \in H_0^1(\Omega). \tag{1.1.8}$$

If the boundary condition is the non-zero Neumann condition $\frac{\partial u}{\partial \mathbf{n}} = g(x,t)$ with $u_0 \in H^1(\Omega)$ and $v_0 \in L^2(\Omega)$, then we have the variational formulation: Find $u(\cdot,t) \in H^1(\Omega)$ such that for $t \in \mathbb{R}^+$

$$\int_\Omega \ddot{u}v \,\mathrm{d}\mathbf{x} + \int_\Omega a\nabla u \cdot \nabla v \,\mathrm{d}\mathbf{x} = \int_{\partial\Omega} ag\, v \,\mathrm{d}\mathbf{s} \quad \forall v \in H^1(\Omega). \tag{1.1.9}$$

Finally if the impedance boundary condition is prescribed on the boundary with $u_0 \in H^1(\Omega)$ and $v_0 \in L^2(\Omega)$, then we have the variational formulation: Find $u(\cdot,t) \in H^1(\Omega)$ such that for $t \in \mathbb{R}^+$

$$\int_\Omega \ddot{u}v \,\mathrm{d}\mathbf{x} + \int_\Omega a\nabla u \cdot \nabla v \,\mathrm{d}\mathbf{x} = -\int_{\partial\Omega} \frac{1}{\gamma} a\dot{u}\, v \,\mathrm{d}\mathbf{s} \quad \forall v \in H^1(\Omega). \tag{1.1.10}$$

All the variational formulations defined above must be coupled with the initial data (1.1.2). If $u_0 \in H^1(\Omega)$ and $v_0 \in L^2(\Omega)$, we can prove the existence and uniqueness of (weak) solution [1] $u(x,t) \in L^2([0,T]; H^1(\Omega))$ for $t \in [0,T], T \leq \infty$ with $\dot{u}(x,t) \in L^2([0,T]; L^2(\Omega))$ and [2] $\ddot{u}(x,t) \in L^2([0,T]; H^{-1}(\Omega)$; see, Chapter 3, Section 8 of [88].

1.1.2 The telegraph equation

The telegraph equation gives the representation of the modified form of the wave equation when the effect of a dissipative force such as friction is non negligible. The equation appears with different names in different areas of applications; for example, it is called the telegraph equation in electricity (transmission line equation) (see, Section 1 – 6 of [63] and Chapter 7 of [42]), heat wave equation in heat conduction problems (see, survey paper [81]) and damped wave equation in mechanical wave

[1] The space $L^p([0,T];V)$ with $1 \leq p \leq \infty$ is referred to as Bochner spaces, with V being a Banach space with norm $\|\cdot\|_V$ [88]

[2] $H^{-1}(\Omega)$ is referred to as the dual of $H^1(\Omega)$ [85].

propagation (see, Chapter 2 of [42]). The equation is given by

$$\ddot{u} - \nabla \cdot (a \nabla u) + \alpha \dot{u} = 0, \tag{1.1.11}$$

where the parameter α denotes the damping constant and the unknown scalar solution u could be the amplitude of a damped wave or propagated temperature in a heat conduction problem.

Similarly, if we set up the damped wave problem with zero Dirichlet boundary condition, we can define the following variational form: Find $u(\cdot, t) \in H_0^1(\Omega)$ such that for $t \in \mathbb{R}^+$

$$\int_\Omega \ddot{u} v + \int_\Omega a \nabla u \cdot \nabla v + \int_\Omega \alpha \dot{u} v = 0 \quad \forall v \in H_0^1(\Omega), \tag{1.1.12}$$

and if the boundary condition is of Neumann type, e.g., $\frac{\partial u}{\partial \mathbf{n}} = g(x, t)$, then we have the weak formulation to be: Find $u(\cdot, t) \in H^1(\Omega)$ such that for $t \in \mathbb{R}^+$

$$\int_\Omega \ddot{u} v \, \mathrm{d}\mathbf{x} + \int_\Omega a \nabla u \cdot \nabla v \, \mathrm{d}\mathbf{x} + \int_\Omega \alpha \dot{u} v = \int_{\partial\Omega} a g \, v \, \mathrm{d}\mathbf{s} \quad \forall v \in H^1(\Omega). \tag{1.1.13}$$

Finally, for impedance boundary condition, the weak formulation is given to be: Find $u(\cdot, t) \in H^1(\Omega)$ such that for $t \in \mathbb{R}^+$

$$\int_\Omega \ddot{u} v \, \mathrm{d}\mathbf{x} + \int_\Omega a \nabla u \cdot \nabla v \, \mathrm{d}\mathbf{x} + \int_\Omega \alpha \dot{u} v = - \int_{\partial\Omega} \frac{1}{\gamma} a \dot{u} \, v \, \mathrm{d}\mathbf{s} \quad \forall v \in H^1(\Omega). \tag{1.1.14}$$

The variational formulations above must be presented with the initial data also with the same assumption as in the undamped wave problem.

1.2 Standard discretizations of time dependent wave equations

The classical methods that have been employed for the spatial discretization of the wave equation are the finite element method (FEM) and the finite difference method (FDM). These methods are often coupled with finite difference time stepping

schemes for the full discretization of wave problems. FDM and FEM are regarded as the standard numerical methods in the scientific community. These classical methods have their advantages as well as limitations but ways have been derived to circumvent their different limitations.

The FDM is regarded as the oldest numerical method for differential equations in history (see [69] [105], [37] and [79]). The method is well known and appreciated for its simplicity and efficiency in approximating wave problems on a uniform mesh. Spatial discretization with FDM often leads to semi-discrete formulations which offer flexibility in choosing any time discretization scheme [69].

Despite the attractive advantages offered by the method, its efficiency is limited when employed to problems on complex geometries or grids with non-conforming boundaries [79]. Methods such as mesh adaption technique [53] and local mesh refinements (see, [40], [38], [39], [90] and [89]) may be necessary to circumvent this issue. Furthermore, stability bottleneck due to CFL restrictions for explicit finite difference schemes usually requires the use of complex techniques such as local time stepping methods [130] to circumvent the problem. Another drawback to the use of FDM for wave problems is the requirement of high-density grid/extremely fine mesh (see [1] and [93]) for a system with high frequency content which can be computationally intolerable. Despite the disadvantages, the method is still well appreciated, applied and studied.

A more flexible spatial discretization method that handles complex geometries effectively is the FEM (see [79], [69] and [93]). The method has been investigated, extended and applied successfully in many areas of applications including wave problem simulations ([74], [37]). Other advantages offered by the method include high order accuracy, possibility of hp-adaptivity and freedom to choose any time discretization scheme for complete discretization [69]. However, the explicit form of the method for wave problems may require a local time stepping or an implicit time stepping technique to circumvent stability issue [60]. Furthermore, the requirement of exceedingly fine mesh [93] or high order methods in both space and time [37] for accurate solution at high frequency can be computationally expensive.

Despite the above limitations, the method is still considered as the most effective classical method for full wave simulations.

1.3 Aims and outline of thesis

In this thesis, we develop and analyse a family of space-time discontinuous Galerkin (DG) methods that can utilize special Trefftz spaces for the solutions of second order wave equations. The main purpose for this is to search for approximations in finite dimensional spaces whose building blocks are waves themselves.

The use of Trefftz spaces offer vital advantages such as systematic discretization in space and time for easy implementation, inclusion of analytical features of the wave equation in the approximation spaces and reduction in computational complexity during implementation.

We organise the thesis as follows. In Chapter 2, we review DG methods as well as Trefftz methods for the wave equation. We also give an overview of space-time methods for the wave equation. We conclude the chapter with reviews of special inequalities that will be needed for analysis in the dissertation. In Chapter 3, the Trefftz space-time DG method is constructed for the acoustic wave equation. The chapter also includes analysis such as the existence of solution in the Trefftz space, the rates of convergence and a best approximation result. We end the chapter with a brief extension of the idea to a wave problem with transparent boundary condition.

In Chapter 4, we extend the Trefftz space-time technique to a damped wave equation. The construction of the Trefftz space with good approximation properties is the major non-trivial work in the chapter. We also give relevant analysis similar to Chapter 3 but rates of convergence are proved only in one spatial dimension. In Chapter 5, we show the performance of the method through a series of numerical experiments. We verify the rates of convergence numerically and we compare the performance of the Trefftz spaces with the polynomial spaces of total degree which has more degrees of freedom per wavelength.

Chapter 2

Literature review and background knowledge

In this chapter we review the discontinuous Galerkin (DG) method; its history and its extension to second order wave problems are briefly discussed. We also present an overview of the Trefftz space method whose emergence in recent years has been a new direction for the development of efficient methods in the time-domain acoustic problems. We shall review its application and its extension in time-harmonic regime as well as in time-domain (transient domain). In Section 2.4, we extend our reviews to the space-time techniques for the wave equation in order to put the developed method in perspective.

At the later part of this chapter, we present important inequalities, error bounds and approximation properties in finite dimensional spaces which are background information for the analysis of the developed method. Finally, we conclude the chapter with a brief note on the software employed in the thesis.

2.1 Review of the DG method

Discontinuous Galerkin (DG) methods are classified as special finite element methods that allow the use of discontinuous functions in their test and trial spaces for the approximation of partial differential equations [47]. This class of methods dates back to 1973 when it was first introduced and applied by Reed and Hill to approximate a

linear hyperbolic PDE called the neutron transport equation [108]. However, it was not until the late 80s that the scientists in the research community started to exploit the attractive properties and approximating power possessed by this class of methods. The advantages include flexibility in changing the degrees of polynomials even locally on some elements in the mesh, high parallel efficiency, high order accuracy, and flexibility in approximating on complex geometry. The DG methods have been found successful in many areas such as electromagnetism [31, 36, 62], meteorology and weather forecasting [35, 56], fluid flow [13, 29, 68] and acoustics [5, 119, 61].

In the early 70s, Aubin Nitsche [101, 102] developed a method similar to the DG method where discontinuous functions were used in approximating an elliptic problems. He introduced a penalty term to enforce continuity in his method. Later in 1973, Babuška [7] merged Nitsche's approach with the basic idea of the DG method into the finite element framework and this evolved to be the origin of different interior penalty DG (IPDG) methods. In the late 70s, the foundation of IPDG method was established by Baker, Wheeler, Arnold, Delves, and Hall [128, 2, 46, 11, 4, 3]. Extension to other versions such as non-symmetric interior penalty DG (NIPG) method and incomplete interior penalty DG (IIPG) was due to the works of Reviere, Oden, Baumann, Babuška, Wheeler, Girault, Dawson, and Sun in the 90s [111, 45, 8, 110, 47].

DG methods have experienced a tremendous development over the years (more than a decade), we shall therefore summarise the development and the extensions of this class of methods with relevant literatures. The first analysis for the DG method was carried out by Lesaint and Raviart in 1974 [87] while the error estimate was improved by Johnson, Nävert, and Pitkäranta in 1986 [80]. Extension to time-dependent hyperbolic PDEs was carried out by Chavent and Cockburn in 1989 [24] and the order of convergence was improved by Cockburn and Shu using the explicit Runge-Kunta time discretization scheme [33]. The extension of the DG methods to hyperbolic conservation laws by Cockburn et al in the series of papers [32, 33, 28, 34] was seen as a great breakthrough in the development of this class of methods where a framework to approximate nonlinear time-dependent hyperbolic problems was

established. We refer the reader to survey and review papers [30, 3, 115, 4] for more on the history and the development of DG methods.

In the context of second order wave problems, Wang and his co-workers in [125] classify the existing variants of the methods into three categories which are, space-only DG methods, time only DG methods and space-time DG methods. Space only DG formulation follows the usual (spatial) semi-discretization but with the DG approach and suitable time-stepping schemes are then employed for its time discretization. Reformulation to a system of first order has been studied in [100, 25]. If an explicit time stepping scheme such as the leapfrog method is employed for the time discretization of the semi-discretrized method, then the derived complete explicit scheme will be under the influence of the CFL restriction [43] and special techniques such as local time stepping methods or implicit time stepping methods maybe required to overcome stability restriction, see, [48], [60] and [49].

The time only DG methods on the other hand involve the use of standard finite element shape functions which are strongly continuous across elements of a single time step but discontinuous across successive time steps [125]. This class of DG methods offers stability advantages as well as the allowance of high-order integration schemes in contrast to the explicit space only DG methods. Concrete examples can be found in [72] for elastodynamics and in [117] for acoustic wave problems .

The third category is the most recent class of DG methods for wave problems in the time-domain. Methods under this class partition the (space-time) domain systematically into space-time elements and basis functions that are continuous within the space-time element but discontinuous across element boundaries in both space and time are employed in the approximation space. This approach often leads to implicit space-time methods, however, careful adjustment in the space-time mesh partitioning may lead to a semi-explicit space-time method [96]. The first space-time DG method that utilizes wave-like basis functions has been studied in [106, 125] and the extension of the idea to Maxwell's equations has been studied in [83, 50].

2.2 Overview of the Trefftz space method

The Trefftz method is a class of numerical methods that involves the use of functions that satisfy the targeted partial differential equation. In the framework of finite elements, this method utilizes functions that satisfy the targeted PDE locally in its approximation space. The origin of this method can be traced back to 1926 when it was first employed by a German mathematician named Erich Trefftz [122] for the Laplace equation. Since then different versions of the method have been developed, analysed, and applied to a different range of PDEs by scientists and engineers [131]. Recently, the attractive advantages and flexibility of the method have attracted scientists in computational mathematics to study the method and different variants are being proposed for different PDEs. The attractive advantages include but are not limited to fewer degrees of freedom to obtain the same accuracy compared to standard polynomial space and flexibility in incorporating the analytical properties of the targeted PDE.

The use of the Trefftz method for the approximation of wave problems have been studied in the literature. In the time-harmonic domain (both acoustic and electromagnetism), the so-called plane-wave DG method [54] and the ultra-weak variational formulation [22] are examples of well-studied Trefftz methods. The use of this method in this regime does not only help in reducing the number of basis functions locally but also helps in reducing the computational complexity that may arise from evaluating integrals over spatial elements in their variational formulations [57, 15]. Moreover, analytical features such as dominant directions and oscillatory character of the wave can be incorporated in the approximating spaces of the variational methods [14].

The excellent performance of the Trefftz method in the frequency regime has motivated the research community to investigate the method in the corresponding time-domain (see Section 2.4).

All the works cited above are in the framework of the DG method and hence we can infer that the DG framework provides the capacity to accommodate Trefftz spaces for the solution of PDE of interest and acoustic wave problems.

2.3 Review of time stepping methods for wave problems

Discretization of the time-dependent wave equation in space either by FEM or DG method is not enough for the full simulation of the problem. One still requires a time stepping method to achieve a fully discretized scheme and for the purpose of developing a time stepping method. The usual practice is the use of finite difference time stepping schemes such as leapfrog, Runge-Kutta, and Crank-Nicolson schemes after splitting the semi-discretized scheme into systems of first order in time (see, [25],[26] and [100]). However, second-order leapfrog or Newmark scheme can be used directly on the semi-discretized formulation without splitting into systems of first order [61].

The type of time stepping method employed determines the form of the fully discretized method which can be either explicit or implicit. Explicit methods solve for the solution of a system at a later time t_{n+1} directly from known values at previous states t_n, t_{n-1}. For example, the semi-discretized formulation of the wave equation (1.1.1) where DG is employed spatially can be written compactly as

$$(\ddot{u}, v)_\Omega + a(u, v) = 0, \tag{2.3.1}$$

where $a(\cdot,\cdot)$ represents a DG bilinear form and v is a test function from a finite dimensional space. Employing the leapfrog time discretization to discretize (2.3.1) in time introduces a second-order central difference scheme which gives

$$\left(\frac{U^{n+1} - 2U^n + U^{n-1}}{k^2}, v\right)_\Omega + a(U^n, v) = 0, \tag{2.3.2}$$

where U^n represents the solution at t_n and $k > 0$ is the time step, and $t_n = nk$. Rearranging now, we have

$$(U^{n+1}, v)_\Omega = 2(U^n, v)_\Omega - (U^{n-1}, v)_\Omega - k^2 a(U^n, v).$$

If we introduce the usual standard basis functions such that the discrete solution U^n can be represented locally as a linear combination of these basis functions, i.e., $U^n = \sum_{i=1}^{P} \alpha_i^n(t)\phi(x)$, then the fully discretized scheme can be written compactly as

$$MU^{n+1} = 2MU^n - MU^{n-1} - k^2 AU^n, \tag{2.3.3}$$

where M represents the global DG mass matrix and A represents the global DG stiffness matrix. Clearly, the solution at each time step is given by a closed form formula involving the solution of previous time steps. Although we still need to solve a linear system at each time step, the idea of mass lumping technique can be introduced in this case to speed up the computation [99, 114, 85]. Explicit methods are cheap, easier to implement, and easier to parallelize compared to implicit methods. However, small time-steps must be chosen in order to avoid numerical instability. Moreover, the smallest element in the mesh during refinement process controls the maximum time-step allowed by the CFL condition [60]. Techniques such as locally implicit time stepping or local time stepping methods are often required to circumvent stability problems [48].

Implicit methods, on the other hand, give the solution of a system at a later time t_{n+1} by solving coupled sets of equations involving the later state and previous states. For example, note that the semi-discretized formulation (2.3.1) can be written as

$$M\ddot{\alpha}(t) + A\alpha(t) = 0, \tag{2.3.4}$$

where we have introduced the standard basis functions locally on each element to derive the above equation. If we split the semi-discretized formulation (2.3.4) into first order system we have

$$\begin{aligned} M\dot{\alpha}(t) &= M\eta(t) \\ M\dot{\eta}(t) + A\alpha(t) &= 0. \end{aligned} \tag{2.3.5}$$

Employing the Crank-Nicolson time discretization gives

$$\begin{aligned} M\left(\frac{\eta_n - \eta_{n-1}}{k_n}\right) + A\left(\frac{\alpha_{n-1} + \alpha_n}{2}\right) &= 0 \\ M\left(\frac{\alpha_n - \alpha_{n-1}}{k_n}\right) &= M\left(\frac{\eta_{n-1} + \eta_n}{2}\right). \end{aligned} \tag{2.3.6}$$

Compactly in block matrix form we have

$$\begin{bmatrix} M & \frac{k_n}{2}M \\ \frac{k_n}{2}A & M \end{bmatrix}\begin{bmatrix} \alpha_n \\ \eta_n \end{bmatrix} = \begin{bmatrix} M & \frac{k_n}{2}M \\ -\frac{k_n}{2}A & M \end{bmatrix}\begin{bmatrix} \alpha_{n-1} \\ \eta_{n-1}. \end{bmatrix} \tag{2.3.7}$$

Obviously, the scheme is implicit and it involves solving a linear system with large matrix at each time step. Moreover, mass lumping technique cannot be employed in this case because of the presence of the stiffness matrix within the block matrix on the LHS. Implicit methods are in general expensive compared to explicit methods but in contrast, they allow larger time steps in implementation.

The shortcomings of the method of lines [3] approach which gives rise to either implicit or explicit methods have motivated the use of space-time discretizations as alternative techniques for wave problems simulation [109, 104, 91]. Space-time DG methods may be implicit, however, a special mesh discretization technique called the tent pitching (see, [123],[96], and [116]) can be employed to transform space-time methods into quasi-explicit methods where only solutions on the space-time front are stored and the rest of the computation is treated explicitly. Hence we could have a reduction in computer memory usage as well as local time stepping advantage.

2.4 Review of space-time methods for the wave equation

In this section, we briefly review some of the works done by practitioners in the field of numerical partial differential equations that are pertinent to the development of our method. These works stand as stepping stone and are crucial to the formulation

[3] Method of lines approach involve discretizing separately usually in space first and later in time

of the Trefftz space-time discontinuous Galerkin method.

The origin of the space-time method can be traced back to the late $80s$ when Hulbert and Hughes [73] developed this class of methods for elastodynamics problems. The framework of the method was based on the success of the time discontinuous Galerkin method for first order systems. This also motivated the extension of the time discontinuous Galerkin method to second-order hyperbolic problems. Stability of the schemes was controlled or enforced by means of stabilizing operators in least-squares form

$$a_{GLS}(u,v) \coloneqq \sum_{n=0}^{N-1} \int_{I_n} (\ddot{u} - \Delta u, \ddot{v} - \Delta v)_\Omega + ([\nabla u], [\nabla v])_\Gamma \mathrm{d}t, \tag{2.4.1}$$

where Γ represents the skeleton of the mesh and the spatial discontinuity across each element edge e along a slab is denoted by the spatial jump $[u]|_e \coloneqq u^+\mathbf{n}^+ + u^-\mathbf{n}^-$. The method allows the use of high-order time integration schemes and stability is less of an issue compared to the semi-discretized approach with explicit time stepping scheme. This class of methods developed in [72, 73] are now classified as the time only DG methods since they allow the shape functions to be discontinuous only across successive time steps [125]. The extension and the success of the approach can also be seen in [118, 117] for the acoustic problem.

In 1993, the time only DG technique was introduced directly into approximating second-order wave equation by Johnson, see, [78] without the inclusion of the Galerkin least square stability. He was able to prove the optimal rate of convergence using polynomial basis functions which are piecewise linear both in space and in time. Costanzo and Huang in 2005 [41] extended the work done by Hulbert and Hughes [72, 73] to completely unstructured meshes. They developed an unconditionally stable space-time method for the elastodynamics problem. Unconditionally stable in this context implies that the Galerkin least square term is not necessary for the stability of the scheme. However, the error analysis was not provided in the paper.

Finally, the use of non-standard basis functions in the DG approximation space as a means to enrich the space of approximation was introduced in [106]. These basis

functions are wave-like analytical solutions of the wave equation defined locally on each space-time element $K \times I_n$, with $I_n = (t_n, t_{n+1})$

$$S_h(K \times I_n) = \left\{ v(x,t) = v_0 + \sum_{\ell=1}^{p-1} \sum_{j=1}^{J} \beta_{\ell j} (ct - x \cdot \alpha_j^{(K,n)})^{\ell}, v_0, v_{\ell,j} \in \mathbb{R}, x \in K, t \in I_n \right\}, \tag{2.4.2}$$

where $\alpha_j^{(K,n)}$, $|\alpha_j^{(K,n)}| = 1$, represent directions of propagation. For example in one spatial dimension, we have two directions of propagation and hence the approximation on a space-time element is

$$\begin{aligned} v(x,t) = v_0 + v_1(ct - x) + v_2(ct - x)^2 + \cdots + v_p(ct - x)^p \\ + v_{p+1}(ct + x) + v_{p+2}(ct + x)^2 + \cdots + v_{2p}(ct + x)^p. \end{aligned} \tag{2.4.3}$$

In higher spatial dimensions, directions of propagation can be incorporated in the construction of local solutions by using the equi-distributed idea of Cessenat and Després in [22], (see Section 3.4 and Subsection 5.3.2).

In contrast to other time-space method discussed above, the basis functions introduced in the approximation space are discontinuous in space along slab and discontinuous in time across the slab. However, the stability of the scheme was enforced by means of a Lagrangian multiplier. Below is the variational formulation of the approach when $a \equiv 1$:

$$\begin{aligned} &\sum_{n=0}^{N-1} \int_{I_n} ((\ddot{u}, \dot{v})_\Omega + (\nabla u, \nabla \dot{v})_\Omega) \, \mathrm{dt} \\ &+ (\llbracket \dot{u}(t_n) \rrbracket, \dot{v}(t_n^+))_\Omega + (\llbracket \nabla u(t_n) \rrbracket, \nabla v(t_n^+))_\Omega + \int_{I_n} ([\dot{v}], \lambda)_\Gamma \mathrm{dt} = 0, \end{aligned}$$

subject to

$$\int_{I_n} ([\dot{u}], \mu)_\Gamma \mathrm{dt} = 0,$$

where $\llbracket u(x,t_n) \rrbracket = u(x,t_n^+) - u(x,t_n^-)$ denotes the temporal jump which measures the pointwise gradient due to the discontinuity across the time slab. Kretzschmar et al in [82] modified the DG formulation of Monk and Richter [96] for Maxwell's problem by introducing the local solutions of the type (2.4.2) in the approximation spaces of their formulation. Extension of the same idea for the wave equation on unstructured

meshes can be found in [94]. This class of methods is generally classified as the space-time DG method because they allow the shape functions to be discontinuous both in space and in time .

For the development of our method, we propose to use the technique of [106] but in contrast to them, we enforce continuity and stability in space by an interior penalty approach instead of the Lagrangian multipliers. This helps us to restrict the number of degrees of freedom only to the number of wave-like basis functions per element and the entries of our linear system are therefore reduced. The choice of the penalty terms introduced in our bilinear form is crucial for the theoretical analysis and for the practical behaviour of the method for the case of lowest order basis functions.

2.5 Important inequalities

In this section, we cover briefly useful inequalities that are usually employed in the analysis of DG methods. These inequalities emerged as a result of various theorems and lemmas which we quote and we refer readers to the texts cited for their proofs. We present the following definitions before we proceed.

Definition 2.5.1 (Shape-regularity)**.** *A family of meshes* $\{\mathcal{T}_h\}$ $h > 0$ *is said to be shape-regular if there exists* ϱ_0 *such that for any* $K \in \mathcal{T}_h$, $\varrho_K = h_K/\rho_K \leq \varrho_0$, *where* ρ_K *denotes the diameter of an inscribed circle in* K *and* $h_K = \mathrm{diam}(K)$ *[6, 47, 55, 51].*

Definition 2.5.2. *A family of meshes* $\{\mathcal{T}_h\}$ *is said to be quasi-uniform iff it is shape regular and* $\exists$ *a* c *such that* $h_K \geq ch \ \forall \ h$ *and* $\forall \ K \in \mathcal{T}_h$.

Definition 2.5.3. *Let* $\mathcal{T}_h$ *be a mesh of* Ω, *we define the finite element space to be* $V_h^p = \{v \in L^2(\Omega) : v|_K \in \mathcal{P}_p(K) \forall K \in \mathcal{T}_h\}$, *where* $\mathcal{P}_p$ *denotes the space of polynomials of order not more than* p.

2.5.1 Broken Sobolev spaces

The idea of broken Sobolev spaces is a very important concept in the analysis of DG methods. We refer readers to [47, 110] for details. To explain the concept in

this thesis, we focus on the Sobolev space $H^1(\Omega)$. Let $\mathcal{T}_h$ be a regular subdivision (or a mesh) of a Lipschitz domain Ω into a set of simplices. We define the Sobolev space $H^1(K)$ for any $K \in \mathcal{T}_h$ to be

$$H^1(K) := \{v \in L^2(K) : \frac{\partial}{\partial x_j} \in L^2(K) \forall j = 1, \ldots, d\}, \tag{2.5.1}$$

and the broken Sobolev space is defined as

$$H^1(\mathcal{T}_h) := \{v \in L^2(\Omega) : \forall K \in \mathcal{T}_h, v|_K \in H^1(K)\}. \tag{2.5.2}$$

The above definition is crucial for the proper definition of the trace inequality in the discontinuous Galerkin context as we shall see in the next subsection.

2.5.2 Trace theorem and inequalities

It is well known that the functions in Sobolev spaces $H^s(\Omega)$, $s \geq 0$, are defined by Lebesgue integrals only up to measure zero and the definition of the restriction of such functions to the boundary is not clear since the boundary has exactly zero measure. A trace theorem however gives answer to the question of defining the restriction of Sobolev functions on the boundary. In the following we state a version of the trace theorem found in [110, 113, 65].

Theorem 2.5.4. *Let Ω be a bounded Lipschitz domain with boundary $\partial\Omega$ and let $\mathbf{n}$ denote the outward normal vector. There exist trace operators $\gamma_0 : H^s(\Omega) \longrightarrow H^{s-1/2}(\partial\Omega)$ and $\gamma_1 : H^s(\Omega) \longrightarrow H^{s-3/2}(\partial\Omega)$ for $s \geq 1$, which define the restriction of function v and the restriction of its normal derivative $\mathbf{n} \cdot \nabla v$ to the boundary $\partial\Omega$ as linear maps from the Sobolev space on Ω to the Sobolev space on the boundary $\partial\Omega$. Furthermore, the operators are surjective and if $v \in \mathcal{C}^1(\bar{\Omega})$, then $\gamma_0 v = v|_{\partial\Omega}$ and $\gamma_1 v = \mathbf{n} \cdot \nabla v|_{\partial\Omega}$.*

Theorem 2.5.5 (Stability estimate [20, 47])**.** *Let Ω be a bounded domain with Lipschitz boundary $\partial\Omega$. Then there exists a constant C such that*

$$\|v\|_{L^2(\partial\Omega)} \leq C \|v\|^{1/2}_{L^2(\Omega)} \|v\|^{1/2}_{H^1(\Omega)}, \quad \forall v \in H^1(\Omega). \tag{2.5.3}$$

In the context of discontinuous Galerkin methods, the continuous trace inequality (2.5.3) can be extended to broken Sobolev spaces based on the definition of the later. We can infer from (2.5.3) that there exists a constant C (which depends on the shape-regularity of the mesh) independent of h such that (see, e.g., [47])

$$\|v\|_{L^2(\partial K)} \leq C\|v\|^{1/2}_{L^2(K)}\|v\|^{1/2}_{H^1(K)}, \quad \forall v \in H^1(\mathcal{T}_h),\, K \in \mathcal{T}_h. \tag{2.5.4}$$

The standard local trace inequalities which incorporate the edge length $|e|$, the local meshsize $h_K = \text{diam}(K)$, and the area $|K|$ of a simplex K in the mesh $\mathcal{T}_h$ are given by [110, 85]

$$\begin{aligned} \|\gamma_0 v\|_{L^2(e)} &\leq C|e|^{1/2}|K|^{-1/2}\big(\|v\|_{L^2(K)} + h_K\|\nabla v\|_{L^2(K)}\big), s \geq 1, \\ \|\gamma_1 v\|_{L^2(e)} &\leq C|e|^{1/2}|K|^{-1/2}\big(\|\nabla v\|_{L^2(K)} + h_K\|\nabla^2 v\|_{L^2(K)}\big), s \geq 2. \end{aligned} \tag{2.5.5}$$

The corresponding discrete version of the inequalities are derived via equivalence of norm in finite dimensional space; (see Chapter 2 of [110]). Let $\mathcal{P}_p(K)$ denote the space of polynomials of degree not more than p, then the discrete trace inequalities are defined as follows

$$\begin{aligned} \|v\|_{L^2(e)} &\leq \tilde{C}|e|^{1/2}|K|^{-1/2}\|v\|_{L^2(K)}, \quad \forall v \in \mathcal{P}_p(K),\ e \in \partial K, \\ \|\nabla v \cdot n\|_{L^2(e)} &\leq \tilde{C}|e|^{1/2}|K|^{-1/2}\|\nabla v\|_{L^2(K)}, \quad \forall v \in \mathcal{P}_p(K),\ e \in \partial K, \end{aligned} \tag{2.5.6}$$

where $\tilde{C}$ is independent of the simplex diameter h_K and the function v but depends on the degree p of the polynomial. The above inequality (2.5.5) and (2.5.6) are very useful in the DG analysis. We refer to [88] for the general theory of trace and [127] for the explicit estimates of the constant $\tilde{C}$ when the mesh elements are intervals, triangles and tetrahedra.

2.5.3 Inverse inequalities

Another important set of inequalities are the inverse inequalities. We state the following theorem to present them (see, [113, 51]).

Theorem 2.5.6. *Let K be a simplex in a quasi-uniform mesh $\mathcal{T}_h$ of Ω with parameter*

ϱ_0 and let e be an edge of K. Then for all $v \in V_h^p$, we have the following inverse estimates:

$$\begin{aligned} \|\nabla v\|_{L^2(K)} &\leq C h_K^{-1} \|v\|_{L^2(K)}, \\ \|v\|^2_{L^2(e)} &\leq C h_K^{-1} \|v\|^2_{L^2(K)}, \quad \forall v \in \mathcal{P}_p(K), \end{aligned} \tag{2.5.7}$$

where the constant C depends on the order of polynomial p, but independent of the element diameter h_K and the function v.

Other forms of inverse inequalities in different norms are given in [113, Theorem 4.76] A more general framework for the construction of inverse inequality is presented in [27].

2.5.4 Cauchy-Schwarz inequality

Another useful inequality that often occurs in the analysis of finite element methods is the Cauchy-Schwarz inequality, (see e.g., [47, 113]). We present the inequality as follows: let $v, w \in L^2(\Omega)$, then

$$\int_\Omega vw \leq \|v\|_{L^2(\Omega)} \|w\|_{L^2(\Omega)}. \tag{2.5.8}$$

2.5.5 Gronwall inequality

Another useful inequality that is usually employed in the analysis of time dependent problems is the Gronwall inequality. We state both the continuous and the discrete forms of the lemma from [110].

Lemma 2.5.7. *Let f, g, h be piecewise continuous non-negative functions defined on $[a, b]$. Assume that g is non-decreasing and that there exists a positive constant C independent of t such that*

$$f(t) + h(t) \leq g(t) + C \int_a^t f(s)\, ds, \quad \forall\, t \in (a, b).$$

Then

$$f(t) + h(t) \leq g(t) e^{C(t-a)}, \quad \forall t \in (a, b). \tag{2.5.9}$$

2.6 L^2-projection and error bound

2.6.1 The orthogonal L^2–projection

The L^2–projection or orthogonal L^2-projection $P_h v$ gives the approximation to v in the form of an average. Let $v \in L^2(K)$, then the L^2-projection $P_h v$ satisfies

$$\int_K (v - P_h v) w = 0 \quad \forall w \in \mathcal{P}_p(K). \tag{2.6.1}$$

The function v being approximated need not to be continuous. Moreover, the L^2-projection also satisfies

$$\|v - P_h v\|_{L^2(K)} \leq \|v - w\|_{L^2(K)} \quad \forall w \in \mathcal{P}_p(K) \tag{2.6.2}$$

and hence is the best approximation to the function v with respect to the L^2–norm.

The following lemma gives the bound on the orthogonal L^2 approximation error (see, [47, 6]).

Lemma 2.6.1. *Let K be a simplex in $\mathcal{T}_h$ and let $P_h v$ denote the $L^2(K)$–orthogonal projection onto $\mathcal{P}_p(K)$. Then for $s \in \{0, \ldots, p+1\}$ and any $v \in H^s(\Omega)$, there holds*

$$\|v - P_h v\|_{H^m(K)} \leq C h_K^{s-m} \|v\|_{H^s(K)} \ \forall\, m \in \{0, \ldots, s\}, \tag{2.6.3}$$

where C is independent of K and mesh size $h = \max_{K \in \mathcal{T}_h} h_K$.

For example if $m = 0$, we have the following error bound [6]

$$\|v - P_h v\|_{L^2(K)} \leq C h_K^s \|v\|_{H^s(K)}. \tag{2.6.4}$$

Shape regularity assumption on $\mathcal{T}_h$ is not required for the estimate (2.6.4) [47].

2.6.2 Approximation properties in polynomial space

In this subsection, we state two theorems that summarise the bounds on the error for any function v in the Sobolev space $H^s(K)$ defined on simplex K and on the

edge e of K. We refer to [9, 10, 112] for the origin of the theorems.

Theorem 2.6.2. *Let K be a simplex with diameter h_K and let $v \in H^s(K)$ for $s \geq 1$ and $p \geq 0$ an integer. Then there exists a constant C independent of v and h_K and a function $\tilde{v} \in \mathcal{P}_p(K)$ such that $\forall\ 0 \leq q \leq s$,*

$$\|v - \tilde{v}\|_{H^q(K)} \leq C h_K^{\min(p+1,s)-q} |v|_{H^s(K)}. \tag{2.6.5}$$

The next theorem gives the bound on the approximation error on the edge e of K see, [112] or [47].

Theorem 2.6.3. *Let K be a simplex with diameter h_K. Let e denote the edge or face of K and n_K denotes the outward normal to ∂K. Then there exists a constant C_1 independent of K and h_K and an approximation $\tilde{v} \in \mathcal{P}_p(K)$ such that*

$$\|\tilde{v} - v\|_{L^2(e)} \leq C_1 h_K^{\min(p+1,s)-1/2} |v|_{H^s(K)}, \tag{2.6.6}$$

and if $s \geq 2$, then there exists a constant C_2 independent of K and h_K such that

$$\|\nabla(\tilde{v} - v)|_K \cdot n_K\|_{L^2(e)} \leq C_2 h_K^{\min(p+1,s)-3/2} |v|_{H^s(K)}. \tag{2.6.7}$$

2.7 MATLAB and Chebfun software

Numerical experiments in this thesis are carried out with the aid of MATLAB. The software provides a well-equipped platform for the creation of meshes, the execution of numerical algorithms as well as viewing the results of simulation graphically. MATLAB also provides a good environment for the use of a package called 'Chebfun2' that we employed in our implementation.

Chebfun2 is MATLAB-based software that computes and approximates functions of two variables over a specified rectangular domain $[a, b] \times [c, d]$ up to machine precision (relatively up to 10^{-15}). The development of the first version of the software was to approximate functions of one variable based on the idea that smooth functions

can be approximated efficiently by polynomial interpolation at Chebyshev points

$$x_j = \cos\left(\frac{j\pi}{n}\right), \quad 0 \le j \le n, \tag{2.7.1}$$

or by expansion in Chebyshev polynomials [107]. This version was released in 2004. In 2013, the extension of the software to compute functions of two variables as well as vector-valued functions was released and named Chebfun2.

The development of Chebfun2 rests on the univariate representations and algorithms of the original Chebfun [120]. Moreover, the salient observation that many functions of two variables can be approximated efficiently by low rank functions motivated the extension [121, 59, 59]. For example, given a real valued function $f(x, y)$ defined on $[-1, 1]\times[-1, 1]$, the value of the function can be obtained optimally using the singular value decomposition (SVD):

$$f(x, y) = \sum_{j=1}^{\infty} \alpha_j \phi_j(x)\psi_j(y), \tag{2.7.2}$$

where ϕ_j, ψ_j, $j = 1\ldots$ are univariate rank 1 functions and are orthogonal in $L^2([-1, 1]^2)$, α_j, $j = 1\ldots\infty$ are non-increasing real sequences of singular values [120]. Now the optimal rank k approximant f_k to f is obtained by truncation

$$f(x, y) \approx f_k(x, y) = \sum_{j=1}^{k} \alpha_j \phi_j(x)\psi_j(y), \tag{2.7.3}$$

where α_j, $j = 1\ldots k$ decays depending on the smoothness of the function being approximated.

This idea is introduced numerically in Chebfun2 by sampling f_k on $n\times n$ Chebychev tensor grid and computing the SVD of the sampled matrix. Then the optimal k rank matrix in the discrete 2−norm is obtained by coupling the first k singular values with the left and right singular vectors.

Chebfun2 offers simple syntax and many MATLAB commands are overloaded in the software. To avoid confusion, the name of the software is Chebfun2 with capital "C" while its MATLAB object is chebfun2. We present below a simple code snippet 2.1 of Chebfun2 operation and its output in Figure 2.2 and in Figure 2.3.

```
x = chebfun2(@(x,y) cos(x), [-1 1 -1 1]);
y = chebfun2(@(x,y) sin(y), [-1 1 -1 1]);
f = sin(x.*y);
plot(f),
xlabel('x','interpreter','latex','FontSize',12)
ylabel('y','interpreter','latex','FontSize',12)
```

Figure 2.1: *Code snippet showing chebfun2 object in MATLAB.*

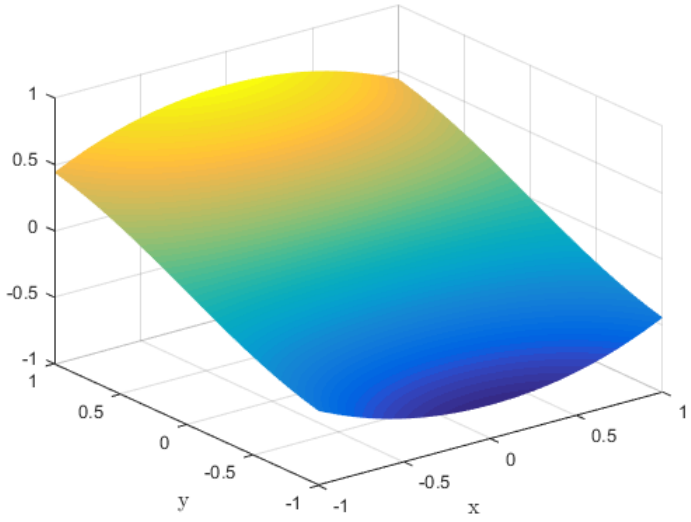


Figure 2.2: *Chebfun2 operation in MATLAB.*

```
x =
   chebfun2 object (1 smooth surface)
       domain                  rank       corner values
[  -1,   1] x [  -1,   1]         1     [0.54 0.54 0.54 0.54]
vertical scale =   1

y =
   chebfun2 object (1 smooth surface)
       domain                  rank       corner values
[  -1,   1] x [  -1,   1]         1     [-0.84 -0.84 0.84 0.84]
vertical scale = 0.84

f =
   chebfun2 object (1 smooth surface)
       domain                  rank       corner values
[  -1,   1] x [  -1,   1]         6     [-0.44 -0.44 0.44 0.44]
vertical scale = 0.75
```

Figure 2.3: *Approximation of higher rank function by two rank* 1 *functions* .

Chapter 3

Analysis of Trefftz space-time DG method for the wave equation

In this chapter, we develop and analyse a space-time discontinuous Galerkin method utilising special non-standard polynomial bases called Trefftz basis functions for the scalar undamped wave equation in second order formulation. The DG method considered is motivated by the class of interior penalty DG (IPDG) methods, as well as by the classical work of Hulbert and Hughes [73, 72]. The choice of penalty terms included in the bilinear form is essential for both the theoretical analysis and for the practical behaviour of the method for the case of lowest order basis functions (not necessary for the practical behaviour of higher order basis functions).

The motivation and objective behind the use of Trefftz basis functions is to reduce the number of degrees of freedom per wavelength required to obtain accurate results. This idea has been found very successful in practice, especially in the frequency domain [22, 95], where the prominent example is the use of plane wave bases in the approximation spaces. The approach brings along some advantages which include efficient and effective method that can approximate problems with energy at high frequencies and simplification of the computational task as implementation can be restricted to the space-time skeleton of the mesh.

The most natural way of including space-time Trefftz functions is within the confines of a space-time DG method. In this work, we restrict the construction of our method to space-time slabs while descritizing the wave equation in primal form

to ensure solvability on each time-step, as well as to aid the presentation and the analysis of the method. However, with minor modifications, completely unstructured space-time meshes could, in principle, be used with the proposed space-time DG frame-work. This construction leads to a stable, dissipative scheme for general polynomial bases.

We organise the rest of this chapter as follow. In Sections 3.1 and 3.2, we construct the space-time IPDG method beginning with space-time polynomial spaces and we prove its stability. We proceed in Section 3.3 to analyse polynomial Trefftz spaces and we prove quasi optimality. Finally in this Chapter, we prove convergence rates for the method in spatial dimension $d = 1, 2, 3$; moreover we also provide $hp-$version a priori bounds for $d = 1$.

3.1 Model problem

We consider the acoustic wave equation

$$\begin{aligned} \ddot{u} - \nabla \cdot (a \nabla u) &= 0 && \text{in } \Omega \times [0,T], \\ u &= 0 && \text{on } \partial\Omega \times [0,T], \\ u(x,0) = u_0(x),\ \dot{u}(x,0) &= v_0(x), && \text{in } \Omega, \end{aligned} \tag{3.1.1}$$

where Ω is a bounded Lipschitz domain in $\mathbb{R}^d$, $\partial\Omega$ its boundary and $0 < c_a \le a(x) \le C_a$, $(x \in \Omega)$ a piecewise constant function. If Ω_1 and Ω_2 are two subsets of Ω with the boundary Γ separating them and with $a \equiv a_1$ in Ω_2 and $a \equiv a_2$ in Ω_2, then if we denote by $u_1 = u|_{\Omega_1}$ and $u_2 = u|_{\Omega_2}$ we further have the transmission conditions

$$u_1 = u_2, \quad \mathbf{n} \cdot a_1 \nabla u_1 = \mathbf{n} \cdot a_2 \nabla u_2 \tag{3.1.2}$$

where $\mathbf{n}$ is the exterior normal to Ω_1 (or Ω_2).

Let $u_0 \in H_0^1(\Omega)$ and $v_0 \in L^2(\Omega)$, then (3.1.1) has a unique weak solution U with

$$U \in L^2([0,T]; H_0^1(\Omega)), \quad \dot{U} \in L^2([0,T]; L^2(\Omega)), \quad \ddot{U} \in L^2([0,T]; H^{-1}(\Omega)), \tag{3.1.3}$$

see [88, Theorem 8.1]. Furthermore, according to [88, Theorem 8.2], the solution is continuous in time with

$$U \in C([0,T]; H_0^1(\Omega)), \quad \dot{U} \in C([0,T]; L^2(\Omega)). \tag{3.1.4}$$

We denote throughout the discussion the space of all solutions of 3.1.1 by

$$\mathcal{X} := \left\{ U \mid U \text{ weak solution of (3.1.1) } \forall u_0 \in H_0^1(\Omega), v_0 \in L^2(\Omega) \right\}, \tag{3.1.5}$$

and the weak solution depends on the coefficient $a(x)$.

3.2 Construction of space-time finite element space

We aim to discretize this problem by a new time-space interior penalty discontinuous Galerkin method. In principle, this could be done on a general time-space mesh, however for the simplicity of presentation (and implementation) we construct a time discretization $0 = t_0 < t_1 < \cdots < t_N = T$ and locally quasi-uniform spatial-meshes $\mathcal{T}_n$ of Ω consisting of open simplices such that $\overline{\Omega} = \cup_{K \epsilon \mathcal{T}_n} \overline{K}$, with $K \cap \widetilde{K} = \varnothing$, for $K, \widetilde{K} \in \mathcal{T}_n$ and $K \neq \widetilde{K}$. Therefore the space-time mesh consists of time-slabs $\mathcal{T}_n \times I_n$, where $I_n = (t_n, t_{n+1})$, $\tau = t_{n+1} - t_n$.

The discrete space-time approximation space will consist of piecewise polynomials on each time-slab, given by the local space-time finite element space:

$$S_n^{h,p} := \left\{ u \in L^2(\Omega \times I_n) \ : \ u|_{K \times I_n} \in \mathcal{P}_p(\mathbb{R}^{d+1}),\ K \in \mathcal{T}_n \right\},$$

where $\mathcal{P}_p(\mathbb{R}^{d+1})$ is the space of polynomials of total degree p in $d+1$ variables; the complete space-time finite element space on $\Omega \times [0,T]$, will be denoted by

$$V^{h,p} := \{ u \in L^2(\Omega \times [0,T]) \ : \ u|_{\Omega \times I_n} \in S_n^{h,p},\ n = 0, 1, \ldots, N-1 \}.$$

We require some notation. The *skeleton* of the mesh, is defined by $\Gamma_n := \cup_{K \epsilon \mathcal{T}} \partial K$ and the interior skeleton by $\Gamma_n^{\text{int}} = \Gamma_n \setminus \partial\Omega$; the time-step counter n will be omitted

for brevity, when confusion is unlikely to occur. Moreover, we define the union of two skeletons of two subsequent meshes by $\hat{\Gamma}_n := \Gamma_{n-1} \cup \Gamma_n$ 3.1.

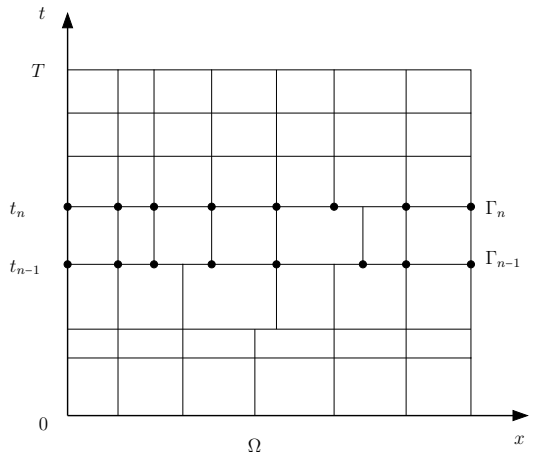


Figure 3.1: An example of a space-time mesh. Skeletons Γ_n and Γ_{n-1} are highlighted with black dots. The union of the two is denoted by $\hat{\Gamma}_n$.

Let us denote by K^+ and K^- two spatial elements sharing a face $e = \bar{K}^+ \cap \bar{K}^- \subset \Gamma_{\text{int}}$, with respective outward normal vectors $\mathbf{n}^+$ and $\mathbf{n}^-$ on e. For $u : \Omega \to \mathbb{R}$ and $\mathbf{v} : \Omega \to \mathbb{R}^d$, let $u^\pm : e \to \mathbb{R}$ and $\mathbf{v}^\pm : e \to \mathbb{R}^d$ be the traces on e with limits taken from $K^\pm$. We define the respective jumps and averages across each face $e \in \Gamma_{\text{int}}$ by

$$\{u\}|_e = \frac{1}{2}(u^+ + u^-), \qquad \{\mathbf{v}\}|_e = \frac{1}{2}(\mathbf{v}^+ + \mathbf{v}^-),$$
$$[u]|_e = u^+\mathbf{n}^+ + u^-\mathbf{n}^-, \quad [\mathbf{v}]|_e = \mathbf{v}^+ \cdot \mathbf{n}^+ + \mathbf{v}^- \cdot \mathbf{n}^-;$$

if $e \subset K^+ \cap \partial\Omega$, we set $\{\mathbf{v}\}|_e = \mathbf{v}^+$ and $[u]|_e = u^+\mathbf{n}^+$. Further, we define the temporal jump by

$$[\![u(t_n)]\!] = u(t_n^+) - u(t_n^-), \qquad [\![u(t_0)]\!] = u(t_0^+).$$

We denote the spatial meshsize by $h : \Omega \times [0,T] \to \mathbb{R}$, defined by $h(x,t) = \text{diam}(K)$ if $x \in K$ for $K \in \mathcal{T}_n$ and $t \in I_n$; when $x \in e = \bar{K}^+ \cap \bar{K}^-$, we set $h(x,t) := \{h\}$ to be the average. Finally, we assume that there exist $c_{\mathcal{T}} > 0$ such that

$$\text{diam}(K)/\rho_K \le c_{\mathcal{T}}, \qquad \forall K \in \mathcal{T}_n,\ n = 0, 1, \ldots, N-1, \tag{3.2.1}$$

where ρ_K is the radius of the inscribed circle of K.

For simplicity of the presentation *only*, we shall, henceforth, assume space-time

shape-regularity $\operatorname{diam} K \sim |I_n|$ for all $K \in \mathcal{T}_n$; this allows us to consider one space-time meshsize $h_K^n = \operatorname{diam}(K \times I_n)$ per space-time element. Finally the broken spatial gradient will be denoted by $\nabla_n v$, given by $(\nabla_n v)|_K := (\nabla v)|_K$ for all $K \in \mathcal{T}_n$, and a $v \in \mathcal{C}(I_n; H_0^1(\Omega)) + S_n^{h,p}$; collectively, we shall denote the broken gradient by $\widetilde{\nabla} v$ defined as

$$(\widetilde{\nabla} v)|_{\Omega \times I_n} := (\nabla_n v)|_{\Omega \times I_n}, \qquad n = 0, \dots, N-1,$$

for $v \in C(\prod_{n=0}^{N-1} I_n; H_0^1(\Omega)) + V^{h,p}$, which means we allow v to be discontinuous both in space and in time.

3.3 Development of a space-time discontinuous Galerkin method

To derive the weak form suitable for DG discretisation we will follow an energy argument. We begin with the assumption that the governing equation (3.1.1) has a smooth solution u and the test function $v \in \mathcal{X} + V^{h,p}$. The standard symmetric interior penalty discontinous Galerkin weak formulation on the time-slab I_n when tested with $\dot{v}$ is given by

$$\begin{aligned}(\ddot{u}, \dot{v})_{\Omega \times I_n} + (a\widetilde{\nabla} u, \widetilde{\nabla} \dot{v})_{\Omega \times I_n} - (\{a \nabla u\}, [\dot{v}])_{\Gamma_n \times I_n} \\ - ([u], \{a \nabla \dot{v}\})_{\Gamma_n \times I_n} + (\sigma_0 [u], [\dot{v}])_{\Gamma_n \times I_n} = 0,\end{aligned} \tag{3.3.1}$$

where

$$\sigma_0(x,t) := C_{\sigma_0} p^2 h(x,t)^{-1}, \tag{3.3.2}$$

for a positive constant C_{σ_0} independent of p and h. This immediately motivates the definition of *discrete energy* $E_h(t,v)$ at $t \in I_n$ along a time-space slab:

$$E_h(t,v) := \frac{1}{2}\|\dot{v}(t)\|_\Omega^2 + \frac{1}{2}\|\sqrt{a}\widetilde{\nabla} v(t)\|_\Omega^2 + \frac{1}{2}\|\sqrt{\sigma_0}[v(t)]\|_{\Gamma_n}^2 - (\{a \nabla v(t)\}, [v(t)])_{\Gamma_n}. \tag{3.3.3}$$

We prove in Lemma 3.3.2 that the energy $E_h(t,v)$ is non-negative.

Now, choosing $v = u$ as test function in (3.3.1) and summing over n, we have

$$\begin{aligned} 0 &= \sum_{n=0}^{N-1} \int_{I_n} \frac{d}{dt}\Big(\frac{1}{2}\|\dot{u}\|_\Omega^2 + \frac{1}{2}\|\sqrt{a}\widetilde{\nabla} u\|_\Omega^2 - (\{a\nabla u\},[u])_{\Gamma_n} + \frac{1}{2}\|\sqrt{\sigma_0}\,[u]\|_{\Gamma_n}^2\Big)dt \\ &= E_h(t_N^-,u) - E_h(t_0^+,u) - \sum_{n=1}^{N-1} [\![E_h(t_n,u)]\!]. \end{aligned}$$

We discover that we need to modify the formulation (3.3.1) in order to allow for discontinuity in time and also to control the term $[\![E_h(t_n,u)]\!]$ which has no definite sign. Therefore, we employ the upwind algebraic product rule

$$\begin{aligned} &[\![f(u(t_n))]\!][\![g(u(t_n))]\!] \\ &= [\![f(u(t_n))]\!]g(u(t_n^+)) + [\![g(u(t_n))]\!]f(u(t_n^+)) - [\![f(u(t_n))g(u(t_n))]\!], \end{aligned} \tag{3.3.4}$$

to each term in (3.3.1) to have

$$\begin{aligned} &\sum_{n=0}^{N-1} (\ddot{u},\dot{v})_{\Omega\times I_n} + ([\![\dot{u}(t_n)]\!],\dot{v}(t_n^+))_\Omega \\ &+ \big(a\widetilde{\nabla} u,\widetilde{\nabla}\dot{v}\big)_{\Omega\times I_n} + \big(a[\![\widetilde{\nabla} u(t_n)]\!],\widetilde{\nabla} v(t_n^+)\big)_\Omega \\ &- (\{a\nabla u\},[\dot{v}])_{\Gamma_n\times I_n} - \big([\![\{a\widetilde{\nabla} u(t_n)\}]\!],[v(t_n^+)]\big)_{\Gamma_n} \\ &- ([u],\{a\nabla\dot{v}\})_{\Gamma_n\times I_n} - ([\![[u(t_n)]]\!],\{a\nabla v(t_n^+)\})_{\Gamma_n} \\ &+ (\sigma_0[u],[\dot{v}])_{\Gamma_n\times I_n} + (\sigma_0[\![[u(t_n)]]\!],[v(t_n^+)])_{\Gamma_n} \\ &+ (\sigma_1[u],[v])_{\Gamma_n\times I_n} + (\sigma_2[a\nabla u],[a\nabla v])_{\Gamma_n\times I_n} = \mathcal{B}^{\text{init}}(v), \end{aligned} \tag{3.3.5}$$

where $\mathcal{B}^{\text{init}}$ is defined by

$$\begin{aligned} \mathcal{B}^{\text{init}}(v) \coloneqq{}& (v_0,\dot{v}(t_0^+))_\Omega + \big(a\widetilde{\nabla} u_0,\widetilde{\nabla} v(t_0^+)\big)_\Omega \\ &- (\{a\nabla u_0\},[v(t_0^+)])_{\Gamma_0} - ([u_0],\{a\nabla v(t_0^+)\})_{\Gamma_0} \\ &+ (\sigma_0[u_0],[v(t_0^+)])_{\Gamma_0}. \end{aligned} \tag{3.3.6}$$

The special penalty terms with non-negative parameters σ_1 and σ_2 in the weak formulation do not affect the consistency of the method; the need for their inclusion as well as the choice of σ_1 and σ_2 will become apparent in the convergence analysis. Note also that the last two terms in the definition of $\mathcal{B}^{\text{init}}$ are zero if the initial data is continuous in space.

Thus, we have arrived at a space-time DG method, which can be thought of in two ways: as a method for obtaining a discrete solution on a fixed space-time domain $\Omega \times [0,T]$ or as a time-stepping method. The former viewpoint will be important for the proof of convergence estimates while the latter will be useful for the implementation of the method. Consequently we define the following three bilinear forms to describe these two viewpoints:

$$\begin{aligned}\mathcal{A}_n(u,v) \coloneqq\ & (\ddot{u},\dot{v})_{\Omega\times I_n} + (\dot{u}(t_n^+),\dot{v}(t_n^+))_\Omega \\ & + (a\nabla u, \nabla\dot{v})_{\Omega\times I_n} + (a\nabla u(t_n^+), \nabla v(t_n^+))_\Omega \\ & - (\{a\nabla u\},[\dot{v}])_{\Gamma_n\times I_n} - (\{a\nabla u(t_n^+)\},[v(t_n^+)])_{\Gamma_n} \\ & - ([u],\{a\nabla\dot{v}\})_{\Gamma_n\times I_n} - ([u(t_n^+)],\{a\nabla v(t_n^+)\})_{\Gamma_n} \\ & + (\sigma_0[u],[\dot{v}])_{\Gamma_n\times I_n} + (\sigma_0[u(t_n^+)],[v(t_n^+)])_{\Gamma_n} \\ & + (\sigma_1[u],[v])_{\Gamma_n\times I_n} + (\sigma_2[a\nabla u],[a\nabla v])_{\Gamma_n\times I_n},\end{aligned} \tag{3.3.7}$$

$$\begin{aligned}\mathcal{B}_n(u,v) \coloneqq\ & (\dot{u}(t_n^-),\dot{v}(t_n^+))_\Omega + \left(a\widetilde{\nabla}u(t_n^-),\widetilde{\nabla}v(t_n^+)\right)_\Omega \\ & - (\{a\nabla u(t_n^-)\},[v(t_n^+)])_{\Gamma_n} - ([u(t_n^-)],\{a\nabla v(t_n^+)\})_{\Gamma_n} \\ & + (\sigma_0[u(t_n^-)],[v(t_n^+)])_{\Gamma_n},\end{aligned} \tag{3.3.8}$$

and

$$\mathcal{A}(u,v) \coloneqq \sum_{n=0}^{N-1}\mathcal{A}_n(u,v) - \sum_{n=1}^{N-1}\mathcal{B}_n(u,v), \tag{3.3.9}$$

which is just the same as the left-hand side of (3.3.5). We present the method with the following definition.

Definition 3.3.1. *Given subspaces $X_n \subseteq S_n^{h,p}$, the time-stepping method is described by: find $u^n \in X_n$, $n = 1,2,\dots,N-1$, such that*

$$\mathcal{A}_n(u^n,v) = \mathcal{B}_n(u^{n-1},v), \qquad \text{for all } v \in X_n, \tag{3.3.10}$$

and

$$\mathcal{A}_0(u^0,v) = \mathcal{B}^{init}(v), \qquad \text{for all } v \in X_0. \tag{3.3.11}$$

Equivalently, given a subspace $X \subseteq V^{h,p}$, the full space-time discrete system can be

presented as: find $u \in X$ such that

$$\mathcal{A}(u,v) = \mathcal{B}^{init}(v), \qquad \text{for all } v \in X. \tag{3.3.12}$$

Before we prove the stability and consistency of the method, we prove the following immediate lemma arising from the bilinear forms and the energy (3.3.3).

Lemma 3.3.2. *The energy $E_h(t,v)$ defined in* (3.3.3) *is non-negative.*

Proof. It suffices to bound the last term of the energy $E_h(t,v)$ and we shall employ the classical inverse inequality $\|v\|^2_{\partial K} \le C_{inv}p^2|\partial K|/|K|\|v\|^2_K$, for all $v \in \mathcal{P}_p(K)$, (see, Chapter 2 for review). Note that

$$\begin{aligned}
(\{a\nabla v(t)\},[v(t)])_{\Gamma_n} &= \int_{\Gamma_n} (\frac{\sigma_0}{2})^{-1/2}\{a\nabla v(t)\}(\frac{\sigma_0}{2})^{1/2}[v(t)] \\
&\le \frac{\kappa}{2}\int_{\Gamma_n}(\frac{\sigma}{2})^{-1}|\{a\nabla v\}|^2\mathrm{ds} + \frac{1}{2\kappa}\int_{\Gamma_n}(\frac{\sigma_0}{2})|[v]|^2\mathrm{ds} \\
&\le C_a\sum_{K\in\mathcal{T}_n}\int_{\partial K}|\nabla v(t)|^2\mathrm{ds} + \frac{1}{2\kappa}\int_{\Gamma_n}\sigma_0|[v(t)]|^2\mathrm{ds} \\
&\le \sum_{K\in\mathcal{T}_n}\frac{c_{\mathcal{T}}C_aC_{inv}p^2}{c_a h_K^{(n)}}\int_K|\sqrt{a}\nabla v(t)|^2\mathrm{dx} + \frac{1}{2\kappa}\int_{\Gamma_n}\sigma_0|[v(t)]|^2\mathrm{ds},
\end{aligned}$$

where we have used the Young's inequality in the second line and the inverse inequality in the last line. If we choose C_{σ_0} large enough, then the energy $E_h(t,v)$ is non-negative not only for smooth v but also for functions in $V^{h,p}$. In particular, choosing $\kappa = 2$ with

$$C_{\sigma_0} \ge c_{\mathcal{T}}^2 C_a C_{inv}/c_a \tag{3.3.13}$$

is sufficient so that

$$|(\{a\nabla v(t)\},[v(t)])_{\Gamma_n}| \le \frac{1}{4}\|\sqrt{\sigma_0}[v(t)]\|^2_{\Gamma_n} + \frac{1}{2}\|\sqrt{a}\widetilde{\nabla}v(t)\|^2_\Omega, \tag{3.3.14}$$

ensuring the non-negativity of the energy. □

Lemma 3.3.3. *It holds that for $w \in \mathcal{X} + V^{h,p}$,*

$$\mathcal{A}_n(w,w) = E_h(t_{n+1}^-,w) + E_h(t_n^+,w) + \|\sqrt{\sigma_1}[w]\|^2_{\Gamma_n\times I_n} + \|\sigma_2[a\nabla w]\|^2_{\Gamma_n\times I_n}, \tag{3.3.15}$$

for $n = 0, 1, \dots, N-1$, *and*

$$\begin{aligned}\mathcal{A}(w,w) = E_h(t_N^-, w) + E_h(t_0^+, w) + \sum_{n=1}^{N-1} \Big(& \frac{1}{2}\|[\![\dot{w}(t_n)]\!]\|_\Omega^2 + \frac{1}{2}\|\sqrt{a}[\![\widetilde{\nabla} w(t_n)]\!]\|_\Omega^2 \\ & - \Big(\big([\![\{a\widetilde{\nabla} w(t_n)\}]\!], [[w(t_n)]]\big)_{\Gamma_n} + \frac{1}{2}\|[\![\sqrt{\sigma_0}\,[w(t_n)]]\!]\|_{\Gamma_n}^2 \Big) \\ & + \sum_{n=0}^{N-1} \Big(\|\sqrt{\sigma_1}\,[w]\|_{\Gamma_n\times I_n}^2 + \|\sqrt{\sigma_2}\,[a\nabla w]\|_{\Gamma_n\times I_n}^2 \Big).\end{aligned} \tag{3.3.16}$$

Proof. The identities follow from the definitions of the bilinear forms and the energy $E_h(t,w)$. We give the proof for the time stepping bilinear form (3.3.7) below

$$\begin{aligned}\mathcal{A}_n(w,w) &= \frac{1}{2}\frac{d}{dt}\|\dot{w}(t)\|_{\Omega\times I_n}^2 + \|\dot{w}(t_n^+)\|_\Omega + \frac{1}{2}\frac{d}{dt}\|a\widetilde{\nabla} w\|_{\Omega\times I_n}^2 \\ &\quad + \|a\widetilde{\nabla} w(t_n^+)\|_\Omega^2 - \frac{d}{dt}\big(\,[w]\,, \{a\nabla w\}\,\big)_{\Gamma_n\times I_n} - 2\big(\,\{a\nabla w(t_n^+)\}\,, [w(t_n^+)]\,\big)_{\Gamma_n} \\ &\quad + \frac{1}{2}\frac{d}{dt}\|\sqrt{\sigma_0}\,[w]\,\|_{\Gamma_n}^2 + \|\sqrt{\sigma_0}\,[w(t_n)]\,\|_{\Gamma_n}^2 + \|\sqrt{\sigma_1}\,[w]\,\|_{\Gamma_n\times I_n}^2 + \|\sqrt{\sigma_2}\,[a\nabla w]\,\|_{\Gamma_n\times I_n}^2 \\ &= \frac{1}{2}\|\dot{w}(t_{n+1}^-)\|_\Omega^2 - \frac{1}{2}\|\dot{w}(t_n^+)\|_\Omega^2 + \|\dot{w}(t_n^+)\|_\Omega^2 + \frac{1}{2}\|a\widetilde{\nabla} w(t_{n+1}^-)\|_\Omega^2 \\ &\quad - \frac{1}{2}\|a\widetilde{\nabla} w(t_n^+)\|_\Omega^2 + \|a\widetilde{\nabla} w(t_n^+)\|_\Omega^2 - \big(\,[w(t_{n+1}^-)]\,, \{a\nabla w(t_{n+1}^-)\}\,\big)_{\Gamma_n} + \big(\,[w(t_n^+)]\,, \{a\nabla w(t_n^+)\}\,\big) \\ &\quad - 2\big(\,\{a\nabla w(t_n^+)\}\,, [w(t_n^+)]\,\big)_{\Gamma_n} + \frac{1}{2}\|\sqrt{\sigma_0}\,[w(t_{n+1}^-)]\,\|_{\Gamma_n}^2 - \frac{1}{2}\|\sqrt{\sigma_0}\,[w(t_n^+)]\,\|_{\Gamma_n}^2 \\ &\quad + \|\sqrt{\sigma_0}\,[w(t_n^+)]\,\|_{\Gamma_n}^2 + \|\sqrt{\sigma_1}\|_{\Gamma_n\times I_n}^2 + \|\sqrt{\sigma_2}\,[a\nabla w]\,\|_{\Gamma_n\times I_n} \\ &= \frac{1}{2}\|\dot{w}(t_{n+1}^-)\|_\Omega^2 + \frac{1}{2}\|\dot{w}(t_n^+)\|_\Omega^2 + \frac{1}{2}\|a\widetilde{\nabla} w(t_{n+1}^-)\|_\Omega^2 + \frac{1}{2}\|a\widetilde{\nabla} w(t_n^+)\|_\Omega^2 \\ &\quad - \big(\,[w(t_{n+1}^-)]\,, \{a\nabla w(t_{n+1}^-)\}\,\big)_{\Gamma_n} - \big(\,[w(t_n^+)]\,, \{a\nabla w(t_n^+)\}\,\big)_{\Gamma_n} + \frac{1}{2}\|\sqrt{\sigma_0}\,[w(t_{n+1}^-)]\,\|_{\Gamma_n}^2 \\ &\quad + \frac{1}{2}\|\sqrt{\sigma_0}\,[w(t_n^+)]\,\|_{\Gamma_n}^2 + \|\sqrt{\sigma_1}\,[w]\,\|_{\Gamma_n\times I_n}^2 + \|\sqrt{\sigma_2}\,[a\nabla w]\,\|_{\Gamma_n\times I_n}^2,\end{aligned} \tag{3.3.17}$$

which leads to

$$\mathcal{A}_n(w,w) = E(t_{n+1}^-, w) + E(t_n^+, w) + \|\sqrt{\sigma_1}\,[w]\|_{\Gamma_n\times I_n}^2 + \|\sqrt{\sigma_2}\,[a\nabla w]\|_{\Gamma_n\times I_n}^2. \tag{3.3.18}$$

The discrete bilinear form can be proved in the same way. □

Note that, when $n = 0$ in particular, we have

$$\mathcal{A}_0(w,w) = E(t_1^-,w) + E(t_0^+,w) + \|\sqrt{\sigma_1}\,[u]\|^2_{\Gamma_0\times I_0} + \|\sqrt{\sigma}_2\,[a\nabla w]\|^2_{\Gamma_0\times I_0} = \mathcal{B}^{\text{init}}(w). \tag{3.3.19}$$

Theorem 3.3.4 (Consistency and Stability)**.** *Let the spaces $S_n^{h,p}$ for $n = 0,\dots,N-1$ be given. Then the following statements hold:*

1. *Let u be the (weak) solution of (3.1.1) with $u_0 \in H_0^1(\Omega)$ and $v_0 \in L^2(\Omega)$. Then u satisfies (3.3.12).*

2. *For C_{σ_0} satisfying (3.3.13) and for any $v \in V^{h,p}$ and $t \in (0,T)$, the energy $E_h(t,v)$ is bounded below as*

$$E_h(t,v) \geq \frac{1}{2}\|\dot{v}(t)\|^2_\Omega + \frac{1}{4}\|\sqrt{a}\widetilde{\nabla}v(t)\|^2_\Omega. \tag{3.3.20}$$

Further, let $u^n \in S_n^{h,p}$, $n = 0,\dots,N-1$, satisfy (3.3.5). Then, $E_h(t_N^-,u^n) \leq E_h(t_1^-,u^0)$.

Proof. The first statement follows from the derivation of the formulation and the regularity of the unique solution u; see (3.1.3) and (3.1.4). To prove the second statement we proceed as follows. Combining (3.3.12) with (3.3.16) gives the energy identity

$$\begin{aligned} E_h(t_N^-,u) = \mathcal{B}^{\text{init}}(u) - E_h(t_0^+,u) &- \sum_{n=1}^{N-1}\frac{1}{2}\|[\![\dot{u}(t_n)]\!]\|^2_\Omega + \frac{1}{2}\|\sqrt{a}[\![\widetilde{\nabla}u(t_n)]\!]\|^2_\Omega \\ &+ \sum_{n=1}^{N-1}\left([\![\{a\widetilde{\nabla}u(t_n)\}]\!],[[u(t_n)]]\right)_{\hat{\Gamma}_n} - \frac{1}{2}\|\sqrt{\sigma_0}[[u(t_n)]]\|^2_{\hat{\Gamma}_n} \\ &- \sum_{n=0}^{N-1}\|\sqrt{\sigma_1}\,[u]\|^2_{\Gamma_n\times I_n} - \sum_{n=0}^{N-1}\|\sqrt{\sigma_2}\,[a\nabla u]\|^2_{\Gamma_n\times I_n}. \end{aligned} \tag{3.3.21}$$

Using (3.3.19) with the above, the energy identity (3.3.21) can be written as

$$\begin{aligned} E_h(t_N^-,u) = E_h(t_1^-,u) - \sum_{n=1}^{N-1}\Big(&\frac{1}{2}\|[\![\dot{u}(t_n)]\!]\|^2_\Omega + \frac{1}{2}\|\sqrt{a}[\![\widetilde{\nabla}u(t_n)]\!]\|^2_\Omega \\ &- \left([\![\{a\widetilde{\nabla}u(t_n)\}]\!],[[u(t_n)]]\right)_{\hat{\Gamma}_n} + \frac{1}{2}\|\sqrt{\sigma_0}[[u(t_n)]]\|^2_{\hat{\Gamma}_n} \\ &+ \|\sqrt{\sigma_1}\,[u]\|^2_{\Gamma_n\times I_n} + \|\sqrt{\sigma_2}\,[a\nabla u]\|^2_{\Gamma_n\times I_n}\Big), \end{aligned} \tag{3.3.22}$$

for u, with $u|_{\Omega\times I_n} \in S_n^{h,p}$, $n = 0, 1, \ldots, N-1$. Arguments used to prove the non-negativity of the discrete energy, also show that the discrete energy decreases at each time-step. □

3.4 Polynomial Trefftz spaces

We consider the discrete space of local polynomial solutions to the wave equation, where we make an additional assumption on the mesh and on $a(x) \in \mathbb{R}^{d\times d}$ that allows us to define the Trefftz spaces. Such polynomial spaces have been employed in the literature; see for example [83, 129, 92].

Assumption 3.4.1. *Let the diffusion coefficient $a(x)$ and the mesh be such that $a(x)$ is constant in each element $K \in \mathcal{T}_n$ for each n.*

Definition 3.4.2 (Polynomial Trefftz spaces)**.** *Let $S^{h,p}_{n,Trefftz} \subseteq S^{h,p}_n$ be a subspace of functions satisfying locally the homogeneous wave equation on any space-time element $K \times I_n$:*

$$S^{h,p}_{n,Trefftz} := \left\{ v \in S^{h,p}_n \ : \ \ddot{v}(t,x) - \nabla\cdot\left(a\,\nabla v\right)(t,x) = 0, \quad t \in I_n, x \in K,\, K \in \mathcal{T}_n \right\}.$$

The space on $\Omega \times [0,T]$ is then defined as

$$V^{h,p}_{Trefftz} = \left\{ u \in L^2(\Omega\times[0,T]) \ : \ u|_{\Omega\times I_n} \in S^{h,p}_{n,Trefftz},\ n = 0, 1\ldots, N-1 \right\} \subseteq V^{h,p}.$$

For example, functions in this space could be polynomial plane waves given by

$$(t + a^{-1/2}\alpha\cdot x)^j, \qquad |\alpha| = 1, \alpha \in \mathbb{R}^d, j \in \{0,\ldots,p\}, \tag{3.4.1}$$

where α is a direction vector.

Proposition 3.4.3. *The local dimension of the Trefftz space in $\mathbb{R}^d$ is given by*

$$\dim(S^{h,p}_{n,Trefftz}(K)) = \begin{cases} 2p+1 & d = 1 \\ (p+1)^2 & d = 2 \\ \frac{1}{6}(p+1)(p+2)(2p+3) & d = 3 \end{cases}.$$

Proof. We shall make use of the polynomial plane waves for the proof of this proposition. For $d = 1$, note that we have maximum of two directions fixed for all polynomial order $j > 0$, hence we have $2j$ polynomial plane waves. Now including piecewise constant term gives $2j + 1$ linearly independent plane waves. For $d = 3$ in [129] it is shown that the dimension of the Trefftz, homogeneous polynomials of degree j is $(j + 1)^2$ (This also corresponds to the total linearly independent plane wave polynomials for chosen order j), hence the total dimension is given by

$$\sum_{j=0}^{p}(j+1)^2 = \tfrac{1}{6}(p+1)(p+2)(2p+3).$$

The case $d = 2$ can be proved in similar way by noticing that the dimension of the Trefftz, homogeneous polynomials of degree j is $2j + 1$. □

The idea of involving the directions of propagation can be introduced in the basis functions. For example, if the dominant directions are not known a priori, the idea of equi-distributed directions of the form

$$\alpha_i = \begin{bmatrix} \cos(\frac{2\pi(i-1)}{m}) \\ \sin(\frac{2\pi(i-1)}{m}) \end{bmatrix}, \tag{3.4.2}$$

where $m = 2j + 1$ can be used to fill up the Trefftz space in two dimensions (see (3.4.1)).

Dim	1D	2D	3D
Poly	$\frac{1}{2}(p+1)(p+2)$	$\frac{(p+1)(p+2)(p+3)}{6}$	$\frac{(p+1)(p+2)(p+3)(p+4)}{24}$
Trefftz	$2p+1$	$(p+1)^2$	$\frac{(p+1)(p+2)(2p+3)}{6}$

Table 3.1: *Local dimensions for Trefftz spaces and polynomial spaces with respect to d spatial dimension.*

3.4.1 Existence and uniqueness in the Trefftz space

In this subsection, we prove the existence and uniqueness of solution in the Trefftz space by investigating the property of the obtained DG energy norm. We discover that the DG energy norm is indeed a norm on the subspace of Trefftz polynomials.

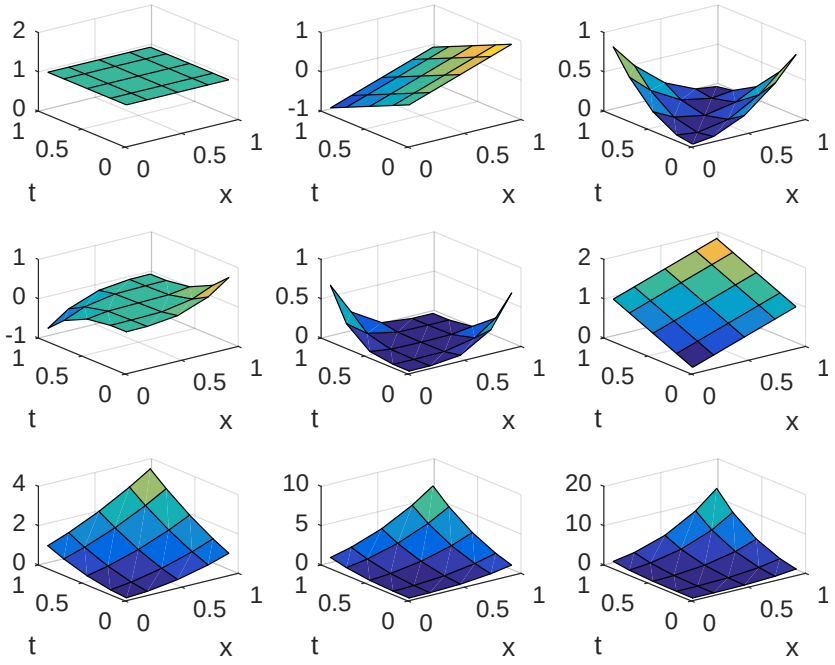


Figure 3.2: *Plot of Trefftz basis functions in* $1-$*dimension on reference space-time element* $(0,1)\times(0,1)$.

This also includes piecewise linear polynomials as $V^{h,p}_{\text{Trefftz}} = V^{h,p}$ for $p = 1$.

Proposition 3.4.4. *With the choice of* C_{σ_0} *as in* (3.3.13) *and* $\sigma_1, \sigma_2 > 0$*, bilinear forms* $\mathcal{A}_n(\cdot,\cdot)$ *and* $\mathcal{A}(\cdot,\cdot)$ *give rise to two semi-norms*

$$|||v|||_n := \big(\mathcal{A}_n(v,v)\big)^{1/2}, \qquad v \in S^{h,p}_n$$

and

$$|||v||| := \big(\mathcal{A}(v,v)\big)^{1/2}, \qquad v \in V^{h,p}.$$

These are in fact norms on Trefftz subspaces $S^{h,p}_{n,\,Trefftz}$ *and* $V^{h,p}_{Trefftz}$.

Proof. Recalling (3.3.18) and using the positivity of the energy $E_h(t,u)$ (4.5.2), we deduce that $|||v|||^2_n \geq 0$ and is hence a semi-norm.

Suppose $|||v|||_n = 0$ for $v \in S^{h,p}_{n,\text{Trefftz}}$. Then, $a\nabla v$ and v have no jumps across the space skeleton and hence v is a weak solution of the homogeneous wave equation on $\Omega \times I_n$ with zero initial and boundary conditions. Uniqueness implies $v \equiv 0$ and hence that $|||\cdot|||_n$ is a norm on this space.

The analysis of $|||\cdot|||$ is similar recalling (3.3.16), which shows that $|||\cdot|||$ is a semi-norm if the stabilization parameter is chosen correctly. Proceeding as in the first

case, shows that it is in fact a norm on the Trefftz spaces. □

Corollary 3.4.5. *Under the conditions of the above proposition and with initial data $u_0 \in H_0^1(\Omega)$, $v_0 \in L^2(\Omega)$, the discrete system* (3.3.12) *with $X = V_{Trefftz}^{h,p}$ has a unique solution.*

Proof. The uniqueness of the solution to (3.3.12) over the Trefftz space $X = V_{\text{Trefftz}}^{h,p}$ follows from $\mathcal{A}(\cdot,\cdot)$ being a norm on this space. Existence of the solution to the linear system follows from uniqueness. □

We present the convergence analysis of the Trefftz based method in the next subsection.

3.4.2 Convergence analysis

The convergence analysis guarantees quasi-optimality of the proposed method. We proceed with the following proposition.

Proposition 3.4.6. *Let $w \in \mathcal{X} + V_{Trefftz}^{h,p}$ and $v \in V_{Trefftz}^{h,p}$, then*

$$|\mathcal{A}(w,v)| \le C_\star |||w|||_\star |||v|||,$$

for some constant $C_\star > 0$ and

$$\begin{aligned}|||w|||_\star^2 =& \frac{1}{2}\sum_{n=1}^{N}\Big(\|\dot{w}(t_n^-)\|_\Omega^2 + \|\sqrt{a}\nabla w(t_n^-)\|_\Omega^2 + \|\sqrt{\sigma_0}\,[w(t_n^-)]\,\|_{\Gamma_n}^2 + \|\sigma_0^{-1/2}\,\{a\nabla w(t_n^-)\}\,\|_{\Gamma_n}^2\Big)\\ &+ \sum_{n=0}^{N-1}\Big(\|\sqrt{\sigma_1}\,[w]\|_{\Gamma_n\times I_n}^2 + \|\sqrt{\sigma_2}\,[a\nabla w]\|_{\Gamma_n\times I_n}^2 + \|\sigma_2^{-1/2}\,\{\dot{w}\}\|_{\Gamma_n^{\text{int}}\times I_n}^2\\ &\qquad + \|\sigma_1^{-1/2}\,\{a\nabla\dot{w}\}\|_{\Gamma_n\times I_n}^2 + \|\sigma_0\sigma_1^{-1/2}\,[\dot{w}]\|_{\Gamma_n\times I_n}^2\Big).\end{aligned}$$

Proof. Integrating by parts first in time and further in space we have

$$\begin{aligned}
&(\ddot{w},\dot{v})_{\Omega\times I_n} + (a\nabla w,\nabla\dot{v})_{\Omega\times I_n}\\
=\ &-(\dot{w},\ddot{v})_{\Omega\times I_n} - (a\nabla\dot{w},\nabla v)_{\Omega\times I_n}\\
&+(\dot{w}(t_{n+1}^-),\dot{v}(t_{n+1}^-))_\Omega - (\dot{w}(t_n^+),\dot{v}(t_n^+))_\Omega\\
&+(a\nabla w(t_{n+1}^-),\nabla v(t_{n+1}^-))_\Omega - (a\nabla w(t_n^+),\nabla v(t_n^+))_\Omega\\
=\ &-([\dot{w}],\{a\nabla v\})_{\Gamma_n\times I_n} - (\{\dot{w}\},[a\nabla v])_{\Gamma_n^{\mathrm{int}}\times I_n}\\
&+(\dot{w}(t_{n+1}^-),\dot{v}(t_{n+1}^-))_\Omega - (\dot{w}(t_n^+),\dot{v}(t_n^+))_\Omega\\
&+(a\nabla w(t_{n+1}^-),\nabla v(t_{n+1}^-))_\Omega - (a\nabla w(t_n^+),\nabla v(t_n^+))_\Omega,
\end{aligned}$$

since $v\in V_{\mathrm{Trefftz}}^{h,p}$, using the identity

$$-(a\nabla\dot{w},\nabla v)_{\Omega\times I_n} = (\dot{w},\nabla\cdot a\nabla v)_{\Omega\times I_n} - ([\dot{w}],\{a\nabla v\})_{\Gamma_n\times I_n} - (\{\dot{w}\},[a\nabla v])_{\Gamma_n^{\mathrm{int}}\times I_n},$$

in the second step. Further integrations by parts in time yield

$$\begin{aligned}
&-([\dot{w}],\{a\nabla v\})_{\Gamma_n\times I_n}\\
=\ &([w],\{a\nabla\dot{v}\})_{\Gamma_n\times I_n} - ([w(t_{n+1}^-)],\{a\nabla v(t_{n+1}^-)\})_{\Gamma_n} + ([w(t_n^+)],\{a\nabla v(t_n^+)\})_{\Gamma_n},
\end{aligned}$$

and

$$\begin{aligned}
&-(\{a\nabla w\},[\dot{v}])_{\Gamma_n\times I_n} + (\sigma_0[w],[\dot{v}])_{\Gamma_n\times I_n}\\
=\ &(\{a\nabla\dot{w}\},[v])_{\Gamma_n\times I_n} - (\sigma_0[\dot{w}],[v])_{\Gamma_n\times I_n}\\
&-(\{a\nabla w(t_{n+1}^-)\},[v(t_{n+1}^-)])_{\Gamma_n} + (\{a\nabla w(t_n^+)\},[v(t_n^+)])_{\Gamma_n}\\
&+(\sigma_0[w(t_{n+1}^-)],[v(t_{n+1}^-)])_{\Gamma_n} - (\sigma_0[w(t_n^+)],[v(t_n^+)])_{\Gamma_n}.
\end{aligned}$$

Substituting these into (3.3.7), we obtain

$$\begin{aligned}\mathcal{A}_n(w,v) = {} & (\{a\nabla\dot{w}\},[v])_{\Gamma_n\times I_n} - (\{a\nabla w(t_{n+1}^-)\},[v(t_{n+1}^-)])_{\Gamma_n} \\ & - (\sigma_0[\dot{w}],[v])_{\Gamma_n\times I_n} + (\sigma_0[w(t_{n+1}^-)],[v(t_{n+1}^-)])_{\Gamma_n} \\ & - (\{\dot{w}\},[a\nabla v])_{\Gamma_n^{\mathrm{int}}\times I_n} - ([w(t_{n+1}^-)],\{a\nabla v(t_{n+1}^-)\})_{\Gamma_n} \\ & + (\dot{w}(t_{n+1}^-),\dot{v}(t_{n+1}^-))_\Omega + (a\nabla w(t_{n+1}^-),\nabla v(t_{n+1}^-))_\Omega \\ & + (\sigma_1[w],[v])_{\Gamma_n\times I_n} + (\sigma_2[a\nabla w],[a\nabla v])_{\Gamma_n\times I_n}.\end{aligned} \tag{3.4.3}$$

Therefore

$$\begin{aligned}\mathcal{A}(w,v) &= \sum_{n=0}^{N-1}\mathcal{A}_n(w,v) - \sum_{n=1}^{N-1}\mathcal{B}_n(w,v) \\ &= \sum_{n=0}^{N-1}\Big((\{a\nabla\dot{w}\},[v])_{\Gamma_n\times I_n} - (\sigma_0[\dot{w}],[v])_{\Gamma_n\times I_n} - (\{\dot{w}\},[a\nabla v])_{\Gamma_n^{\mathrm{int}}\times I_n} \\ &\quad + (\sigma_1[w],[v])_{\Gamma_n\times I_n} + (\sigma_2[a\nabla w],[a\nabla v])_{\Gamma_n\times I_n}\Big) \\ &\quad + \sum_{n=1}^{N}\Big((\dot{w}(t_n^-),[\![\dot{v}(t_n)]\!])_\Omega + (a\nabla w(t_n^-),[\![\nabla v(t_n)]\!])_\Omega \\ &\quad - (\{a\nabla w(t_n^-)\},[\![[v(t_n)]]\!])_{\Gamma_n} - ([w(t_n^-)],[\![\{a\nabla v(t_n)\}]\!])_{\Gamma_n} \\ &\quad + (\sigma_0[w(t_n^-)],[\![[v(t_n)]]\!])_{\Gamma_n}\Big),\end{aligned} \tag{3.4.4}$$

where we have adopted the notational convention

$$[\![f(t_N^-)]\!] := f(t_N^-).$$

It is now clear how to estimate most of the terms to obtain the stated result using the Cauchy-Schwarz inequality. The first two terms on the right hand side in the above sum are estimated as follows

$$\Big(\sigma_1^{-1/2}(\{a\nabla\dot{w}\}-\sigma_0[\dot{w}]),\sqrt{\sigma_1}[v]\Big)_{\Gamma_n\times I_n} \le \|\sigma_1^{-1/2}(\{a\nabla\dot{w}\}-\sigma_0[\dot{w}])\|_{\Gamma_n\times I_n}\|\sqrt{\sigma_1}[v]\|_{\Gamma_n\times I_n};$$

for the third term, we have

$$(\{\dot{w}\},[a\nabla v])_{\Gamma_n^{\mathrm{int}}\times I_n} \le \|\sigma_2^{-1/2}\{\dot{w}\}\|_{\Gamma_n^{\mathrm{int}}\times I_n}\|\sqrt{\sigma_2}[a\nabla v]\|_{\Gamma_n^{\mathrm{int}}\times I_n}.$$

□

Remark 3.4.7. *Note that* (3.4.3) *shows that for Trefftz functions the bilinear form can be evaluated without computing integrals over the volume terms* $\Omega \times I_n$. *This can bring considerable savings, especially in higher spatial dimensions.*

Theorem 3.4.8. *Let* $U \in V^{h,p}_{Trefftz}$ *be the discrete solution of the Trefftz space-time discontinuous Galerkin method and let* $u \in \mathcal{X}$ *be the exact solution. Then*

$$|||U-u||| \leq \inf_{V \in V^{h,p}_{Trefftz}} \big(C_\star |||V-u|||_\star + |||V-u||| \big),$$

where $|||\cdot|||_\star$ is defined in Proposition 3.4.6.

Proof. By Galerkin orthogonality

$$\mathcal{A}(V-U,v) = \mathcal{A}(V-u,v),$$

for any $V, v \in V^{h,p}_{\mathrm{Trefftz}}$. Hence, by Proposition 3.4.6,

$$|||V-U|||^2 = \mathcal{A}(V-U,V-U) = \mathcal{A}(V-u,V-U) \leq C_\star |||V-u|||_\star |||V-U|||$$

giving

$$|||U-u||| \leq |||V-U||| + |||V-u||| \leq C_\star |||V-u|||_\star + |||V-u|||.$$

□

To conclude this subsection we show that in the case of Trefftz polynomials, the discrete norm can be bounded below by an L^2-temporal norm. For simplicity of the presentation *only*, we shall, henceforth, make use of the following assumption.

Assumption 3.4.9. *We assume that* $\mathrm{diam}(K \times I_n)/\rho_{K\times I_n} \leq c_{\mathcal{T}}$, *for all* $K \in \mathcal{T}_n$, $n = 0, 1, \ldots, N-1$.

Proposition 3.4.10. *For any* $v \in S^{h,p}_{n,Trefftz}$ *it holds*

$$\|\dot{v}\|^2_{\Omega\times I_n} + \|a^{1/2}\widetilde{\nabla} v\|^2_{\Omega\times I_n} \leq (t_{n+1}-t_n) e^{\widetilde{C}(t_{n+1}-t_n)/h} \left(\|a^{1/2}\widetilde{\nabla} v(t_n^+)\|^2_\Omega + \|\dot{v}(t_n^+)\|^2_\Omega \right),$$

where

$$\widetilde{C} = c_{\mathcal{T}} C_{inv} p^2 C_a \quad \textit{and} \qquad h = \min_{x\in\Omega} h(x,t),\ t \in I_n.$$

The same estimate holds with t_n^+ replaced by t_{n+1}^-. Consequently, with the Assumption 3.4.9 we have

$$\|\dot{V}\|^2_{\Omega\times(0,T)} + \|\sqrt{a}\widetilde{\nabla} V\|^2_{\Omega\times(0,T)} \le C e^{\widetilde{C} c_{\mathcal{T}} \tau} |||V|||^2_\star$$

for all $V \in V^{h,p}_{Trefftz}$ with a constant $C > 0$ independent of the meshsize and $\tau = \max_n \tau_n$.

Proof. Note that for an element K with exterior normal ν

$$\begin{aligned}
\frac{d}{dt}\left(\tfrac{1}{2}\|\dot{v}(t)\|^2_K + \tfrac{1}{2}\|a^{1/2}\nabla v(t)\|^2_K\right) &= \left(\ddot{v}(t), \dot{v}(t)\right)_K + \left(a\nabla v(t), \nabla\dot{v}(t)\right)_K \\
&= \left(\nu\cdot a\nabla v(t), \dot{v}(t)\right)_{\partial K} \\
&\le \|\nu\cdot a\nabla v(t)\|_{\partial K}\|\dot{v}(t)\|_{\partial K} \\
&\le \tfrac{1}{2}\|\nu\cdot a\nabla v(t)\|^2_{\partial K} + \tfrac{1}{2}\|\dot{v}(t)\|^2_{\partial K} \\
&\le C_{\mathrm{inv}} p^2 |\partial K|/|K| \left(\tfrac{1}{2}\|a\nabla v(t)\|^2_K + \tfrac{1}{2}\|\dot{v}(t)\|^2_K\right), \\
&\le C_a C_{\mathrm{inv}} p^2 c_{\mathcal{T}} h_K^{-1} \left(\tfrac{1}{2}\|\sqrt{a}\nabla v(t)\|^2_K + \tfrac{1}{2}\|\dot{v}(t)\|^2_K\right),
\end{aligned}$$

where we have used the Cauchy-Schwarz inequality in the third line, Young's inequality in the fourth line and the discrete trace inequality (2.5.6) in the fifth line. The Gronwall inequality (2.5.9) now gives us

$$\begin{aligned}
\tfrac{1}{2}\|a^{1/2}\nabla v(t)\|^2_K + \tfrac{1}{2}\|\dot{v}(t)\|^2_K &\le e^{C_{K,\max}(t-t_n)/h_K}\left(\tfrac{1}{2}\|\sqrt{a}\nabla v(t_n^+)\|^2_K + \tfrac{1}{2}\|\dot{v}(t_n^+)\|^2_K\right) \\
&\le e^{C_{K,\max}(t_{n+1}-t_n)/h_K}\left(\tfrac{1}{2}\|\sqrt{a}\nabla v(t_n^+)\|^2_K + \tfrac{1}{2}\|\dot{v}(t_n^+)\|^2_K\right)
\end{aligned}$$

as well as

$$\tfrac{1}{2}\|\sqrt{a}\nabla v(t)\|^2_K + \tfrac{1}{2}\|\dot{v}(t)\|^2_K \le e^{C_{K,\max}(t_{n+1}-t_n)/h_K}\left(\tfrac{1}{2}\|\sqrt{a}\nabla v(t_{n+1}^-)\|^2_K + \tfrac{1}{2}\|\dot{v}(t_{n+1}^-)\|^2_K\right)$$

for all $t \in [t_n, t_{n+1}]$. Integrating in time and summing over all K gives the required result.

□

The above two results allow us to conclude that we can also bound the error in a more standard norm.

Corollary 3.4.11. *Under the hypothesis of Theorem 3.4.8 and under Assumption 3.4.9, we have*

$$\|\dot{U}-\dot{u}\|_{\Omega\times[0,T]}+\|\sqrt{a}\widetilde{\nabla}(U-u)\|_{\Omega\times[0,T]}\leq C\inf_{V\in V^{h,p}_{Trefftz}}\Big(\sqrt{\tau}|||V-u|||_{\star}+\|\dot{V}-\dot{u}\|_{\Omega\times[0,T]}+\|\sqrt{a}\widetilde{\nabla}(V-u)\|_{\Omega\times[0,T]}\Big),$$

for some constant C independent of u, U and the mesh parameters.

Proof. Using the triangle inequality and the above proposition we have that for any $V\in V^{h,p}_{\text{Trefftz}}$

$$\begin{aligned}\|\dot{U}-\dot{u}\|_{\Omega\times[0,T]}+\|\sqrt{a}\widetilde{\nabla}(U-u)\|_{\Omega\times[0,T]}&\leq\|\dot{U}-\dot{V}\|_{\Omega\times[0,T]}+\|\sqrt{a}\widetilde{\nabla}(U-V)\|_{\Omega\times[0,T]}\\&\quad+\|\dot{V}-\dot{u}\|_{\Omega\times[0,T]}+\|\sqrt{a}\widetilde{\nabla}(V-u)\|_{\Omega\times[0,T]}\\&\leq C\sqrt{\tau}|||U-V|||_{\star}+\|\dot{V}-\dot{u}\|_{\Omega\times[0,T]}+\|\sqrt{a}\widetilde{\nabla}(V-u)\|_{\Omega\times[0,T]}.\end{aligned}$$

□

3.5 A priori error bounds

The special form of the exact solutions to the wave equation give rise to special approximation results, for which, as we will show, the Trefftz basis is sufficient to deliver the expected rates of convergence for the proposed method. The quasi-optimality estimate below does not depend on the use of the Gronwall's inequality that is often used for the a priori bound for time-dependent problems (see, [23, 126, 77, 124, 84, 71]).

Lemma 3.5.1. *Let the setting of Theorem 3.4.8 hold, let $V\in V^{h,p}_{Trefftz}$ be an arbitrary*

function in the discrete space, and let $\eta = u - V$. *Then*

$$\begin{aligned}|||U-u|||^2 \le C\sum_{n=0}^{N-1}\sum_{K\in\mathcal{T}^n} C_a\Big(&\frac{p^2}{\tau_n}\Big(\min\{1,\frac{\tau_n^2}{h_K^2}\}\|\dot\eta\|^2_{K\times I_n}+\|\nabla\eta\|^2_{K\times I_n}\Big)\\ &+\tau_n\Big(\|\nabla\dot\eta\|^2_{K\times I_n}+\min\{1,\frac{h_K^2}{\tau_n^2}\}p^{-1}\|D^2\eta\|^2_{K\times I_n}\Big)\\ &+\frac{(h_K^2\tau_n)}{p^4}\|D^2\dot\eta\|^2_{K\times I_n}+\frac{p^4}{h_K^2\tau_n}\|\eta\|^2_{K\times I_n}\Big),\end{aligned} \tag{3.5.1}$$

where $U \in V^{h,p}_{Trefftz}$ *is the discrete solution.*

Proof. Theorem 3.4.8 implies

$$|||U-u||| \le C_\star|||\eta|||_\star + |||\eta|||.$$

We shall now estimate each term of the norms on the right-hand side. We shall repeatedly use the standard trace estimate $\|v\|^2_{\partial\omega} \le C(\mathrm{diam}(\omega)^{-1}\|v\|^2_\omega + \mathrm{diam}(\omega)\|\nabla v\|^2_\omega)$, for $v \in H^1(\omega)$, where ω is a subset of $\mathbb{R}^k$, $k = 1\ldots, d+1$. We proceed as follows

$$\begin{aligned}\sum_{n=1}^{N}\|\dot\eta(t_n^-)\|^2_\Omega &= \sum_{n=1}^{N}\sum_{K\in\mathcal{T}^{n-1}}\|\dot\eta\|^2_{K\times\{t_n^-\}} \le C\sum_{n=0}^{N-1}\sum_{K\in\mathcal{T}^n}\Big(\frac{C_a p}{\tau_n}\|\dot\eta\|^2_{K\times I_n}+\frac{\tau_n}{C_a p}\|\ddot\eta\|^2_{K\times I_n}\Big)\\ &\le C\sum_{n=0}^{N-1}\sum_{K\in\mathcal{T}^n}\Big(\frac{C_a p}{\tau_n}\|\dot\eta\|^2_{K\times I_n}+\frac{\tau_n}{C_a p}\|\nabla\cdot a(\cdot)\nabla\eta\|^2_{K\times I_n}\Big)\\ &\le C\sum_{n=0}^{N-1}\sum_{K\in\mathcal{T}^n} C_a\Big(\frac{p}{\tau_n}\|\dot\eta\|^2_{K\times I_n}+\frac{\tau_n}{p}\|\Delta\eta\|^2_{K\times I_n}\Big).\end{aligned}$$

We prefer to retain an explicit dependence on the polynomial degree p at this point, as it will be of relevance in the error analysis for $d = 1$. In analogous fashion, we also have

$$\sum_{n=1}^{N}\|\sqrt{a}\nabla\eta(t_n^-)\|^2_\Omega \le C\sum_{n=0}^{N-1}\sum_{K\in\mathcal{T}^n} C_a\Big(\frac{p}{\tau_n}\|\nabla\eta\|^2_{K\times I_n}+\frac{\tau_n}{p}\|\nabla\dot\eta\|^2_{K\times I_n}\Big).$$

Next, we estimate the penalty term:

$$\begin{aligned}\sum_{n=1}^{N}\|\sqrt{\sigma_0}\,[\eta(t_n^-)]\,\|_{\Gamma_n}^2 &\le C\sum_{n=0}^{N-1}\sum_{K\in\mathcal{T}^n} C_a\Big(\frac{p^3}{\tau_n h_K}\|\eta\|_{\partial K\times I_n}^2 + p\frac{\tau_n}{h_K}\|\dot\eta\|_{\partial K\times I_n}^2\Big)\\ &\le C\sum_{n=0}^{N-1}\sum_{K\in\mathcal{T}^n} C_a\Big(\frac{p^4}{\tau_n h_K^2}\|\eta\|_{K\times I_n}^2 + \frac{p^2}{\tau_n}\|\nabla\eta\|_{K\times I_n}^2\\ &\qquad + \frac{p^2\tau_n}{h_K^2}\|\dot\eta\|_{K\times I_n}^2 + \tau_n\|\nabla\dot\eta\|_{K\times I_n}^2\Big).\end{aligned}$$

Similarly, we also have

$$\begin{aligned}\sum_{n=1}^{N}\|\sigma_0^{-1/2}\,\{a\nabla\eta(t_n^-)\}\,\|_{\Gamma_n}^2 &\le C\sum_{n=0}^{N-1}\sum_{K\in\mathcal{T}^n} C_a\Big(\frac{h_K}{p\tau_n}\|\nabla\eta\|_{\partial K\times I_n}^2 + \frac{\tau_n h_K}{p^3}\|\nabla\dot\eta\|_{\partial K\times I_n}^2\Big)\\ &\le C\sum_{n=0}^{N-1}\sum_{K\in\mathcal{T}^n} C_a\Big(\frac{1}{\tau_n}\|\nabla\eta\|_{K\times I_n}^2 + \frac{h_K^2}{p^2\tau_n}\|D^2\eta\|_{K\times I_n}^2\\ &\qquad + \frac{\tau_n}{p^2}\|\nabla\dot\eta\|_{K\times I_n}^2 + \frac{h_K^2\tau_n}{p^4}\|D^2\dot\eta\|_{K\times I_n}^2\Big).\end{aligned}$$

Next, we choose $\sigma_1|_{\partial K\cap\Gamma_n\times I_n} = C_a p^3/(h\tau_n)$, and we have

$$\sum_{n=0}^{N-1}\|\sqrt{\sigma_1}\,[\eta]\|_{\Gamma_n\times I_n}^2 \le C\sum_{n=0}^{N-1}\sum_{K\in\mathcal{T}^n} C_a\Big(\frac{p^4}{\tau_n h_K^2}\|\eta\|_{K\times I_n}^2 + \frac{p^2}{\tau_n}\|\nabla\eta\|_{K\times I_n}^2\Big).$$

Further, we choose $\sigma_2 = h_K/(C_a\tau_n)$, to have

$$\sum_{n=0}^{N-1}\|\sqrt{\sigma_2}\,[a\nabla\eta]\|_{\Gamma_n\times I_n}^2 \le C\sum_{n=0}^{N-1}\sum_{K\in\mathcal{T}^n} C_a\Big(\frac{p^2}{\tau_n}\|\nabla\eta\|_{K\times I_n}^2 + \frac{h_K^2}{p^2\tau_n}\|D^2\eta\|_{K\times I_n}^2\Big).$$

The next term is treated as follows:

$$\sum_{n=0}^{N-1}\|\sigma_2^{-1/2}\,\{\dot\eta\}\|_{\Gamma_n^{\mathrm{int}}\times I_n}^2 \le C\sum_{n=0}^{N-1}\sum_{K\in\mathcal{T}^n} C_a\Big(\frac{p^2\tau_n}{h_K^2}\|\dot\eta\|_{K\times I_n}^2 + \frac{\tau_n}{p^2}\|\nabla\dot\eta\|_{K\times I_n}^2\Big).$$

Continuing, we have

$$\sum_{n=0}^{N-1}\|\sigma_1^{-1/2}\,\{a\nabla\dot\eta\}\|_{\Gamma_n\times I_n}^2 \le C\sum_{n=0}^{N-1}\sum_{K\in\mathcal{T}^n} C_a\Big(\frac{\tau_n}{p^2}\|\nabla\dot\eta\|_{K\times I_n}^2 + \frac{\tau_n h_K^2}{p^4}\|D^2\dot\eta\|_{K\times I_n}^2\Big).$$

Finally, we estimate

$$\sum_{n=0}^{N-1}\|\sigma_0\sigma_1^{-1/2}\,[\dot u]\|_{\Gamma_n\times I_n}^2 \le C\sum_{n=0}^{N-1}\sum_{K\in\mathcal{T}^n} C_a\Big(\frac{p^2\tau_n}{h_K^2}\|\dot\eta\|_{K\times I_n}^2 + \tau_n\|\nabla\dot\eta\|_{K\times I_n}^2\Big).$$

The remaining terms in $|||\eta|||$ are treated completely analogously. □

To complete the error analysis, we need to prove the existence of an appropriate approximation $u_h \in V_{\text{Trefftz}}^{h,p}$ of the exact solution. If the exact solution is sufficiently many times continuously differentiable within each space-time element, we can construct such an approximation locally with the following proposition.

Proposition 3.5.2. *Let $J \subset \mathbb{R}^{d+1}$ be a star-shaped with respect to a ball $B \subset J$. Then there exists a projector*

$$\Pi^p : H^{p+1}(J) \to \mathcal{P}_p(J)$$

such that for any $v \in H^{p+1}(J)$

$$\|D^\beta(v - \Pi^p v)\|_J \leq C(\operatorname{diam}(J))^{p+1-\beta}\|v\|_{H^{p+1}(J)}, \quad |\beta| \leq p, \tag{3.5.2}$$

and further if v satisfies the wave equation $\ddot{v} - \nabla \cdot a\nabla v = 0$ in J then so does $\Pi^p v$. The constant C depends on p and on the shape of J.

Proof. We can define $\Pi^p v$ to be the averaged Taylor polynomial of order p centered at y and evaluated at x, i.e.,

$$\Pi^p v(x) = \sum_{|\alpha| \leq p} \frac{1}{\alpha!} \int_B D^\alpha v(y)(x-y)^\alpha \phi(y) dy, \tag{3.5.3}$$

where $\phi \in \mathcal{C}_0^\infty(\mathbb{R}^{d+1})$ is an arbitrary cut-off function satisfying $\int_B \phi = 1$ and $\operatorname{supp}\phi = \overline{B}$; [see Chapter 4 of [20]]. Then the Bramble-Hilbert lemma gives us the approximation property required, [see Lemma 4.3.8 of [20]]. Therefore, it only remains to show that $\Pi^p v$ satisfies the wave equation if v does. For the cases $p \leq 1$, the proof is clear. For the cases, $p \geq 2$, the result follows from the property of averaged Taylor polynomials

$$D^\alpha \Pi^p v = \Pi^{p-|\alpha|} D^\alpha v, \quad |\alpha| \leq p.$$

□

Applying such a projector to the exact solution and combining this with Lemma 3.5.1 gives us a proof of the convergence order of the discrete scheme.

Theorem 3.5.3. *Let the exact solution $u \in \mathcal{X}$ be such that for each space time element $K \times I_n$, $u|_{K\times I_n} \in H^{s+1}(K \times I_n)$ for some $0 \le s \le p$. Then*

$$|||U - u||| \le C\left(\sum_{n=0}^{N-1}\sum_{K\in\mathcal{T}^n} (h_K^{(n)})^{2s-1}\|u\|^2_{H^{s+1}(K\times I_n)}\right)^{1/2} \le C(u)h^{s-1/2}, \tag{3.5.4}$$

where $h = \max_{K,n} h_K^{(n)}$ and

$$C(u) = \left(\sum_{n=0}^{N-1}\sum_{K\in\mathcal{T}^n} \|u\|^2_{H^{s+1}(K\times I_n)}\right)^{1/2}.$$

The above estimate gives the upper bound for a standard space-time energy norm (see Chapter 5).

3.5.1 *hp*-version error analysis for $d = 1$

We show in this subsection that the Trefftz basis is sufficient to deliver the expected *hp*-version a priori error bounds for $d = 1$, along with a proof of the exponential convergence of the *p*-version space-time DG method for the case of analytic exact solutions.

To discuss the Trefftz-basis case for $d = 1$, let $K = [x_0, x_1]$, and start from the basic observation that the exact solution to the wave equation on each space time element is of the form

$$u(x,t)|_{K\times I_n} = F^1_{n,K}(a^{-1/2}x + t) + F^2_{n,K}(a^{-1/2}x - t), \tag{3.5.5}$$

where we can define F^1 and F^2 by

$$F^1_{n,K}(a^{-1/2}x + t) = \tfrac{1}{2}u(x,t) + \tfrac{1}{2}v(x,t), \qquad (x,t) \in K \times I_n$$

and

$$F^2_{n,K}(a^{-1/2}x - t) = \tfrac{1}{2}u(x,t) - \tfrac{1}{2}v(x,t), \qquad (x,t) \in K \times I_n,$$

where

$$v(x,t) = a^{1/2}\int_{t_n}^{t} u_x(x,\tau)d\tau + a^{-1/2}\int_{x_0}^{x} u_t(x',t_n)dx'.$$

It is not difficult to see that these are well-defined, i.e., that the right-hand sides indeed depend only on $a^{-1/2}x \pm t$ by virtue of satisfying the equations $a^{1/2}f_x \mp f_t = 0$ respectively.

For $\hat{I} := (-1,1)$, we define the H^1-projection operator $\hat{\lambda}_p : H^1(\hat{I}) \to \mathcal{P}_p(\hat{I})$, $p \geq 1$, defined by setting, for $\hat{u} \in H^1(\hat{I})$,

$$(\hat{\lambda}_p \hat{u})(x) := \int_{-1}^{x} \hat{\pi}_{p-1}(\hat{u}')(\eta)\,\mathrm{d}\eta + \hat{u}(-1), \qquad x \in \hat{I},$$

with $\hat{\pi}_{p-1}$ being the L^2-orthogonal projection operator onto $\mathcal{P}_{p-1}(\hat{I})$.

Now, upon considering the linear scalings $\psi^1_{n,K} : \hat{I} \to J^1_{n,K}$, $K \in \mathcal{T}_n$, such that

$$J^1_{n,K} := \big(\min_{(x,t)\in K\times I_n}\{x+ct\}, \max_{(x,t)\in K\times I_n}\{x+ct\}\big),$$

and $\psi^2_{n,K} : \hat{I} \to J^2_{n,K}$, $K \in \mathcal{T}_n$, such that

$$J^2_{n,K} := \big(\min_{(x,t)\in K\times I_n}\{x-ct\}, \max_{(x,t)\in K\times I_n}\{x-ct\}\big),$$

we define the univariate space-time elemental projection operators λ^i_p, $i = 1,2$, piecewise by

$$(\lambda^i_p F)|_{J^i_{n,K}} := \hat{\lambda}^i_p((F \circ \psi^i_{n,K})|_{\hat{I}}), \quad K \in \mathcal{T}_n,\ n = 0,1,\ldots,N-1.$$

Using these, we can now define the *Trefftz projection* $\Pi_p u$ of a function u of the form (3.5.5) element-wise by

$$(\Pi_p u)|_{K\times I_n} := \lambda^1_p F^1_{n,K}(x+ct) + \lambda^2_p F^2_{n,K}(x-ct), \tag{3.5.6}$$

$K \in \mathcal{T}_n$, $n = 0,1,\ldots,N-1$. The approximation properties of Π_p follow from the respective properties of λ^i_p, $i = 1,2$. Space-time shape regularity implies $J^i_{n,K} \sim h^{(n)}_K$, $i = 1,2$.

We denote by $\Phi(p,s)$ the quantity $\Phi(p,s) := (\Gamma(p-s+1)/\Gamma(p+s+1))^{\frac{1}{2}}$, with p, s real numbers such that $0 \leq s \leq p$ and $\Gamma(\cdot)$ being the Gamma function; we also adopt the standard convention $\Gamma(1) = 0! = 1$. Making use of *Stirling's formula*,

$\Gamma(n) \sim \sqrt{2\pi} n^{n-\frac{1}{2}} e^{-n}$, $n > 0$, we have, $\Phi(p,s) \leq Cp^{-s}$, for $p \geq 1$, with $0 \leq s \leq p$ and $C > 0$ constant depending only on s.

We have the following *hp*-approximation results for λ_p^i, $i = 1, 2$.

Lemma 3.5.4. *Let $v \in H^{k+1}(J)$, for $k \geq 1$, and let $h = \mathrm{diam}(J)$ with J; finally let λ_p be any of the λ_p^i, $i = 1, 2$. Then the following error bounds hold:*

$$\|v - \lambda_p v\|_J \leq Cp^{-1}\Phi(p,s)h^{s+1}|v|_{s+1,J}, \tag{3.5.7}$$

and

$$\|v' - (\lambda_p v)'\|_J \leq C\Phi(p,s)h^{s}|v|_{s+1,J}, \tag{3.5.8}$$

with $0 \leq s \leq \min\{p, k\}$, $p \geq 1$.

Also, let $v \in H^{k+1}(J)$, with $k \geq 2$. Then, the following bound holds:

$$\|v'' - (\lambda_p v)''\|_J \leq Cp^{3/2}\Phi(p,m)h^{m-1}|v|_{m+1,J}, \tag{3.5.9}$$

with $1 \leq m \leq \min\{p, k\}$. Finally, let $v \in H^{k+1}(J)$, with $k \geq 3$. Then, the following bound holds:

$$\|v''' - (\lambda_p v)'''\|_J \leq Cp^{7/2}\Phi(p,l)h^{l-2}|v|_{l+1,J}, \tag{3.5.10}$$

with $2 \leq l \leq \min\{p-1, k\}$.

Proof. The proof of (3.5.7) and (3.5.8) for the H^1-projection λ_p^i can be found in [Theorem 3.17 of [113]]. For the proof of (3.5.9) [see, Theorem 4.2 of [55]] while the proof of (3.5.10) follows along the same lines as in the proof of (3.5.9) from [55]. □

These *hp*-approximation estimates imply the following bound.

Theorem 3.5.5. *Let $u|_{K\times I_n} \in H^{k+1}(K \times I_n)$, for $k \geq 3$ be the exact solution to* (3.1.1)*. Then, for space-time meshes satisfying Assumption 3.4.9, the following error bounds hold:*

$$|||U - u|||^2 \leq Cp^3\Phi^2(p,s) \sum_{n=0}^{N-1} \sum_{K\in\mathcal{T}^n} \mathrm{diam}(K \times I_n)^{2s-1}|u|^2_{s+1,K\times I_n}, \tag{3.5.11}$$

for $3 \leq s \leq \min p+1, k$ *and* $h = \max_{K,n}\{\operatorname{diam}(K \times I_n)\}$, *with* $C(u) > 0$ *constant, independent of* p, h u, *and* U. *Moreover, if* u *is analytic on a neighbourhood of* Ω, *there exists* $r > 0$, *depending on the analyticity region of* u *in a neighbourhood of* $\Omega \times (0,T)$, *such that*

$$|||U-u|||^2 \leq C(u)p^3 \exp(-rp) \sum_{n=0}^{N-1} \sum_{K\in\mathcal{T}^n} |K \times I_n| \operatorname{diam}(K \times I_n)^{2s-1}. \tag{3.5.12}$$

Proof. The proof of (3.5.11) follows by combining the *hp*-approximation bounds from (3.5.4) with Lemma 3.5.1.

For (3.5.12), we work as follows. Analyticity of u implies that there exists a $d > 0$, such that for all $s \geq 0$,

$$|u|_{s,K\times I_n} \leq C d^s \Gamma(s+1) |K \times I_n|^{1/2}. \tag{3.5.13}$$

Using this, setting $s = \gamma p$ for some $0 < \gamma < 1$, along with Stirling's formula, we arrive at the bound

$$\Phi^2(p,\gamma p)|u|^2_{\gamma p+1,K\times I_n} \leq C\left((2\gamma d)^{2\gamma}\frac{(1-\gamma)^{1-\gamma}}{(1+\gamma)^{1+\gamma}}\right)^p |K \times I_n|,$$

with the precise choice of γ remaining at our disposal. The function

$$F(\gamma) := (2\gamma d)^{2\gamma}\frac{(1-\gamma)^{1-\gamma}}{(1+\gamma)^{1+\gamma}},$$

has a minimum at $\gamma_{\min} := (1+4d^2)^{-1/2}$, giving $F(\gamma_{\min}) < 1$. Setting, now $r = 1/2|\log F(\gamma_{\min})|$, the result follows. □

Remark 3.5.6. *The bound* (3.5.11) *is suboptimal in* p *by one order. This is a standard feature of hp-version DG methods whose analysis requires the use of hp-type inverse estimates. It is possible to slightly improve on this result and obtaining only* $1/2$ *order* p*-suboptimal bounds, using the classical hp-approximation results from [21, 9], instead of the* H^1*-projection operator as done above. These results, however, are not suitable for the proof of the exponential rate of* p*-convergence.*

3.6 Space-time DG with transparent boundary condition

In this section we extend the analysis of the space-time DG technique to the case of mixed boundary conditions. We define the undamped wave problem with Dirichlet boundary condition and transparent boundary condition. We begin with the development of the scheme for this case and we proceed with a basic a priori analysis which includes the existence and uniqueness analysis as well as consistency and stability of the scheme. The numerical simulations which show the behaviour of the method in line with theoretical findings are presented in Chapter 5.

3.6.1 Model problem

We consider the wave problem

$$\begin{aligned} \ddot{u} - \nabla \cdot (a \nabla u) &= 0 && \text{in } \Omega \times [0,T], \\ u &= 0 && \text{on } \Gamma_D \times [0,T], \\ \partial_n u + \dot{u} &= 0 && \text{on } \Gamma_T \times [0,T], \\ u(x,0) = u_0(x),\ \dot{u}(x,0) &= v_0(x), && \text{in } \Omega, \end{aligned} \tag{3.6.1}$$

where Γ_D and Γ_T denote the Dirichlet and transparent boundaries respectively and the whole boundary $\partial\Omega = \Gamma_D \cup \Gamma_T$. Also $\Gamma_D \cap \Gamma_T = \varnothing$.

For the development of the approximating scheme, we follow the energy argument as before. We assume again that the exact solution u of the system is smooth and we let $v \in \mathcal{X} + V^{h,p}$. The corresponding standard symmetric DG formulation when the system is tested by $\dot{v}$ is given by

$$\begin{aligned} &(\ddot{u}, \dot{v})_{\Omega \times I_n} + (a \widetilde{\nabla} u, \widetilde{\nabla} \dot{v})_{\Omega \times I_n} - (\{a \nabla u\}, [\dot{v}])_{\Gamma^n_{int} \cup \Gamma_D \times I_n} \\ &-([u], \{a \nabla \dot{v}\})_{\Gamma^n_{int} \cup \Gamma^n_D \times I_n} - (\sigma_0 [u], [\dot{v}])_{\Gamma^n_{int} \cup \Gamma^n_D \times I_n} + (a\dot{u}, \dot{v})_{\Gamma^n_T \times I_n} = 0. \end{aligned} \tag{3.6.2}$$

Note that the Dirichlet boundary is penalised and stabilised along with the interior

interfaces while the transparent boundary is separated. Choosing $v = u$ in (3.6.2) motivates the use of the following discrete *energy*

$$E_h(t,u) = \frac{1}{2}\|\dot{u}\|^2_{L^2(\Omega)} + \frac{1}{2}\|\sqrt{a}\widetilde{\nabla}u\|^2_{L^2(\Omega)} - (\{\nabla u\},[u])_{\Gamma^n_{int}\cup\Gamma^n_D} + \frac{1}{2}\|\sqrt{\sigma_0}\,[u]\,\|^2_{(\Gamma^n_{int}\cup\Gamma_D)}. \tag{3.6.3}$$

Now following the same steps as in Section 3.3 leads to the following weak formulation for the problem

$$\begin{aligned}
&\sum_{n=0}^{N-1}(\ddot{u},\dot{v})_{\Omega\times I_n} + (\llbracket\dot{u}(t_n)\rrbracket,\dot{v}(t_n^+))_\Omega \\
&+ (a\widetilde{\nabla}u,\widetilde{\nabla}\dot{v})_{\Omega\times I_n} + (\llbracket a\widetilde{\nabla}u(t_n)\rrbracket,\widetilde{\nabla}v(t_n^+))_\Omega \\
&- (\{a\nabla u\},[\dot{v}])_{\Gamma^n_{int}\cup\Gamma^n_D\times I_n} - (\llbracket\{\widetilde{\nabla}u(t_n)\}\rrbracket,[v])_{\Gamma^n_{int}\cup\Gamma^n_D} \\
&- ([u],\{a\nabla\dot{v}\})_{\Gamma^n_{int}\cup\Gamma^n_D\times I_n} - (\llbracket[u(t_n)]\rrbracket,\{a\nabla v\})_{\Gamma^n_{int}\cup\Gamma^n_D} \\
&- (\sigma_0[u],[\dot{v}])_{\Gamma^n_{int}\cup\Gamma^n_D\times I_n} + (\sigma_0\llbracket[u(t_n)]\rrbracket,[v(t_n^+)])_{\Gamma^n_{int}\cup\Gamma^n_D} \\
&+ (\sigma_1[u(t_n)],[v(t_n)])_{\Gamma^n_{int}\times I_n} + (\sigma_2[\nabla u(t_n)],[\nabla v(t_n)])_{\Gamma^n_{int}\times I_n} + (a\dot{u},\dot{v})_{\Gamma^n_T\times I_n} = \mathcal{B}^{\text{init}}(v),
\end{aligned} \tag{3.6.4}$$

where

$$\begin{aligned}
\mathcal{B}^{\text{init}}(v) = {}& (v_0,\dot{v}(t_0^+))_\Omega + \left(\widetilde{\nabla}u_0,\widetilde{\nabla}v(t_0^+)\right)_\Omega \\
&- (\{\nabla u_0\},[v(t_0^+)])_{\Gamma^n_{int}\cup\Gamma^n_D} - ([u_0],\{\nabla v(t_0^+)\})_{\Gamma^n_{int}\cup\Gamma^n_D} \\
&+ (\sigma_0[u_0],[v(t_0^+)])_{\Gamma^n_{int}\cup\Gamma^n_D}.
\end{aligned} \tag{3.6.5}$$

We have again arrived at a space-time discrete scheme for the stated wave problem which can be described in two ways: as a method for obtaining a discrete solution on a fixed space-time domain $\Omega\times[0,T]$ or as a time-stepping method. We define

the following bilinear forms to present the two viewpoints:

$$\begin{aligned}\mathcal{A}_n(u,v) &= (\ddot{u},\dot{v})_{\Omega\times I_n} + (\dot{u}(t_n^+),\dot{v}(t_n^+))_\Omega \\ &+ (a\widetilde{\nabla}u,\widetilde{\nabla}\dot{v})_{\Omega\times I_n} + (\widetilde{\nabla}u(t_n^+),\widetilde{\nabla}v(t_n^+))_\Omega \\ &- (\{a\nabla u\},[\dot{v}])_{\Gamma_{int}^n\cup\Gamma_D^n\times I_n} - (\{\nabla u(t_n^+)\},[v])_{\Gamma_{int}^n\cup\Gamma_D^n} \\ &- ([u],\{a\nabla\dot{v}\})_{\Gamma_{int}^n\cup\Gamma_D^n\times I_n} - ([u(t_n^+)],\{\nabla v\})_{\Gamma_{int}^n\cup\Gamma_D^n} \\ &- (\sigma_0[u],[\dot{v}])_{\Gamma_{int}^n\cup\Gamma_D^n\times I_n} + (\sigma_0[u(t_n^+)],[v(t_n^+)])_{\Gamma_{int}^n\cup\Gamma_D^n} \\ &+ (\sigma_1[u(t_n)],[v(t_n)])_{\Gamma_D^n\cup\Gamma_{int}^n\times I_n} + (\sigma_2[\nabla u(t_n)],[\nabla v(t_n)])_{\Gamma_{int}^n\times I_n} + (\dot{u},\dot{v})_{\Gamma_T^n\times I_n}\end{aligned} \tag{3.6.6}$$

$$\begin{aligned}\mathcal{B}_n(u,v) &= (\dot{u}(t_n^-),\dot{v}(t_n^+))_\Omega + (\nabla u(t_n^-),\nabla v(t_n^+))_\Omega \\ &- (\{\nabla u(t_n^-)\},[v(t_n^+)])_{\Gamma_{int}^n\cup\Gamma_D^n} - ([u(t_n^-)],\{a\nabla v(t_n^+)\})_{\Gamma_{int}^n\cup\Gamma_D^n} \\ &+ (\sigma_0[u(t_n^-)],[v(t_n^+)])_{\Gamma_{int}^n\cup\Gamma_D^n}\end{aligned} \tag{3.6.7}$$

and

$$\mathcal{A}(u,v) = \sum_{n=0}^{N-1} a_n(u,v) - \sum_{n=1}^{N-1} b_n(u,v), \tag{3.6.8}$$

which is equal to the left-hand side in (3.6.4).

Definition 3.6.1. *Given subspaces $X_n \subseteq S_n^{h,p}$, the time stepping method is described by: find $u^n \in X_n, n = 0,\dots,N-1$ such that*

$$\mathcal{A}_n(u^n,v) = \mathcal{B}_n(u^{n-1},v), \qquad \text{for all } v \in X_n, \tag{3.6.9}$$

and

$$\mathcal{A}_0(u^0,v) = \mathcal{B}^{init}(v), \qquad \text{for all } v \in X_0. \tag{3.6.10}$$

Equivalently, given a subspace $X \subseteq V^{h,p}$, the full space-time discrete system is presented as : find $u \in X$ such that

$$\mathcal{A}(u,v) = \mathcal{B}^{init}(v), \qquad \text{for all } v \in X. \tag{3.6.11}$$

Lemma 3.6.2. *The following identities hold for $u \in \mathcal{X} + V^{h,p}$:*

$$\mathcal{A}_n(u,u) = E_h(t_{n+1}^-,u) + E_h(t_n^+,u) + \|\sqrt{\sigma_1}\,[u]\,\|^2_{\Gamma^n\times I_n} + \|\sqrt{\sigma_2}\,[u]\,\|^2_{\Gamma^n\times I_n} + \|\sqrt{a}\dot{v}\|^2_{\Gamma_T\times I_n}, \tag{3.6.12}$$

for $n = 0,\dots,N-1$, $\Gamma^n = \Gamma^n_{int} \cup \Gamma^n_D$ and

$$\begin{aligned}\mathcal{A}(u,u) &= E_h(t_N^-,u) + E_h(t_0^+,u) + \sum_{n=1}^{N-1}\frac{1}{2}\|[\![\dot{u}(t_n)]\!]\|^2_\Omega + \sum_{n=1}^{N-1}\frac{1}{2}\|[\![a\widetilde{\nabla}u(t_n)]\!]\|^2_\Omega \\ &+ \sum_{n=1}^{N-1}([\![\{a\widetilde{\nabla}u(t_n)\}]\!],[[u(t_n)]])_{\hat{\Gamma_n}} + \sum_{n=1}^{N-1}\frac{1}{2}\|[\![\sqrt{\sigma_0}\,[u(t_n)]]\!]\|^2_{\Gamma^n} \\ &+ \sum_{n=0}^{N-1}\|\sigma_1\,[u]\,\|^2_{\Gamma^n\times I_n} + \sum_{n=0}^{N-1}\|\sqrt{\sigma_2}\,[\nabla u]\,\|^2_{\Gamma^n\times I_n} + \sum_{n=0}^{N-1}\|\dot{u}(t_n)\|^2_{\Gamma_T\times I_n}\end{aligned} \tag{3.6.13}$$

Proof. The identities follow from the definition of the bilinear forms and the energy $E_h(t,u)$, see 3.3.3. □

Theorem 3.6.3 (Consistency and Stability)**.** *Let the spaces $S_n^{h,p}$ for $n = 0,\dots,N-1$ be given. Then the following statements hold:*

1. *Let u be the (weak) solution of (3.6.1) with $u_0 \in H^1_0(\Omega)$ and $v_0 \in L^2(\Omega)$. Then u satisfies (3.6.11).*

2. *For sufficiently large σ_0, and for any $v \in S_n^{h,p}$ and $t \in I_n$, the energy $E_h(t,v)$ is bounded below as*

$$E_h(t,v) \geq \frac{1}{2}\|\dot{v}(t)\|^2_\Omega + \frac{1}{4}\|\nabla v(t)\|^2_\Omega. + \frac{1}{4}\|\sqrt{\sigma_0}\,[v(t)]\|^2_{\Gamma_n}. \tag{3.6.14}$$

Further, let $u^n \in S_n^{h,p}$, $n = 0,\dots,N-1$, satisfy (3.6.4). Then

$$E_h(t_N^-,u^n) \leq E_h(t_1^-,u^0).$$

Proof. Statement 1 follows from the derivation of the formulation and the regularity of the unique solution u. We discover that the jump terms vanish when u is substituted into the scheme and we are left with the original problem (3.6.1). We have already shown that the energy $E_h(t,v)$ is non-negative and bounded below by same quantity in 3.3.2. For the rest of the proof, we follow the same argument in 3.3.4 to

arrive at

$$
\begin{aligned}
E_h(t_N^-) = E_h(t_1^-, u) - \sum_{n=1}^{N-1} \Bigg(& \frac{1}{2}\|[\![\dot{u}(t_n)]\!]\|_\Omega^2 + \frac{1}{2}\|[\![\sqrt{a}\widetilde{\nabla} u(t_n)]\!]\|_\Omega^2 \\
& - ([\![\{a\nabla u(t_n)\}]\!], [\![[u(t_n)]]\!])_{\Gamma^n} + \frac{1}{2}\|[\![\sqrt{\sigma_0}\,[u(t_n)]]\!]\|_{\Gamma^n}^2 \\
& + \|\sqrt{\sigma_1}\,[u]\,\|_{\Gamma^n \times I_n}^2 + \|\sqrt{\sigma_2}\,[\nabla u]\,\|_{\Gamma_{int}^n \times I_n}^2 \Bigg) - \sum_{n=1}^{N-1} \|\dot{u}(t_n)\|_{\Gamma_T^n \times I_n}^2
\end{aligned} \tag{3.6.15}
$$

which also shows that the method is dissipative. □

Proposition 3.6.4. *With σ_0 chosen sufficiently large and $\sigma_1, \sigma_2 > 0$, the bilinear forms $a_n(\cdot,\cdot)$ and $a(\cdot,\cdot)$ derived for* (3.6.1) *give rise to two semi-norms*

$$
|||v|||_n := \big(a_n(v,v)\big)^{1/2}, \qquad v \in S_n^{h,p}
$$

and

$$
|||v||| := \big(a(v,v)\big)^{1/2}, \qquad v \in V^{h,p}.
$$

These are in fact norms on Trefftz subspaces $S_{n,Trefftz}^{h,p}$ and $V_{Trefftz}^{h,p}$.

Proof. We note that the bilinear form $a_n(v,v)$ can be expressed as

$$
a_n(v,v) = E_h(t_{n+1}^-, v) + E_h(t_n^+, v) + \|\sqrt{\sigma_1}\,[v]\,\|_{\Gamma_n \times I_n}^2 + \|\sqrt{\sigma_2}\,[v]\,\|_{\Gamma_{int} \times I_n}^2 + \|\sqrt{a}\dot{v}\|_{\Gamma_T \times I_n}^2 \geq 0. \tag{3.6.16}
$$

This immediately implies that $a_n(v,v)$ is a semi-norm. Now suppose $|||v|||_n = 0$ for $v \in S_{n,\mathrm{Trefftz}}^{h,p}$, then v solves the homogeneous wave equation with zero initial condition, zero Dirichlet boundary condition and zero impedance boundary. Therefore with $v \equiv 0$, implies $|||v|||_n$ is a norm on $S_{n,\mathrm{Trefftz}}^{h,p}$. The analysis for $a(v,v)$ can be shown in similar way, see Subsection 3.4.1.

□

The convergence result can be established by following the same argument as used in Subsection 3.4.2.

Chapter 4

Analysis of Trefftz space-time DG method for the damped wave equation

In this chapter, we extend the idea of the Trefftz space-time DG method to a scalar damped wave equation in second order formulation. In comparison with the established analysis of undamped wave problems, the Trefftz space-time analysis for the damped wave problem is a bit different and complicated as the wavelike basis functions are not readily available in terms of polynomials.

To resolve this non-trivial aspect of the analysis, we employ the solution (analytical) formula of the PDE in the approximation space, i.e., the discrete (local) solutions are derived by propagating polynomial initial data using the solution formula of the PDE.

This approach guarantees local functions with good approximation properties and by construction we still preserve the already proven local dimensions of the Trefftz space for general linear wave equations. In addition, this approach establishes a general method for generating local Trefftz basis functions for general linear wave equations in time domain.

We organise this chapter as follows. In the next section we introduce the damped wave model and define necessary notation for settings in Sobolev spaces. In Section 4.3, we construct the Trefftz space with the idea of a particular solution for-

mula. We proceed in Section 4.4 to present the space-time IPDG for the damped wave equation and we show its stability. Finally, we prove convergence for the case $d = 1$ spatial dimension.

4.1 Model problem

We consider the telegraph equation

$$\begin{aligned} \ddot{u} - \nabla \cdot (a \nabla u) + \alpha \dot{u} &= 0 \qquad && \text{in } \Omega \times [0,T], \\ u(0) &= 0 \qquad && \text{on } \partial\Omega \times [0,T], \\ u(x,0) = u_0(x), \quad \dot{u}(x,0) &= v_0(x), && \text{in } \Omega, \end{aligned} \tag{4.1.1}$$

where Ω is a bounded Lipschitz domain in $\mathbb{R}^d$, $\partial\Omega$ its boundary, $\alpha \in \mathbb{R}_{>0}$ and $0 < c_a \leq a(x) \leq C_a$ a piecewise constant function. We also have that if Ω_1 and Ω_2 are two subsets of Ω with the boundary Γ_{12} separating them and with $a \equiv a_1 \in \Omega_1$ and $a \equiv a_2 \in \Omega_2$, then if we denote by $u_1 = u|_{\Omega_1}$ and $u_2 = u|_{\Omega_2}$, we have the transmission conditions

$$u_1 = u_2, \quad a_1 \partial_n u_1 = a_2 \partial_n u_2, \quad \text{on } \Gamma_{12}, \tag{4.1.2}$$

where $\mathbf{n}$ denotes the exterior normal to Ω_1 or (Ω_2). The existence of a unique (weak) solution for the damped wave equation, and even for non-linear damped wave problems, has been studied in literature. If $u_0 \in H_0^1(\Omega)$ and $v_0 \in L^2(\Omega)$, then (4.1.1) has a unique (weak) solution u

$$u \in L^2([0,T]; H_0^1(\Omega)), \ \dot{u} \in L^2([0,T]; L^2(\Omega)), \ \ddot{u} \in L^2([0,T]; H^{-1}(\Omega)). \tag{4.1.3}$$

Furthermore, the solution is continuous in time with

$$u \in \mathcal{C}([0,T]; H_0^1(\Omega)), \quad \dot{u} \in \mathcal{C}([0,T]; L^2(\Omega)). \tag{4.1.4}$$

We refer readers to Chapter 7 of [52] and the following papers with other works cited therein for reference [75, 19, 67, 18, 17].

4.2 Space-time finite element (polynomial) space

In this section we discuss the construction of the time-space discretization of the problem just as in Chapter 3. We also present a polynomial space of approximation as the method can accommodate any space of approximation including orthogonal polynomial spaces.

For simplicity of presentation and implementation, we construct a time discretization $0 = t_0 < t_1 < \cdots < t_N = T$ and locally quasi-uniform spatial meshes $\mathcal{T}_n$ of Ω consisting of open simplexes such that $\Omega = \cup_{K\in\mathcal{T}_n}\overline{K}$. Therefore the space-time mesh consists of time-slabs $\mathcal{T}_n \times I_n$, where $I_n = (t_n, t_{n+1})$ and $\tau_n = t_{n+1} - t_n$. For the purpose of comparison, we recall the local discrete space of piecewise polynomials on each space-time slab defined as:

$$S_n^{h,p} = \{u \in L^2(\Omega \times I_n) : u|_{K\times I_n} \in \mathcal{P}_p(\mathbb{R}^{d+1}),\ K \in \mathcal{T}_n\}, \tag{4.2.1}$$

where $\mathcal{P}_p$ is the space of polynomials of total degree p and the complete space-time polynomial space on $\Omega \times [0,T]$ defined as:

$$V^{h,p} = \{u \in L^2(\Omega \times [0,T]) : u|_{\Omega\times I_n} \in S_n^{h,p}, n = 0, 1, \ldots, N-1\}. \tag{4.2.2}$$

We maintain the notation for the skeleton of the mesh as $\Gamma_n := \cup_{K\in\mathcal{T}_n}\partial K$ as well as the interior skeleton $\Gamma_{int} = \Gamma_n \setminus \partial\Omega$. The definition of averages and jumps as well as other mesh parameters remain the same as in Chapter 3.

4.3 Construction of Trefftz spaces for the damped wave equation

To construct Trefftz spaces with good approximation properties for this problem is not trivial. However, flexibility of the DG method allows us to construct Trefftz spaces using the analytical solution of the damped wave problem. We begin with a vital assumption on $a(x)$.

Assumption 4.3.1. *Let $a(\cdot)$ and the mesh be such that $a(\cdot)$ is constant in each element $K \in \mathcal{T}_n$ for each n.*

Definition 4.3.2 (Non-polynomial Trefftz spaces)**.** *Let $S^{h,p}_{n,Trefftz}$ be a subspace of functions satisfying locally the homogeneous damped wave equation on any space-time element $K \times I_n$:*

$$\begin{aligned} S^{h,p}_{n,Trefftz} := \big\{ v \in L^2(\Omega \times I_n) \ : \ v|_{K\times I_n} \text{satisfies } \ddot{v}(t,x) - \nabla \cdot (a\,\nabla v)(t,x) + \alpha \dot{v} = 0, \\ \text{with } v(x,t_n) \in \mathcal{P}_p(\mathbb{R}^d),\ \dot{v}(x,t_n) \in \mathcal{P}_{p-1}(\mathbb{R}^d),\ t \in I_n,\ x \in K,\ K \in \mathcal{T}_n \big\}, \end{aligned} \tag{4.3.1}$$

where $\mathcal{P}_p$ and $\mathcal{P}_{p-1}$ denote spaces of polynomials of total degrees p and $p-1$ respectively. The space on $\Omega \times [0,T]$ is now defined as

$$V^{h,p}_{Trefftz} = \left\{ v \in L^2(\Omega \times [0,T]) \ : \ v|_{\Omega\times I_n} \in S^{h,p}_{n,Trefftz},\ n = 0, 1 \ldots, N-1 \right\}.$$

Unlike the undamped wave problem studied in Chapter 3, the Trefftz functions for the damped wave equation are not readily available in terms of polynomials. Therefore we generate the local solutions by propagating polynomial initial data in time using a solution formula. For example in one spatial dimension, we discover that the d'Alembert-type progressive wave solution formula of the form (see[98, 64]),

$$\begin{aligned} u_h(x,t) = {} & \frac{1}{2}\left[u_0^h(x-ct) + u_0^h(x+ct)\right]\exp(-t\alpha/2) \\ & + \frac{\alpha}{4c}\exp(-t\alpha/2)\int_{x-ct}^{x+ct} u_0^h(s)\left\{ I_0\left(\rho(s)\frac{t\alpha}{2}\right) + \frac{1}{\rho(s)} I_1\left(\rho(s)\frac{t\alpha}{2}\right)\right\} \mathrm{ds} \\ & + \frac{1}{2c}\exp(-t\alpha/2)\int_{x-ct}^{x+ct} v_0^h(s) I_0\left(\rho(s)\frac{t\alpha}{2}\right) \mathrm{ds}, \end{aligned} \tag{4.3.2}$$

can be used to propagate initial data $u_0^h(x) \in \mathcal{P}_p$ and $v_0^h(x) \in \mathcal{P}_{p-1}$ in time, where $\rho(s) = \rho(s;x,t) = \sqrt{1-(x-s)^2/(ct)^2}$ and I_ν is the modified Bessel function of the first kind of order ν. Note that u_0^h and v_0^h are traces of the Trefftz space on the space-time slab $\Omega \times I_h$, $I_h = [0,h]$ and $u_h(x,t)$ propagates the initial data in time. Moreover the propagated (local) initial data could be monomials (see Table 4.1), orthogonal polynomials or scaled polynomials.

$$\begin{array}{cc} u_0^h = 1 & v_0^h = 0 \\ x & 0 \\ \vdots & \vdots \\ x^p & 0 \\ 0 & 1 \\ 0 & x \\ \vdots & \vdots \\ 0 & x^{p-1} \end{array} \tag{4.3.4}$$

Table 4.1: Example of polynomial initial data

The local degrees of freedom still remains $2p+1$ for the case of spatial dimension $d = 1$ based on the derivation of u_0 and v_0 from polynomial spaces $\mathcal{P}_p$ and $\mathcal{P}_{p-1}$ respectively. In general, the local dimension of the Trefftz spaces for spatial dimensions $d = \{1, 2, 3\}$ can be determined by adding up the number of propagated initial data in $\mathcal{P}_p(K)$ and $\mathcal{P}_{p-1}(K)$, i.e.,

$$\dim S^{h,p}_{n,\text{Trefftz}} = \dim \mathcal{P}_p(K) + \dim \mathcal{P}_{p-1}(K). \tag{4.3.3}$$

Therefore, the local degrees of freedom in spatial dimensions $d = 2$ and $d = 3$ remains $(p+1)^2$ and $(p+1)(p+2)(2p+3)/6$ respectively.

The two integrals in (4.3.2) can be approximated by quadrature and hence we rewrite

$$\int_{x-ct}^{x+ct} u_0^h(s) \left\{ I_0\left(\rho(s)\frac{t\alpha}{2}\right) + \frac{1}{\rho(s)} I_1\left(\rho(s)\frac{t\alpha}{2}\right) \right\} \mathrm{ds} \tag{4.3.5}$$

as

$$ct \int_{-1}^{1} u_0^h(ctq + x) \left\{ I_0\left(\rho((ctq+x))t\alpha/2\right) + \frac{1}{\rho(ctq+x)} I_1\left(\rho((ctq+x))t\alpha/2\right) \right\} \mathrm{dq}, \tag{4.3.6}$$

where we have used the relation

$$q = \frac{2s - a - b}{b - a} \tag{4.3.7}$$

of the Gaussian quadrature defined on $[-1, 1]$. We can now explicitly define $\rho(s)$ as

a function of q by the change of variable:

$$\rho(q) = \sqrt{1-q^2}. \tag{4.3.8}$$

Therefore the integral can be written explicitly

$$ct \int_{-1}^{1} u_0^h(ctq+x) \left\{ I_0\left(\sqrt{1-q^2}\frac{t\alpha}{2}\right) + \frac{1}{\sqrt{1-q^2}} I_1\left(\sqrt{1-q^2}\frac{t\alpha}{2}\right)\right\} \mathrm{dq}. \tag{4.3.9}$$

In similar way, the second integral becomes

$$ct \int_{-1}^{1} v_0^h(qct+x) I_0\left(\sqrt{1-q^2}\frac{t\alpha}{2}\right) \mathrm{dq}. \tag{4.3.10}$$

The derived local solutions are dependent on the size of the physical space-time element $K \times I_n$ and hence transformation to a reference element cannot be employed in this case. In our implementation procedure, we compute the basis functions in advance which is expensive. In order to speed up the implementation process, we take the advantage of the low rank property in Chebfun2 to make the implementation process faster, (see, Snippet 4.1 for Chebfun2 code and Table 4.2 for the ranks of the basis functions). The MATLAB function 'solbasis' in the code snippet computes each i^{th} basis function in advance. We present the plot of the Trefftz basis functions in 1–dimension up to order 4 below in Figure 4.2.

```
function Z=computebasischeb(p,a,h)
P=2*p+1;
c=1; N=20; domx=[0 h];
dom2=[domx,0 h];% space-time local domain
Z=cell(P,1);
for i=1:P % Chebfun2 in use
    Z{i}=chebfun2(@(x,t)solbasis(i,x*h,t*h,h,c,p,a,N),dom2);
end
```

Figure 4.1: *Code snippet for the basis functions.*

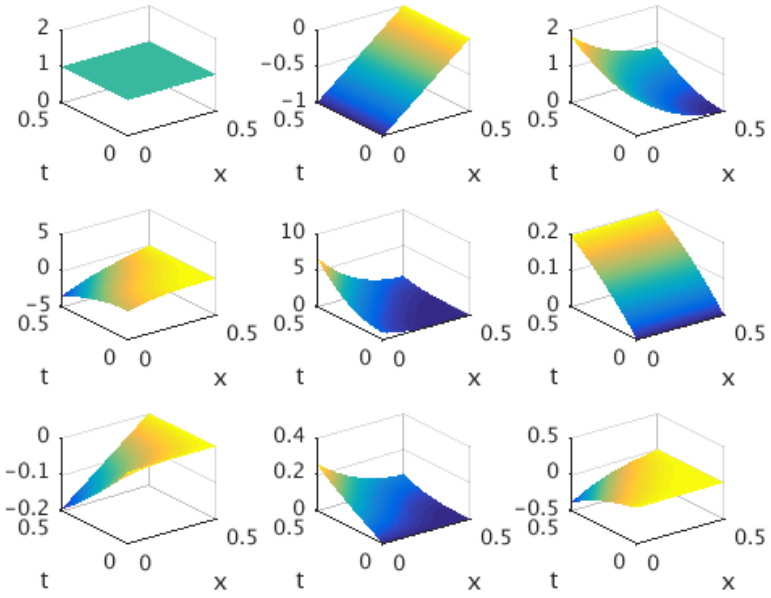


Figure 4.2: *Plot of Trefftz basis functions in* 1*−dimension on reference space-time element* $(0,h)\times(0,h)$, $h=0.5$.

Basis functions	Rank
$Z\{1\}$	1
$Z\{2\}$	1
$Z\{3\}$	2
$Z\{4\}$	2
$Z\{5\}$	3
$Z\{6\}$	1
$Z\{7\}$	1
$Z\{8\}$	2
$Z\{9\}$	2

Table 4.2: *Ranks of the constructed* 1*−dimensional Trefftz Basis functions (see Figure 4.2).*

4.4 Space-time DG method for the damped wave equation

We shall follow an energy argument just as in Chapter 3 in deriving the weak form for the PDE. Assume that u is a smooth solution of (4.1.1) and let $v \in \mathcal{X} + V^{h,p}$ or $v \in \mathcal{X} + V^{h,p}_{\text{Trefftz}}$. The standard symmetric IPDG weak formulation on the time-slab I_n when tested with $\dot{v}$ is given by

$$\begin{aligned}(\ddot{u}, \dot{v})_{\Omega\times I_n} + (\widetilde{\nabla} u, \widetilde{\nabla}\dot{v})_{\Omega\times I_n} - (\{\nabla u\}, [\dot{v}])_{\Gamma_n\times I_n} - ([u], \{\nabla\dot{v}\})_{\Gamma_n\times I_n} \\ -(\sigma_0 [u], [\dot{v}])_{\Gamma_n\times I_n} + (\alpha\dot{u}, \dot{v})_{\Omega\times I_n} = 0,\end{aligned} \tag{4.4.1}$$

where

$$\sigma_0(x,t) \coloneqq C_{\sigma_0} h(x,t)^{-1}, \tag{4.4.2}$$

and C_{σ_0} is a positive constant to be determined later. Note that if $v = u$ in (4.4.1), then we have the following:

$$\frac{1}{2}\|\dot{u}\|^2_{L^2(\Omega)} + \frac{1}{2}\|\nabla u\|^2_{L^2(\Omega)} - (\{\nabla u\}, [u])_\Gamma + \frac{1}{2}\|\sigma_0^{1/2}[u]\|^2_{(\Gamma)} + \|\alpha^{1/2}\dot{u}\|^2_{(\Omega\times I_n)}, \tag{4.4.3}$$

which suggests the same *discrete energy* as in Chapter 3, (see, equation (3.3.3))

$$E_h(u, t_n) = \frac{1}{2}\|\dot{u}(t_n)\|^2_{L^2(\Omega)} + \frac{1}{2}\|\widetilde{\nabla} u(t_n)\|^2_{L^2(\Omega)} - (\{\nabla u\}, [u])_\Gamma + \frac{1}{2}\|\sigma_0^{1/2}[u(t_n)]\|^2_{(\Gamma)}. \tag{4.4.4}$$

Now choosing as test function $v = u$ in (4.4.1), summing over n and employing the algebraic identity (3.3.4) just as in Chapter 3 leads to the following space-time weak

formulation:

$$\begin{aligned}
&\sum_{n=0}^{N-1} (\ddot{u}, \dot{v})_{\Omega\times I_n} + (\llbracket \dot{u}(t_n) \rrbracket, \dot{v}(t_n^+))_\Omega \\
&+ \left(a\widetilde{\nabla} u, \widetilde{\nabla}\dot{v}\right)_{\Omega\times I_n} + \left(\llbracket a\widetilde{\nabla} u(t_n) \rrbracket, \widetilde{\nabla} v(t_n^+)\right)_\Omega \\
&- \left(\{a\nabla u\}, [\dot{v}]\right)_{\Gamma_n\times I_n} - \left(\llbracket \{a\widetilde{\nabla} u(t_n)\} \rrbracket, [v(t_n^+)]\right)_{\hat{\Gamma}_n} \\
&- \left([u], \{a\nabla\dot{v}\}\right)_{\Gamma_n\times I_n} - \left(\llbracket [u(t_n)] \rrbracket, \{a\nabla v(t_n^+)\}\right)_{\hat{\Gamma}_n} \\
&+ \left(\sigma_0 [u], [\dot{v}]\right)_{\Gamma_n\times I_n} + \left(\sigma_0 \llbracket [u(t_n)] \rrbracket, [v(t_n^+)]\right)_{\hat{\Gamma}_n} \\
&+ \left(\sigma_1 [u], [v]\right)_{\Gamma_n\times I_n} + \left(\sigma_2 [a\nabla u], [a\nabla v]\right)_{\Gamma_n\times I_n} + (\alpha\dot{u}, \dot{v})_{\Omega\times I_n} = \mathcal{B}^{\text{init}}(v),
\end{aligned} \tag{4.4.5}$$

where,

$$\begin{aligned}
\mathcal{B}^{\text{init}}(v) = &(v_0, \dot{v}(t_0^+))_\Omega + \left(a\widetilde{\nabla} u_0, \widetilde{\nabla} v(t_0^+)\right)_\Omega - \left(\{a\nabla u_0\}, [v(t_0^+)]\right)_{\Gamma_0} \\
&- \left([u_0], \{a\nabla v(t_0^+)\}\right)_{\Gamma_0} + \left(\sigma_0 [u_0], [v(t_0^+)]\right)_{\Gamma_0}.
\end{aligned} \tag{4.4.6}$$

Thus, we have arrived at a space-time discrete method for the damped wave equation, which we present in two viewpoints: as a method for obtaining a discrete solution on a fixed space-time domain $\Omega\times[0,T]$ or as a time-stepping method. Just as in Chapter 3, the former viewpoint will be used for the convergence analysis while the latter will be used for the implementation of the method. In order to present the two viewpoints compactly, we define three bilinear forms:

$$\begin{aligned}
\mathcal{A}_n(u,v) \coloneqq\ &(\ddot{u}, \dot{v})_{\Omega\times I_n} + (\dot{u}(t_n^+), \dot{v}(t_n^+))_\Omega \\
&+ \left(a\widetilde{\nabla} u, \widetilde{\nabla}\dot{v}\right)_{\Omega\times I_n} + \left(a\widetilde{\nabla} u(t_n^+), \widetilde{\nabla} v(t_n^+)\right)_\Omega \\
&- \left(\{a\nabla u\}, [\dot{v}]\right)_{\Gamma_n\times I_n} - \left(\{a\nabla u(t_n^+)\}, [v(t_n^+)]\right)_{\Gamma_n} \\
&- \left([u], \{a\nabla\dot{v}\}\right)_{\Gamma_n\times I_n} - \left([u(t_n^+)], \{a\nabla v(t_n^+)\}\right)_{\Gamma_n} \\
&+ \left(\sigma_0 [u], [\dot{v}]\right)_{\Gamma_n\times I_n} + \left(\sigma_0 [u(t_n^+)], [v(t_n^+)]\right)_{\Gamma_n} \\
&+ \left(\sigma_1 [u], [v]\right)_{\Gamma_n\times I_n} + \left(\sigma_2 [a\nabla u], [a\nabla v]\right)_{\Gamma_n\times I_n} + (\alpha\dot{u}, \dot{v})_{\Omega\times I_n},
\end{aligned} \tag{4.4.7}$$

$$\begin{aligned}
\mathcal{B}_n(u,v) \coloneqq\ &(\dot{u}(t_n^-), \dot{v}(t_n^+))_\Omega + \left(a\widetilde{\nabla} u(t_n^-), \widetilde{\nabla} v(t_n^+)\right)_\Omega \\
&- \left(\{a\nabla u(t_n^-)\}, [v(t_n^+)]\right)_{\Gamma_n} - \left([u(t_n^-)], \{a\nabla v(t_n^+)\}\right)_{\Gamma_{n-1}} \\
&+ \left(\sigma_0 [u(t_n^-)], [v(t_n^+)]\right)_{\hat{\Gamma}_n},
\end{aligned} \tag{4.4.8}$$

and

$$\mathcal{A}(u,v) := \sum_{n=0}^{N-1} \mathcal{A}_n(u,v) - \sum_{n=1}^{N-1} \mathcal{B}_n(u,v). \tag{4.4.9}$$

Definition 4.4.1. *Given subspace $X_n \subseteq S_n^{h,p}$ or $X_n \subseteq S_{n,Trefftz}^{h,p}$, the time stepping method for the damped wave equation is described by: find $u^n \in X_n$, $n = 1, 2, \ldots, N-1$, such that*

$$\mathcal{A}_n(u^n, v) = \mathcal{B}_n(u^{n-1}, v), \quad \text{for all } v \in X_n \tag{4.4.10}$$

and

$$\mathcal{A}_0(u^n, v) = \mathcal{B}^{init}(v), \quad \text{for all } v \in X_0. \tag{4.4.11}$$

Equivalently, given a subspace $X \subseteq V^{h,p}$ or $X \subseteq V_{Trefftz}^{h,p}$, the full space-time discrete system is presented as: find $u \in X$ such that

$$\mathcal{A}(u,v) = \mathcal{B}^{init}(v), \quad \text{for all } v \in X. \tag{4.4.12}$$

Before we proceed to the a priori analysis, we need to establish the following lemma.

Lemma 4.4.2. *It holds that for $u \in \mathcal{X} + V_{Trefftz}^{h,p}$,*

$$\mathcal{A}_n(u,u) = E_h(t_{n+1}^-, u) + E_h(t_n^+, u) + \|\sqrt{\sigma_1}\,[u]\|_{\Gamma_n \times I_n}^2 + \|\sigma_2\,[a\nabla u]\|_{\Gamma_n \times I_n}^2 + \|\alpha^{1/2}\dot{u}\|_{\Omega \times I_n}, \tag{4.4.13}$$

for $n = 0, 1, \ldots, N-1$, and

$$\begin{aligned}\mathcal{A}(u,u) = E_h(t_N^-, u) + E_h(t_0^+, u) + \sum_{n=1}^{N-1} \Big(& \frac{1}{2}\|[\![\dot{u}(t_n)]\!]\|_\Omega^2 + \frac{1}{2}\|\sqrt{a}[\![\widetilde{\nabla} u(t_n)]\!]\|_\Omega^2 \\ & - \Big(\big([\![\{ a\widetilde{\nabla} u(t_n) \}]\!], [[u(t_n)]] \big)_{\Gamma_n} + \frac{1}{2}\|[\![\sqrt{\sigma_0}\,[u(t_n)]]\!]\|_{\Gamma_n}^2 \Big) \\ & + \sum_{n=0}^{N-1} \Big(\|\sqrt{\sigma_1}\,[u]\|_{\Gamma_n \times I_n}^2 + \|\sqrt{\sigma_2}\,[a\nabla u]\|_{\Gamma_n \times I_n}^2 \Big) + \sum_{n=0}^{N-1} \|\sqrt{\alpha}\dot{u}\|_{\Omega \times I_n}^2. \end{aligned} \tag{4.4.14}$$

Proof. let $v = u$ and using the time stepping bilinear (4.4.7) , we have

$$\begin{aligned}\mathcal{A}_n(u,u) = &\frac{1}{2}\frac{d}{dt}\|\dot{u}(t)\|^2_{\Omega\times I_n} + \|\dot{u}(t_n^+)\|^2_{\Omega} + \frac{1}{2}\frac{d}{dt}\|a\nabla u\|^2_{\Omega\times I_n}\\ &+ \|a\nabla u(t_n^+)\|^2_{\Omega} - \frac{d}{dt}([u],\{a\nabla u\})_{\Gamma\times I_n} - 2(\{a\nabla u(t_n^+)\},[v(t_n^+)])_{\Gamma_n}\\ &+ \frac{1}{2}\frac{d}{dt}\|\sqrt{\sigma_0}\,[u(t_n^+)]\,\|^2_{\Gamma_n} + \|\sqrt{\sigma_1}\,[u]\,\|^2_{\Gamma_n\times I_n} + \|\sqrt{\sigma_2}[\![a\nabla u]\!]\|^2_{\Gamma_n\times I_n} + \|\sqrt{\alpha}\dot{u}\|_{\Omega\times I_n}.\end{aligned} \tag{4.4.15}$$

Further analysis gives

$$\begin{aligned}\mathcal{A}_n(u,u) = &\frac{1}{2}\|\dot{u}(t_{n+1}^-)\|^2_{\Omega} + \frac{1}{2}\|\dot{u}(t_n^+)\|^2_{\Omega} + \frac{1}{2}\|a\nabla u(t_{n+1}^-)\|^2_{\Omega} + \frac{1}{2}\|a\nabla u(t_n^+)\|^2_{\Omega}\\ &-([u(t_{n+1}^-)],\{a\nabla u(t_{n+1}^-)\})_{\Gamma_n} - ([u(t_n^+)],\{a\nabla u(t_n^+)\}) + \frac{1}{2}\|\sqrt{\sigma_0}\,[u(t_{n+1}^-)]\,\|^2_{\Gamma_n}\\ &+ \frac{1}{2}\|\sqrt{\sigma_0}\,[u(t_n^+)]\,\|^2_{\Gamma_n} + \|\sqrt{\sigma_1}\,[u]\,\|^2_{\Gamma_n\times I_n} + \|\sqrt{\sigma_2}\,[a\nabla u]\,\|^2_{\Gamma_n\times I_n} + \|\sqrt{\alpha}\dot{u}\|^2_{\Omega\times I_n}.\end{aligned} \tag{4.4.16}$$

Hence from the definition of the energy (4.4.4) we have the result.

In similar way as above, let $v = u$ in the discrete bilinear form. Cancelling out the time integrals with the time derivative and re-arranging with the definition of the energy $E_h(t,u)$ we have

$$\begin{aligned}\mathcal{A}(u,u) = &E_h(t_N^-,u) + E_h(t_0^+,u) - \sum_{n=1}^{N-1}\frac{1}{2}[\![\|\dot{u}(t_n)\|^2_{\Omega}]\!] - \sum_{n=1}^{N-1}\frac{1}{2}[\![\|\sqrt{a}\nabla u\|^2_{\Omega}]\!]\\ &+ \sum_{n=1}^{N-1}([\![\dot{u}(t_n)]\!],\dot{u}(t_n^+))_{\Omega} + \sum_{n=1}^{N-1}([\![a\nabla u(t_n)]\!],\nabla u(t_n^+))_{\Omega} + \sum_{n=1}^{N-1}[\![(\{a\nabla u(t_n)\},[u(t_n^+)])_{\Gamma_n}]\!]\\ &- \sum_{n=1}^{N-1}([\![\{a\nabla u(t_n)\}]\!],[v(t_n^+)])_{\Gamma_n} - \sum_{n=1}^{N-1}([\![[u(t_n)]]\!],\{a\nabla u(t_n^+)\})_{\Gamma_n} - \sum_{n=1}^{N-1}\frac{1}{2}[\![\|\sqrt{\sigma_0}\,[u(t_n)]\,\|^2_{\Gamma_n}]\!]\\ &+ \sum_{n=0}^{N-1}\|\sqrt{\sigma_1}\,[u]\,\|^2_{\Gamma_n\times I_n} + \sum_{n=1}^{N-1}(\sigma_0[\![[u(t_n)]]\!],[u(t_n^+)])_{\Gamma_n}\\ &+ \sum_{n=0}^{N-1}\|\sqrt{\sigma_2}\,[a\nabla u]\,\|^2_{\Gamma_n\times I_n} + \sum_{n=0}^{N-1}\|\sqrt{\alpha}\dot{u}\|^2_{\Omega\times I_n}.\end{aligned} \tag{4.4.17}$$

Finally, using the the algebraic upwinding (3.3.4) completes the proof.

□

4.5 A priori analysis for the damped wave problem

In this section we proceed to give relevant a priori analysis which include consistency and stability of the derived method, quasi-optimality and rate of convergence for one dimensional case. Note that the Trefftz functions for the damped wave problem are not in the space of polynomials, hence the analysis procedure in this chapter is slightly different from what we studied in Chapter 3. We shall employ the following lemma from [20].

Lemma 4.5.1. *Let $\rho h \le \operatorname{diam} K \le h$ where $0 \le h \le 1$, and let $\mathcal{S}$ be a finite dimensional subspace of Sobolev space $H^l(K) \cap H^m(K)$ where $0 \le m \le l$. Then there exists $C := C(\mathcal{S}, K, l, \rho)$ such that for all $v \in \mathcal{S}$, we have*

$$\|v\|_{H^l(K)} \le C h^{m-l} \|v\|_{H^m(K)}. \tag{4.5.1}$$

Theorem 4.5.2. *Let the spaces $S^{h,p}_{n,Trefftz}$ for $n = 0, \dots, N-1$ be given. Then the following statements hold:*

1. *Let u be the (weak) solution of (4.1.1) with $u_0 \in H^1_0(\Omega)$ and $v_0 \in L^2(\Omega)$. Then u satisfies (4.4.12).*

2. *For C_{σ_0} chosen big enough, (see (4.4.2)) and for any $v \in S^{h,p}_{n,Trefftz}$ and $t \in I_n$, the energy $E_h(t, v)$ is bounded below as*

$$E_h(t, v) \ge \frac{1}{2}\|\dot{v}(t)\|^2_\Omega + \frac{1}{4}\|\sqrt{a}\widetilde{\nabla} v(t)\|^2_\Omega. \tag{4.5.2}$$

   where C_{σ_0} is independent of $h(x,t)$ but may depend on the polynomial degree p. Further, let $u^n \in S^{h,p}_{n,Trefftz}$, $n = 0, \dots, N-1$, satisfy (4.4.5). Then, $E_h(t_N^-, u^n) \le E_h(t_1^-, u^0)$.

Proof. The first statement follows from the derivation of the scheme and the regularity of the unique solution u. To continue with the rest of the statement, we need to show that if σ_0 is chosen big enough, then the energy (4.4.4) is a non-negative

quantity. Since the functions in the local Trefftz space $S^{h,p}_{n,\text{Trefftz}}$ are not polynomials, the discrete inverse inequality does not apply directly. However, the standard trace inequality of the form (see Theorem $A.11$ in [113])

$$\|\gamma_0 u\|_{\partial K} \leq C(K)\|u\|_{H^1(K)} \tag{4.5.3}$$

can be used to derive a suitable inequality. Now, we make use of Lemma 4.5.1 and we choose $l = 1$, $m = 0$ and $S = S^{h,p}_{n,\text{Trefftz}}$ is a finite dimensional subspace of $H^1(K)$. Then we have

$$\|v\|_{H^1(K)} \leq Ch^{-1}\|v\|_{L^2(K)} \ \forall \ v \ \in \ \mathcal{S}. \tag{4.5.4}$$

Now the trace inequality (4.5.3) above can be used and hence we have

$$\|\gamma_0 u\|_{L^2(\partial K)} \leq Ch^{-1}\|u\|_{L^2(K)} \text{ for } u \ \in \ H^1(K). \tag{4.5.5}$$

Therefore we can write

$$\|u\|_{L^2(\partial K)} \leq Ch^{-1}\|u\|_{L^2(K)} \text{ for } u \ \in \ H^1(K). \tag{4.5.6}$$

To show the positivity of the energy (4.4.4), it suffices to bound the last term in the expression. Using the Cauchy-Schwarz and Young's inequality we have

$$\begin{aligned} |(\{a\nabla v(t)\}, [v(t)])_{\Gamma_n}| &\leq |\big(\frac{\sigma_0}{2}\big)^{-1/2}\{a\nabla v(t)\}, \big(\frac{\sigma_0}{2}\big)^{1/2}[v(t)])| \\ &\leq \int_{\Gamma_n} \sigma_0^{-1}|\{a^{1/2}\nabla v(t)\}|^2 \mathrm{ds} + \frac{\sigma_0}{4}\int_{\Gamma_n}|[v(t)]|^2 \mathrm{ds}. \end{aligned} \tag{4.5.7}$$

Using the trace inequality (4.5.6) we bound the first term in the above to have

$$\begin{aligned} \int_{\Gamma_n} \sigma_0^{-1}|\{a^{1/2}\nabla v(t)\}|^2 \mathrm{ds} &\leq \sum_{K\in\mathcal{T}_n}\int_{\partial K} \sigma_0^{-1}|n\cdot a^{1/2}\nabla v(t)|^2 \mathrm{ds} \\ &\leq \sum_{K\in\mathcal{T}_n} C\sigma_0^{-1}\frac{1}{h}\big(\|a^{1/2}\nabla v(t)\|_K^2\big). \end{aligned} \tag{4.5.8}$$

Now choosing $C_{\sigma_0} \geq 2C$ is enough to write

$$|(\{a\nabla v(t)\}, [v(t)])_{\Gamma_n}| \leq \frac{1}{2}\|a^{1/2}\nabla v\|_\Omega^2 + \frac{1}{4}\|\sigma_0^{1/2}[v(t)]\|_{\Gamma_n}, \tag{4.5.9}$$

which makes the energy bounded below by

$$E_h(t,v) \geq \frac{1}{2}\|\dot{v}(t)\|_\Omega^2 + \frac{1}{4}\|\sigma_0^{1/2}\,[v(t)]\,\|_{\Gamma_n}^2. \tag{4.5.10}$$

To complete the rest of the proof, note that from (4.4.13) we have

$$\mathcal{A}_0(u,u) = E_h(t_1^-,u)+E_h(t_0^+,u)+\|\sqrt{\sigma_1}\,[u]\,\|_{\Gamma_0\times I_0}^2+\|\sqrt{\sigma_2}\,[a\nabla u]\,\|_{\Gamma_0\times I_0}^2+\|\sqrt{\alpha}\dot{u}\|_{\Omega\times I_0}^2 = \mathcal{B}^{\text{init}}(u). \tag{4.5.11}$$

Therefore we can write

$$\begin{aligned} E_h(t_N^-,u) = E_h(t_1^{-1},u) - \sum_{n=1}^{N-1}\Big(&\frac{1}{2}\|[\![u(t_n)]\!]\|_\Omega^2 + \frac{1}{2}\|[\![\sqrt{a}\nabla u(t_n)]\!]\|_\Omega^2 - ([\![\{a\nabla u(t_n)\}]\!],[\![[u(t_n)]]\!])_{\Gamma_n} \\ &+ \frac{1}{2}\|\sqrt{\sigma_0}[\![u(t_n)]\!]\|_{\Gamma_n}^2 + \|\sqrt{\sigma_1}\,[u]\,\|_{\Gamma_n\times I_n}^2 + \|\sqrt{\sigma_2}\,[a\nabla u]\,\|_{\Gamma_n\times I_n}^2\Big) - \sum_{n=0}^{N-1}\|\sqrt{\alpha}\dot{u}\|_{\Omega\times I_n}^2. \end{aligned} \tag{4.5.12}$$

□

4.5.1 Existence and uniqueness of solution

In this subsection we prove the existence and uniqueness of the solution in the Trefftz space. We investigate the DG bilinear forms (4.4.7) and (4.4.9) and we show that they give rise to norms on respective Trefftz spaces $S^{h,p}_{n,\text{Trefftz}}$ and $V^{h,p}_{\text{Trefftz}}$.

Proposition 4.5.3. *With σ_0 chosen large enough and $\sigma_1,\sigma_2 > 0$, bilinear forms $a_n(\cdot,\cdot)$ and $a(\cdot,\cdot)$ give rise to two norms*

$$|||v|||_n := \big(\mathcal{A}_n(v,v)\big)^{1/2}, \qquad v \in S^{h,p}_{n,\mathit{Trefftz}}$$

and

$$|||v||| := \big(\mathcal{A}(v,v)\big)^{1/2}, \qquad v \in V^{h,p}_{\mathit{Trefftz}}.$$

Proof. Note that if $|||v|||_n = 0$ for $v \in S^{h,p}_{n,\text{Trefftz}}$, then v solves the homogeneous damped wave equation with zero initial and boundary conditions. Uniqueness implies $v \equiv 0$ and hence that $|||\cdot|||_n$ is a norm on this Trefftz space. The proof for $|||\cdot|||$ follows the same way by recalling (4.4.14) and proceeding as in the first case, shows that it is

in fact a norm on the Trefftz space $V^{h,p}_{\text{Trefftz}}$. □

Corollary 4.5.4. *Under the conditions of the above proposition and with initial data $u_0 \in H^1_0(\Omega)$, $v_0 \in L^2(\Omega)$, the discrete system (4.4.12) with $X = V^{h,p}_{Trefftz}$ has a unique solution.*

We present the convergence analysis of the Trefftz based method in the next subsection.

4.5.2 Convergence analysis

We now establish the quasi-optimality of the proposed method for the damped wave problem. We begin with the following vital proposition.

Proposition 4.5.5. *Let $w \in \mathcal{X} + V^{h,p}_{Trefftz}$ and $v \in V^{h,p}_{Trefftz}$, then*

$$\mathcal{A}(w,v) \le C_\star |\!|\!| w |\!|\!|_\star |\!|\!| v |\!|\!|, \tag{4.5.13}$$

for some constant $C_\star > 0$ and

$$\begin{aligned}
|\!|\!| w |\!|\!|_\star^2 =& \frac{1}{2}\sum_{n=1}^{N}\Big(\|\dot{w}(t_n^-)\|_\Omega^2 + \|\sqrt{a}\nabla w(t_n^-)\|_\Omega^2 + \|\sqrt{\sigma_0}\,[w(t_n^-)]\,\|_{\Gamma_n}^2 + \|\sigma_0^{-1/2}\{a\nabla w(t_n^-)\}\,\|_{\Gamma_n}^2\Big) \\
&+ \sum_{n=0}^{N-1}\Big(\|\sqrt{\sigma_1}\,[w]\|_{\Gamma_n\times I_n}^2 + \|\sqrt{\sigma_2}\,[a\nabla w]\|_{\Gamma_n\times I_n}^2 + \|\sigma_2^{-1/2}\{\dot{w}\}\|_{\Gamma_n^{\text{int}}\times I_n}^2 \\
&\qquad + \|\sigma_1^{-1/2}\{a\nabla\dot{w}\}\|_{\Gamma_n\times I_n}^2 + \|\sigma_0\sigma_1^{-1/2}[\dot{w}]\|_{\Gamma_n\times I_n}^2\Big) + \sum_{n=0}^{N-1}\|2^{1/2}\dot{w}\|_{\Omega\times I_n}^2.
\end{aligned}$$

Proof. Analogously to the proof of Proposition 3.4.6, we have that

$$\begin{aligned}
\mathcal{A}(w,v) &= \sum_{n=0}^{N-1} \mathcal{A}_n(w,v) - \sum_{n=1}^{N-1} \mathcal{B}_n(w,v) \\
&= \sum_{n=0}^{N-1} \Big(\left(\{a\nabla \dot{w}\}, [v]\right)_{\Gamma_n \times I_n} - \left(\sigma_0 [\dot{w}], [v]\right)_{\Gamma_n \times I_n} - \left(\{\dot{w}\}, [a\nabla v]\right)_{\Gamma_n^{\mathrm{int}} \times I_n} \\
&\quad + \left(\sigma_1 [w], [v]\right)_{\Gamma_n \times I_n} + \left(\sigma_2 [a\nabla w], [a\nabla v]\right)_{\Gamma_n \times I_n} \Big) \\
&\quad - \sum_{n=1}^{N} \Big((\dot{w}(t_n^-), [\![\dot{v}(t_n)]\!])_\Omega + (a\nabla w(t_n^-), [\![\nabla v(t_n)]\!])_\Omega \\
&\quad - \left(\{a\nabla w(t_n^-)\}, [\![[v(t_n)]]\!]\right)_{\Gamma_n} - \left([w(t_n^-)], [\![\{a\nabla v(t_n)\}]\!]\right)_{\Gamma_n} \\
&\quad + \left(\sigma_0 [w(t_n^-)], [\![[v(t_n)]]\!]\right)_{\Gamma_n} \Big) + \sum_{n=0}^{N-1} (2\alpha \dot{w}, \dot{v})_{\Omega \times I_n}.
\end{aligned} \tag{4.5.14}$$

The rest of the proof as before can be completed using the Cauchy-Schwarz inequality. □

Theorem 4.5.6. *Let $U \in V_{Trefftz}^{h,p}$ be the discrete solution of the Trefftz space-time discontinuous Galerkin method and let $u \in \mathcal{X}$ be the exact solution. Then*

$$|||U - u||| \leq \inf_{V \in V_{Trefftz}^{h,p}} \left(C_\star |||u - V|||_\star + |||u - V||| \right) \tag{4.5.15}$$

Proof. The proof follows the same analogy as in Theorem 3.4.8 using Proposition 4.5.5, the fact that discrete bilinear form $a(\cdot,\cdot)$ is a norm on the Trefftz space, and the Galerkin orthogonality. □

4.6 A priori error bounds in one spatial dimension

The Trefftz basis employed for the case of damped wave problem is sufficient to deliver the expected rates of convergence for the proposed method.

Lemma 4.6.1. *Let the setting of Theorem 4.5.6 hold, let $u_h \in V_{Trefftz}^{h,p}$ be an arbitrary*

function in the discrete space, and let $\eta = u - u_h$. *Then*

$$\begin{aligned}|||U-u|||^2 \le C\sum_{n=0}^{N-1}\sum_{K\in\mathcal{T}^n} C_a\Bigg(&\frac{1}{\tau_n}\Big(\|\nabla\eta\|^2_{K\times I_n} + \min\{1,\frac{\tau_n^2}{h_K^2}\}\|\dot\eta\|^2_{K\times I_n}\Big)\\ &+\tau_n\Big(\|\nabla\dot\eta\|^2_{K\times I_n} + \min\{1,\frac{h_K^2}{\tau_n^2}\}\|D^2\eta\|^2_{K\times I_n}\Big)\\ &+h_K^2\tau_n\|D^2\dot\eta\|^2_{K\times I_n} + \frac{1}{h_K^2\tau_n}\|\eta\|^2_{K\times I_n} + \tau_n\|\Delta\eta-\alpha\dot\eta\|^2_{K\times I_n}\Bigg)\end{aligned} \tag{4.6.1}$$

where $U \in V^{h,p}_{Trefftz}$ *is the discrete solution.*

Proof. The proof follows the same analogy as in Section 3.5 by employing the result of Theorem 4.5.6 and repeated use of the trace estimate $\|v\|^2_{\partial\omega} \le C(\mathrm{diam}(\omega)^{-1}\|v\|^2_\omega + \mathrm{diam}(\omega)\|\nabla v\|^2_\omega)$, for $v \in H^1(\omega)$, where $\omega \subset \mathbb{R}^d$ or $\mathbb{R}^{d+1}$. Note that the associated constants with the stability parameters have been taken to be C_a which is independent of h but may depend on the polynomial degree p.

□

To complete the error analysis we need to investigate the existence of an appropriate approximation $u_h \in V^{h,p}_{\mathrm{Trefftz}}$ of the exact solution. We shall prove existence with a proposition and a lemma. The convergence analysis will later be finalised with a theorem. Let us begin with the following definition.

Definition 4.6.2. *Let* u *be a smooth enough solution of* $\ddot u - \nabla\cdot(a\nabla u) + \alpha\dot u = 0$ *in* $\Omega\times I$, *with initial data* $u_0 = \gamma u|_{t=0}$ *and* $v_0 = \gamma\dot u|_{t=0}$, $h = 1/M$, $I = (0,h)$ *and let* $x_j = jh$, $j = 0\dots M$. *The projections* $\pi^p u_0$ *and* $\pi^{p-1}v_0$ *are defined for* $x \in (x_j, x_{j+1})$ *as follows:*

$$\begin{aligned}\pi^p u_0 &= u_0(x_j) + u_0'(x_j)(x-x_j) + \cdots + \frac{u_0^{(p)}(x_j)}{p!}(x-x_j)^p,\\ \pi^{p-1}v_0 &= v_0(x_j) + v_0'(x_j)(x-x_j) + \cdots + \frac{v_0^{(p-1)}(x_j)}{(p-1)!}(x-x_j)^{p-1}.\end{aligned} \tag{4.6.2}$$

Proposition 4.6.3. *Let* $\Omega = (0,1)$, $h = 1/M > 0$, *and* $I = (0,h)$ *and let* $x_j = jh$, $j = 0\dots M$, $K = (x_j, x_{j+1})$. *Let* u *be given such that* $\ddot u - \Delta u + \alpha\dot u = 0$ *in* $\Omega\times I$ *and* $u(0,t) = u(1,t) = 0$, $t \in I$. *Let* $u_0 = u(\cdot,0)$, $v_0 = \dot u(\cdot,0)$ *and define* $u_0(x) = -u(-x)$ *for* $x \in (-1,0)$ *and* $u_0(x) = -u_0(-x)$ *for* $x \in (1,2)$ *and correspondingly for* v_0. *For* $x < -1$

and $x > 2$ define both u_0 and v_0 to be identically equal to 0. Let U be the solution of

$$\ddot{U} - \Delta U + \alpha \dot{U} = 0, \quad \text{on } \mathbb{R} \times I$$
$$U(\cdot, 0) = U_0 = \pi^p u_0,\ \dot{U}(\cdot, 0) = V_0 = \pi^{p-1} v_0 \quad \text{on } \mathbb{R}.$$

Then

$$\begin{aligned} \|u - U\|^2_{L^2(K\times I)} \leq C h^{2p+3} \Big(& \|u_0^{(p+1)}\|^2_{\infty,(x_{j-1},x_{j+2})} + \|v_0^{(p)}\|^2_{\infty,(x_{j-1},x_{j+2})} \\ & + \max_{l\in\{j-1,j,j+1\}} |v_0^{(p)}(x_l)|^2 + \|v_0^{(p+1)}\|^2_{\infty,(x_{j-1},x_{j+2})} \Big), \end{aligned} \tag{4.6.3}$$

where $\|\cdot\|_\infty$ denotes the classical maximum norm.

Proof. Note that $u_0 = U_0 + R_{p+1}(x)$ and $v_0 = V_0 + R_p(x)$ where the remainders are defined as

$$R_{p+1}(x) = (x - x_j)^{p+1} \frac{u_0^{(p+1)}(\xi_x)}{(p+1)!},$$
$$R_p(x) = (x - x_j)^p \frac{v_0^{(p)}(\xi_x)}{(p)!}, \quad \xi_x \in (x_j, x), \quad x \in (x_j, x_{j+1}).$$

Note that the solution of

$$\ddot{\tilde{u}} - \Delta \tilde{u} + \alpha \dot{\tilde{u}} = 0, \quad \text{on } \mathbb{R} \times I,$$
$$\tilde{u}(\cdot, 0) = u_0, \quad \dot{\tilde{u}}(\cdot, 0) = v_0, \quad \text{on } \mathbb{R},$$

given by the solution formula

$$\begin{aligned} \tilde{u}(x,t) = & \frac{1}{2}[u_0(x-t) + u_0(x+t)]e^{-\alpha t/2} \\ & + \frac{1}{4}\alpha e^{-\alpha t/2} \int_{x-t}^{x+t} \tilde{u}_0(s) \left(I_0\left(\rho(s)\frac{t}{2}\alpha\right) + \frac{1}{\rho(s)} I_1\left(\rho(s)\frac{t}{2}\alpha\right)\right) ds \\ & + \frac{1}{2} e^{-\alpha t/2} \int_{x-t}^{x+t} \tilde{v}_0(s) I_0(\rho(s)\frac{t}{2}\alpha) ds, \end{aligned} \tag{4.6.4}$$

(where $\rho(s) = \sqrt{1 - (x-s)^2/t^2}$, and I_v is the modified Bessel function of the first kind of order ν) satisfies $\tilde{u} \equiv u$ on $\Omega \times I$ since they both have the same initial data and solve the same equation. The odd extension of the initial data ensures that $\tilde{u}$ satisfies the boundary conditions, i.e., $\tilde{u}(0,t) = \tilde{u}(1,t) = 0$ (Note that the integral of

an odd function over a symmetric interval is zero).

The advantage of the above consideration is that both u and u^h can be represented by the solution formula on $\mathbb{R}$ when restricted to $\Omega \times I$. Therefore the error $e = u - U$ is given by the same formula after a change of variable for the integrals as:

$$\begin{aligned} e(x,t) = {} & \frac{1}{2}\left(e_0(x-t) + e_0(x+t)\right) e^{-\alpha t/2} \\ & + \frac{1}{4}\alpha t e^{-\alpha t/2} \int_{-1}^{1} e_0(tq+x)\left(I_0\left(\sqrt{1-q^2}\frac{t\alpha}{2}\right) + \frac{1}{\sqrt{1-q^2}} I_1\left(\sqrt{1-q^2}\frac{t\alpha}{2}\right)\right) dq \\ & + \frac{1}{2} t e^{-t\alpha/2} \int_{-1}^{1} e_1(tq+x) I_0\left(\sqrt{1-q^2}\frac{t\alpha}{2}\right) dq, \end{aligned} \tag{4.6.5}$$

where

$$e_0 = u_0 - U_0, \qquad e_1 = v_0 - V_0.$$

We require the bound on the L^2 norm of the error, i.e.,

$$\|e\|^2_{L^2(K\times I)} = \int_0^h \int_{x_j}^{x_{j+1}} |e(x,t)|^2 dx dt. \tag{4.6.6}$$

Let us begin with the first term, thus

$$\begin{aligned} & \int_0^h e^{-\alpha t} \int_{x_j}^{x_{j+1}} \left|e_0(x-t) + e_0(x+t)\right|^2 dx dt \\ & \le 2 \int_0^h e^{-\alpha t} \left(\int_{x_j}^{x_{j+1}} \left|e_0(x-t)\right|^2 dx + \int_{x_j}^{x_{j+1}} \left|e_0(x+t)\right|^2 dx \right) dt \\ & \le 2 \int_0^h e^{-\alpha t} \left(\|e_0\|^2_{L^2(x_{j-1},x_{j+2})} + \|e_0\|^2_{L^2(x_j,x_{j+2})} \right) dt \\ & \le 4 \|e_0\|^2_{L^2(x_{j-1},x_{j+2})} \int_0^h e^{-\alpha t} dt \\ & = \frac{4}{\alpha}(1 - e^{-\alpha h}) \|e_0\|^2_{L^2(x_{j-1},x_{j+2})}. \end{aligned} \tag{4.6.7}$$

Before we continue, we note that there exists $C_0 > 0$ such that

$$I_0(x) \le C_0, \qquad x \in (0, h\alpha/2)$$

and

$$\frac{1}{\sqrt{1-q^2}} I_1(\sqrt{1-q^2}x) \le C_0, \quad x \in (0, h\alpha/2),\ q \in (-1,1).$$

The above estimate can be seen by inspection of the formula

$$I_n(z) = \left(\frac{1}{2}z\right)^n \sum_{k=0}^{\infty} \frac{\left(\frac{1}{4}z^2\right)^k}{k!(n+k)!},\ n \in \{0,1\}. \tag{4.6.8}$$

Note also that $I_0(x) \ge 0, I_1(x) \ge 0$ for $x \ge 0$.

We now proceed with the bound on the second term in (4.6.5):

$$\begin{aligned}
&\int_0^h t^2 e^{-\alpha t} \int_{x_j}^{x_{j+1}} \left| \int_{-1}^{1} e_0(tq+x) \left(I_0\left(\sqrt{1-q^2}\frac{t\alpha}{2}\right) + \frac{1}{\sqrt{1-q^2}} I_1\left(\sqrt{1-q^2}\frac{t\alpha}{2}\right)\right) dq \right|^2 dxdt \\
&\le 4C_0^2 \int_0^h t^2 e^{-\alpha t} \int_{x_j}^{x_{j+1}} \left(\int_{-1}^{1} |e_0(tq+x)| dq \right)^2 dxdt \\
&\le 16 C_0^2 \int_0^h t^2 e^{-\alpha t} \int_{x_j}^{x_{j+1}} \int_{-1}^{1} |e_0(tq+x)|^2 dq dx dt \\
&\le 16 C_0^2 \int_0^h t^2 e^{-\alpha t} dt \|e_0\|^2_{L^2(x_{j-1},x_{j+2})} \le \frac{32}{\alpha^3}\left(1 - e^{-\alpha h}\left(\frac{h^2\alpha^2}{2} + h\alpha + 1\right)\right) C_0^2 \|e_0\|^2_{L^2(x_{j-1},x_{j+2})}.
\end{aligned}$$

For the final term, we note that there exists a constant C_1 and a $h_0 > 0$ such that

$$|I_0(x) - 1| \le C_1 h^2, \qquad 0 \le h \le h_0.$$

Then we have

$$\begin{aligned}
&\int_0^h t^2 e^{-\alpha t} \int_{x_j}^{x_{j+1}} \left| \int_{-1}^{1} e_1(tq+x) I_0\left(\sqrt{1-q^2}\frac{t\alpha}{2}\right) dq \right|^2 dxdt \\
&\le\ 2 \int_0^h t^2 e^{-\alpha t} \int_{x_j}^{x_{j+1}} \left(\left| \int_{-1}^{1} e_1(tq+x) dq \right|^2 \right. \\
&\left. + \left| \int_{-1}^{1} e_1(tq+x)(I_0\left(\sqrt{1-q^2}\frac{t\alpha}{2}\right) - 1) dq \right|^2 \right) dxdt \\
&\le 2 \int_0^h t^2 e^{-\alpha t} \int_{x_j}^{x_{j+1}} \left| \int_{-1}^{1} e_1(tq+x) dq \right|^2 dxdt \\
&+ C_1^2 h^4 \frac{16}{\alpha^3}\left(1 - e^{-\alpha h}\left(\frac{h^2\alpha^2}{2} + h\alpha + 1\right)\right) \|e_1\|^2_{L^2(x_{j-1},x_{j+2})} \\
&\le \frac{2}{\alpha}(1 - e^{-\alpha h}) \max_{t\in(0,h)} \int_{x_j}^{x_{j+1}} \left| \int_{x-t}^{x+t} e_1(q) dq \right|^2 dx \\
&+ C_1^2 h^4 \frac{16}{\alpha^3}\left(1 - e^{-\alpha h}\left(\frac{h^2\alpha^2}{2} + h\alpha + 1\right)\right) \|e_1\|^2_{L^2(x_{j-1},x_{j+2})}.
\end{aligned}$$

We shall now make use of the definition of our projector (4.6.2) for the term

$$\int_{x_j}^{x_{j+1}} \left| \int_{x-t}^{x+t} e_1(q)dq \right|^2 dx$$

in our analysis.

We note that

$$\left| \int_{x-t}^{x+t} e_1(q)dq \right| \leq \left| \int_{(x-t,x+t)\cap(x_{j-1},x_j)} e_1(q)dq \right| + \left| \int_{(x-t,x+t)\cap(x_j,x_{j+1})} e_1(q)dq \right| + \left| \int_{(x-t,x+t)\cap(x_{j+1},x_{j+2})} e_1(q)dq \right| .$$

Let $(a,b) = (x-t, x+t) \cap (x_l, x_{l+1})$ where $l \in \{j-1, j, j+1\}$. Making use of the remainder now we have

$$\begin{aligned}
&\left| \int_a^b e_1(q)dq \right| = \left| \int_a^b R_p(q)dq \right| = \left| \int_a^b \frac{(q-x_l)^p}{p!} v_0^{(p)}(\xi_q)dq \right| \leq |v_0^{(p)}(x_l)| \left| \int_a^b \frac{(q-x_l)^p}{p!} dq \right| \\
&+ \left| \int_a^b \frac{(q-x_l)}{p!} (v_0^{(p)}(\xi_q) - v_0^{(p)}(x_k))dq \right| \leq Ch^{p+1}|v_0^{(p)}(x_l)| + C \max_{y\in(x_l,x_{l+1})} |v_0^{(p)}(y) - v_0^{(p)}(x_l)|h^p \\
&\leq Ch^{p+1} \max_{l\in\{j-1,j,j+1\}} |v_0^{(p)}(x_l)|^2 + Ch^{p+1} \|v_0^{(p+1)}\|_{\infty,(x_l,x_{l+1})}.
\end{aligned}$$

To complete the proof, we note that $\|e_0\|_{L^2(x_{j-1},x_{j+2})}$ and $\|e_1\|_{L^2(x_{j-1},x_{j+2})}$ can be estimated by bounding the remainders in classical maximum norm, i.e.,

$$\begin{aligned}
\|e_0\|^2_{L^2(x_{j-1},x_{j+2})} &= \|R_{p+1}\|^2_{L^2(x_{j-1},x_{j+2})} \leq Ch^{2p+2} \|u_0^{(p+1)}\|^2_{\infty,(x_{j-1},x_{j+2})} \\
\|e_1\|^2_{L^2(x_{j-1},x_{j+2})} &= \|R_p\|^2_{L^2(x_{j-1},x_{j+2})} \leq Ch^{2p} \|v_0^{(p)}\|^2_{\infty,(x_{j-1},x_{j+2})}.
\end{aligned} \tag{4.6.9}$$

We note that the coefficient $(1 - e^{-\alpha h})$ is of size h and we can verify that $(1 - e^{-\alpha h}(h^2\alpha^2/2 + h\alpha + 1))$ is of size h^3 by noticing that $1 + h\alpha + h^2\alpha^2/2 + O(h^3)$ is a Taylor expansion of $e^{\alpha x}$, $x \in (0, h)$. Combining the analysis now, we write

$$\begin{aligned}
\|e\|^2_{L^2(K\times I)} \leq Ch^{2p+3} \Bigg[&\left(\frac{4}{\alpha} + \frac{32}{\alpha^3} C_0^2 h^2 \right) \|u_0^{(p+1)}\|^2_{\infty,(x_{j-1},x_{j+2})} + \left(C_1 h^4 \frac{16}{\alpha^3} \right) \|v_0^{(p)}\|^2_{\infty,(x_{j-1},x_{j+2})} \\
&+ \left(\frac{2}{\alpha} \right) \|v_0^{(p+1)}\|^2_{\infty,(x_{j-1},x_{j+2})} + \max_{l\in\{j-1,j,j+1\}} |v_0^{(p)}(x_l)|^2 \Bigg].
\end{aligned} \tag{4.6.10}$$

Note that the right hand side of the estimate above presents the error in terms of the initial data with norms over spatial domain (x_{j-1}, x_{j+2}). □

Based on the above proposition, we now state and proof the theorem that generalises the local error estimate when the (weak) derivative D^α, $|\alpha| \le p$, is involved.

Theorem 4.6.4. *Let the setting of Proposition 4.6.3 hold and in addition, let $u \in \mathcal{C}^{p+1}(K \times I)$ with u_h defined as in Proposition 4.6.3. Then*

$$\begin{aligned}\|D^\beta(u-u^h)\|^2_{L^2(K\times I)} \le Ch^{2p+3-2|\beta|}\Big(&\|u_0^{(p+1)}\|^2_{\infty,(x_{j-1},x_{j+2})} + \|v_0^{(p)}\|^2_{\infty,(x_{j-1},x_{j+2})} \\ &+ \max_{l\in\{j-1,j,j+1\}} |v_0^{(p)}(x_l)|^2 + \|v_0^{(p+1)}\|^2_{\infty,(x_{j-1},x_{j+2})}\Big),\end{aligned} \tag{4.6.11}$$

where β represents space-time multi-index, $|\beta| \le p$.

Proof. Note that u^h is derived by propagating πu_0 and $\pi^{p-1}v_0$ in time by the solution formula (4.3.2). Hence, it suffices to investigate the bound on the derivatives of the error formula

$$\begin{aligned}e(x,t) = &\frac{1}{2}\left(e_0(x-t)+e_0(x+t)\right)e^{-\alpha t/2} \\ &+ \frac{1}{4}\alpha t e^{-\alpha t/2}\int_{-1}^{1} e_0(tq+x)\left(I_0\left(\sqrt{1-q^2}\frac{t\alpha}{2}\right) + \frac{1}{\sqrt{1-q^2}}I_1\left(\sqrt{1-q^2}\frac{t\alpha}{2}\right)\right)dq \\ &+ \frac{1}{2}te^{-t\alpha/2}\int_{-1}^{1} e_1(tq+x)I_0\left(\sqrt{1-q^2}\frac{t\alpha}{2}\right)dq\end{aligned} \tag{4.6.12}$$

as in Proposition 4.6.3. Also the functions $I_\nu(x)$, $\nu \in \{0,1\}$ and their derivatives in the error formula remain bounded and add no new information to the analysis.

For the spatial derivative, we discover that the number of terms remain unchanged in the error formula when continuously differentiated and similar analysis as in Proposition 4.6.3 can be employed. Suppose $e(x,t)$ is differentiated l times in space, then the following terms will occur in the error bound of the spatial derivatives:

$$\|e_0^{(l)}\|^2_{L^2(x_{j-1},x_{j+2})},\ \max_{l\in\{j-1,j,j+1\}}|v_0^{(p-l)}|^2,\ \text{and}\ \|v_0^{(p+1-l)}\|^2_{\infty,(x_{j-1},x_{j+2})}. \tag{4.6.13}$$

Now, note that

$$e_0(x) = u_0(x) - u_0^h(x) = R_{p+1}(x) = \frac{(x-x_j)^{p+1}}{(p+1)!}u_0^{(p+1)}(\xi), \quad \xi \in (x,x_j). \tag{4.6.14}$$

Therefore, we have the following estimates on the remainders:

$$\begin{aligned}\|e_0^{(l)}\|^2_{L^2(x_{j-1},x_{j+2})} &\leq Ch^{2p+2-2l}\|u_0^{(p+1)}\|^2_{\infty,(x_{j-1},x_{j+2})},\\ \|e_1^{(l)}\|^2_{L^2(x_{j-1},x_{j+2})} &\leq Ch^{2p-2l}\|v_0^{(p)}\|^2_{\infty,(x_{j-1},x_{j+2})}.\end{aligned} \tag{4.6.15}$$

Following the same approach as in Proposition 4.6.3 with the new estimates above shows that the same loss of order h can be derived, remembering that $(1-e^{-\alpha h})$ and $(1-e^{-\alpha h}\left(1+h\alpha+\frac{h^2\alpha^2}{2}\right))$ are of size h and h^3 respectively.

For the time derivatives, we notice that the number of terms increases as we compute the derivatives of $e(x,t)$ in time. However, by inspection, many of the terms can be analysed in similar way to Proposition 4.6.3.

Let us consider the time derivative of the third term in the error formula (4.6.5) which is the most interesting part of the analysis, i.e.,

$$\begin{aligned}&\frac{\partial}{\partial t}\left(\frac{1}{2}te^{-t\alpha/2}\int_{-1}^{1}e_1(tq+x)I_0\left(\sqrt{1-q^2}\frac{t\alpha}{2}\right)\right)\\ &=\left[\frac{1}{2}e^{-\alpha t/2}-\frac{t\alpha}{4}e^{-\alpha t/2}\right]\int_{-1}^{1}e_1(tq+x)I_0\left(\sqrt{1-q^2}\frac{t\alpha}{2}\right)dq\\ &+\frac{1}{2}te^{-\alpha t/2}\int_{-1}^{1}\left(e_1'(tq+x)I_0\left(\sqrt{1-q^2}\frac{t\alpha}{2}\right)+e_1(tq+x)\sqrt{1-q^2}\frac{\alpha}{2}I_1\left(\sqrt{1-q^2}\frac{t\alpha}{2}\right)\right)dq.\end{aligned} \tag{4.6.16}$$

We now need to carefully analyse the L^2 bound of the above expression. If we split the first term on the right into two, we discover that the first part can be analysed as follows

$$\begin{aligned}\int_0^h e^{-\alpha t}\int_{x_j}^{x_{j+1}}\left|e_1(tq+x)I_0\left(\sqrt{1-q^2}t\frac{\alpha}{2}\right)dq\right|^2dxdt &\leq C_0^2\int_0^h e^{-\alpha t}\int_{x_j}^{x_{j+1}}\left(\int_{-1}^{1}|e_1(tq+x)|dq\right)^2dxdt\\ &\leq \frac{4}{\alpha}C_0^2(1-e^{-\alpha h})\|e_1\|^2_{L^2(x_{j-1},x_{j+2})},\end{aligned} \tag{4.6.17}$$

where we have used Cauchy-Schwarz in the second step. The analysis of the second

part has been carefully done under Proposition 4.6.3, thus:

$$\begin{aligned}&\int_0^h t^2 e^{-\alpha t}\int_{x_j}^{x_{j+1}}\left|\int_{-1}^{1} e_1(tq+x)I_0\left(\sqrt{1-q^2}t\alpha/2\right)dq\right|^2 dxdt \quad \leq \\ &C_1^2 h^4\frac{16}{\alpha^3}\left(1-e^{-\alpha h}\left(\frac{h^2\alpha^2}{2}+h\alpha+1\right)\right)\|e_1\|^2_{L^2(x_{j-1},x_{j+2})} \\ &+\frac{2}{\alpha}(1-e^{-h\alpha})\max_{t\in(0,h)}\int_{x_j}^{x_{j+1}}\left|\int_{-1}^{1} e_1(tq+x)dq\right|^2 dx.\end{aligned} \tag{4.6.18}$$

In a similar way, the second term on the right of (4.6.16) can be split into two and analysed separately. Following the same approach as in Proposition 4.6.3, we have

$$\begin{aligned}&\int_0^h t^2 e^{-\alpha t}\int_{x_j}^{x_{j+1}}\left|\int_{-1}^{1} e_1'(tq+x)qI_0\left(\sqrt{1-q^2}t\frac{\alpha}{2}\right)dq\right|^2 dxdt \\ &\leq\frac{2}{\alpha}(1-e^{-\alpha h})\max_{t\in(0,h)}\int_{x_j}^{x_{j+1}}\left|\int_{-1}^{1} e_1'(tq+x)dq\right|^2 dx \\ &+C_1^2 h^4\frac{16}{\alpha^3}\left(1-e^{-\alpha h}\left(\frac{h^2\alpha^2}{2}+h\alpha+1\right)\right)\|e_1'\|^2_{L^2(x_{j-1},x_{j+2})}.\end{aligned} \tag{4.6.19}$$

The second part has been analysed already in the proof of the proposition.

We can now complete the proof by using the definition of our projection (4.6.2) together with standard estimates. The case of higher temporal derivatives as well as mixed derivatives can be analysed in the same fashion.

□

Theorem 4.6.5. *Let the exact solution $u \in \mathcal{X}$ be such that for each space time element $K \times I_n$, $u|_{K\times I} \in \mathcal{C}^{s+1}(K\times I)$ for some $0 \leq s \leq p$. With the settings of Proposition 4.6.3 and the validity of Theorem 4.6.4 we have*

$$|||U-u||| \leq C(u)h^{s-1/2}, \tag{4.6.20}$$

where $h = \max_{K,n} h_K^{(n)}$ and

$$C(u) = \left(\max_{\substack{x\in\Omega\\ t\in[0,T]}}|\frac{d^{s+1}}{dx^{s+1}}u|^2 + \max_{\substack{x\in\Omega\\ t\in[0,T]}}|\frac{d^s}{dx^s}\dot{u}|^2 + \max_{\substack{x\in\Omega,\\ t\in[0,T]}}|\frac{d^{s+1}}{dx^{s+1}}\dot{u}|^2\right)^{1/2}. \tag{4.6.21}$$

Proof. Applying the result of Lemma 4.6.1 and taking note that Theorem 4.6.4

gives us the estimate on a space-time slab with $t \in [0, h]$. Figure 4.3 illustrates the propagation on subsequent space-time slabs. Hence we have

$$|||U - u|||^2 \leq C(h_K^{(n)})^{2s} \left(\sum_{n=0}^{N-1} \max_{x \in \Omega} |u^{(s+1)}(x, t_n)|^2 + \max_{x \in \Omega} |\dot{u}^{(s)}(x, t_n)|^2 + \max_{\substack{x \in \Omega \\ t \in [0,T]}} |\dot{u}^{(s+1)}(x, t_n)|^2 \right), \tag{4.6.22}$$

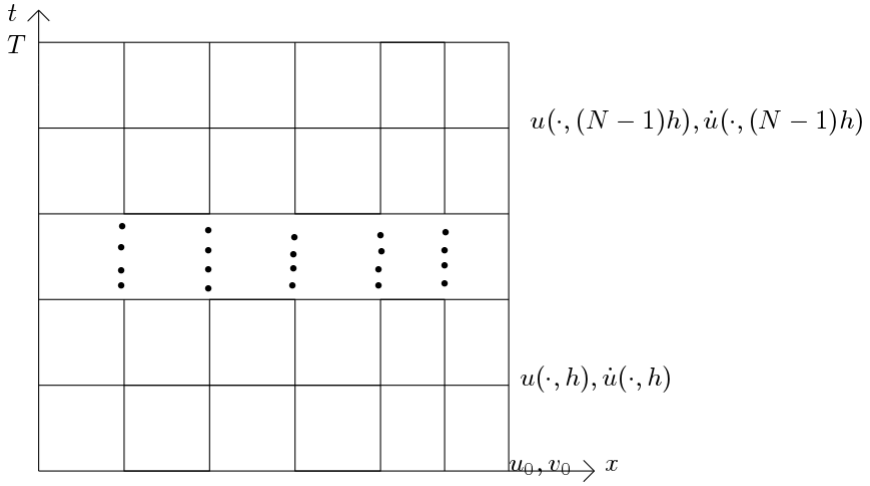


Figure 4.3: *Graphical illustration of propagation from each space-time slab*

where $t_n = nh$. Continuing with the bound we have

$$|||U - u|||^2 \leq Nh^{2s} \left(\max_{\substack{t \in [0,T] \\ x \in \Omega}} |u^{(s+1)}(x, t_n)|^2 + \max_{\substack{t \in [0,T] \\ x \in \Omega}} |\dot{u}^{(s)}(x, t_n)|^2 + \max_{\substack{x \in \Omega, \\ t \in [0,T]}} |\dot{u}^{(s+1)}(x, t_n)|^2 \right), \tag{4.6.23}$$

where with $N = T/h$ completes the proof.

□

Chapter 5

Numerical experiments

In this chapter we perform numerical experiments to assess the performance of the new space-time interior penalty scheme. We consider two types of wave problems for the case of one spatial dimension; the first is in the form of a standing wave while the second is Gaussian in nature due to exponential initial data. We assume that the speed $c = 1$ and that $\Omega = (0, 1)$. The first problem is considered to assess attributes such as convergence and dissipation while the second problem is considered to check the capability of the method in approximating systems with high frequency content.

In two dimensions, we consider a deformation or deflection of membrane problem. We make use of another special Trefftz functions which are capable of accommodating the directions of propagation of the wave problem. We also show in this chapter that equivalent order of convergence can be obtained with the Trefftz spaces with fewer number of degrees of freedom compared to standard polynomial spaces. We conclude the chapter with numerical experiments for the damped wave problem

In general, the cost of implementation, especially the computation of volume integrals increases with the number of degrees of freedom as the order of approximation is increased. However, the terms in the new space-time method that we develop, can be reduced to integrals over space and skeleton see, Remark 3.4.7. This can be counted as a considerable reduction in computational complexity as we shall see in the higher dimensional implementation of the Trefftz based method.

In each experiment, the spatial meshes are kept fixed $\mathcal{T} = \mathcal{T}_n$ and a uniform time-step is used. In the one-dimensional $(d = 1)$ examples the spatial mesh is a uniform

set of intervals, whereas for $d = 2$, the spatial mesh is a quasi-uniform triangulation. The resulting linear systems at each time-step are solved by standard sparse direct solvers.

5.1 Numerical experiments in one spatial dimension

Let us represent the undamped wave operator by $\Box = \partial_t^2 - \partial_x^2$ where we have fixed the diffusion coefficient $a \equiv 1$. Let us define $\hat{u}(x,t) = u\left(\frac{x-x_j}{h}, \frac{t-t_n}{h}\right)$ on a reference Trefftz element $(0,h) \times (0,h)$. This allows us to define each basis function on the reference Trefftz element as $\hat{v} = v\left(\frac{x-x_j}{h}, \frac{t-t_n}{h}\right)$ so that $\Box\hat{v} = \frac{1}{h^2} \Box v = 0$ on each element.

5.1.1 Approximation of standing wave problem

Example 5.1.1. *We consider the following wave equation defined on the spatial domain* $\Omega = (0,1)$ *with initial data*

$$u(x,0) = \sin(5\pi x) + 2\sin(7\pi x), \qquad \dot{u}(x,0) = 0. \tag{5.1.1}$$

The problem is solved analytically by the method of separation of variables:

$$u(x,t) = \sin(5\pi x)\cos(5\pi t) + 2\sin(7\pi x)\cos(7\pi t),$$

and the numerically obtained convergence orders are computed using the formula

$$\text{Convergence order} = \log_2\left(\frac{\text{error}_h}{\text{error}_{h/2}}\right). \tag{5.1.2}$$

The wave-like basis functions in our Trefftz spaces are $(x \pm t)^i$, $i = 0, \dots p$ where p is the highest order of interest. The approximation of the problem with final time $T = 1$ using the Trefftz space of order $p = 2$ and polynomial space of the same order is shown in Figure 5.1. We also investigate the convergence in the full DG norm $|||\cdot|||$

and we discover that the numerical experiment verifies the result of the convergence analysis as we lose half an order of convergence; see Table 5.1. The convergence plots with respect to number of degrees of freedoms and with respect to the mesh-size are presented in Figure 5.3 and 5.2. Finally, the plot of the error with CPU time (see, Figure 5.4) together with Table 5.2 shows that computation can be made faster with Trefftz spaces compared to polynomial spaces of total degrees.

N	$p=2$	$p=3$	$p=4$
20	1.07	2.50	3.50
40	1.39	2.53	3.53
80	1.48	2.52	3.55
160	1.47	2.51	3.52
320	1.49	2.50	3.54

N	$p=2$	$p=3$	$p=4$
20	0.93	2.06	3.40
40	1.28	2.19	3.44
80	1.44	2.31	3.48
160	1.45	2.39	3.49
320	1.49	2.45	3.50

Table 5.1: *Numerically obtained convergence orders in the DG norm $|||\cdot|||$ for Trefftz spaces on the left and for polynomial space on the right for the standing wave problem.*

We also investigate the convergence of the error in the wave energy norm at the final time-step:

$$\text{error} = \Big(\frac{1}{2}\|\dot{u}(\cdot,T) - \dot{u}_h(\cdot,T^-)\|_\Omega^2 + \frac{1}{2}\|\nabla u(\cdot,T) - \nabla u_h(\cdot,T^-)\|_\Omega^2\Big)^{1/2}. \tag{5.1.3}$$

This experiment shows optimal convergence as we do not lose half an order of convergence. This is because unlike the DG norm, this measure does not accumulate the error over all time-steps. The result is shown in Table 5.3

5.1.2 Wave with high energy content

Example 5.1.2. *We consider another wave problem with initial data*

$$u(x,0) = e^{-\left(\frac{x-5/8}{\delta}\right)^2}, \qquad \dot{u}(x,0) = 0, \tag{5.1.4}$$

N	$p=2$	$p=3$	$p=4$
20	2.02	2.45	3.57
40	3.98	5.90	7.14
80	13.44	19.79	23.55
160	45.74	61.95	79.38
320	178.80	226.74	263.40

N	$p=2$	$p=3$	$p=4$
20	6.80	17.90	40.10
40	17.5	46.60	113.50
80	57.60	138.60	278.40
160	203.70	449.40	811.30
320	670.50	1470.70	2783.90

Table 5.2: *Time elapsed in seconds for the computation of errors in energy DG norm $|||\cdot|||$ for Trefftz spaces (left) and for polynomial spaces (right).*

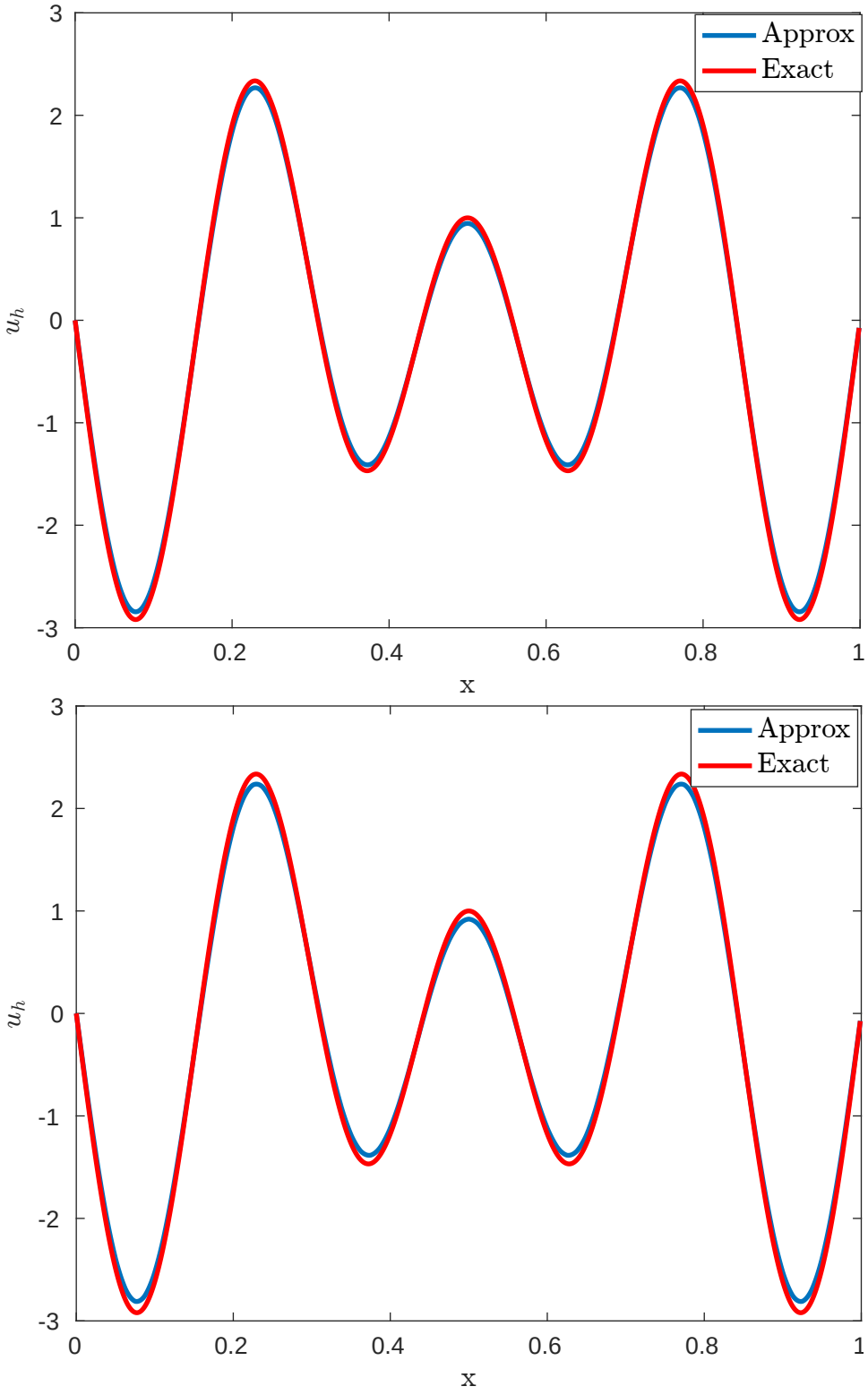


Figure 5.1: *Trefftz space approximation(upper) and polynomial space (lower).*

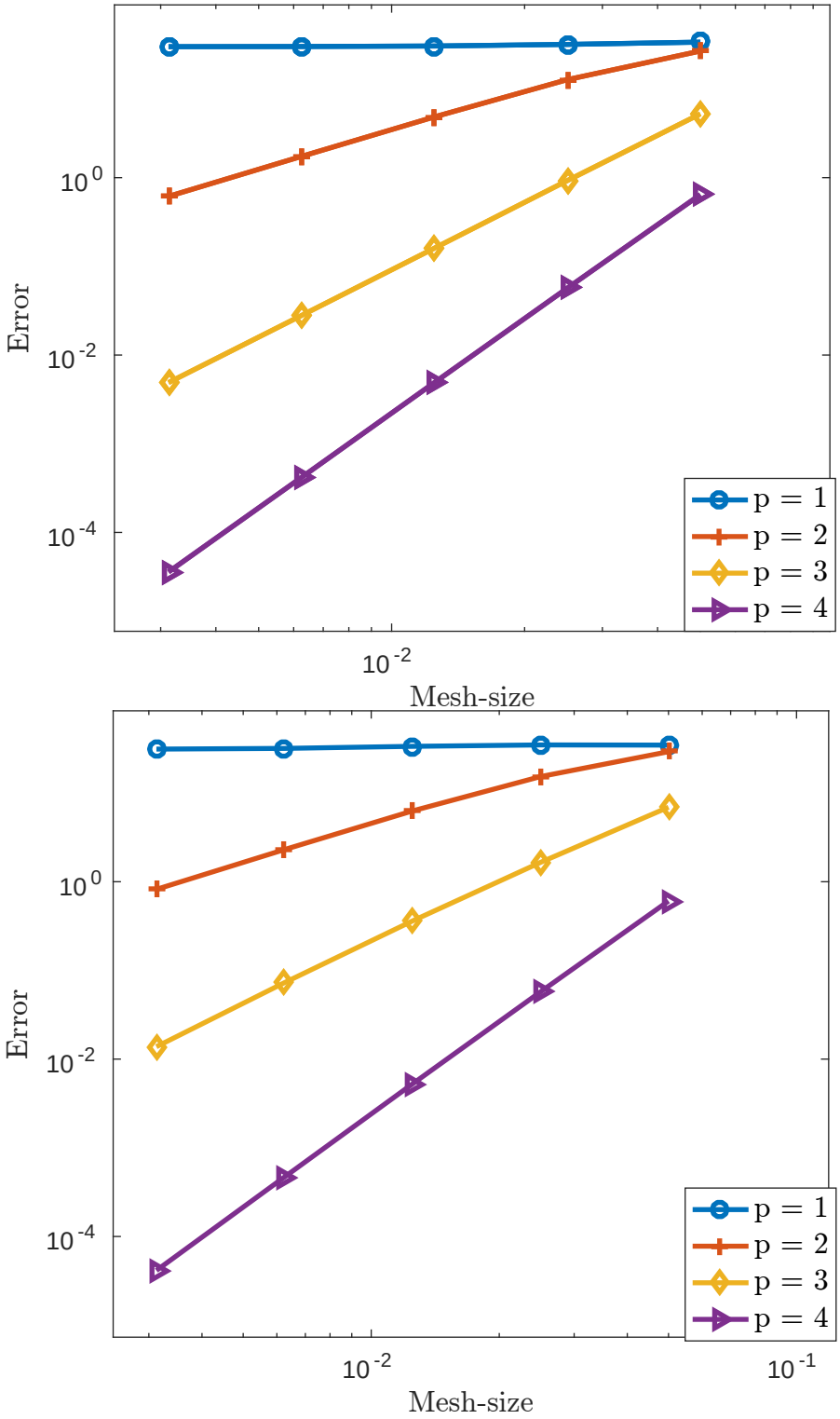


Figure 5.2: *Convergence of the error in the DG norm $|||\cdot|||$ with mesh size for Trefftz space(upper) and for polynomial space (lower) for the standing wave problem.*

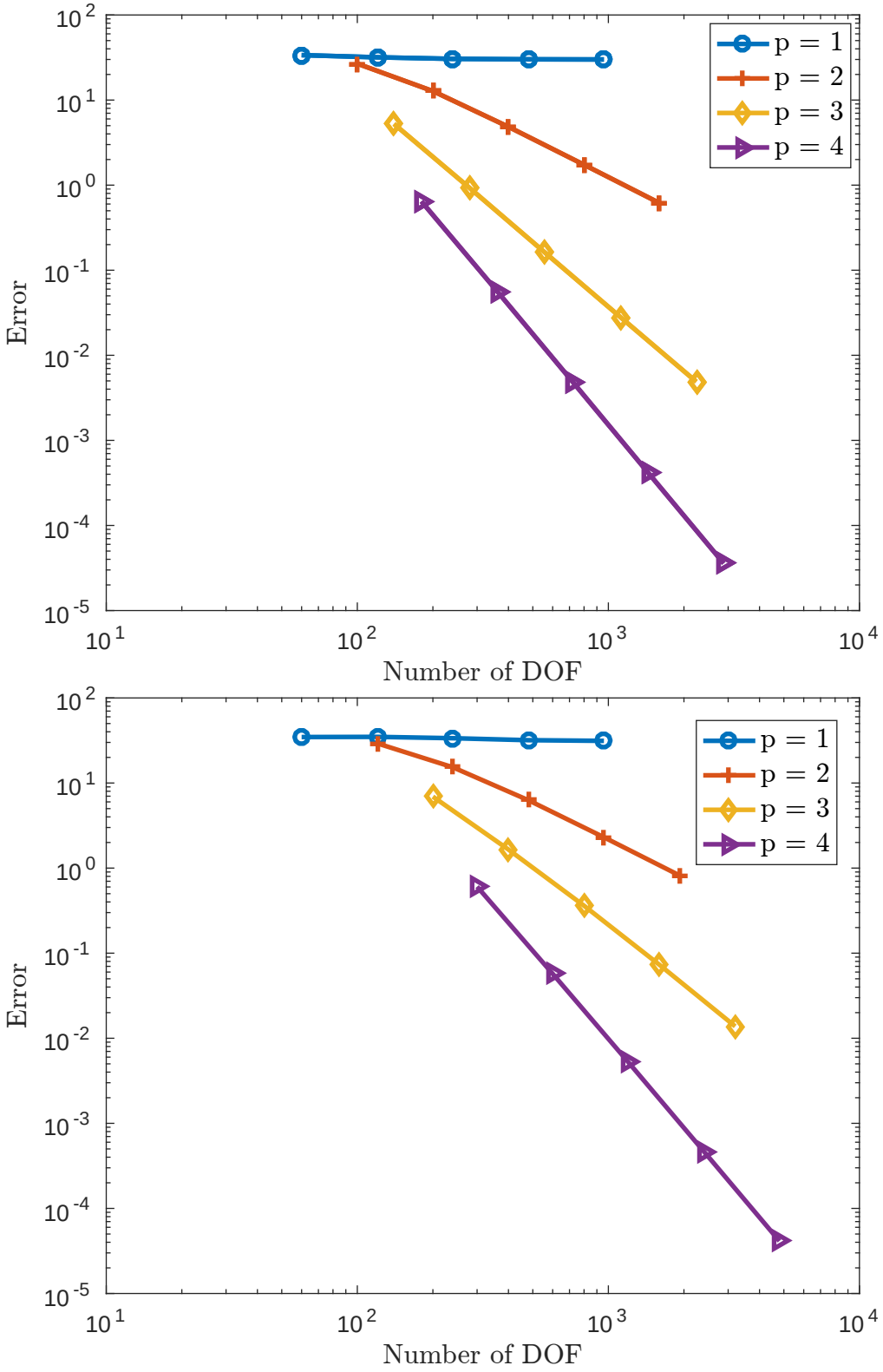


Figure 5.3: *Convergence of the error in the DG norm $|||\cdot|||$ with degrees of freedom for Trefftz space(upper) and for polynomial space (lower) for the standing wave problem.*

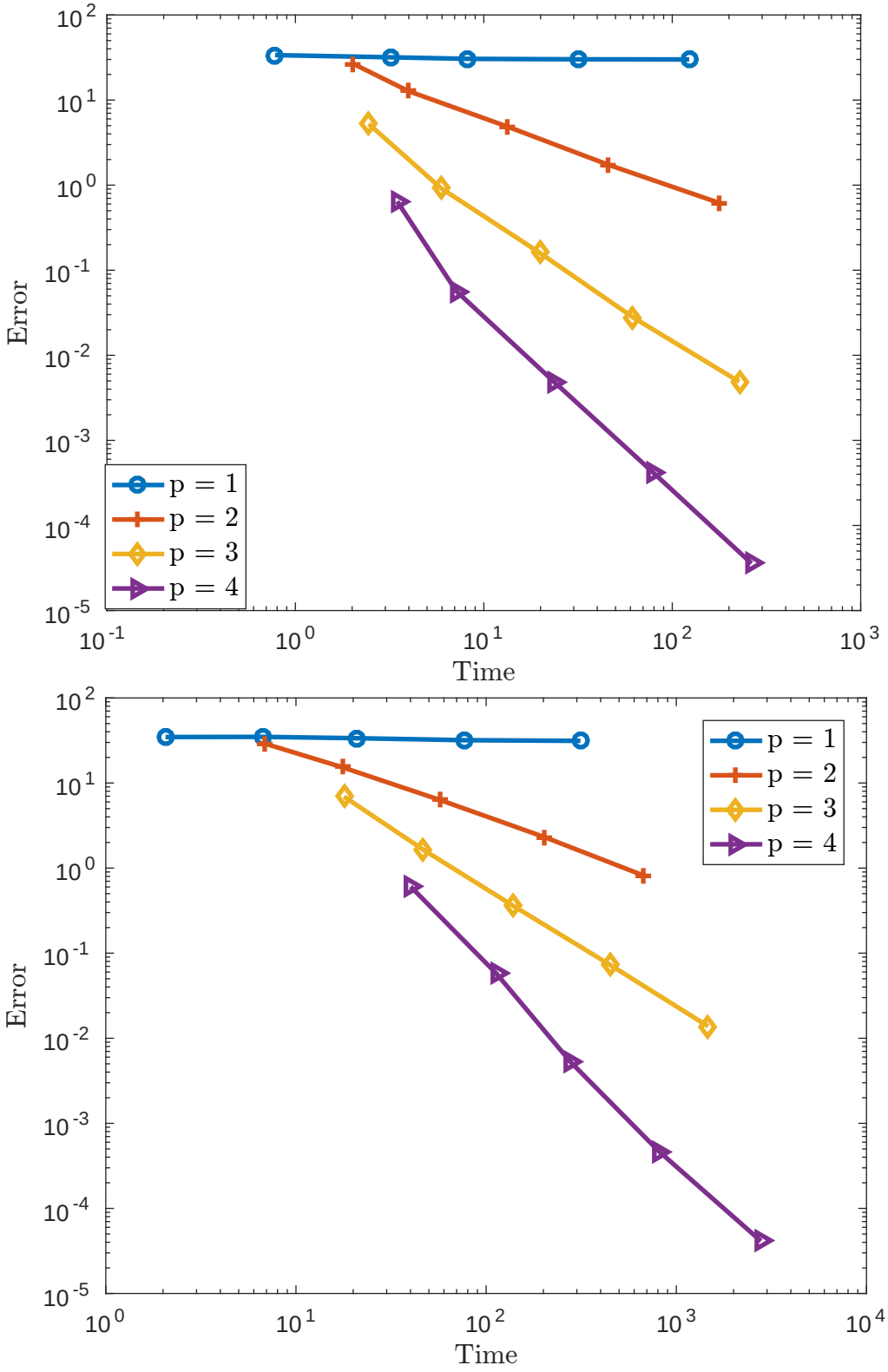


Figure 5.4: *Semilog plot of the error with CPU time in seconds for Trefftz spaces (upper) and polynomial spaces (lower).*

N	$p=2$	$p=3$	$p=4$
10	0.37	3.29	4.16
20	2.26	3.14	3.97
40	2.95	3.07	4.01
80	2.75	3.07	4.00
160	2.17	3.03	4.00

N	$p=2$	$p=3$	$p=4$
10	0.47	2.90	4.29
20	1.66	2.73	4.01
40	2.49	2.86	4.03
80	2.88	3.27	4.02
160	2.76	3.40	4.00

Table 5.3: *Numerically obtained convergence orders in the wave norm for Trefftz spaces on the left and for polynomial space on the right for the standing wave problem.*

where $\delta \leq \delta_0 = 7.5 \times 10^{-2}$. Note that the initial data are not exactly zero at the boundary, but are less than 10^{-11} in the range of parameter δ that we consider. This slight discrepancy with the boundary condition does not influence in any visible way our numerical results. Since the energy of the exact solution stays constant it is given for all times by

$$\text{exact energy} = \tfrac{1}{2}\|u_x(x,0)\|_\Omega^2 \approx 2\delta^{-1}\int_{-\infty}^{\infty} y^2 e^{-2y^2} dy = \delta^{-1}\frac{\sqrt{\pi}}{2\sqrt{2}},$$

where the approximation in the second step is of the order of 10^{-11} for reasons given above and the final equality is obtained by using integration by parts to reduce it to the Gaussian integral [58]. The error is computed in the discrete energy norm

$$\text{error} = |||u - u_h|||.$$

Since the exact solution is smooth, note that we have

$$|||u|||^2 = \mathcal{A}(u,u) = 2 \times \text{exact energy},$$

see (3.3.16).

We investigate the convergence order of the numerical method with Trefftz space and we compare with polynomial space of total degree. We choose $\delta_0 = 7.5 \times 10^{-2}$ and $T = 1/4$. Note that we choose such a small time interval in order to reach the asymptotic regime earlier, this is especially important for lower orders. In Figure 5.5 and Tables 5.4 and 5.5, the convergence curves and numerically computed convergence orders are presented. These confirm the theoretical results that we prove in

N	$p=2$	$p=3$	$p=4$	$p=5$
5	0.98	1.85	3.64	5.07
10	1.37	2.10	3.57	5.06
20	1.38	2.28	3.52	4.77
40	1.46	2.42	3.51	4.76
80	1.49	2.51	3.51	4.63

N	$p=2$	$p=3$	$p=4$	$p=5$
5	0.90	1.85	3.74	5.56
10	1.17	2.11	3.39	5.05
20	1.34	2.26	3.41	4.31
40	1.44	2.38	3.41	4.91
80	1.45	2.54	3.46	4.79

Table 5.4: Numerically obtained orders of convergence of the error in the DG norm $|||\cdot|||$ for Trefftz spaces on the left and for polynomial space on the right

Chapter 3. Note that the errors obtained for the full polynomial space and for the Trefftz spaces are very similar for the same order, but the Trefftz spaces require fewer degrees of freedom and cheaper implementation; see Remark 3.4.7 and Figure 5.6. We also discover that the higher order approximations converge without the two extra stabilization terms, i.e., with $\sigma_1 = \sigma_2 = 0$, but with the piecewise linear functions it stagnates. Recall that the dimensions of the Trefftz space and the polynomial space are equal when $p = 1$ (see, Subsection 3.4.1). Table 5.5 shows that both spaces have equivalent convergence orders. Another important discovery is the loss of stability at $p = 4$ for the polynomial space during the experiment which does not occur in the case of Trefftz space. The instability for the polynomial space is easily rectified by increasing the stability parameter σ_0. We also present the exponential convergence of the scheme as investigated in Chapter 3.

No of Elements	80	160	320	640	1280	2560	5120
Trefftz/Polynomial (p=1)	0.08	0.11	0.19	0.29	0.38	0.44	0.47

Table 5.5: *Numerically obtained convergence orders for linear elements.*

K	$p=1$	$p=2$	$p=3$	$p=4$	$p=5$
10	30	50	70	90	110
20	60	100	140	180	220
40	120	200	280	360	440
80	240	400	560	720	880
160	480	800	1120	1440	1760

K	$p=1$	$p=2$	$p=3$	$p=4$	$p=5$
10	30	60	100	150	210
20	60	120	200	300	420
40	120	240	400	600	480
80	240	480	800	1200	1680
160	480	960	1600	2400	3360

Table 5.6: *Number of degrees of freedom for chosen number of elements for Trefftz spaces on the left and for polynomial spaces on the right.*

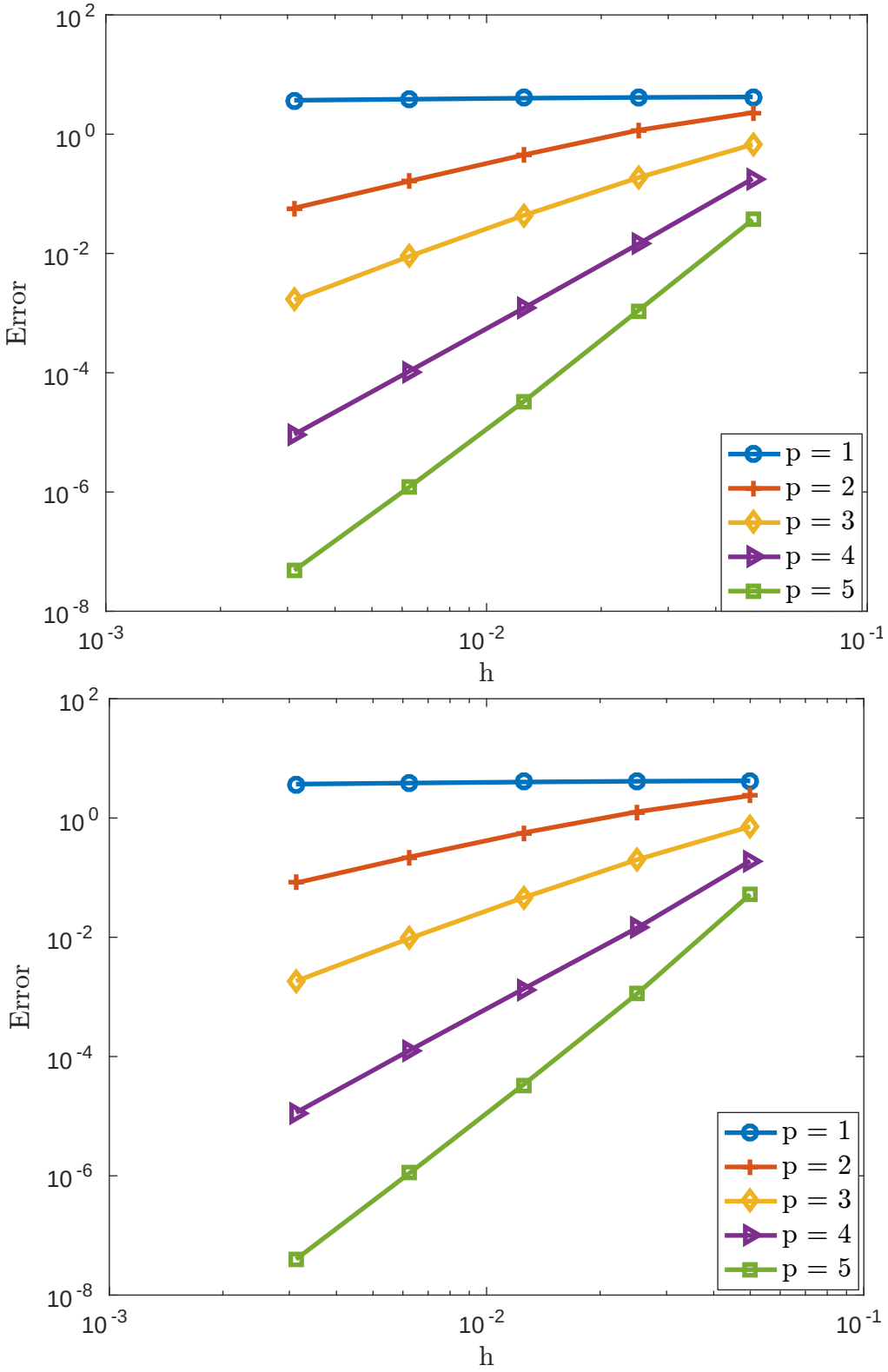


Figure 5.5: *Convergence of the error in the DG norm $|||\cdot|||$ for Trefftz (upper), and polynomial (lower), space-time DG method of order p. The error is plotted against the uniform mesh width in time and space $h = T/N$.*

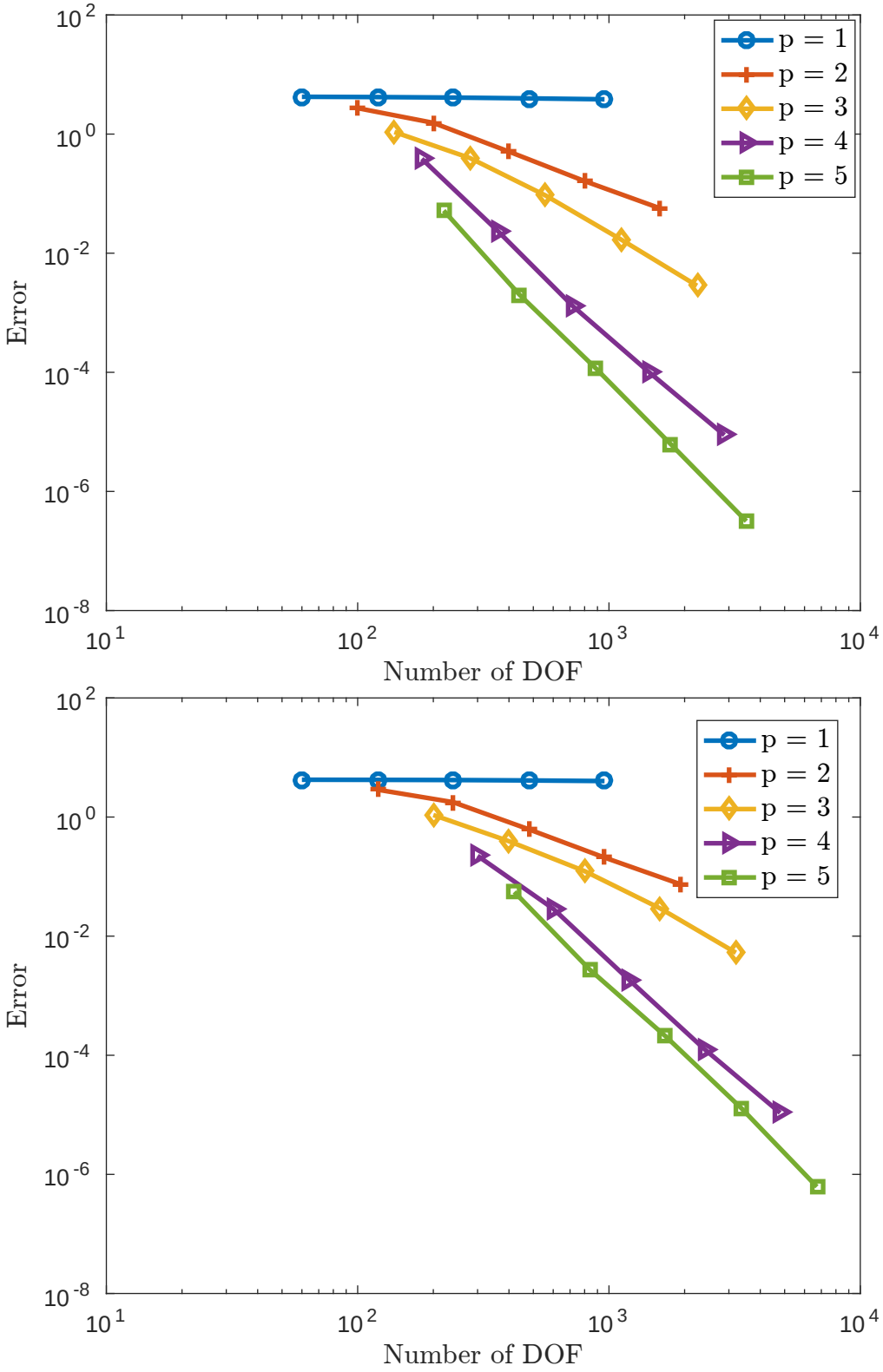


Figure 5.6: *Convergence of the error in the DG norm $|||\cdot|||$ for Trefftz spaces (upper), and polynomial spaces (lower). The error is plotted against number of degrees of freedom.*

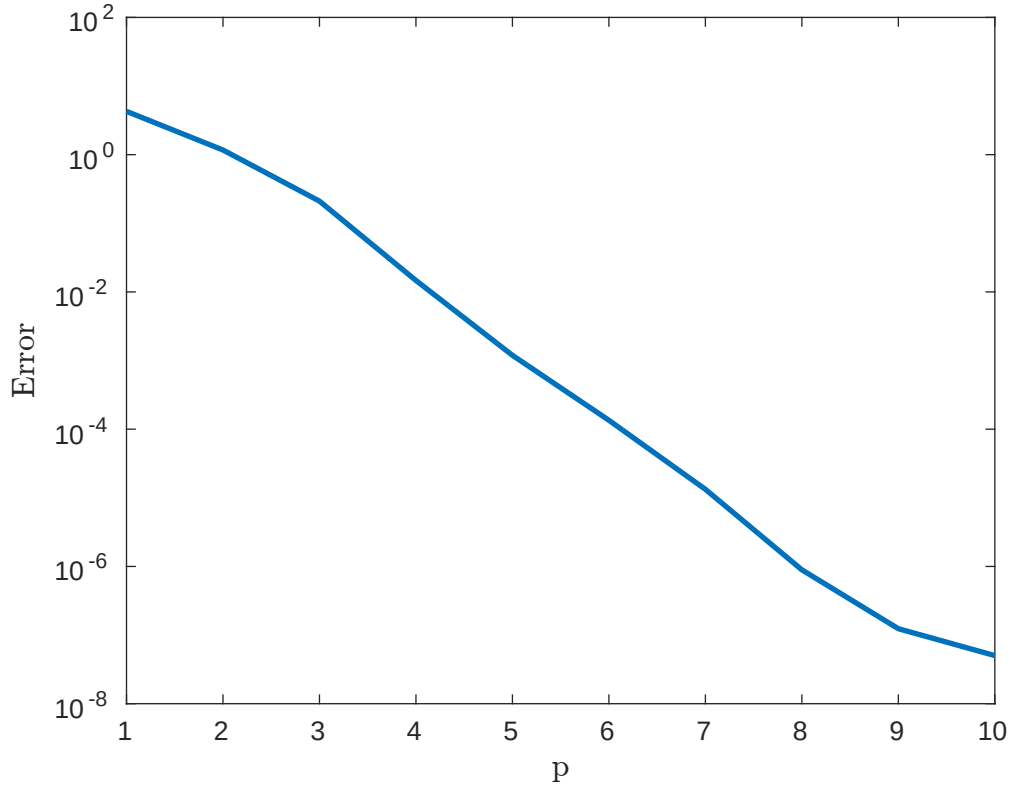


Figure 5.7: Convergence of the Trefftz method with fixed mesh width $h = 1/40$ and increasing polynomial order p.

N	$p = 2$	$p = 3$	$p = 4$
5	2.28×10^{0}	6.67×10^{-1}	1.30×10^{-1}
10	1.11×10^{0}	1.34×10^{-1}	1.10×10^{-2}
20	4.41×10^{-1}	2.37×10^{-2}	1.14×10^{-3}
40	1.60×10^{-1}	4.16×10^{-3}	1.07×10^{-4}
80	5.67×10^{-2}	7.29×10^{-4}	9.49×10^{-6}

N	$p = 2$	$p = 3$	$p = 4$
5	2.34×10^{0}	6.98×10^{-1}	2.23×10^{-1}
10	1.26×10^{0}	1.44×10^{-1}	2.88×10^{-2}
20	5.47×10^{-1}	2.53×10^{-2}	1.85×10^{-3}
40	2.08×10^{-1}	4.22×10^{-3}	1.27×10^{-4}
80	7.51×10^{-2}	7.14×10^{-4}	1.08×10^{-5}

Table 5.7: *Errors for Trefftz spaces on the left and polynomial spaces on the right.*

5.1.3 Long-time energy behaviour

The space-time DG method that we developed is dissipative, so we expect the energy

$$E(t) = \frac{1}{2}\|\dot{u}_h\|^2 + \frac{1}{2}\|\nabla u_h\|^2 \tag{5.1.5}$$

and the discrete energy E_h defined in (3.3.3) to decay over time. However, if the accuracy of the approximation is high we expect this decay to be very slow. This is indeed what the numerical experiments show in Figure 5.8, where we compute up to time $T = 5$ with $\delta_0/4$. We go further to show that the order $p = 4$ for both Trefftz space and polynomial space actually decay by plotting the semilog plot of the energy at long time $T = 200$ (see Figure 5.9). The non-monotone nature of the case $p = 2$ in Figure 5.9 can be controlled by increasing the penalty parameter σ_0 and σ_1.

5.1.4 Waves with energy at high-frequences

Note that if we decrease the parameter $\delta > 0$ in the definition of the initial data (5.1.4), the Gaussian becomes narrower and energy at higher frequences is excited. In the following set of experiments we investigate the error while decreasing both $\delta > 0$ and the mesh-width $h > 0$. In an ideal case, $h \propto \delta$ would be sufficient to obtain a constant relative error which we define as

$$\text{error}_\delta = \left(\frac{\delta}{2}\|\dot{u}(\cdot,T) - \dot{u}_h(\cdot,T^-)\|_\Omega^2 + \frac{\delta}{2}\|\nabla u(\cdot,T) - \nabla u_h(\cdot,T^-)\|_\Omega^2\right)^{1/2}. \tag{5.1.6}$$

As the Tables 5.8, 5.9, and 5.10 indicate, the lower order methods are far from this ideal, whereas order 4 Trefftz method for the set of experiments we performed comes very close to it. The same phenomenon is shown with Figure 5.10.

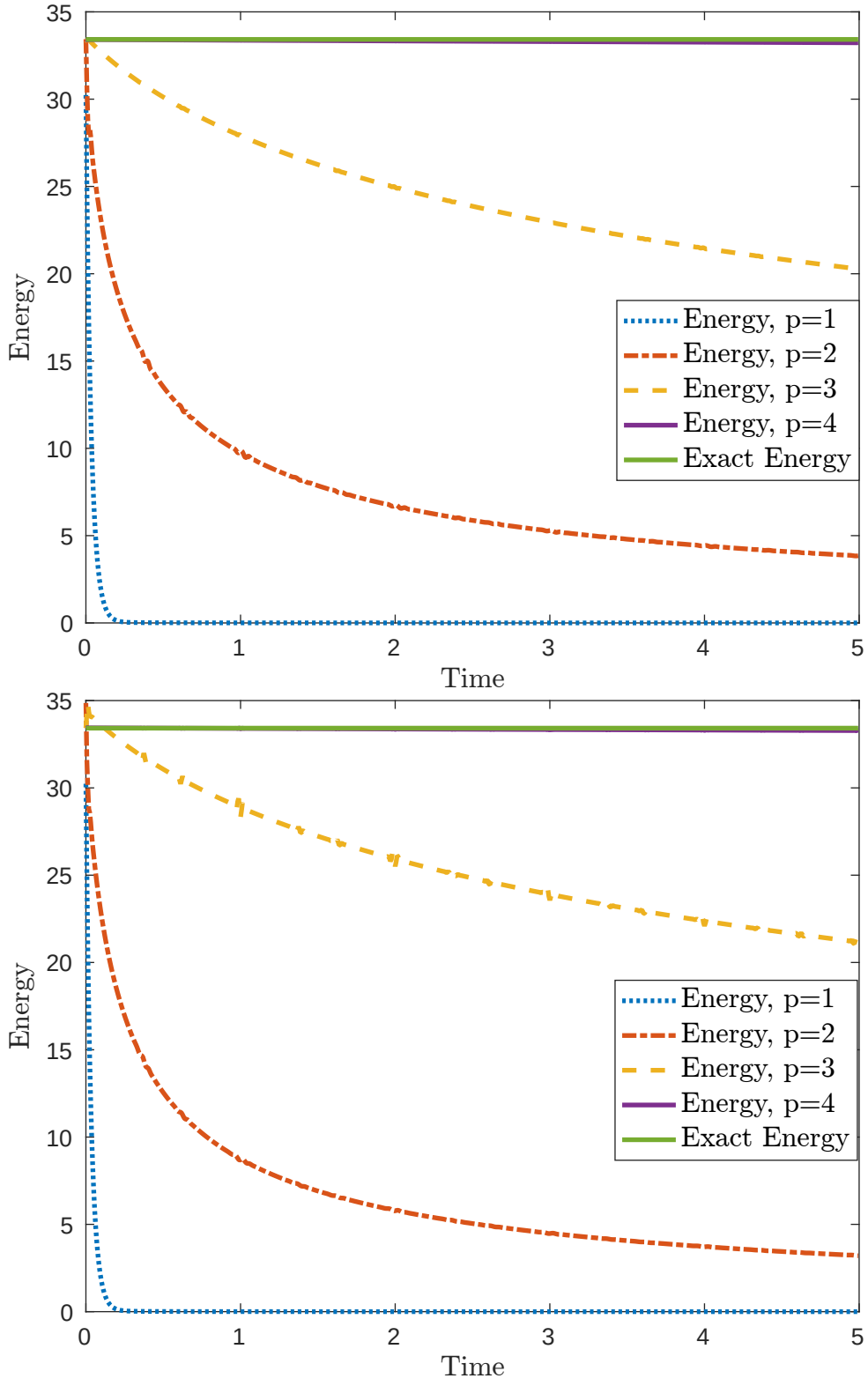


Figure 5.8: *Energy* $E(t) = \frac{1}{2}\|\dot{u}_h\|^2 + \frac{1}{2}\|\nabla u_h(t)\|^2$ *computed with different polynomial orders for the Trefftz spaces (up) and polynomial spaces (down). Note that the line corresponding to* $p = 4$ *is not visible as it is covered by the line for the exact energy. Plotting* E_h *instead of* E *essentially produce the same results.*

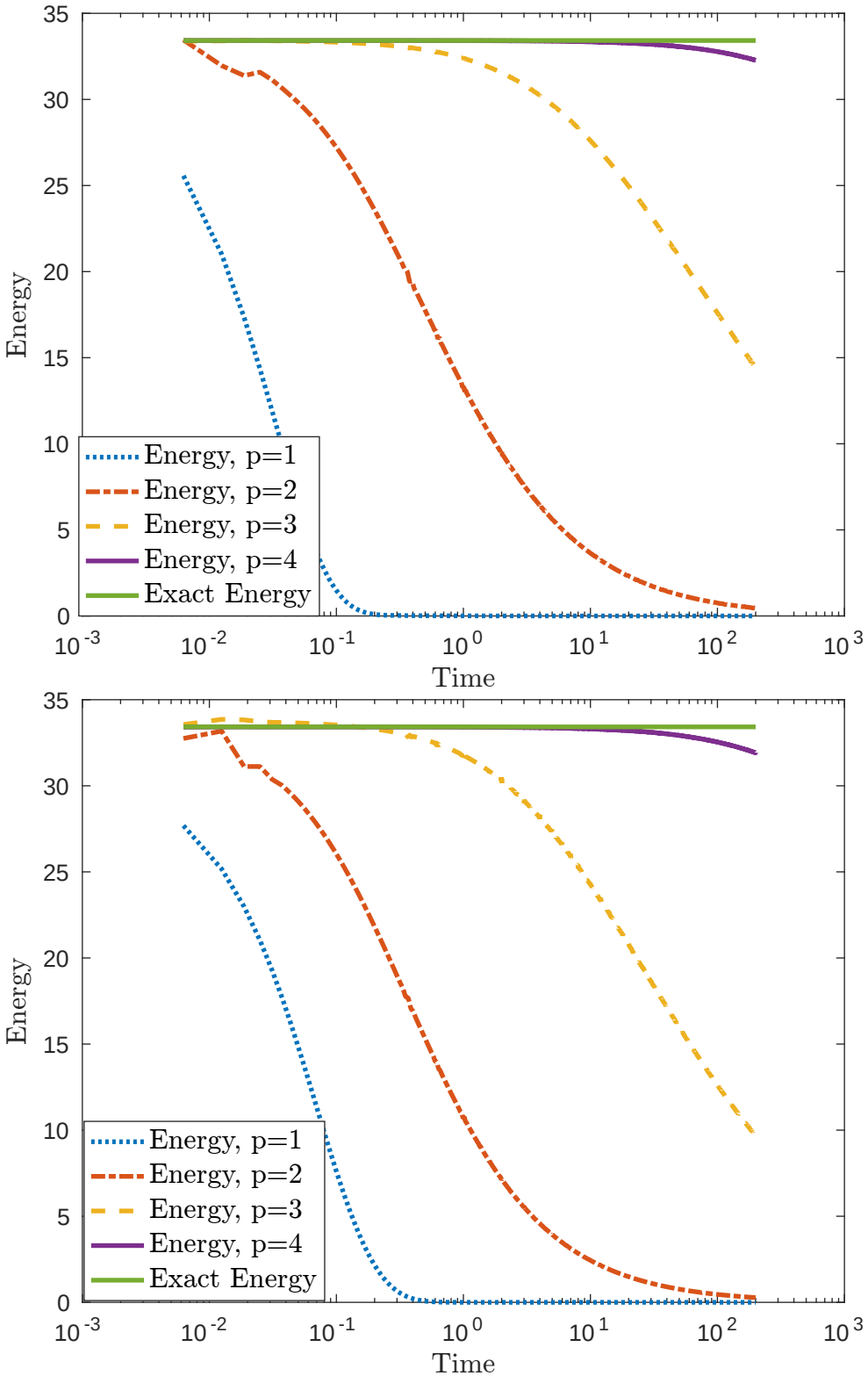


Figure 5.9: *Semilog plot of Energy* $E(t) = \frac{1}{2}\|\dot{u}_h(t)\|^2 + \frac{1}{2}\|\nabla u_h(t)\|^2$ *with time (Final time* $T = 200$*) computed with different polynomial orders for the Trefftz spaces (up) and polynomial spaces (down). Plotting* E_h *instead of* E *essentially produce the same results.*

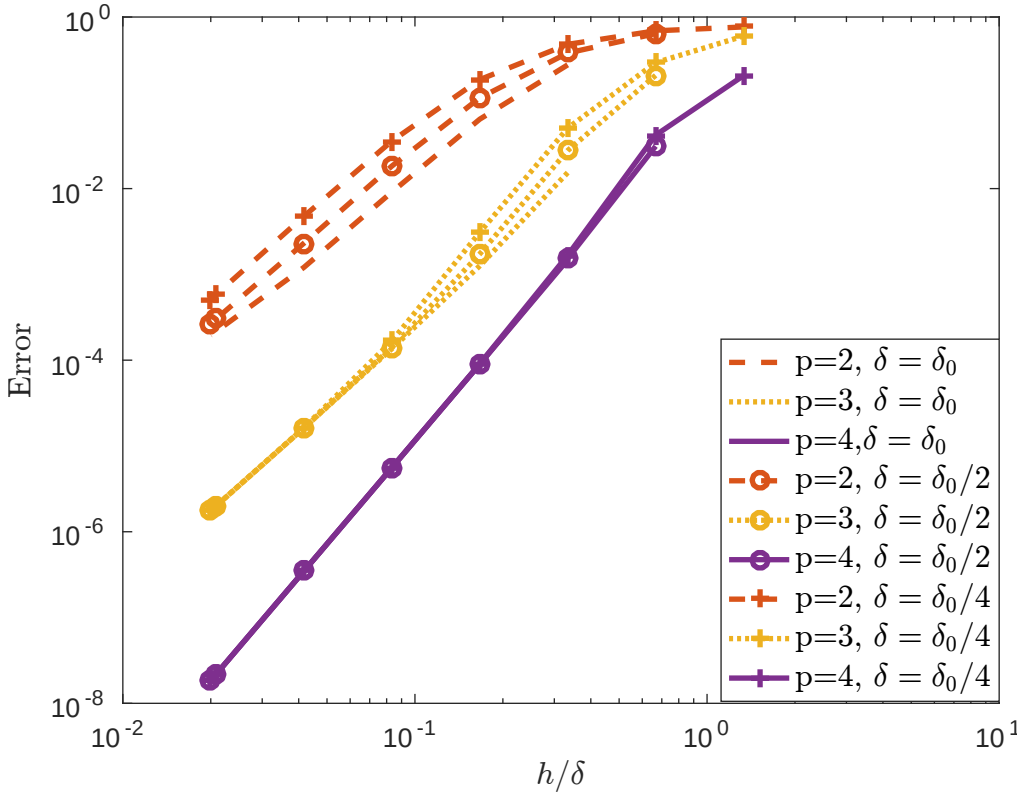


Figure 5.10: *The plot of Scaled error see* (5.1.6)*, against* h/δ*, for Trefftz space.The final time is chosen to be* $T = 1$ *and* $K_{max} = round(1/(2 \times 10^{-2} \times \delta))$

h/δ	$p = 3$	$p = 4$
0.3333	2.85×10^{-2}	1.32×10^{-3}
0.1666	1.54×10^{-3}	9.72×10^{-5}
0.0833	1.32×10^{-4}	6.84×10^{-6}
0.0416	1.59×10^{-5}	4.45×10^{-7}
0.0208	1.98×10^{-6}	2.81×10^{-8}

h/δ	$p = 3$	$p = 4$
0.3333	4.24×10^{-2}	1.53×10^{-3}
0.1666	3.22×10^{-3}	8.79×10^{-5}
0.0833	2.96×10^{-4}	5.38×10^{-6}
0.0416	2.87×10^{-5}	3.33×10^{-7}
0.0208	2.76×10^{-6}	2.09×10^{-8}

Table 5.8: Scaled error with h/δ fixed, see (5.1.6), for Trefftz on the left and polynomial spaces on the right with $\delta = \delta_0$.

h/δ	$p = 3$	$p = 4$
0.6666	3.12×10^{-1}	3.70×10^{-2}
0.3333	5.08×10^{-2}	1.50×10^{-3}
0.1666	2.63×10^{-3}	9.82×10^{-5}
0.0833	1.49×10^{-4}	6.85×10^{-6}
0.0416	1.60×10^{-5}	4.45×10^{-7}

h/δ	$p = 3$	$p = 4$
0.6666	3.49×10^{-1}	5.37×10^{-2}
0.3333	7.46×10^{-2}	1.86×10^{-3}
0.1666	5.47×10^{-3}	8.93×10^{-5}
0.0833	3.43×10^{-4}	5.39×10^{-6}
0.0416	2.90×10^{-5}	3.33×10^{-7}

Table 5.9: Scaled error with h/δ fixed, see (5.1.6), for Trefftz on the left and polynomial spaces on the right with $\delta = \delta_0/2$.

h/δ	$p = 3$	$p = 4$
1.3333	6.64×10^{-1}	2.71×10^{-1}
0.6666	3.92×10^{-1}	5.79×10^{-2}
0.3333	8.55×10^{-2}	2.07×10^{-3}
0.1666	5.04×10^{-3}	1.00×10^{-4}
0.0833	2.08×10^{-4}	6.86×10^{-6}

h/δ	$p = 3$	$p = 4$
1.3333	6.81×10^{-1}	3.39×10^{-1}
0.6666	4.39×10^{-1}	8.26×10^{-2}
0.3333	1.24×10^{-1}	2.70×10^{-3}
0.1666	1.03×10^{-2}	9.32×10^{-5}
0.0833	5.11×10^{-4}	5.41×10^{-6}

Table 5.10: Scaled error with h/δ fixed, see (5.1.6), for Trefftz on the left and polynomial spaces on the right with $\delta = \delta_0/4$.

5.2 Numerical experiments with transparent conditions

In this section we carry out numerical simulation for a wave problem in one space dimension with reflecting (Dirichlet) boundary condition at the left, and a transparent condition at the right. See for example [76]. Initial data are chosen so that the solution splits into a left and right travelling wave. The right travelling wave passes through the transparent boundary, while the left travelling wave is reflected by the left hand boundary before passing through the right hand boundary. The solution is zero from then onwards. The numerical results below illustrate the performance of the new space-time DG method, and they accurately reproduce the behaviour of the exact solution.

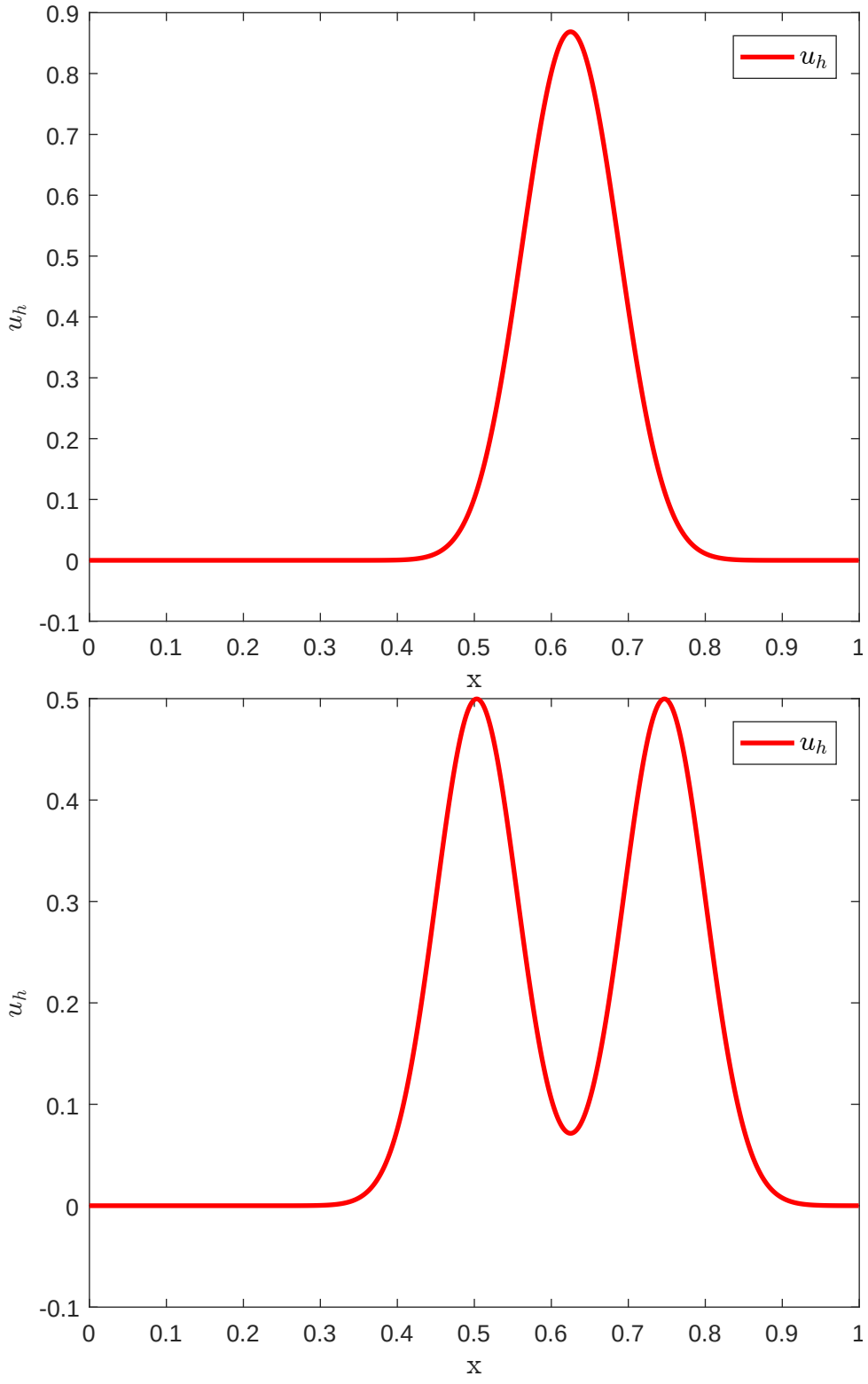


Figure 5.11: *Simulation at* $T = 1/32$ *(upper) and at* $T = 1/8$*(lower).*

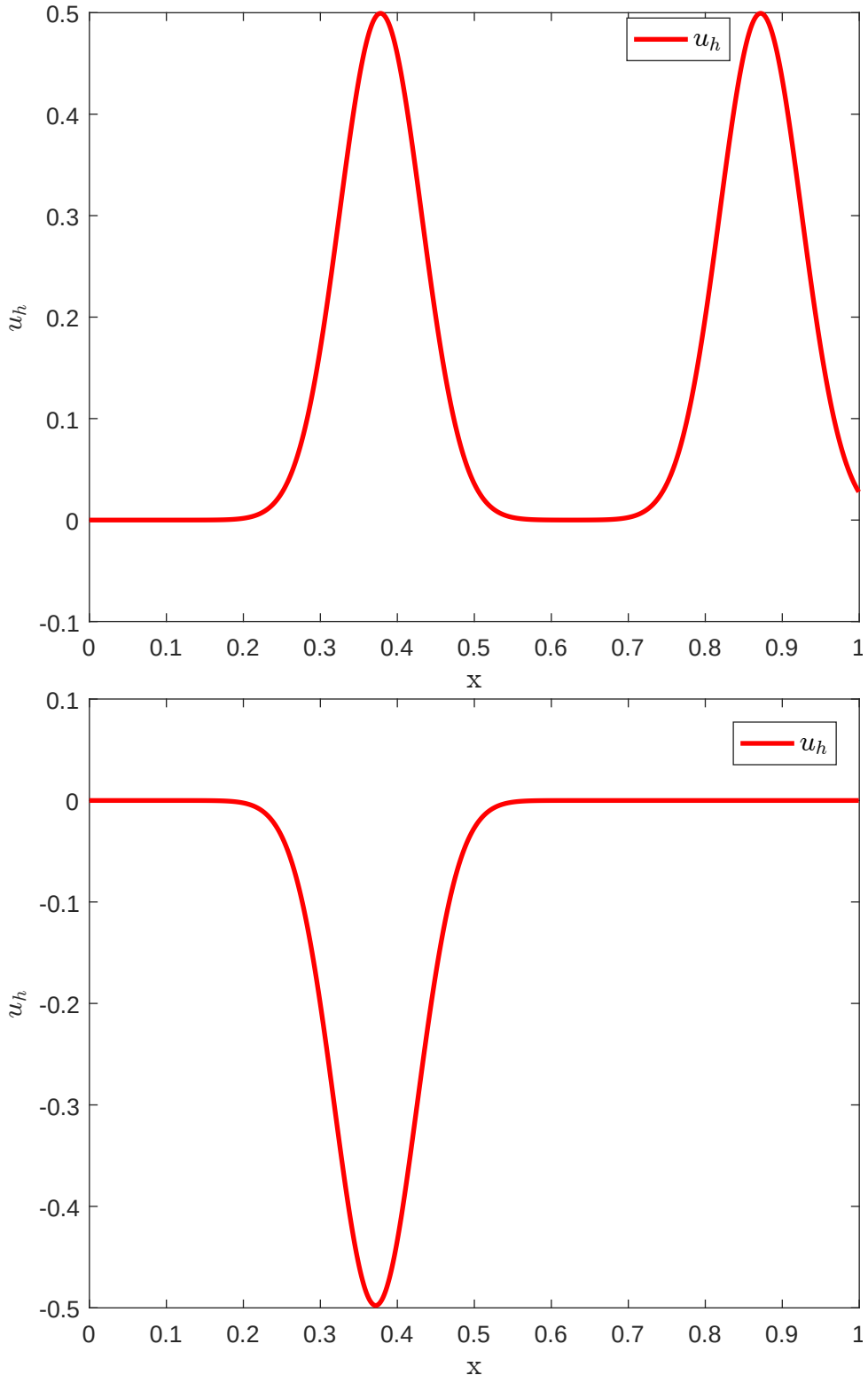


Figure 5.12: *Simulation at* $T = 1/4$ *(upper) and at* $T = 1$ *(lower).*

5.3 Higher dimensional implementation

In this section we present the implementation results of the Trefftz space-time DG scheme in higher spatial dimension. We focus particularly on two spatial dimensions, however implementation in three spatial dimensions is possible.

We generate two types of Trefftz basis functions; directional type and the set generated by Taylor's expansion. The directional basis is of the form $(t-\alpha\cdot\mathbf{x})^i,\ i = 0\dots p$, where α is a direction vector and p is the maximum order of the polynomial. We investigate the convergence order of the Trefftz based method numerically and we compare with polynomial spaces of total degree.

5.3.1 Derivation of Trefftz basis functions from truncation of Taylor polynomial

In this section we briefly discuss the first type of Trefftz basis functions used in the implementation of our Trefftz based method. We study in particular the technique developed by Artur Maciąg and Jörg Wauer in [92]. We assume that $u \in \mathcal{C}^{N+1}$ in the neighborhood of (x_0, y_0, t_0) and we expand the solution u in Taylor's series. The coefficients of the terms in the expansion up to the interested order p give us the required Trefftz basis functions after the elimination of $\frac{\partial^2 u}{\partial t^2}$ in the expansion. We illustrate the technique with examples when $N = 2, 3$ and 4.

Let $N = 2$ and expand the solution $u(x, y, t)$ in Taylor's series to have

$$\begin{aligned} u(x,y,t) = u(x_0,y_0,t_0) &+ \frac{\partial u}{\partial x}\hat{x} + \frac{\partial u}{\partial y}\hat{y} + \frac{\partial u}{\partial t}\hat{t} + \frac{\partial^2 u}{\partial x^2}\frac{\hat{x}^2}{2} \\ &+ \frac{\partial^2 u}{\partial y^2}\frac{\hat{y}^2}{2} + \frac{\partial^2 u}{\partial t^2}\frac{\hat{t}^2}{2!} + \frac{\partial^2 u}{\partial x\partial y}\hat{x}\hat{y} + \frac{\partial^2 u}{\partial x\partial t}\hat{x}\hat{t} + \frac{\partial^2 u}{\partial y\partial t}\hat{t}\hat{y} + R_3, \end{aligned} \tag{5.3.1}$$

where $\hat{x} = x - x_0$, $\hat{y} = y - y_0$, $\hat{t} = t - t_0$ and R_3 represent the remainder. We eliminate $\frac{\partial^2 u}{\partial t^2}$ in the above equation by substituting $\frac{\partial^2 u}{\partial^2 t} = \frac{\partial^2 u}{\partial^2 x} + \frac{\partial^2 u}{\partial y^2}$ to get

$$\begin{aligned} u(x,y,t) = u(x_0,y_0,t_0,) &+ \frac{\partial u}{\partial x}\hat{x} + \frac{\partial u}{\partial y}\hat{y} + \frac{\partial u}{\partial t}\hat{t} + \frac{\partial^2 u}{\partial x^2}\left(\frac{\hat{x}^2}{2} + \frac{\hat{t}^2}{2}\right) \\ &+ \frac{\partial^2 u}{\partial y^2}\left(\frac{\hat{y}^2}{2} + \frac{\hat{t}^2}{2}\right) + \frac{\partial u}{\partial x\partial y}\hat{x}\hat{y} + \frac{\partial^2 u}{\partial x\partial t}\hat{x}\hat{t} + \frac{\partial^2 u}{\partial y\partial t}\hat{y}\hat{t} + R_3. \end{aligned} \tag{5.3.2}$$

The basis functions required are the coefficients of the equation above. We arrange the Trefftz polynomial basis functions up to order $p = 2$ as follow:

$$\left\{1, \hat{x}, \hat{y}, \hat{t}, \frac{\hat{x}^2}{2} + \frac{\hat{t}^2}{2}, \frac{\hat{y}^2}{2} + \frac{\hat{t}^2}{2}, \hat{x}\hat{y}, \hat{x}\hat{t}, \hat{y}\hat{t}\right\}. \tag{5.3.3}$$

To derive the Trefftz basis functions when $N = 3$, we focus only on the higher order Taylor expansion of function u and we have

$$\begin{aligned}&\frac{\partial^3 u}{\partial x^3}\frac{\hat{x}^3}{3!} + \frac{\partial^3 u}{\partial x^2 \partial y}\frac{\hat{x}^2}{2!}\hat{y} + \frac{\partial^3 u}{\partial x^2 \partial t}\frac{\hat{x}^2}{2!}\hat{t} + \frac{\partial^3 u}{\partial x \partial y \partial t}\hat{x}\hat{y}\hat{t} \\ &+ \frac{\partial^3 u}{\partial y^2 \partial x}\frac{\hat{y}^2}{2}\hat{x} + \frac{\partial u}{\partial^2 y \partial t}\frac{\hat{y}^2}{2!}\hat{t} + \frac{\partial u}{\partial t^2 \partial x}\frac{\hat{t}^2}{2!}\hat{x} + \frac{\partial^3 u}{\partial t^2 \partial y}\frac{\hat{t}^2}{2!}\hat{y} + \frac{\partial^3 u}{\partial y^3}\frac{\hat{y^3}}{3!} + \frac{\partial^3}{\partial t^3}\frac{\hat{t}^3}{3!} + \frac{\partial^3 u}{\partial x \partial y \partial t}\hat{x}\hat{y}\hat{t}.\end{aligned} \tag{5.3.4}$$

Eliminating $\frac{\partial^u}{\partial t^2}$ by substituting $\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}$ we have

$$\begin{aligned}&\frac{\partial^3 u}{\partial x \partial y \partial t}\hat{x}\hat{y}\hat{t} + \frac{\partial^3 u}{\partial x^3}\left(\frac{\hat{x}^3}{3!} + \frac{\hat{t}^2}{2!}\hat{x}\right) + \frac{\partial^3 u}{\partial x^2 \partial y}\left(\frac{\hat{x}^2}{2!}\hat{y} + \frac{\hat{t}^2}{2!}\hat{y}\right) \\ &+ \frac{\partial^u}{\partial x^2 \partial t}\left(\frac{\hat{x}^2}{2!}\hat{t} + \frac{\hat{t}^3}{3!}\right) + \frac{\partial^3 u}{\partial y^2 \partial x}\left(\frac{\hat{y}^2}{2!}\hat{x} + \frac{\hat{t}^2}{2!}\hat{x}\right) + \frac{\partial^3 u}{\partial y^2 \partial t}\left(\frac{\hat{y}^2}{2}\hat{t} + \frac{\hat{t}^3}{3!}\right) + \frac{\partial^3}{\partial y^3}\left(\frac{\hat{t}^2}{2!}\hat{y} + \frac{\hat{y}^3}{3!}\right).\end{aligned} \tag{5.3.5}$$

As before, the coefficients in the above expansion collated with the Trefftz basis for $p = 2$ give us the required Trefftz basis functions up to order $p = 3$. We arrange the basis functions as follow:

$$\begin{aligned}\Bigg\{1, \hat{x}, \hat{y}, \hat{t}, \frac{\hat{x}^2}{2} + \frac{\hat{t}^2}{2}, \frac{\hat{y}^2}{2} + \frac{\hat{t}^2}{2}, \hat{x}\hat{y}, \hat{x}\hat{t}, \hat{y}\hat{t}, \hat{x}\hat{y}\hat{t}, \frac{\hat{x}^3}{3!} + \frac{\hat{t}^2}{2}\hat{x}, \frac{\hat{x}^2}{2!}\hat{y} + \frac{\hat{t}^2}{2!}\hat{y}, \\ \frac{\hat{x}^2}{2!}\hat{t} + \frac{\hat{t}^3}{3!}, \frac{\hat{y}^2}{2!}\hat{x} + \frac{\hat{t}^2}{2!}\hat{x}, \frac{\hat{y^2}}{2}\hat{t} + \frac{\hat{t}^3}{3!}, \frac{\hat{t}^2}{2!}\hat{y} + \frac{\hat{y}^3}{3!}\Bigg\},\end{aligned} \tag{5.3.6}$$

and further arrangement gives

$$\begin{aligned}&\left\{1, \hat{x}, \hat{y}, \hat{t}, \hat{x}\hat{y}, \hat{x}\hat{t}, \hat{y}\hat{t}, \hat{x}\hat{y}\hat{t}, \hat{y}^2 + \hat{t}^2 + \hat{x}^2 + \hat{t}^2, \hat{x}^3 + 3\hat{t}^2\hat{x},\right. \\ &\left.\hat{x}^2\hat{y} + \hat{t}^2\hat{y}, 3\hat{x}^2\hat{t} + \hat{t}^3, \hat{y}^2\hat{x} + \hat{t}^2\hat{x}, 3\hat{y^2}\hat{t} + \hat{t}^3, 3\hat{t}^2\hat{y} + \hat{y}^3\right\}.\end{aligned} \tag{5.3.7}$$

In the same way, we generate the basis functions when $N = 4$.

$$\begin{aligned}
&\frac{\partial^4 u}{\partial x^4}\frac{\hat{x}^4}{4!} + \frac{\partial^4 u}{\partial x^3 \partial y}\frac{\hat{x}^3}{3!}\hat{y} + \frac{\partial^4 u}{\partial x^3 \partial t}\frac{\hat{x}^3}{3!}\hat{t} + \frac{\partial^4 u}{\partial x^2 \partial y^2}\frac{\hat{x}^2}{2!}\frac{\hat{y}^2}{2!} + \frac{\partial^4 u}{\partial x^2 \partial t^2}\frac{\hat{x}^2}{2!}\frac{\hat{t}^2}{2!} + \frac{\partial^4 u}{\partial x^2 \partial y \partial t}\frac{\hat{x}^2}{2!}\hat{y}\hat{t} \\
&+ \frac{\partial^4 u}{\partial y^3 \partial x}\frac{\hat{y}^3}{3!}\hat{x} + \frac{\partial^4 u}{\partial x^2 \partial y^4}\frac{\hat{y}^4}{4!} + \frac{\partial^4 u}{\partial y^3 \partial t}\frac{\hat{y}^3}{3!}\hat{t} + \frac{\partial^4 u}{\partial y^2 \partial t^2}\frac{\hat{y}^2}{2!}\frac{\hat{t}^2}{2!} + \frac{\partial^4 u}{\partial y^2 \partial x \partial t}\frac{\hat{y}^2}{2!}\hat{x}\hat{t} \\
&+ \frac{\partial^4 u}{\partial t^4}\frac{\hat{t}^4}{4!} + \frac{\partial^4 u}{\partial t^3 \partial x}\frac{\hat{t}^3}{3!}\hat{x} + \frac{\partial^4 u}{\partial t^3 \partial y}\frac{\hat{t}^3}{3!}\hat{y} + \frac{\partial^4 u}{\partial t^2 \partial x \partial y}\frac{\hat{t}^2}{2!}\hat{x}\hat{y}.
\end{aligned} \tag{5.3.8}$$

Eliminating $\frac{\partial^2 u}{\partial t^2}$ and factorising we have

$$\begin{aligned}
&\frac{\partial^4 u}{\partial x^4}\left(\frac{\hat{x}^2}{4!} + \frac{\hat{x}^2\hat{t}^2}{4} + \frac{\hat{t}^4}{4!}\right) + \frac{\partial^4 u}{\partial y^3 \partial y}\left(\frac{\hat{x}^3\hat{y}}{3!} + \frac{\hat{t}^2\hat{x}\hat{y}}{2!}\right) + \frac{\partial^4 u}{\partial x^3 \partial t}\left(\frac{\hat{x}^3\hat{t}}{3!} + \frac{\hat{t}^3\hat{x}}{3!}\right) \\
&+ \frac{\partial^4 u}{\partial x^2 \partial y^2}\left(\frac{\hat{x}^2\hat{y}^2}{4} + \frac{\hat{y}^2\hat{x}^2}{4} + \frac{\hat{x}^2\hat{t}^2}{4} + \frac{\hat{t}^4}{12}\right) + \frac{\partial^4}{\partial x^2 \partial y \partial t}\left(\frac{\hat{x}^2\hat{y}\hat{t}}{2!} + \frac{\hat{t}^3\hat{y}}{3!}\right) + \frac{\partial^4 u}{\partial y^3 \partial x}\left(\frac{\hat{y}^3\hat{x}}{3!} + \frac{\hat{t}^2\hat{x}\hat{y}}{2!}\right) \\
&+ \frac{\partial^4 u}{\partial y^4}\left(\frac{\hat{y}^4}{4!} + \frac{\hat{t}^4}{4!}\right) + \frac{\partial^4 u}{\partial y^3 \partial t}\left(\frac{\hat{y}^3\hat{t}}{3!} + \frac{\hat{t}^3\hat{y}}{3!}\right) + \frac{\partial^4 u}{\partial y^2 \partial x \partial t}\left(\frac{\hat{y}^2\hat{x}\hat{t}}{2!} + \frac{\hat{t}^3\hat{x}}{3!}\right).
\end{aligned} \tag{5.3.9}$$

Just as above, we finally have the following Trefftz basis functions up to order $p = 4$

$$\begin{aligned}
&\left\{1,\ \hat{x},\ \hat{y},\ \hat{t},\ \hat{x}\hat{y},\ \hat{x}\hat{t},\ \hat{y}\hat{t},\ \hat{x}^2 + \hat{t}^2,\ \hat{y}^2 + \hat{t}^2,\ \ \hat{x}\hat{y}\hat{t},\ \hat{x}^3 + 3\hat{t}^2\hat{x},\right. \\
&\hat{x}^2\hat{y} + \hat{t}^2\hat{y},\ 3\hat{x}^2\hat{t} + \hat{t}^3,\ \hat{y}^2\hat{x} + \hat{t}^2\hat{x}, 3\hat{y}\hat{t} + \hat{t}^3,\ 3\hat{t}^2\hat{y} + \hat{y}^3,\ \hat{x}^4 + 6\hat{x}^2\hat{t}^2 + \hat{t}^4, \\
&\hat{x}^3\hat{y} + 3\hat{t}^2\hat{x}\hat{y},\ \hat{x}^3\hat{t} + \hat{t}^3\hat{x},\ 6\hat{x}^2\hat{y}^2 + 6\hat{y}^2\hat{t}^2 + 6\hat{x}^2\hat{t}^2 + 2\hat{t}^4,\ 3\hat{x}^2\hat{y}\hat{t} + \hat{t}^3\hat{y},\ \hat{y}^3\hat{x} + 3\hat{t}^2\hat{x}\hat{y},\ \hat{y}^4 + \hat{t}^4 + 6\hat{y}^2\hat{t}^2, \\
&\left.\hat{y}^3\hat{t} + \hat{t}^3\hat{y},\ 3\hat{y}^2\hat{x}\hat{t} + \hat{t}^3\hat{x}\right\}.
\end{aligned} \tag{5.3.10}$$

5.3.2 Directional Trefftz space

Alternatively, We can generate local solutions by considering the wave-like functions of the form $(t + \alpha \cdot \mathbf{x})^j$ $j = 0, \ldots, p$, where α is a vector of directions and p is the highest order of polynomials in the space. We have already shown in Chapter 3 that for any order j, there are at most $2j + 1$ linearly independent wave functions; hence the total degrees of freedom per element can be computed by the formula

$$P = \sum_{j=0}^{p}(2j + 1) = p(p + 2) + 1. \tag{5.3.11}$$

If the dominant direction of propagation is not known, then the idea of equi-distributed directions in a unit circle as suggested in the paper [22] can be used to fill up the whole Trefftz space,i.e

$$\alpha_i = \begin{bmatrix} \cos\left(\frac{2\pi(i-1)}{m}\right) \\ \\ \sin\left(\frac{2\pi(i-1)}{m}\right) \end{bmatrix} \; i = 1, \ldots, m, \tag{5.3.12}$$

where $m = 2j + 1$. The maximum number of direction vectors α is connected with the number of linearly independent local solutions $2j + 1$ and this suggests that if the dominant directions of propagation are known locally on each element then the number of degrees of freedom can be reduced. However, if the dominant directions of the wave are known a priori, then only the basis functions in the wave direction are enough to approximate the problem, see Example 5.3.2. Table 5.11 shows the maximum directions with respect to chosen order j of the Trefftz space.

Order	Maximum directions
$j = 1$	3
$j = 2$	5
$j = 3$	7
$j = 4$	9

Table 5.11: *Maximum number of directions with respect to order of local solutions.*

5.3.3 Implementation and algorithm

Before we present the computer implementation of the Trefftz based method, we exploit a vital advantage of the Trefftz basis functions by rewriting the formulation without the space-time volume integral. This is done in order to reduce the cost of computation and also to make the implementation faster.

We analyse the third term of the discrete bilinear form (3.3.7) as follows:

$$\begin{aligned} (a\nabla u, \nabla \dot{v})_{\Omega\times I_n} &= ([a\nabla u\ \dot{v}])_{\Gamma_n\times I_n} - (\nabla\cdot a\nabla u, \dot{v})_{\Omega\times I_n} \\ &= ([a\nabla u], \{\dot{v}\})_{\Gamma_n\times I_n} + (\{a\nabla u\}, [\dot{v}])_{\Gamma_n\times I_n} - (\nabla\cdot a\nabla u, \dot{v})_{\Omega\times I_n}. \end{aligned} \tag{5.3.13}$$

Substituting into (3.3.7) we have the reduced form

$$\begin{aligned}\mathcal{A}_n(u,v) \coloneqq & \left(\dot{u}(t_n^+), \dot{v}(t_n^+)\right)_\Omega + \left(a\nabla u(t_n^+), \nabla v(t_n^+)\right)_\Omega \\ & - \left(\left[a\nabla u\right], \left\{\dot{v}\right\}\right)_{\Gamma_n\times I_n} - \left(\left\{a\nabla u(t_n^+)\right\}, \left[v(t_n^+)\right]\right)_{\Gamma_n} \\ & - \left(\left[u\right], \left\{a\nabla \dot{v}\right\}\right)_{\Gamma_n\times I_n} - \left(\left[u(t_n^+)\right], \left\{a\nabla v(t_n^+)\right\}\right)_{\Gamma_n} \\ & + \left(\sigma_0\left[u\right], \left[\dot{v}\right]\right)_{\Gamma_n\times I_n} + \left(\sigma_0\left[u(t_n^+)\right], \left[v(t_n^+)\right]\right)_{\Gamma_n} \\ & + \left(\sigma_1\left[u\right], \left[v\right]\right)_{\Gamma_n\times I_n} + \left(\sigma_2\left[a\nabla u\right], \left[a\nabla v\right]\right)_{\Gamma_n\times I_n}.\end{aligned} \tag{5.3.14}$$

We emphasise here that discrete bilinear form (5.3.14) can only be used when the basis functions in the implementation are Trefftz basis functions.

Now for the implementation of the scheme, we define the basis function on a physical element to be

$$u_j(\mathbf{x}, t) = f_j\left(\frac{(\mathbf{x}-\mathbf{x}_G)}{\Delta t}, \frac{t-t_n}{\Delta t}\right), \tag{5.3.15}$$

where $\mathbf{x}_G$ is the barycentric coordinate of the physical element and Δt denotes the time step τ_n. We shall assume $\Delta t = h$ in our implementation and the physical coordinate $\mathbf{x}$ is computed via the transformation map $F_k : \hat{K} \longrightarrow K$ explicitly defined as

$$F_k(\xi) = \begin{pmatrix} x_2 - x_1 & x_3 - x_1 \\ y_2 - y_1 & y_3 - y_1 \end{pmatrix}\begin{pmatrix} r \\ s \end{pmatrix} + \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \tag{5.3.16}$$

where the Jacobian of the transformation is readily obtained to be

$$J = \begin{pmatrix} x_2 - x_1 & x_3 - x_1 \\ y_2 - y_1 & y_3 - y_1 \end{pmatrix}. \tag{5.3.17}$$

Now the space integrals in the formulation can be computed using quadrature. For example let us consider the first term in (5.3.14) which resembles the usual mass

matrix integral. We compute the integral as follows:

$$\begin{aligned}\int_K \dot{u}_j(\mathbf{x},t)\dot{u}_i(\mathbf{x},t)dx &= \int_K \frac{1}{(\Delta t)^2}\partial_t f_j\left(\frac{\mathbf{x}-\mathbf{x}_G}{\Delta t},0\right)\partial_t f_i\left(\frac{\mathbf{x}-\mathbf{x}_G}{\Delta t},0\right)dx \\ &= |detJ_k|\int_{\hat{K}} \frac{1}{(\Delta t)^2}\partial_t f_j\left(\frac{\mathbf{F}_k z-\mathbf{x}_G}{\Delta t},0\right)\partial_t f_i\left(\frac{\mathbf{F}_k z-\mathbf{x}_G}{\Delta t},0\right)dz.\end{aligned}$$

For the terms that involve integration over space-time skeleton, we employ tensor-product Gauss quadrature where one of the integrals is a line integral. We present the algorithm for computing the space integrals in our scheme as follows:

Algorithm 1 Algorithm for element stiffness matrix

1: *Let n_t denote the number of elements in a mesh $\mathcal{T}$ with point matrix P and connectivity matrix T.*
2: *Let $n_f = (k+1)^2$ be the number of basis functions per element with k being the highest order specified.*
3: *Initialise the global stiffness matrix A of size $n_f n_t \times n_f n_t$*
4: *Initialise the quadrature points Q_1, Q_2 and weights W_q*
5: **for** $K = 1$ to n_t **do**
6: *Compute the barycentric coordinates $\mathbf{x}_G$ of K*
7: *Compute the global variables $\mathbf{x}$ and the determinant of the Jacobian $detJ$*
8: *Initialise local matrix A_K*
9: *Compute the gradients $\nabla\phi_i$ and the time derivatives $\dot{\phi}_i$ $i = 1,\dots n_f$*
10: *Compute*

$$A_K = \frac{1}{(\Delta t)^2}\times W_q \times detJ\left(\times\begin{bmatrix}\nabla\phi_1\nabla\phi_1 & \dots & \nabla\phi_{n_f}\nabla\phi_1\\ \vdots & \dots & \vdots\\ \nabla\phi_1\nabla\phi_{n_f} & \dots & \nabla\phi_{n_f}\nabla\phi_{n_f}\end{bmatrix}+\begin{bmatrix}\dot{\phi}_1\dot{\phi}_1 & \dots & \dot{\phi}_{n_f}\dot{\phi}_1\\ \vdots & \dots & \vdots\\ \dot{\phi}_1\dot{\phi}_{n_f} & \dots & \dot{\phi}_{n_f}\dot{\phi}_{n_f}\end{bmatrix}\right) \tag{5.3.18}$$

11: *Set the degrees of freedom map: $dofs = (1:n_f) + n_f \times (K-1)$*
12: *Set $A(dofs, dofs) = A(dofs, dofs) + A_K$*
13: **end for**

For the computation and arrangement of the skeleton terms in the scheme, we consider two elements K^+ and K^- sharing an edge e. Locally, each element contains P degrees of freedom. We take advantage of this and we form a $2P\times 1$ vector of basis functions below. Recall also that the space jump is defined to be $[v] = n^+\phi^+ + n^-\phi^-$ where $n^- = -n^+$. Hence we have the jump of the basis functions below

We now define the matrices SE and PE which are $2P\times 2P$ in size to be the local edge stiffness and jump matrices respectively. At the boundary, the jump matrix reduces to $2P\times 1$ and the $\frac{1}{2}$ vanishes since there is no average at the boundary. This makes the size of the matrices PE and SE at the boundary to be $P\times P$. For the

$$\phi = \begin{bmatrix} \phi_1^+ \\ \phi_2^+ \\ \vdots \\ \phi_P^+ \\ \phi_1^- \\ \phi_2^- \\ \vdots \\ \phi_P^- \end{bmatrix}.$$

$$[\phi] = n^+ \begin{bmatrix} \phi_1^+ \\ \phi_2^+ \\ \vdots \\ \phi_P^+ \\ -\phi_1^- \\ -\phi_2^- \\ \vdots \\ -\phi_P^- \end{bmatrix}.$$

terms which correspond to the edge flux matrices, the outer product $[\dot{\phi}]\{\nabla\phi\}$ or $[\phi]\{\nabla\phi\}$ is the integrand, while the outer product $[\phi][\dot{\phi}]$ or $[\phi][\phi]$ is the integrand for the penalty terms. For the actual integration, we employ Gaussian quadrature mapped from $[0,1]$ onto each edge e to approximate both the line integrals and the temporal integral. We present the algorithm as follow:

5.3.4 Numerical simulations and experiments

Example 5.3.1. *We consider the wave problem defined on two dimensional domain* $\Omega = (0,1) \times (0,1)$ *with initial data :*

$$u(\boldsymbol{x}, 0) = \sin(\pi x)\sin(\pi y), \qquad \dot{u}(\boldsymbol{x}, 0) = 0. \tag{5.3.19}$$

The analytical solution is obtained by separation of variables method

$$u(x, y, t) = \cos(\sqrt{(}2)\pi t)\sin(\pi x)\sin(\pi y). \tag{5.3.20}$$

Figure 5.13 below shows the result of the simulation with time $T = 1$, and time step $N = 10$. We assume the same mesh-width $\Delta t = T/N$ for all triangles in the mesh.

Algorithm 2 Algorithm for edge flux and edge stiffness matrices

Let n_t denote the number of elements in a mesh $\mathcal{T}$
with point matrix P and connectivity matrix T.
Let $n_f = (k+1)^2$ be the number of basis functions per element
with k being the highest order specified.
Initialise the global edge flux PF and stiffness matrices SF of size $n_f n_t \times n_f n_t$
Initialise the quadrature points x_q and weights w_q
Loop over elements:
for $K = 1$ to n_t **do**
    Find each neighbour K^- of the current element K^+
    Compute the barycentric coordinates of K^+ and K^-
    Compute the global variables t, x and y on K^- and K^+
    Initialise local matrices PE, SE of sizes $P \times P$
    Compute the values of time derivatives and gradient of the basis functions
    Compute the local matrices PE and SE
    Set degrees of freedom $dofs = [(1:n_f)+n_f\times(K-1) \quad (1:n_f)+n_f\times(K^- -1)]$
    Set $PE = PE(1:P)$ *and* $SE = SE(1:P)$ *at the boundary*
    Set $PF(dofs, dofs) = PF(dofs, dofs) + PE$
    And set $SF(dofs, dofs) = SF(dofs, dofs) + SE$
end for

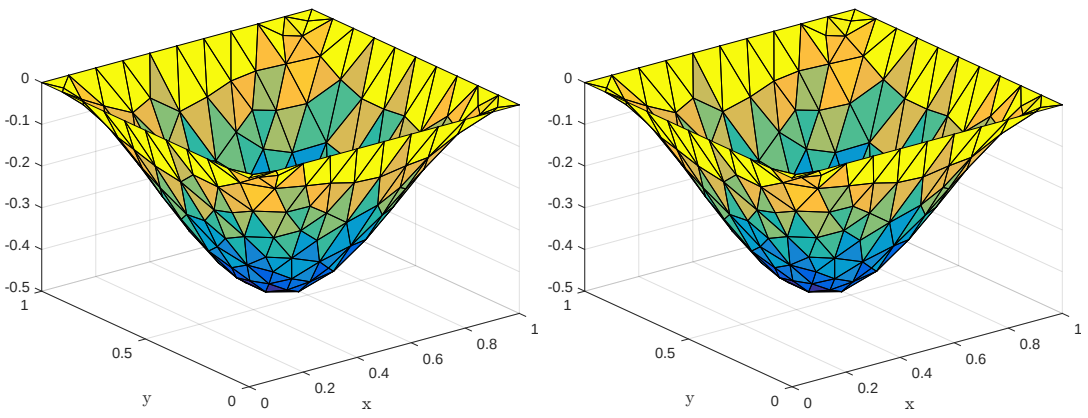


Figure 5.13: *Trefftz space approximation with $p = 2$ (left) and exact solution (right).*

5.3.5 Convergence results

We investigate the convergence of the error in the full DG norm $|||\cdot|||$ as well as in the wave energy norm at the final time-step

$$\text{error} = \left(\frac{1}{2}\|\dot{u}(\cdot,T) - \dot{u}_h(\cdot,T^-)\|_\Omega^2 + \frac{1}{2}\|\nabla u(\cdot,T) - \nabla u_h(\cdot,T^-)\|_\Omega^2\right)^{1/2}. \tag{5.3.21}$$

We discover that higher order convergence is obtainable both with the Trefftz space and the polynomial space. We also discover that with respect to (5.3.21), we do not lose half an order of convergence as when computing the error in the discrete norm $|||\cdot|||$. It is not however surprising as, unlike the DG norm, this error measure does not accumulate the errors over all time-steps. In Figure 5.14 and Tables 5.13 and 5.14, We present the convergence plots, the numerically computed convergence orders and the errors with respect to (5.3.21) .

N	$p=1$	$p=2$	$p=3$	$p=4$
10	0.046	2.039	2.219	3.935
20	0.115	1.990	2.539	3.977
40	0.160	1.979	2.819	3.996

N	$p=1$	$p=2$	$p=3$	$p=4$
10	0.046	1.897	2.133	3.956
20	0.115	1.896	2.248	3.967
40	0.160	1.935	2.612	3.982

Table 5.12: *Numerically obtained convergence orders of the error* (5.3.21) *for Trefftz spaces on the left and for polynomial space on the right.*

N	$p=2$	$p=3$	$p=4$
10	1.49	2.05	3.52
20	1.55	2.24	3.48
40	1.53	2.36	3.50

N	$p=2$	$p=3$	$p=4$
10	1.47	1.98	3.52
20	1.52	2.19	3.50
40	1.52	2.33	3.50

Table 5.13: *Numerically obtained convergence orders in the norm* $|||\cdot|||$ *for Trefftz spaces on the left and for polynomial space on the right.*

N	$p=2$	$p=3$	$p=4$
10	8.68×10^{-2}	6.16×10^{-3}	2.47×10^{-4}
20	2.11×10^{-2}	1.32×10^{-3}	1.61×10^{-5}
40	5.31×10^{-3}	2.27×10^{-4}	1.02×10^{-6}
80	1.35×10^{-3}	3.22×10^{-5}	6.41×10^{-8}

N	$p=2$	$p=3$	$p=4$
10	9.49×10^{-2}	5.92×10^{-3}	2.86×10^{-4}
20	2.55×10^{-2}	1.35×10^{-3}	1.84×10^{-5}
40	6.84×10^{-3}	2.84×10^{-4}	1.17×10^{-6}
80	1.79×10^{-3}	4.65×10^{-5}	7.46×10^{-8}

Table 5.14: *Errors for Trefftz on the left and polynomial spaces on the right computed with* (5.3.21).

For the directional Trefftz case, we consider the wave problem problem defined on $\Omega = (0,1)\times(0,1)$ with initial and boundary data defined below:

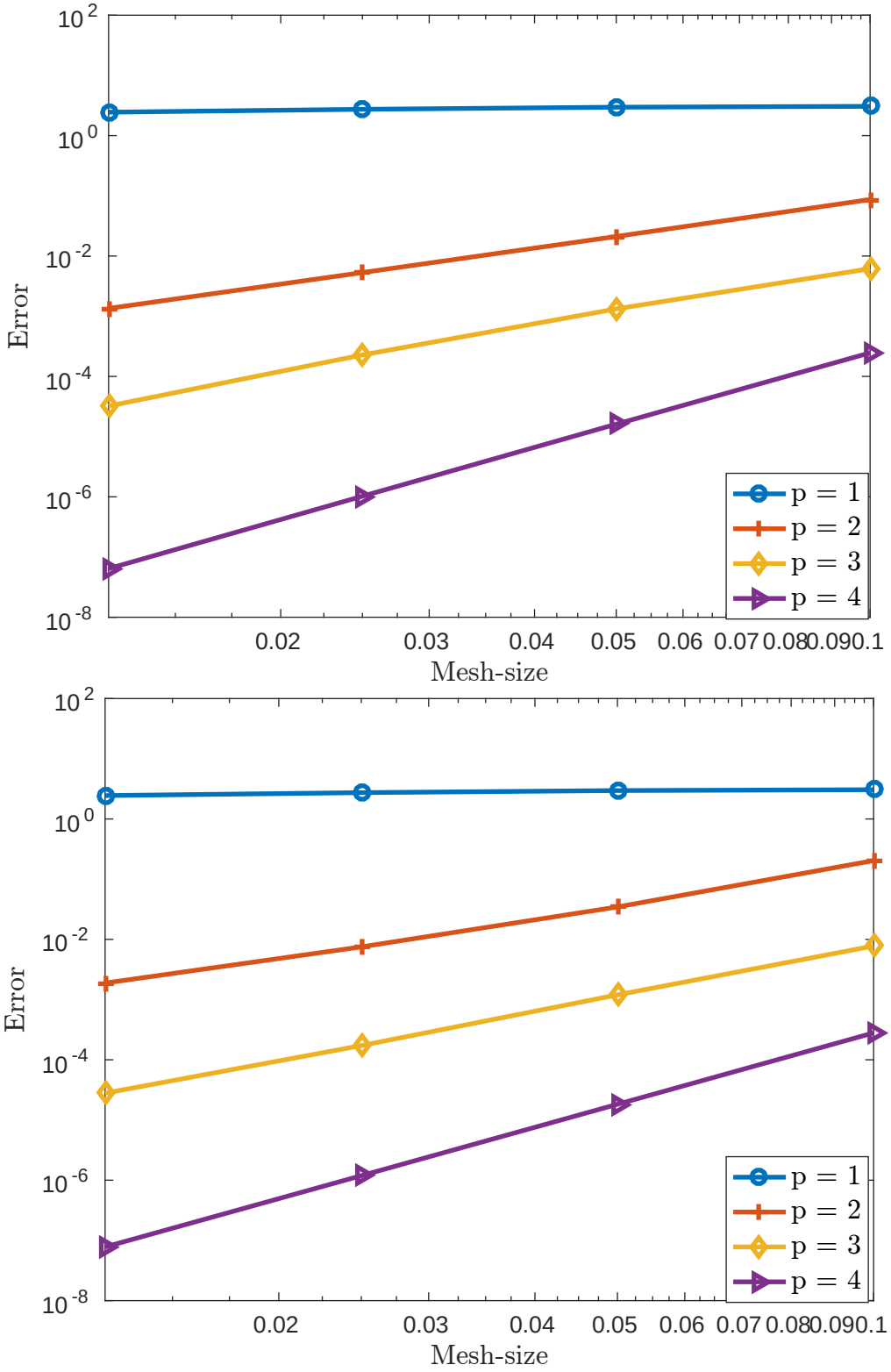


Figure 5.14: *Convergence of the Trefftz (upper) and Polynomial space (lower).*

Example 5.3.2.

$$\begin{aligned} u(\boldsymbol{x},0) &= \sin(2\pi x), \qquad \dot{u}(\boldsymbol{x},0) = 0. \\ u(\boldsymbol{x},0) &= 0 \quad on \quad \Gamma_D, \quad \partial_n u = 0 \quad on \quad \Gamma_N. \end{aligned} \tag{5.3.22}$$

We make use of the Trefftz basis functions in the wave directions to approximate the above problem since the direction of propagation is known a priori. Figure 5.15 shows the result of the simulation using the directional Trefftz space of order $p = 2$, with directions $m = 2$, at final time $T = 1$, and time steps $N = 10$. In this case, a 2−dimensional problem is approximated using 1− dimensional Trefftz basis functions with total degrees of freedom $P = 2p + 1$. If the idea of equi-distributed directions is employed, then the full Trefftz space will be involved in the approximation which will not be different from what we carried out in Example 5.3.1.

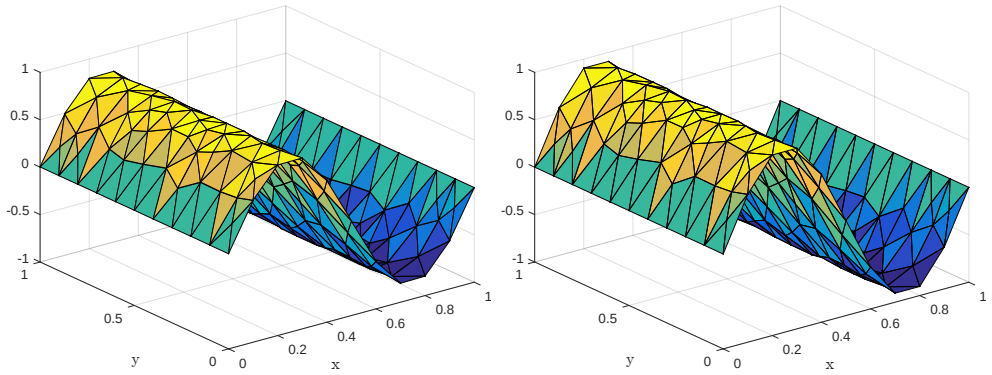


Figure 5.15: *Approximation with directional Trefftz space with mixed boundary conditions.*

5.4 Numerical experiment for the telegraph problem

Recall in Section 5.1 that we represent the undamped wave operator by $\Box = \partial_t^2 - \partial_x^2$ where the diffusion coefficient $a \equiv 1$. We discover that we cannot define $\hat{v} = v\left(\frac{x - x_j}{h}, \frac{t - t_n}{h}\right)$ for the damped wave equation on a reference element $(0, h) \times (0, h)$ since

$$\Box\hat{v} + \alpha\dot{\hat{v}} = \frac{1}{h^2}\Box v + \frac{\alpha}{h}\dot{v} \neq 0. \tag{5.4.1}$$

Therefore we define $\hat{u}(x,t) = u(x - x_j, t - t_n)$ so that each basis now satisfies

$$\Box\hat{v} + \alpha\dot{\hat{v}} = \Box v + \alpha\dot{v} = 0 \tag{5.4.2}$$

on each element.

Example 5.4.1. *We consider the telegraph equation defined on spatial domain* $\Omega = (0,1)$ *and time interval* $[0,T]$ *with the following initial and boundary data*

$$\begin{aligned} u(x,0) = \sin(\pi x),\ \ \dot{u}(x,0) = 0, \\ u = 0 \quad \text{on } \partial\Omega. \end{aligned} \tag{5.4.3}$$

The analytical solution is given by

$$u(x,t) = e^{(-\alpha t/2)}\sin(\pi x)\left[\cos\sqrt{\pi^2 - \frac{\alpha^2}{4}}t + \frac{\alpha}{2\sqrt{\pi^2 - \alpha^2/4}}\sin\sqrt{\pi^2 - \frac{\alpha^2}{4}}t\right]. \tag{5.4.4}$$

The approximation with Trefftz space of order $p = 2$ is presented graphically in 5.16.

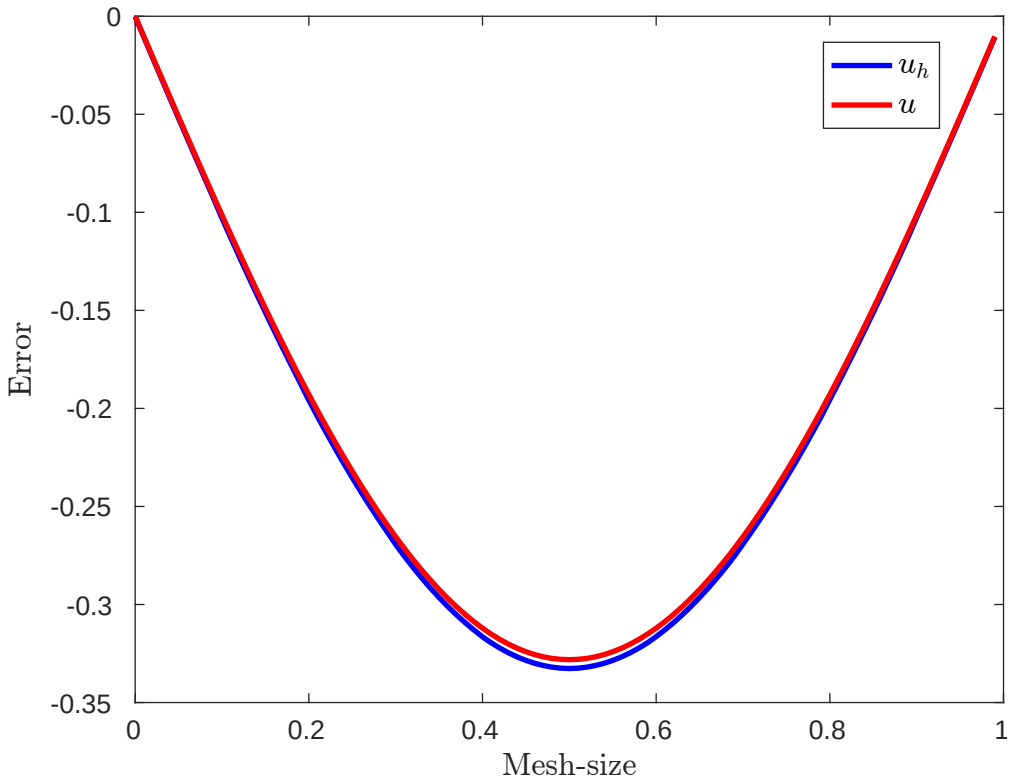


Figure 5.16: *Solution at final time with* $p = 2$, $\alpha = 2$, $K = 10$, *and* $T = 1$.

K	$p=1$	$p=2$	$p=3$	$p=4$
10	0.25	1.87	3.18	4.01
20	0.42	1.93	3.12	4.01
40	0.59	1.96	3.06	4.00

K	$p=1$	$p=2$	$p=3$	$p=4$
10	0.25	2.48	3.06	4.17
20	0.42	2.30	3.27	4.05
40	0.59	2.13	3.21	4.02

Table 5.15: *Numerically obtained convergence order in the wave energy norm (see* (5.4.5)*) for the Trefftz spaces (left) and polynomial spaces (right).*

K	$p=2$	$p=3$	$p=4$
10	1.59	2.48	3.64
20	1.53	2.49	3.58
40	1.52	2.50	3.54

K	$p=2$	$p=3$	$p=4$
10	1.49	2.47	3.54
20	1.50	2.49	3.51
40	1.50	2.50	3.51

Table 5.16: *Numerically obtained convergence order in the DG energy norm* $|||\cdot|||$ *(see* (4.4.14)*) for the Trefftz spaces (left) and polynomial spaces (right).*

5.4.1 Convergence results

We investigate the convergence of the error in the full DG norm $|||\cdot|||$ (see, (4.4.14)) as well as in the following energy norm

$$E(u(t)) = \sqrt{\frac{1}{2}\|\dot{u}(x,T)\|_{\Omega}^2 + \frac{1}{2}\|\nabla u(x,T)\|_{\Omega}^2}. \tag{5.4.5}$$

We discover numerically that the convergence orders are optimal with respect to the above energy norm, see Figure 5.17 and Table 5.15 but we lose approximately half an order in the rate of convergence for the case of the full DG norm (see Figure 5.19) and Table 5.16. We also discover that equivalent orders of convergence are obtainable with Trefftz spaces when compared with polynomial spaces of total degrees which has more degrees of freedom, see 5.18. Finally, we present the table of errors in Tables 5.18 and 5.19 for both Trefftz space and polynomial space for the purpose of comparison.

K	$p=1$	$p=2$	$p=3$	$p=4$
10	30	50	70	90
20	60	100	140	180
40	120	200	280	360
80	240	400	560	720
160	480	800	1120	1440

K	$p=1$	$p=2$	$p=3$	$p=4$
10	30	60	100	150
20	60	120	200	300
40	120	240	400	600
80	240	480	800	1200
160	480	960	1600	2400

Table 5.17: *Number of degrees of freedom for the Trefftz spaces (left) and polynomial spaces(right) used in the experiments.*

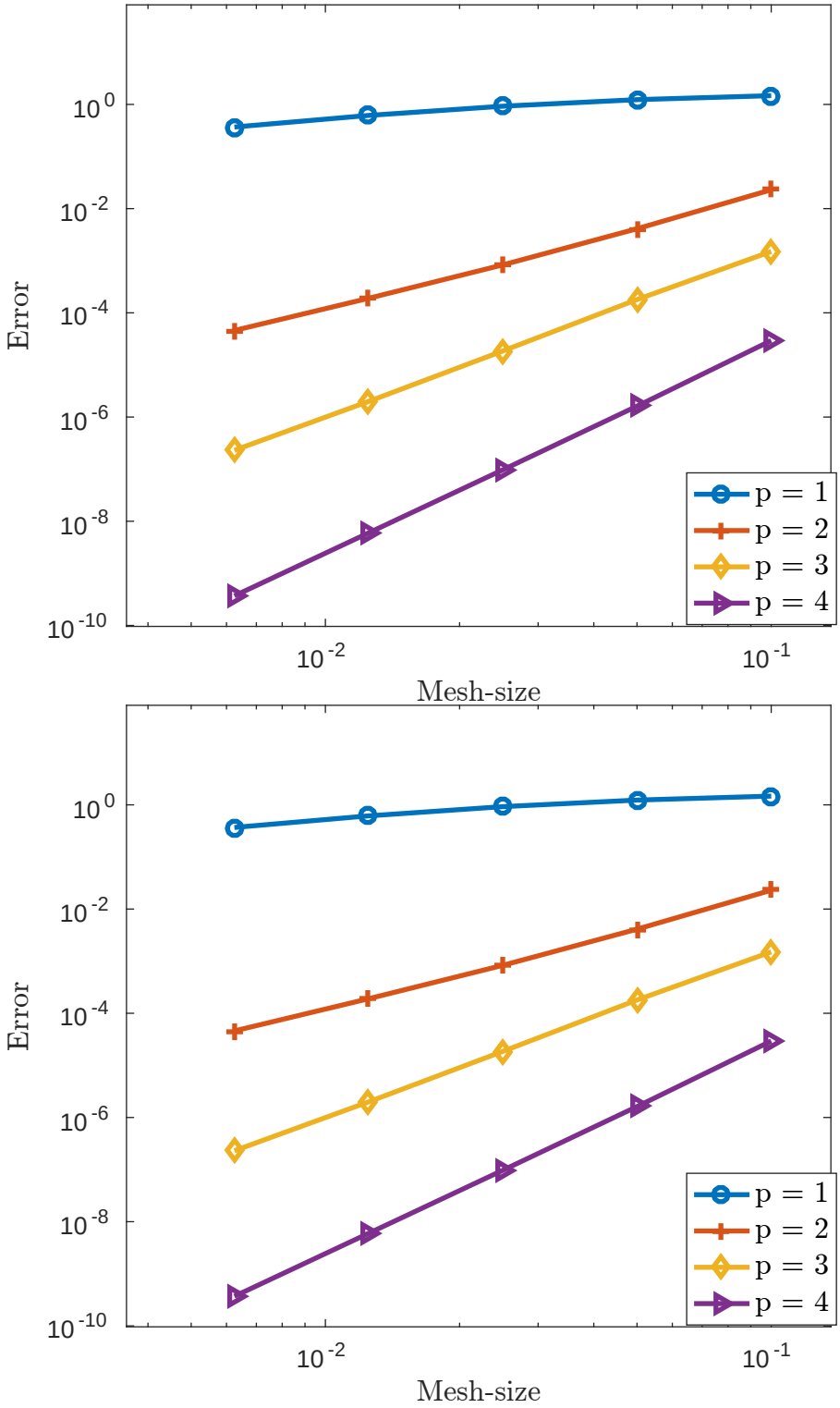


Figure 5.17: *Plot of convergence of the errors in the energy norm (see,* (5.4.5)*) for Trefftz space (upper) and polynomial space (lower).*

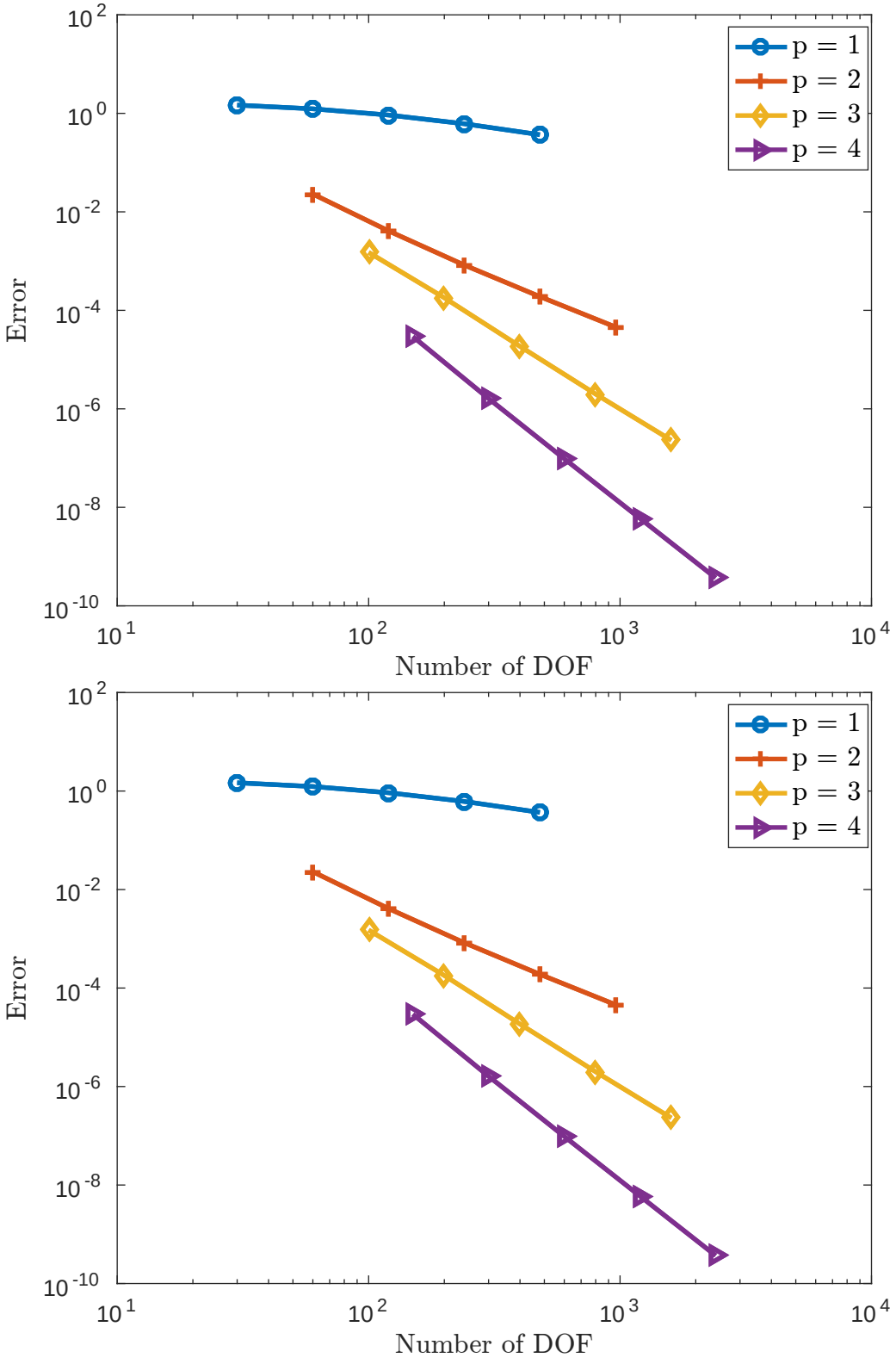


Figure 5.18: *Plot of convergence of the error in the energy norm (see,(5.4.5)) against the degrees of freedom for the Trefftz spaces (upper) and polynomial spaces (lower).*

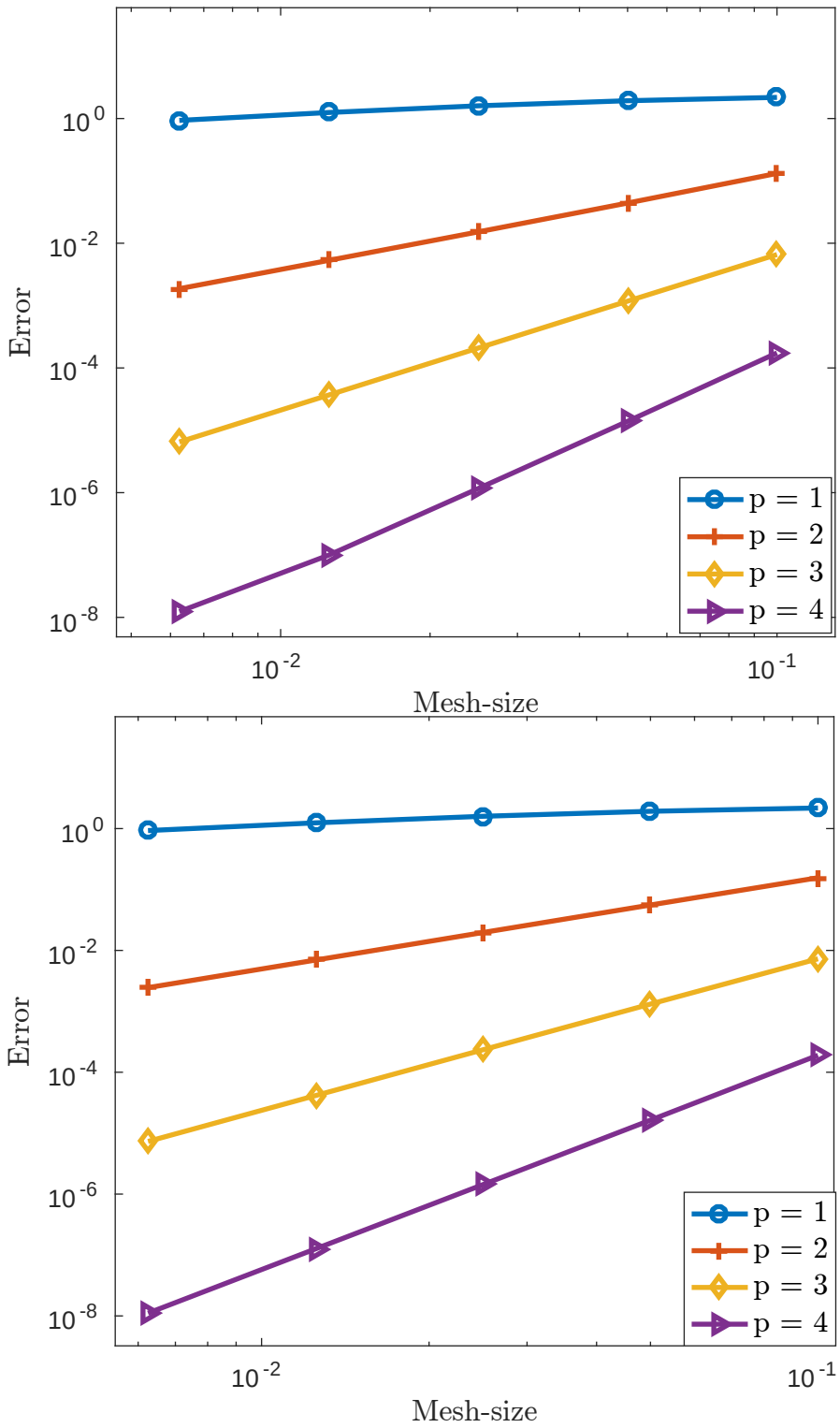


Figure 5.19: *Plot of convergence of the errors in the DG energy norm* $|||\cdot|||$ *(see,* (4.4.14)*) for Trefftz spaces (upper) and polynomial spaces (lower).*

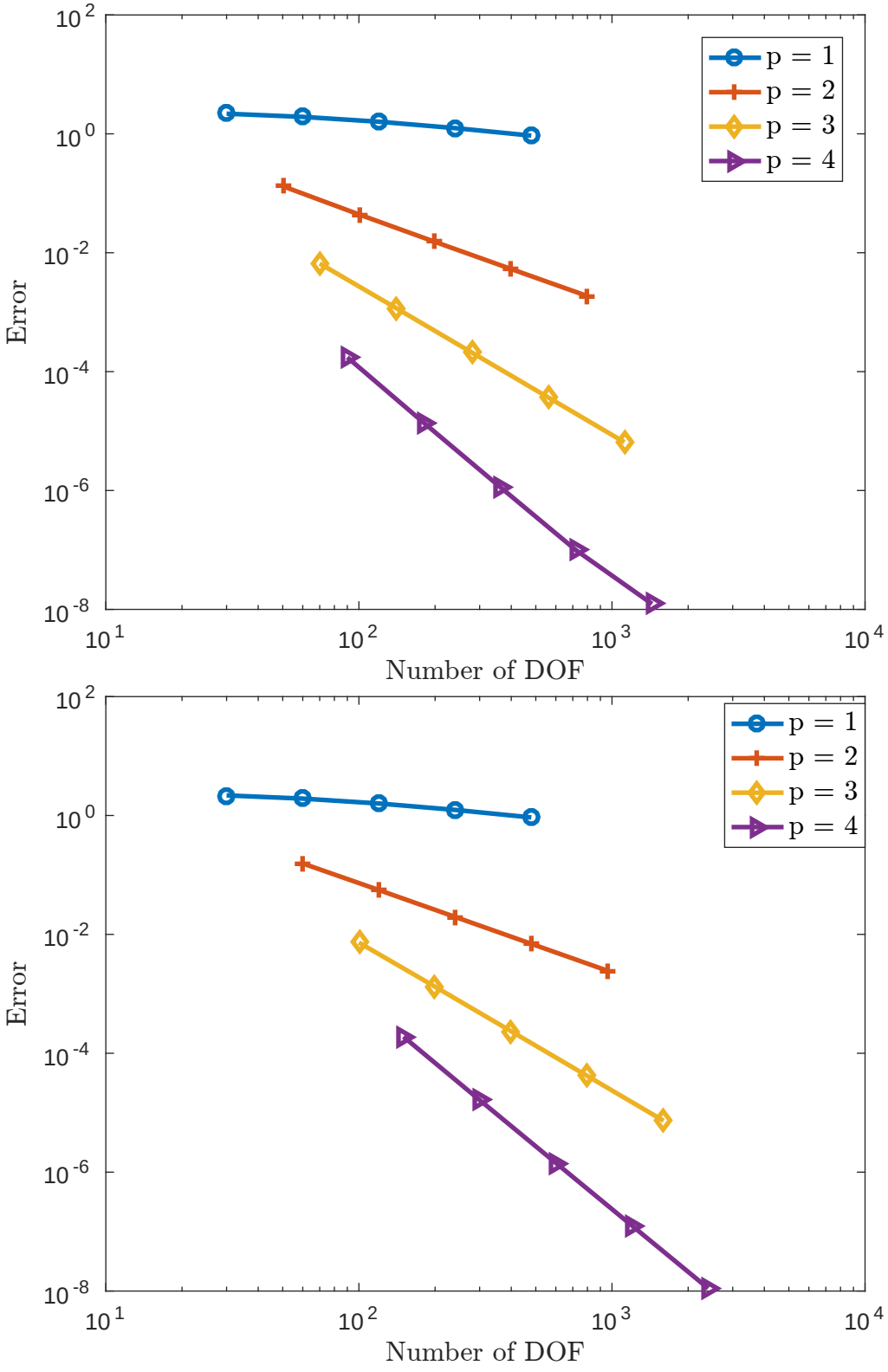


Figure 5.20: *Plot of convergence of the error in the DG energy norm (see,(4.4.14)) against the degrees of freedom for the Trefftz space (upper) and polynomial space(lower).*

N	$p=2$	$p=3$	$p=4$
10	1.47×10^{-2}	6.52×10^{-4}	2.56×10^{-5}
20	4.03×10^{-3}	7.22×10^{-5}	1.59×10^{-6}
40	1.05×10^{-3}	8.30×10^{-6}	9.88×10^{-8}
80	2.71×10^{-4}	9.95×10^{-7}	6.15×10^{-9}

N	$p=2$	$p=3$	$p=4$
10	2.29×10^{-2}	1.50×10^{-3}	2.93×10^{-5}
20	4.10×10^{-3}	1.79×10^{-4}	1.61×10^{-6}
40	8.32×10^{-4}	1.85×10^{-5}	9.70×10^{-8}
80	1.90×10^{-4}	1.99×10^{-6}	5.98×10^{-9}

Table 5.18: *Errors for Trefftz on the left and polynomial spaces on the right computed with* (5.4.5).

N	$p=2$	$p=3$	$p=4$
10	1.31×10^{-1}	6.58×10^{-3}	1.74×10^{-4}
20	4.41×10^{-2}	1.17×10^{-3}	1.39×10^{-5}
40	1.52×10^{-2}	2.08×10^{-4}	1.16×10^{-6}
80	5.31×10^{-3}	3.68×10^{-5}	9.95×10^{-8}

N	$p=2$	$p=3$	$p=4$
10	1.55×10^{-1}	7.28×10^{-3}	1.89×10^{-4}
20	5.56×10^{-2}	1.32×10^{-3}	1.63×10^{-5}
40	1.96×10^{-2}	2.34×10^{-4}	1.43×10^{-6}
80	6.94×10^{-3}	4.13×10^{-5}	1.25×10^{-7}

Table 5.19: *Errors for Trefftz on the left and polynomial spaces on the right computed with* $|||\cdot|||$.

Chapter 6

Conclusion and Future Work

In this thesis we have developed, analysed and implemented a special space-time method for the second order wave equation without splitting the system into systems of first order. The formulation of the method follows the interior penalty discontinuous Galerkin approach together with the classical work of Hulbert and Hughes [73, 72] and the core of the new method is the special Trefftz spaces introduced in the method. The new DG method falls into the framework of space-time DG methods for which a priori analysis such as consistency, stability as well as energy dissipation can be proven without specifying the approximation space in detail. The method is constructed to accommodate any polynomial basis functions provided that they have good approximating properties. However, the use of Trefftz space fulfils our primary goal of approximating wave problems efficiently with a reduced number of degrees of freedom per element, even at high frequencies.

For the undamped wave equation, we construct the Trefftz space from the space of polynomials. We prove the existence of solutions as well as the best approximation property in the Trefftz space. Rates of convergence in the full DG norm are proven in any dimension and numerically verified in spatial dimensions $d = 1$ and $d = 2$.

In Chapter 4, we extend the new time-space DG technique to approximating the telegraph or damped wave equation. The construction of the Trefftz space with good approximating properties for this problem is not trivial. After many trials, we use non-polynomial analytical solution of the problem with polynomial initial data to construct the Trefftz space.

In Chapter 5, we present numerical experiments that highlight the effectiveness of the Trefftz spaces compared to polynomial spaces with more degrees of freedom. We discover that equivalent optimal rate of convergence in the wave energy norm is achievable with the Trefftz and polynomial spaces. Apart from reduction in number of degrees of freedom per element, the Trefftz space also offers a considerable savings advantage over the polynomial space (especially in higher dimension) as evaluation of space-time volume integrals can be avoided in its implementation. With the Trefftz space, analytical features of the problem can be embedded in the approximation space. This advantage is exploited under directional implementation as well as in the construction of Trefftz space for the telegraph problem.

Although the Trefftz space-time DG method studied in this thesis is implicit, further work can be done to make the scheme locally explicit which can allow a reasonable comparison with existing explicit methods. In comparison with higher-order spectral methods, introduction of higher order approximations (as well as implementations) is easier and straightforward in the context of the Trefftz space-time DG method.

The space-time method developed in this thesis looks very promising and has opened doors for more research works. We propose to embark on the following work in future:

(i) Development of non-dissipative or conservative time-space method: We have an idea that if the algebraic upwind identity introduced in the formulation is adjusted, this could lead to a conservative scheme.

(ii) Space-time a posteriori error estimation and adaptivity for greater efficiency.

(iii) Directional adaptivity: It was convenient to introduce direction of propagation with the directional plane wave basis functions by considering equi-distributed directions in plane. The next step is to allow the method to detect dominant directions automatically. We believe this can be done by considering local directions of propagation on each space-time element.

(iv) Higher dimensional implementation of the method for the damped wave prob-

lem.

(v) Implementation of the method on unstructured meshes and converting the method to a semi-explicit form using the ideas in [123, 96].

Bibliography

[1] R. Alford, K. Kelly, and D. M. Boore. Accuracy of finite-difference modeling of the acoustic wave equation. *Geophysics*, 39(6):834–842, 1974.

[2] D. N. Arnold. An interior penalty finite element method with discontinuous elements. *SIAM J. Numer. Anal.*, 19(4):742–760, 1982.

[3] D. N. Arnold, F. Brezzi, B. Cockburn, and D. Marini. Discontinuous Galerkin methods for elliptic problems. In *Discontinuous Galerkin Methods (Newport, RI, 1999)*, volume 11 of *Lect. Notes Comput. Sci. Eng.*, pages 89–101. Springer, Berlin, 2000.

[4] D. N. Arnold, F. Brezzi, B. Cockburn, and L. D. Marini. Unified analysis of discontinuous Galerkin methods for elliptic problems. *SIAM J. Numer. Anal.*, 39(5):1749–1779, 2001/02.

[5] H. Atkins and C.-W. Shu. Quadrature-free implementation of the discontinuous Galerkin method for hyperbolic equations. *NASA Langley Technical Report Server*, 1996.

[6] K. Atkinson and W. Han. *Theoretical Numerical Analysis*. Texts in Applied Mathematics, Volume 39. Springer, Dordrecht, third edition, 2009.

[7] I. Babuška. The finite element method with penalty. *Math. Comp.*, 27:221–228, 1973.

[8] I. Babuška, C. E. Baumann, and J. T. Oden. A discontinuous *hp* finite element method for diffusion problems: 1-D analysis. *Comput. Math. Appl.*, 37(9):103–122, 1999.

[9] I. Babuška and M. Suri. The *h-p* version of the finite element method with quasi-uniform meshes. *RAIRO Modél. Math. Anal. Numér.*, 21(2):199–238, 1987.

[10] I. Babuška and M. Suri. The optimal convergence rate of the *p*-version of the finite element method. *SIAM J. Numer. Anal.*, 24(4):750–776, 1987.

[11] G. A. Baker. Finite element methods for elliptic equations using nonconforming elements. *Math. Comp.*, 31(137):45–59, 1977.

[12] L. Banjai, E. Georgoulis, and L. Oluwaseun. A Trefftz polynomial space-time discontinuous Galerkin method for the second order wave equation. *SIAM J. Numer. Anal.*, 55(1):63–86, 2017.

[13] F. Bassi and S. Rebay. A high-order accurate discontinuous finite element method for the numerical solution of the compressible Navier-Stokes equations. *J. Comput. Phys.*, 131(2):267–279, 1997.

[14] T. Betcke and J. Philips. Adaptive plane wave discontinuous Galerkin methods for Helmholtz problems. In *Proceedings of the 10th International Conference on the Mathematical and Numerical Aspects of Waves, Vancouver,Canada*, pages 261–264, 2011.

[15] T. Betcke and J. Phillips. Approximation by dominant wave directions in plane wave methods. Preprint submitted to Journals of Sound and Vibration. PDF available at http://discovery.ucl.ac.uk/1342769/, 2012.

[16] J. Billingham and A. C. King. *Wave motion*. Cambridge Texts in Applied Mathematics. Cambridge University Press, Cambridge, 2000.

[17] L. Bociu. Local and global wellposedness of weak solutions for the wave equation with nonlinear boundary and interior sources of supercritical exponents and damping. *Nonlinear Anal.*, 71(12):e560–e575, 2009.

[18] L. Bociu and I. Lasiecka. Uniqueness of weak solutions for the semilinear wave equations with supercritical boundary/interior sources and damping. *Discrete Contin. Dyn. Syst.*, 22(4):835–860, 2008.

[19] L. Bociu, P. Radu, and D. Toundykov. Regular solutions of wave equations with super-critical sources and exponential-to-logarithmic damping. *Evol. Equ. Control Theory*, 2(2):255–279, 2013.

[20] S. C. Brenner and L. R. Scott. *The Mathematical Theory of Finite Element Methods*. Texts in Applied Mathematics, Volume 15. Springer, New York, third edition, 2008.

[21] C. Canuto and A. Quarteroni. Approximation results for orthogonal polynomials in Sobolev spaces. *Math. Comp.*, 38(157):67–86, 1982.

[22] O. Cessenat and B. Després. Application of an ultra weak variational formulation of elliptic PDEs to the two-dimensional Helmholtz problem. *SIAM J. Numer. Anal.*, 35(1):255–299, 1998.

[23] J. Chan, R. J. Hewett, and T. Warburton. Weight-adjusted discontinuous Galerkin methods: Wave propagation in heterogeneous media. *arXiv preprint arXiv:1608.01944*, 2016.

[24] G. Chavent and B. Cockburn. The local projection P^0P^1-discontinuous-Galerkin finite element method for scalar conservation laws. *RAIRO Modél. Math. Anal. Numér.*, 23(4):565–592, 1989.

[25] E. T. Chung and B. Engquist. Optimal discontinuous Galerkin methods for wave propagation. *SIAM J. Numer. Anal.*, 44(5):2131–2158, 2006.

[26] E. T. Chung and B. Engquist. Optimal discontinuous Galerkin methods for the acoustic wave equation in higher dimensions. *SIAM J. Numer. Anal.*, 47(5):3820–3848, 2009.

[27] P. G. Ciarlet. *The Finite Element Method for Elliptic Problems.* Classics in Applied Mathematics, Volume 40. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2002.

[28] B. Cockburn, S. Hou, and C.-W. Shu. The Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws. IV. The multidimensional case. *Math. Comp.*, 54(190):545–581, 1990.

[29] B. Cockburn, G. Kanschat, and D. Schotzau. A locally conservative LDG method for the incompressible Navier-Stokes equations. *Math. Comp.*, 74(251):1067–1095 (electronic), 2005.

[30] B. Cockburn, G. E. Karniadakis, and C.-W. Shu. The development of discontinuous Galerkin methods. In *Discontinuous Galerkin methods (Newport, RI, 1999)*, volume 11 of *Lect. Notes Comput. Sci. Eng.*, pages 3–50. Springer, Berlin, 2000.

[31] B. Cockburn, F. Li, and C.-W. Shu. Locally divergence-free discontinuous Galerkin methods for the Maxwell equations. *J. Comput. Phys.*, 194(2):588–610, 2004.

[32] B. Cockburn, S. Y. Lin, and C.-W. Shu. TVB Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws. III. One-dimensional systems. *J. Comput. Phys.*, 84(1):90–113, 1989.

[33] B. Cockburn and C.-W. Shu. TVB Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws. II. General framework. *Math. Comp.*, 52(186):411–435, 1989.

[34] B. Cockburn and C.-W. Shu. The Runge-Kutta discontinuous Galerkin method for conservation laws. V. Multidimensional systems. *J. Comput. Phys.*, 141(2):199–224, 1998.

[35] B. Cockburn and C.-W. Shu. Runge-Kutta discontinuous Galerkin methods for convection-dominated problems. *J. Sci. Comput.*, 16(3):173–261, 2001.

[36] G. Cohen, X. Ferrieres, and S. Pernet. A spatial high-order hexahedral discontinuous Galerkin method to solve Maxwell's equations in time domain. *J. Comput. Phys.*, 217(2):340–363, 2006.

[37] G. C. Cohen. *Higher-order Numerical Methods for Transient Wave Equations.* Springer-Verlag, Berlin, 2002.

[38] F. Collino, T. Fouquet, and P. Joly. A conservative space-time mesh refinement method for the 1-D wave equation. I. Construction. *Numer. Math.*, 95(2):197–221, 2003.

[39] F. Collino, T. Fouquet, and P. Joly. A conservative space-time mesh refinement method for the 1-D wave equation. II. Analysis. *Numer. Math.*, 95(2):223–251, 2003.

[40] F. Collino, T. Fouquet, and P. Joly. Conservative space-time mesh refinement methods for the FDTD solution of Maxwell's equations. *J. Comput. Phys.*, 211(1):9–35, 2006.

[41] F. Costanzo and H. Huang. Proof of unconditional stability for a single-field discontinuous Galerkin finite element formulation for linear elasto-dynamics. *Comput. Methods Appl. Mech. Engrg.*, 194(18-20):2059–2076, 2005.

[42] C. A. Coulson. *Waves. A Mathematical Account of the Common Types of Wave Motion.* Oliver and Boyd, Edinburgh; Interscience Publishers, Inc., New York, 1941.

[43] R. Courant, K. Friedrichs, and H. Lewy. Über die partiellen Differenzengleichungen der mathematischen Physik. *Math. Ann.*, 100(1):32–74, 1928.

[44] R. Dautray and J.-L. Lions. *Mathematical Analysis and Numerical Methods for Science and Technology. Vol. 6.* Springer-Verlag, Berlin, 1993. Evolution problems. II, With the collaboration of Claude Bardos, Michel Cessenat, Alain Kavenoky, Patrick Lascaux, Bertrand Mercier, Olivier Pironneau, Bruno Scheurer and Rémi Sentis, Translated from the French by Alan Craig.

[45] C. Dawson, S. Sun, and M. F. Wheeler. Compatible algorithms for coupled flow and transport. *Comput. Methods Appl. Mech. Engrg.*, 193(23-26):2565–2580, 2004.

[46] L. M. Delves and C. A. Hall. An implicit matching principle for global element calculations. *J. Inst. Math. Appl.*, 23(2):223–234, 1979.

[47] D. A. Di Pietro and A. Ern. *Mathematical Aspects of Discontinuous Galerkin Methods*. Mathématiques & Applications (Berlin) [Mathematics & Applications],Vol 69. Springer, Heidelberg, 2012.

[48] J. Diaz and M. J. Grote. Energy conserving explicit local time stepping for second-order wave equations. *SIAM J. Sci. Comput.*, 31(3):1985–2014, 2009.

[49] M. Dumbser, M. Käser, and E. F. Toro. An arbitrary high-order discontinuous Galerkin method for elastic waves on unstructured meshes-v. local time stepping and p-adaptivity. *Geophysical Journal International*, 171(2):695–717, 2007.

[50] H. Egger, F. Kretzschmar, S. M. Schnepp, and T. Weiland. A space-time discontinuous Galerkin Trefftz method for time dependent Maxwell's equations. *SIAM J. Sci. Comput.*, 37(5):B689–B711, 2015.

[51] A. Ern and J.-L. Guermond. *Theory and Practice of Finite Elements*. Applied Mathematical Sciences, Vol 159. Springer-Verlag, New York, 2004.

[52] L. C. Evans. *Partial Differential Equations*. Graduate Studies in Mathematics, Vol. 19. American Mathematical Society, Providence, RI, second edition, 2010.

[53] F. A. Fernandea and L. Kulas. A simple finite difference approach using unstructured meshes from FEM mesh generators. *IEEE*, 2:585–588, 2004.

[54] G. Gabard. Discontinuous Galerkin methods with plane waves for the displacement-based acoustic equation. *Internat. J. Numer. Methods Engrg.*, 66(3):549–569, 2006.

[55] E. H. Georgoulis and P. Houston. Discontinuous Galerkin methods for the biharmonic problem. *IMA J. Numer. Anal.*, 29(3):573–594, 2009.

[56] F. X. Giraldo, J. S. Hesthaven, and T. Warburton. Nodal high-order discontinuous Galerkin methods for the spherical shallow water equations. *J. Comput. Phys.*, 181(2):499–525, 2002.

[57] C. J. Gittelson, R. Hiptmair, and I. Perugia. Plane wave discontinuous Galerkin methods: analysis of the h-version. *M2AN Math. Model. Numer. Anal.*, 43(2):297–331, 2009.

[58] I. S. Gradshteyn and I. M. Ryzhik. *Table of Integrals, Series, and Products*. Academic Press Inc., San Diego, CA, 2000.

[59] L. Grasedyck, D. Kressner, and C. Tobler. A literature survey of low-rank tensor approximation techniques. *GAMM-Mitt.*, 36(1):53–78, 2013.

[60] M. J. Grote and T. Mitkova. High-order explicit local time-stepping methods for damped wave equations. *J. Comput. Appl. Math.*, 239:270–289, 2013.

[61] M. J. Grote, A. Schneebeli, and D. Schötzau. Discontinuous Galerkin finite element method for the wave equation. *SIAM J. Numer. Anal.*, 44(6):2408–2431, 2006.

[62] M. J. Grote, A. Schneebeli, and D. Schötzau. Interior penalty discontinuous Galerkin method for Maxwell's equations: optimal L^2-norm error estimates. *IMA J. Numer. Anal.*, 28(3):440–468, 2008.

[63] R. B. Guenther and J. W. Lee. *Partial Differential Equations of Mathematical Physics and Integral Equations*. Dover Publications, Inc., Mineola, NY, 1996. Corrected reprint of the 1988 original.

[64] R. B. Guenther and J. W. Lee. *Partial Differential Equations of Mathematical Physics and Integral Equations*. Dover Publications, Inc., Mineola, NY, 1996. Corrected reprint of the 1988 original.

[65] W. Hackbusch. *Elliptic Differential Equations*. Springer Series in Computational Mathematics, Vol. 18. Springer-Verlag, Berlin, english edition, 2010. Theory and numerical treatment, Translated from the 1986 corrected German edition by Regine Fadiman and Patrick D. F. Ion.

[66] I. Harari. A survey of finite element methods for time-harmonic acoustics. *Comput. Methods Appl. Mech. Engrg.*, 195(13-16):1594–1607, 2006.

[67] A. Haraux. A new characterization of weak solutions to the damped wave equations. *Funkcial. Ekvac.*, 31(3):471–482, 1988.

[68] R. Hartmann and P. Houston. Symmetric interior penalty DG methods for the compressible Navier-Stokes equations. I. Method formulation. *Int. J. Numer. Anal. Model.*, 3(1):1–20, 2006.

[69] J. S. Hesthaven and T. Warburton. *Nodal Discontinuous Galerkin Methods*. Texts in Applied Mathematics, Vol. 54. Springer, New York, 2008. Algorithms, analysis, and applications.

[70] R. Hiptmair, A. Moiola, and I. Perugia. Trefftz discontinuous Galerkin methods for acoustic scattering on locally refined meshes. *Appl. Numer. Math.*, 79:79–91, 2014.

[71] P. Houston, C. Schwab, and E. Süli. Discontinuous *hp*-finite element methods for advection-diffusion-reaction problems. *SIAM J. Numer. Anal.*, (6):2133–2163, 2002.

[72] T. J. R. Hughes and G. M. Hulbert. Space-time finite element methods for elastodynamics: formulations and error estimates. *Comput. Methods Appl. Mech. Engrg.*, 66(3):339–363, 1988.

[73] G. M. Hulbert and T. J. R. Hughes. Space-time finite element methods for second-order hyperbolic equations. *Comput. Methods Appl. Mech. Engrg.*, 84(3):327–348, 1990.

[74] F. Ihlenburg. *Finite Element Analysis of Acoustic Scattering*. Applied Mathematical Sciences, Vol 132. Springer-Verlag, New York, 1998.

[75] R. Ikehata. Stable and unstable sets for evolution equations of parabolic and hyperbolic type. *Hiroshima Math. J.*, 26(3):475–491.

[76] D.-C. Ionescu and H. Igel. Transparent boundary conditions for wave propagation on unbounded domains. In *Computational science—ICCS 2003. Part III*, volume 2659 of *Lecture Notes in Comput. Sci.*, pages 807–816. Springer, Berlin, 2003.

[77] E. W. Jenkins, B. Rivière, and M. F. Wheeler. A priori error estimates for mixed finite element approximations of the acoustic wave equation. *SIAM J. Numer. Anal.*, (5):1698–1715 (electronic), 2002.

[78] C. Johnson. Discontinuous Galerkin finite element methods for second order hyperbolic problems. *Comput. Methods Appl. Mech. Engrg.*, 107(1-2):117–129, 1993.

[79] C. Johnson. *Numerical Solution of Partial Differential Equations by the Finite Element Method*. Dover Publications, Inc., Mineola, NY, 2009. Reprint of the 1987 edition.

[80] C. Johnson, U. Nävert, and J. Pitkäranta. Finite element methods for linear hyperbolic problems. *Comput. Methods Appl. Mech. Engrg.*, 45(1-3):285–312, 1984.

[81] D. D. Joseph and L. Preziosi. Heat waves. *Rev. Modern Phys.*, 61(1):41–73, 1989.

[82] F. Kretzschmar, A. Moiola, I. Perugia, and S. M. Schnepp. A priori error analysis of space–time Trefftz discontinuous Galerkin methods for wave problems. page drv064, 2015.

[83] F. Kretzschmar, A. Moiola, I. Perugia, and S. M. Schnepp. A priori error analysis of space-time Trefftz discontinuous Galerkin methods for wave problems, 2015.

[84] F. Kretzschmar, A. Moiola, I. Perugia, and S. M. Schnepp. *A priori* error analysis of space-time Trefftz discontinuous Galerkin methods for wave problems. *IMA J. Numer. Anal.*, (4):1599–1635, 2016.

[85] M. G. Larson and F. Bengzon. *The Finite Element Method: Theory, Implementation, and Applications*. Texts in Computational Science and Engineering, Vol. 10. Springer, Heidelberg, 2013.

[86] S. Larsson and V. Thomée. *Partial Differential Equations with Numerical Methods*, volume 45 of *Texts in Applied Mathematics*. Springer-Verlag, Berlin, 2003.

[87] P. Lasaint and P.-A. Raviart. On a finite element method for solving the neutron transport equation. In *Mathematical aspects of finite elements in partial differential equations (Proc. Sympos., Math. Res. Center, Univ. Wisconsin, Madison, Wis., 1974)*, pages 89–123. Publication No. 33. Math. Res. Center, Univ. of Wisconsin-Madison, Academic Press, New York, 1974.

[88] J.-L. Lions and E. Magenes. *Non-homogeneous Boundary Value Problems and Applications. Vol. I.* Springer-Verlag, New York-Heidelberg, 1972. Translated from the French by P. Kenneth, Die Grundlehren der mathematischen Wissenschaften, Band 181.

[89] V. Lisitsa, G. Reshetova, and V. Tcheverda. Local time-space mesh refinement for finite difference simulation of waves. In *Numerical Mathematics and Advanced Applications 2009*, pages 609–616. Springer, 2010.

[90] V. Lisitsa, G. Reshetova, and V. Tcheverda. Finite-difference algorithm with local time-space grid refinement for simulation of waves. *Computational geosciences*, 16(1):39–54, 2012.

[91] R. B. Lowrie, P. L. Roe, and B. van Leer. Space-time methods for hyperbolic conservation laws. In *Barriers and challenges in computational fluid dynamics (Hampton, VA, 1996)*, volume 6 of *ICASE/LaRC Interdiscip. Ser. Sci. Eng.*, pages 79–98. Kluwer Acad. Publ., Dordrecht, 1998.

[92] A. Maciąg and J. Wauer. Solution of the two-dimensional wave equation by using wave polynomials. *J. Engrg. Math.*, 51(4):339–350, 2005.

[93] K. J. Marfurt. Accuracy of finite-difference and finite-element modeling of the scalar and elastic wave equations. *Geophysics*, 49(5):533–549, 1984.

[94] A. Moiola. Trefftz discontinuous Galerkin methods on unstructured meshes for the wave equation. In *Proceedings of XXIV CEDYA / XIV CMA conference*, pages 387–395, 2015.

[95] A. Moiola, R. Hiptmair, and I. Perugia. Plane wave approximation of homogeneous Helmholtz solutions. *Z. Angew. Math. Phys.*, 62(5):809–837, 2011.

[96] P. Monk and G. R. Richter. A discontinuous Galerkin method for linear symmetric hyperbolic systems in inhomogeneous media. *J. Sci. Comput.*, 22/23:443–477, 2005.

[97] P. M. Morse and H. Feshbach. *Methods of Theoretical Physics.* McGraw-Hill Book Co., Inc., New York-Toronto-London, 1953.

[98] P. M. Morse and H. Feshbach. *Methods of Theoretical Physics.* McGraw-Hill Book Co., Inc., New York-Toronto-London, 1953.

[99] W. A. Mulder. Higher-order mass-lumped finite elements for the wave equation. *Journal of Computational Acoustics*, 9(02):671–680, 2001.

[100] N. C. Nguyen, J. Peraire, and B. Cockburn. High-order implicit hybridizable discontinuous Galerkin methods for acoustics and elastodynamics. *J. Comput. Phys.*, 230(10):3695–3718, 2011.

[101] J. Nitsche. Über ein Variationsprinzip zur Lösung von Dirichlet-Problemen bei Verwendung von Teilräumen, die keinen Randbedingungen unterworfen

sind. *Abh. Math. Sem. Univ. Hamburg*, 36:9–15, 1971. Collection of articles dedicated to Lothar Collatz on his sixtieth birthday.

[102] J. Nitsche. On Dirichlet problems using subspaces with nearly zero boundary conditions. In *The mathematical foundations of the finite element method with applications to partial differential equations (Proc. Sympos., Univ. Maryland, Baltimore, Md., 1972)*, pages 603–627. Academic Press, New York, 1972.

[103] I. V. Oseledets. Constructive representation of functions in low-rank tensor formats. *Constr. Approx.*, 37(1):1–18, 2013.

[104] J. Palaniappan, R. B. Haber, and R. L. Jerrard. A spacetime discontinuous Galerkin method for scalar conservation laws. *Comput. Methods Appl. Mech. Engrg.*, 193(33-35):3607–3631, 2004.

[105] J. Peiró and S. Sherwin. Finite difference, finite element and finite volume methods for partial differential equations. In *Handbook of materials modeling*, pages 2415–2446. Springer, 2005.

[106] S. Petersen, C. Farhat, and R. Tezaur. A space-time discontinuous Galerkin method for the solution of the wave equation in the time domain. *Internat. J. Numer. Methods Engrg.*, 78(3):275–295, 2009.

[107] R. B. Platte and L. N. Trefethen. Chebfun: a new kind of numerical computing. In *Progress in industrial mathematics at ECMI 2008*, volume 15 of *Math. Ind.*, pages 69–87. Springer, Heidelberg, 2010.

[108] W. Reed and T. Hill. Triangular mesh methods for the neutron transport equation. Technical Report L.A-UR-73-479, Los Alamos Scientific Laboratory, 1973.

[109] G. R. Richter. An explicit finite element method for the wave equation. *Appl. Numer. Math.*, 16(1-2):65–80, 1994. A Festschrift to honor Professor Robert Vichnevetsky on his 65th birthday.

[110] B. Rivière. *Discontinuous Galerkin Methods for Solving Elliptic and Parabolic Equations*. Frontiers in Applied Mathematics, Vol. 35. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2008. Theory and implementation.

[111] B. Rivière, M. F. Wheeler, and V. Girault. Improved energy estimates for interior penalty, constrained and discontinuous Galerkin methods for elliptic problems. I. *Comput. Geosci.*, 3(3-4):337–360 (2000), 1999.

[112] B. Rivière, M. F. Wheeler, and V. Girault. A priori error estimates for finite element methods based on discontinuous approximation spaces for elliptic problems. *SIAM J. Numer. Anal.*, 39(3):902–931, 2001.

[113] C. Schwab. *p- and hp-Finite Element Methods*. Numerical Mathematics and Scientific Computation. The Clarendon Press, Oxford University Press, New York, 1998. Theory and applications in solid and fluid mechanics.

[114] R. Shamasundar and W. A. Mulder. Improving the accuracy of mass-lumped finite-elements in the first-order formulation of the wave equation by defect correction. *J. Comput. Phys.*, 322:689–707, 2016.

[115] C.-W. Shu. Discontinuous Galerkin method for time-dependent problems: survey and recent developments. In *Recent developments in discontinuous Galerkin finite element methods for partial differential equations*, volume 157 of *IMA Vol. Math. Appl.*, pages 25–62. Springer, Cham, 2014.

[116] S. Thite. Adaptive spacetime meshing for discontinuous Galerkin methods. *Comput. Geom.*, 42(1):20–44, 2009.

[117] L. L. Thompson and P. M. Pinsky. A space-time finite element method for structural acoustics in infinite domains. I. Formulation, stability and convergence. *Comput. Methods Appl. Mech. Engrg.*, 132(3-4):195–227, 1996.

[118] L. L. Thompson and P. M. Pinsky. A space-time finite element method for structural acoustics in infinite domains. II. Exact time-dependent non-

reflecting boundary conditions. *Comput. Methods Appl. Mech. Engrg.*, 132(3-4):229–258, 1996.

[119] I. Toulopoulos and J. A. Ekaterinaris. High-order discontinuous galerkin discretizations for computational aeroacoustics in complex domains. *AIAA journal*, 44(3):502–511, 2006.

[120] A. Townsend and L. N. Trefethen. An extension of Chebfun to two dimensions. *SIAM J. Sci. Comput.*, 35(6):C495–C518, 2013.

[121] L. N. Trefethen. *Approximation Theory and Approximation Practice*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2013.

[122] E. Trefftz. Ein gegenstuck zum ritzschen verfahren. *Proc. 2nd Int. Cong. Appl. Mech., Zurich, 1926*, pages 131–137, 1926.

[123] A. Üngör and A. Sheffer. Pitching tents in space-time: mesh generation for discontinuous Galerkin method. *Internat. J. Found. Comput. Sci.*, 13(2):201–221, 2002. Volume and surface triangulations.

[124] T. Vdovina and S. E. Minkoff. An a priori error analysis of operator upscaling for the acoustic wave equation. *Int. J. Numer. Anal. Model.*, 5(4):543–569, 2008.

[125] D. Wang, R. Tezaur, and C. Farhat. A hybrid discontinuous in space and time Galerkin method for wave propagation problems. *Internat. J. Numer. Methods Engrg.*, 99(4):263–289, 2014.

[126] T. Warburton. A low-storage curvilinear discontinuous Galerkin method for wave problems. *SIAM J. Sci. Comput.*, (4):A1987–A2012, 2013.

[127] T. Warburton and J. S. Hesthaven. On the constants in *hp*-finite element trace inverse inequalities. *Comput. Methods Appl. Mech. Engrg.*, 192(25):2765–2773, 2003.

[128] M. F. Wheeler. An elliptic collocation-finite element method with interior penalties. *SIAM J. Numer. Anal.*, 15(1):152–161, 1978.

[129] E. T. Whittaker. On the partial differential equations of mathematical physics. *Math. Ann.*, 57(3):333–355, 1903.

[130] F. Zheng and Z. Chen. Numerical dispersion analysis of the unconditionally stable 3-D ADI-FDTD method. *Microwave Theory and Techniques, IEEE Transactions on*, 49(5):1006–1009, 2001.

[131] O. Zienkiewicz. Trefftz type approximation and the generalized finite element method- history and development. 4(3):305–316, 1997.

[132] O. C. Zienkiewicz. Achievements and some unsolved problems of the finite element method. *Internat. J. Numer. Methods Engrg.*, 47(1-3):9–28, 2000. Richard H. Gallagher Memorial Issue.