UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

UNIVERSITY OF PRETORIA

Convergence analysis of finite element approximations of solutions of hyperbolic equations

by

Belinda Stapelberg

Submitted in partial fulfillment for the degree of Magister Scientiae

in the

Faculty of Natural and Agricultural Sciences
Department of Mathematics and Applied Mathematics

January 2015

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

UNIVERSITY OF PRETORIA

Summary

Faculty of Natural and Agricultural Sciences
Department of Mathematics and Applied Mathematics

Magister Scientiae

by Belinda Stapelberg

The vibration of elastic bodies and structures consisting of elastic bodies is an active research field in engineering and applied mathematics. Typically, a mathematical model is a complex system of partial differential equations. However, a model problem may not have a solution in the classical sense and the rate of convergence of numerical approximations depends on the "smoothness" of a solution.

The aim of this research is to investigate the disparity noticed in the theory between the existence of solutions and the regularity assumed on these solutions for convergence of the Galerkin finite element method. In the articles considered, substantially more differentiability properties for the solution are assumed than obtained in existence theory. These assumptions are very restrictive; the solution is required to be smoother than even a classical solution.

The theory of existence of a solution to a general linear vibration problem that appeared in an article published in 2002, was considered first. To compare, alternative theories on existence of solutions to hyperbolic partial differential equations, were also studied. The existence results, improved regularity of solutions and compatibility conditions, which are highly restrictive, are presented.

In 2013 an article appeared wherein convergence is proved, but with weaker assumptions than the other articles considered. This is achieved by splitting the error into the semi-discreet and fully discreet errors. However, it is still necessary to assume higher regularity of the solution. The focus in this dissertation was to compare the article to other research results, and to highlight significant parts of the proofs in the article. Also, minor improvements were made and it was proved that the results obtained from existence theory are sufficient for convergence, but no result on the order of convergence could be obtained.

A recent article (2011) on the continuous Galerkin method, where the model problem considered includes strong damping, was also analysed. The results from this article is proved in great detail, and possible oversights or omissions discovered are either rectified or reported.

The discontinuous Galerkin (DG) finite element method is also included in the research, with the aim to determine whether the assumptions made on the regularity needed for convergence are less restrictive than those made for the continuous Galerkin method. Disappointingly, the results offer no significant improvement. The semi-discreet and fully discreet DG error estimates are from articles published in 2006 and 2009 respectively. The results are proven in greater detail in this dissertation.

Interesting phenomena obtained from numerical experiments are observed and to a large degree the theory and experiments agree. However, there are indications that the order of convergence may in some cases be better than predicted by the theory.

The main conclusion is that there is a problem when applying theoretical results to real world problems. Further research is required to prove results where error estimates are derived without restrictive assumptions on boundary and initial data.

Acknowledgements

The financial assistance of the National Research Foundation (NRF) (South Africa) towards the research is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the author and are not necessarily to be attributed to the NRF.

Contents

Chapter 1

Model problems in variational form

1.1 Introduction

The vibration of elastic bodies and structures consisting of elastic bodies is an active research field in engineering and applied mathematics. A mathematical model for such a vibrating system is a complex system of partial differential equations. Numerical approximation of solutions to these partial differential equations is inevitable. The finite element method proves to be ideal here, and it is used for steady-state problems, eigenvalue problems and dynamic problems.

It has become common practice to use numerical methods (often pre-programmed computer software) to "solve" problems. The motion of a system can be simulated or the natural frequencies calculated. As computers improve, people attempt to solve more complex model problems, often leading to unexpected difficulties.

It is important to realise that a model problem may not have a solution in the classical sense and the rate of convergence of numerical approximations depends on the "smoothness" of a solution. For this reason existence theory (although theoretical) is of great practical importance. These remarks are especially relevant for the field of partial differential equations of hyperbolic type, which is investigated in this dissertation.

The existence of solutions for problems involving partial differential equations can have a number of possibilities. For a start, there is a distinction between so called weak solutions and classical solutions. But there are different definitions of weak solutions. For example, what is referred to as a weak solution in the book of [Eva98], is referred to as a mild solution in the book of [Paz83] (the definitions are in Section 2.4). Either way, the existence of a weak solution or a mild solution is not sufficient for the theory of the finite element method (see Section 6.1 for an example), as we explain below.

Convergence of the finite element approximation for the multi-dimensional wave equation is given in [OR76, Section 9.6]. It is assumed that the solution has a fourth order time derivative. The differentiability properties of the solution obtained from regularity theory require that serious restrictions be imposed on the initial states (see Section 2.4 of this dissertation). One of the most recent articles that still deals with the continuous Galerkin method is the article [Kar11a]. In a preliminary investigation of this article, it was noted that he cites an existence result (from the book of [LM72]) for the existence of a weak solution, but in proving convergence, assumes more differentiability properties for the solution, without giving any reference. It is natural to investigate this problem by consulting more articles.

Numerical experiments are carried out in [Wu03] and [GSS06] to calculate the order of convergence. In both articles smooth solutions were considered as well as solutions that do not satisfy the regularity requirements of the theory. The authors reported that the theory matched the numerical experiments for the smooth solutions, but in other cases results were found that could not be fully explained by the theory.

The aim of the research of this research is to investigate the following::

- convergence of the finite element approximation;
- error estimates;
- existence of solutions to partial differential equations of hyperbolic type;
- regularity properties of solutions.

The ultimate aim is to determine the extent of the disparity between the results obtained from existence theory, the compatibility conditions required for higher regularity and the regularity assumptions made for convergence of the finite element method. Details of the assumptions made for higher regularity are provided in Section 2.4.

1.2 The multi-dimensional wave equation with weak damping

Consider the wave equation in an n-dimensional bounded domain ($n = 2$ or 3) denoted by Ω. The boundary of Ω is denoted by $\partial\Omega$ and the unit outer normal vector to Ω at $\partial\Omega$ by $\mathbf{n}$.

Problem MW

Let Σ be part of the boundary $\partial\Omega$. We have different boundary conditions on Σ and $\partial\Omega - \Sigma$. Given functions f, u_0 and u_1, find w defined on $\bar{\Omega} \times [0, T]$ such that

$$\begin{aligned} \rho\partial_t^2 w &= \nabla \cdot (A\nabla w) - k\partial_t w + f \quad \text{in} \quad \Omega \times (0, T), \\ w &= 0 \quad \text{on} \quad \partial\Omega - \Sigma, \\ (A\ \nabla w) \cdot \mathbf{n} &= 0 \quad \text{on} \quad \Sigma, \end{aligned}$$

while $w(\cdot, 0) = u_0$ and $\partial_t w(\cdot, 0) = u_1$.

The given parameters in the problem are the matrix of functions $A = (a_{ij})$ and the functions k and ρ.

Assumptions on the parameters

1. $a_{ij} \in C(\bar{\Omega}) \cap C^1(\Omega)$.
2. The matrix A is uniformly positive definite, i.e. there exists a constant $\alpha > 0$ such that
$$\sum_{i,j=1}^{n} a_{ij}(x)\xi_i\xi_j \geq \alpha \sum_{i=1}^{n} \xi_i^2,$$
for all $x \in \bar{\Omega}$ and all $(\xi_1, \xi_2, ..., \xi_n) \in \mathbb{R}^n$.
3. There exist positive constants c_1, c_2 and c_3 such that
$c_1 \leq \rho \leq c_2 \quad \text{and} \quad 0 \leq k \leq c_3$.

Remark The differential operator $L = \nabla \cdot (A\nabla)$ is referred to as uniformly strongly elliptic.

The vibration of a membrane

In the two-dimensional case the wave equation models the vibration of a membrane [Inm94, Section 6.6]. We have the partial differential equation

$$\rho\partial_t^2 w \quad = \quad \tau\,\nabla^2 w - k\partial_t w,$$

where ρ is the mass per unit area, τ the constant tension per unit length, and k a damping constant.

The acoustic wave equation

In the three-dimensional case the wave equation models the propagation of sound waves. The following derivation follows [PR05, Section 1.4.2] with some modification.

Consider a gas at rest, i.e. the velocity $\mathbf{v} = 0$. Suppose the pressure is p_0 and the density is ρ_0. A small disturbance leads to motion in the gas, and the pressure p and density ρ are no longer constant. The linear approximation for the continuity equation is

$$\partial_t\,\rho^* + \rho_0\nabla\cdot\mathbf{v} = 0, \tag{1.2.1}$$

where $\rho^* = \rho - \rho_0$.

The linear approximation for the equation of motion for an ideal gas is

$$\rho_0\partial_t\,\mathbf{v} + \nabla p^* = 0, \tag{1.2.2}$$

where $p^* = p - p_0$ and where modified pressure is considered.

Using the approximation

$$\rho^* = f(p) - f(p_0) \doteqdot f'(p_0)p^*$$

in Equation (1.2.1) yields

$$c\partial_t\,p^* + \rho_0\nabla\cdot\mathbf{v} = 0,$$

where $c = f'(p_0)$. It follows that

$$c\partial_t^2\,p^* + \rho_0\partial_t\nabla\cdot\mathbf{v} = 0. \tag{1.2.3}$$

Also, from Equation (1.2.2) we have

$$\rho_0 \partial_t \nabla \cdot \mathbf{v} + \nabla^2 p^* = 0. \tag{1.2.4}$$

Subtracting Equation (1.2.4) from Equation (1.2.3) results in the acoustic wave equation

$$c\partial_t^2 p^* - \nabla^2 p^* = 0.$$

Heat conduction

The multi-dimensional wave equation also models hyperbolic heat conduction, see Section 1.3.

1.2.1 Variational form

Theorem 1.2.1. *[Apo67, p. 457]* ***Gauss' theorem***
If $F_i \in C^1(\bar{\Omega})$ for $i = 1, 2, 3,$ then

$$\iiint_\Omega \nabla \cdot F \, dV = \iint_{\partial\Omega} F \cdot \mathbf{n} \, dA,$$

where $\nabla \cdot F$ is the divergence of the vector F and $\mathbf{n}$ is the unit outward normal.

Proposition 1.2.2. ***Green's formula***
If $u \in C^2(\bar{\Omega})$ and $v \in C^1(\bar{\Omega})$ then

$$\iiint_\Omega -(\nabla \cdot (A\nabla u))v \, dV = \iiint_\Omega A\nabla u \cdot \nabla v \, dV - \iint_{\partial\Omega} v(A\nabla u) \cdot \mathbf{n} \, dA.$$

Proof. We have that $\nabla \cdot ((A\nabla u)v) = \nabla \cdot (A\nabla u)v + A\nabla u \cdot \nabla v$ and therefore

$$\iiint_\Omega -(\nabla \cdot ((A\nabla u))v \, dV = \iiint_\Omega A\nabla u \cdot \nabla v \, dV - \iiint_\Omega \nabla \cdot ((A\nabla u)v) \, dV.$$

From Gauss' theorem (Theorem 1.2.1) we have

$$\iiint_\Omega \nabla \cdot ((A\nabla u)v) dV = \iint_{\partial\Omega} v(A\nabla u) \cdot \mathbf{n} dA$$

and therefore we have the result. □

Now, if w is a solution to Problem MW, then for each $v \in C^1(\bar{\Omega})$, and following from the boundary conditions on Σ

$$\begin{aligned}
\iiint_\Omega \rho \partial_t^2 w v \, dV &= \iiint_\Omega \nabla \cdot (A\nabla w) v \, dV - \iiint_\Omega k \partial_t w v \, dV + \iiint_\Omega f v \, dV \\
&= -\iiint_\Omega A\nabla w \cdot \nabla v \, dV + \iint_{\partial\Omega} v(A\nabla w) \cdot \mathbf{n} \, dA \\
&\qquad - \iiint_\Omega k \partial_t w v \, dV + \iiint_\Omega f v \, dV \\
&= -\iiint_\Omega A\nabla w \cdot \nabla v \, dV + \iint_{\partial\Omega - \Sigma} v(A\nabla w) \cdot \mathbf{n} \, dA \\
&\qquad - \iiint_\Omega k \partial_t w v \, dV + \iiint_\Omega f v \, dV. \qquad (1.2.5)
\end{aligned}$$

Test functions

$$\mathcal{T}(\Omega) := \{v \in C^1(\bar{\Omega}) : v = 0 \text{ on } \partial\Omega - \Sigma\}.$$

Following now from the definition of the test functions, we have from equation (1.2.5) that, if w is a solution to Problem MW, then for each $v \in \mathcal{T}(\Omega)$

$$\iiint_\Omega \rho \partial_t^2 w v \, dV = -\iiint_\Omega A\nabla w \cdot \nabla v \, dV - \iiint_\Omega k \partial_t w v \, dV + \iiint_\Omega f v \, dV. \qquad (1.2.6)$$

Bilinear forms for Problem MW

$$\begin{aligned}
b(u,v) &= \iiint_\Omega A\nabla u \cdot \nabla v \, dV \\
c(u,v) &= \iiint_\Omega \rho u v \, dV \\
a(u,v) &= \iiint_\Omega k u v \, dV
\end{aligned}$$

Notation

$$(f,v)_\Omega = \iiint_\Omega f v \, dV$$

We can now write Problem MW in variational form.

Problem MWV

Find w such that for each $t > 0$, $w(\cdot,t) \in \mathcal{T}(\Omega)$ and

$$c(\partial_t^2 w(\cdot,t), v) + a(\partial_t w(\cdot,t), v) + b(w(\cdot,t), v) = (f(\cdot,t), v)_\Omega$$

for each $v \in \mathcal{T}(\Omega)$.

1.3 Heat conduction

1.3.1 The classical heat equation

In this section we consider heat conduction in a solid. Definitions, units and mechanics of energy transport can be found in Chapter 8 of [BSL60].

Notation

1. The density of the material: ρ. $[kg\ m^{-3}]$
2. The specific heat of the material: c_p. $[J\ kg^{-1}K^{-1}]$
3. Temperature: T. $[K \text{ or } {}^\circ C]$
4. Heat flux: $\mathbf{q}$. $[W\ m^{-2}]$

Mathematical model

Consider an arbitrary region $\mathcal{D}$ in space with boundary $\mathcal{E}$. The choice of zero temperature is arbitrary. The quantity of heat energy to raise the temperature of the material in $\mathcal{D}$ from 0 to T is

$$\iiint_{\mathcal{D}} \rho c_p T\ dV.$$

The flux of heat energy into the region $\mathcal{D}$ is

$$-\iint_{\mathcal{E}} \mathbf{q}\cdot\mathbf{n}\ dS,$$

where $\mathbf{n}$ is the outward unit normal vector to $\mathcal{D}$.

Conservation of heat energy

$$\frac{d}{dt}\int_{\mathcal{D}} \rho c_p T\ dV = -\int_{\mathcal{E}} \mathbf{q}\cdot\mathbf{n}\ dS.$$

This is the basic assumption for the theory. It is often referred to as the energy balance. It follows from the conservation of heat energy and the divergence theorem that

$$\rho c_p \partial_t T = -\nabla\cdot\mathbf{q}. \tag{1.3.1}$$

Another assumption is necessary for the mathematical model. It is referred to as Fourier's law of heat conduction:

$$\mathbf{q} = -k\nabla T. \tag{1.3.2}$$

The constant k is the thermal conductivity.

Combining Equations (1.3.1) and (1.3.2), we arrive at the following partial differential equation, called the classical heat equation:

$$\partial_t T = c^2 \nabla^2 T, \quad \text{where} \quad c^2 = \frac{k^2}{\rho^2 c_p^2}.$$

1.3.2 Hyperbolic heat conduction

Cattaneo [Cat48] and Vernotte [Ver58] indepently proposed a modification to Fourier's law (called the Cattaneo-Vernotte model), with the constitutive equation being:

$$\mathbf{q} + \tau \partial_t \mathbf{q} = -k\nabla T, \tag{1.3.3}$$

where τ is the time delay. This model gives rise to a (weakly) damped wave equation (known as the hyperbolic heat conduction equation, HHCE) when combined with the energy conservation equation (1.3.1). Taking the divergence of (1.3.3) we get

$$\nabla \cdot \mathbf{q} + \tau \nabla \cdot \partial_t \mathbf{q} = -k\nabla \cdot \nabla T. \tag{1.3.4}$$

Differentiating (1.3.1) yields

$$\rho c_p \partial_t^2 T = -\partial_t \nabla \cdot \mathbf{q}. \tag{1.3.5}$$

Combining Equations (1.3.4) and (1.3.5) we obtain:

$$\tau \rho c_p \partial_t^2 T - \nabla \cdot \mathbf{q} = k\nabla^2 T. \tag{1.3.6}$$

Substituting (1.3.1) in (1.3.6) we obtain the hyperbolic heat conduction equation:

$$\tau \rho c_p \partial_t^2 T + \rho c_p \partial_t T = k\nabla^2 T. \tag{1.3.7}$$

Consider a domain Ω as in Section 1.2.

Problem HHCE

Given functions f, T_0 and T_1, find T defined on $\bar{\Omega} \times [0, T]$ such that

$$\begin{aligned} \gamma_2 \partial_t^2 T + \gamma_1 \partial_t T - \nabla^2 T &= 0 \quad \text{in} \quad \Omega \times (0, T), \\ T &= 0 \quad \text{on} \quad \partial\Omega - \Sigma, \\ \nabla T \cdot \mathbf{n} &= 0 \quad \text{on} \quad \Sigma, \end{aligned} \tag{1.3.8}$$

where $\gamma_1 = \dfrac{\rho c_p}{k}$ and $\gamma_2 = \dfrac{\tau \rho c_p}{k}$. Initial conditions are $T(\cdot, 0) = T_0$ and $\partial_t T(\cdot, 0) = T_1$.

Remark If the flux $\mathbf{q} \cdot \mathbf{n}$ is zero on Σ, then $\partial_t \mathbf{q} \cdot \mathbf{n}$ is also zero on Σ. From (1.3.3) it follows that

$$\mathbf{0} = -k \nabla T \cdot \mathbf{n},$$

which gives the boundary condition on Σ.

1.3.3 Variational form

Problem HHCE is a special case of Problem MW. For convenience we give the variational form of Problem HHCE here.

Test functions

$$\mathcal{T}(\Omega) := \{v \in C^1(\bar{\Omega}) : v = 0 \text{ on } \partial\Omega - \Sigma\}.$$

We have for every $v \in \mathcal{T}(\Omega)$,

$$\iiint_\Omega \gamma_1 \partial_t^2 T v \, dV + \iiint_\Omega \gamma_1 \partial_t T v \, dV + \iiint_\Omega \nabla T \cdot \nabla v \, dV = 0.$$

Now we define our bilinear forms.

Bilinear forms for Problem HHCE

$$\begin{aligned} b(u, v) &= \iiint_\Omega \nabla u \cdot \nabla v \, dV \\ c(u, v) &= \iiint_\Omega \gamma_1 u v \, dV \\ a(u, v) &= \frac{1}{\tau} c(u, v) \end{aligned}$$

We can now write Problem HHCE in variational form.

Problem HHCEV

Find T such that for each $t > 0$, $T(\cdot, t) \in \mathcal{T}(\Omega)$ and

$$c(\partial_t^2 T(\cdot, t), v) + a(\partial_t T(\cdot, t), v) + b(T(\cdot, t), v) = 0$$

for each $v \in \mathcal{T}(\Omega)$.

Remark Problem HHCE is a special case of Problem MW.

1.3.4 Dual-Phase-Lag model

Tzou [Tzo95] extended the hyperbolic heat conduction model to the Single-Phase-Lag (SPL) model:

$$\mathbf{q}(\bar{r}, t + \tau_q) = -k\nabla T(\bar{r}, t), \tag{1.3.9}$$

where τ_q is the phase-lag in the heat flux $\mathbf{q}$, and further extended it to include a phase-lag in the temperature gradient (∇T) as well:

$$\mathbf{q}(\bar{r}, t + \tau_q) = -k\nabla T(\bar{r}, t + \tau_T), \tag{1.3.10}$$

known as the Dual-Phase-Lag (DPL) model. If a first-order Taylor expansion is performed on equation (1.3.9), the result is:

$$\mathbf{q}(\bar{r}, t) + \tau_q \partial_t \mathbf{q}(\bar{r}, t) \cong -k\nabla T(\bar{r}, t). \tag{1.3.11}$$

Similarly, for Equation (1.3.10), the result is:

$$\mathbf{q}(\bar{r}, t) + \tau_q \partial_t \mathbf{q}(\bar{r}, t) \cong -k\left[\nabla T(\bar{r}, t) + \tau_T \partial_t \left(\nabla T(\bar{r}, t)\right)\right] \tag{1.3.12}$$

Taking the divergence of Equation (1.3.12) and combing it with (1.3.1) results in

$$-\tau_q \partial_t \left(\nabla \cdot \mathbf{q}\right)(\bar{r}, t) = k\left[\nabla^2 T(\bar{r}, t) + \tau_T \partial_t \left(\nabla^2 T(\bar{r}, t)\right)\right]. \tag{1.3.13}$$

Using equation (1.3.1) again in (1.3.13) we obtain:

$$\tau_q \rho c_p \partial_t^2 T - \rho c_p \partial_t T = -k\nabla^2 T - \tau_T \nabla^2 (\partial_t T) \tag{1.3.14}$$

1.3.5 Generalised Dual-Phase-Lag model

In this section we consider a general form of the DPL model. This generalised model is the focus in [Kar11a], where it is stated without explanation (although a reference is given).

Suppose A is a symmetric matrix that satisfies the same assumptions as for Problem MW in Section 1.2. Consider a generalisation of (1.3.10):

$$\mathbf{q}(\bar{r}, t + \tau_q) = -A\nabla T(\bar{r}, t + \tau_T). \tag{1.3.15}$$

If a first-order Taylor expansion is performed on Equation (1.3.15), then

$$\mathbf{q}(\bar{r}, t) + \tau_q \partial_t \mathbf{q}(\bar{r}, t) \cong -A\nabla T(\bar{r}, t) + \tau_T A \partial_t \nabla T(\bar{r}, t). \tag{1.3.16}$$

Using this together with Equation (1.3.1), we obtain

$$\gamma_2 \partial_t^2 T + \gamma_1 \partial_t T - \nabla \cdot (Q\nabla(\partial_t T)) - \nabla \cdot (A\nabla T) = 0. \tag{1.3.17}$$

where

$$\gamma_2 = \tau_q \rho c_p,\ \gamma_1 = \rho c_p \text{ and } Q = \tau_T A. \tag{1.3.18}$$

Problem DPL

Given functions $\tilde{f}, T_0$ and T_1, find T defined on $\bar{\Omega} \times [0, T]$ such that

$$\begin{aligned} \gamma_2 \partial_t^2 T + \gamma_1 \partial_t T - \nabla \cdot (Q\nabla(\partial_t T)) - \nabla \cdot (A\nabla T) &= \tilde{f} \quad \text{in} \quad \Omega \times (0, T), \\ T &= 0 \quad \text{on} \quad \partial\Omega - \Sigma, \\ A\nabla T \cdot \mathbf{n} &= 0 \quad \text{on} \quad \Sigma, \end{aligned}$$

while $T(\cdot, 0) = T_0$ and $\partial_t T(\cdot, 0) = T_1$.

The function $\tilde{f}$ is a source term.

Remark If the heat flux $\mathbf{q} \cdot \mathbf{n}$ is zero on Σ, then we have $\mathbf{0} = -A\nabla T \cdot \mathbf{n}$.

We consider this model problem in Chapter 4.

1.3.6 Variational form

The derivation of the variational form is similar to that of the multi-dimensional wave equation, using Proposition 1.2.2.

Test functions

$$\mathcal{T}(\Omega) := \{v \in C^1(\bar{\Omega}) : v = 0 \text{ on } \partial\Omega - \Sigma\}.$$

Definition 1.3.1. Bilinear forms for Problem DPL

$$\begin{aligned} b(u,v) &= \iiint_\Omega A\nabla u \cdot \nabla v \, dV \\ c(u,v) &= \iiint_\Omega \gamma_2 uv \, dV \\ a(u,v) &= \iiint_\Omega \gamma_1 uv + Q\nabla u \cdot \nabla v \, dV \end{aligned}$$

The variational form of Problem DPL is therefore given by the following.

Problem DPLV

Find T such that for each $t > 0$, $T(\cdot,t) \in \mathcal{T}(\Omega)$ and

$$c(\partial_t^2 T(\cdot,t), v) + a(\partial_t T(\cdot,t), v) + b(T(\cdot,t), v) = (\tilde{f}(\cdot,t), v)_\Omega$$

for each $v \in \mathcal{T}(\Omega)$.

1.4 Vibration of a Reissner-Mindlin plate

Initially the focus of the dissertation was on the multi-dimensional wave equation. However, the investigation of the article [BV13] (see Chapter 3), which considers a general linear vibration problem, lead to the consideration of other applications. The two applications are the vibration of a Reissner-Mindlin plate model (the focus of this section and Section 6.2), and linear elasto-dynamics (Section 6.3).

1.4.1 Equations of motion

Consider small transverse vibrations of a thin plate with thickness h and density ρ. The reference configuration for the plate is a domain Ω in the plane. The transverse displacement of $\mathbf{x}$ at time t is denoted by $w(\mathbf{x},t)$. Also let $\boldsymbol{\psi}(\mathbf{x},t)$ be the angle between a "material line" and a perpendicular to the plane, and $\phi(\mathbf{x},t)$ the angle between the projection of the material line in the plane and the unit vector $\mathbf{e}_1$ (see [Rei88, Sec 3.2, Sec 3.5]). The angle $\boldsymbol{\psi}$ is approximated by

$$\boldsymbol{\psi} = [\psi_1 \;\; \psi_2]^T = [\psi \cos\phi \;\; \psi \sin\phi]^T.$$

The stresses result in a force density $\mathbf{Q}$ and a moment density $M = \begin{bmatrix} M_{11} & M_{12} \\ M21 & M_{22} \end{bmatrix}$. The external load on plate is denoted by q.

Before proceeding, note that

$$\text{div } M = \begin{bmatrix} \partial_1 M_{11} + \partial_2 M_{12} \\ \partial_1 M21 + \partial_2 M_{22} \end{bmatrix}.$$

The equations of motion of the plate (see [Min51] and [Rei88, p.152]) are given by

$$\rho h \partial_t^2 w \;=\; \text{div } \mathbf{Q} + q, \tag{1.4.1}$$

$$\rho I \partial_t^2 \boldsymbol{\psi} \;=\; \text{div } M - \mathbf{Q}, \tag{1.4.2}$$

where $I = \dfrac{h^3}{12}$ is the length moment of inertia.

1.4.2 Constitutive equations

The constitutive equations for the plate model are derived from Hooke's law (see [Rei88, p.61] and [Min51]).

$$\mathbf{Q} = \kappa^2 G h(\nabla w + \psi), \tag{1.4.3}$$

where G is the shear modulus and κ^2 a correction factor.

$$M = \frac{1}{2} D \begin{bmatrix} 2\big(\partial_1\psi_1 + \nu\partial_2\psi_2\big) & (1-\nu)\big(\partial_1\psi_2 + \partial_2\psi_1\big) \\ (1-\nu)\big(\partial_1\psi_2 + \partial_2\psi_1\big) & 2\big(\partial_2\psi_2 + \nu\partial_1\psi_1\big) \end{bmatrix}, \tag{1.4.4}$$

where $D = \dfrac{EI}{1-\nu^2}$ is a measure of stiffness of the plate. E is Young's modulus and ν Poisson's ratio.

The equations of motion and the constitutive equations above are known as the **Reissner-Mindlin** plate model. It is convenient to derive the variational form of this plate model directly from the equations of motion and the constitutive equations. However, the constitutive equations may be substituted into the equations of motion, leading to a system of three partial differential equations (see [Rei88, p.152] and [Min51]).

1.4.3 Dimensionless form

We derive the dimensionless form of the Reissner-Mindlin plate model in the usual way. We introduce the dimensionless variables

$$\tau = \frac{t}{t_0}, \quad \xi_1 = \frac{x_1}{\ell} \quad \text{and} \quad \xi_2 = \frac{x_2}{\ell},$$

where ℓ is a suitable length and t_0 must still be specified.

The dimensionless variables, with $\mathbf{x} = (x_1, x_2)$ and $\boldsymbol{\xi} = (\xi_1, \xi_2)$, are

$$w^*(\boldsymbol{\xi}, \tau) = \left(\frac{1}{\ell}\right) w(\mathbf{x}, t), \quad \boldsymbol{\psi}^*(\boldsymbol{\xi}, \tau) = \boldsymbol{\psi}(\mathbf{x}, t),$$

$$\boldsymbol{Q}^*(\boldsymbol{\xi}, \tau) = \left(\frac{1}{\ell G\kappa^2}\right) \boldsymbol{Q}(\mathbf{x}, t), \quad M^*(\boldsymbol{\xi}, \tau) = \left(\frac{1}{\ell^2 G\kappa^2}\right) M(\mathbf{x}, t)$$
$$\text{and} \quad q^*(\boldsymbol{\xi}, \tau) = \left(\frac{1}{G\kappa^2}\right) q(\mathbf{x}, t).$$

The dimensionless constants that are used are given by

$$h^* = \frac{h}{\ell}, \quad I^* = \frac{(h^*)^3}{12} \quad \text{and} \quad \beta = \frac{\ell^3 G\kappa^2}{EI}.$$

Choose $t_0 = \ell\sqrt{\dfrac{\rho}{G\kappa^2}}$. Using the original notation for the corresponding dimensionless quantities (for convenience), the equations of motion and constitutive equations in dimensionless form are presented below.

Reissner-Mindlin plate model

$$h\partial_t^2 w = \text{div } \boldsymbol{Q} + q, \tag{1.4.5}$$
$$I\partial_t^2 \boldsymbol{\psi} = \text{div } M - \boldsymbol{Q}, \tag{1.4.6}$$
$$\boldsymbol{Q} = h\big(\nabla w + \boldsymbol{\psi}\big), \tag{1.4.7}$$

$$M = \frac{1}{2\beta(1-\nu^2)} \begin{bmatrix} 2\Big(\partial_1\psi_1 + \nu\partial_2\psi_2\Big) & (1-\nu)\Big(\partial_1\psi_2 + \partial_2\psi_1\Big) \\ (1-\nu)\Big(\partial_1\psi_2 + \partial_2\psi_1\Big) & 2\Big(\partial_2\psi_2 + \nu\partial_1\psi_1\Big) \end{bmatrix}. \tag{1.4.8}$$

1.4.4 Model problem

Problem RM

The reference configuration for the plate is the domain Ω. The model consists of

the equations of motion (1.4.5) and (1.4.6),

the constitutive equations (1.4.7) and (1.4.8), and

the boundary conditions $w = \psi_1 = \psi_2 = 0$.

Note that the boundary conditions on $\partial\Omega$ are conventional homogeneous boundary conditions where the plate is clamped. The initial conditions are

$$w(\cdot, 0) = w_0, \quad \boldsymbol{\psi}(\cdot, 0) = \boldsymbol{\psi}_0, \quad \partial_t w(\cdot, 0) = w_1 \quad \text{and} \quad \partial_t\boldsymbol{\psi}(\cdot, 0) = \boldsymbol{\psi}_1.$$

This model problem is considered in [Wu05] and [Wu06].

1.4.5 Variational form

The variational form is is derived directly from the equations of motion and the constitutive equations.

Using Proposition 1.2.2 we have:

$$\iint_\Omega (\text{div } \mathbf{Q})v\, dA \quad = \quad -\iint_\Omega \mathbf{Q}\cdot\nabla\, v\, dA + \int_{\partial\Omega} (\mathbf{Q}\cdot\mathbf{n})v\, ds. \tag{1.4.9}$$

Also, for any vector valued function $\boldsymbol{\phi} = [\phi_1\ \phi_2]^T$,

$$\iint_\Omega \text{div } M\cdot\boldsymbol{\phi}\, dA = -\iint_\Omega \text{tr}(M\Phi)\, dA + \int_{\partial\Omega} M\mathbf{n}\cdot\boldsymbol{\phi}\, ds. \tag{1.4.10}$$

where $\Phi = \begin{bmatrix} \partial_1\phi_1 & \partial_2\phi_1 \\ \partial_1\phi_2 & \partial_2\phi_2 \end{bmatrix}$ and "tr" denotes the trace of the matrix.

Multiply equation (1.4.5) by an arbitrary scalar valued function v and integrate. Using (1.4.9), we find that

$$h \iint_\Omega \partial_t^2 wv\, dA + \iint_\Omega \mathbf{Q} \cdot \nabla\, v\, dA - \int_{\partial\Omega} (\mathbf{Q} \cdot \mathbf{n}) v\, ds = \iint_\Omega qv\, dA. \tag{1.4.11}$$

From equation (1.4.6) we have

$$I \partial_t^2 \boldsymbol{\psi} \cdot \boldsymbol{\phi} = \text{div }\ M \cdot \boldsymbol{\phi} - Q \cdot \boldsymbol{\phi},$$

where $\boldsymbol{\phi}$ is an arbitrary vector valued function. Using (1.4.10) we find that

$$\begin{aligned} I \iint_\Omega \partial_t^2 \boldsymbol{\psi} \cdot \boldsymbol{\phi}\, dA + \iint_\Omega \text{tr}(M\Phi)\, dA - \int_{\partial\Omega} M\mathbf{n} \cdot \boldsymbol{\phi}\, ds \\ + \iint_\Omega \boldsymbol{Q} \cdot \boldsymbol{\phi}\, dA = 0. \end{aligned} \tag{1.4.12}$$

Test functions

Choose two spaces of test functions $\mathcal{T}_1(\Omega)$ and $\mathcal{T}_2(\Omega)$:

$$\begin{aligned} \mathcal{T}_1(\Omega) &= \{v \in C^1(\bar{\Omega}) \mid v = 0 \text{ on } \partial\Omega\}, \\ \mathcal{T}_2(\Omega) &= \{\boldsymbol{\phi} = [\phi_1\ \phi_2]^T \mid \phi_1, \phi_2 \in C^1(\bar{\Omega}),\ \boldsymbol{\phi} = \mathbf{0} \text{ on } \partial\Omega\}. \end{aligned}$$

Remark Other boundary conditions can be accommodated using different spaces of test functions.

From (1.4.11)

$$h \iint_\Omega \partial_t^2 wv\, dA + \iint_\Omega \mathbf{Q} \cdot \nabla\, v\, dA = \iint_\Omega qv\, dA, \tag{1.4.13}$$

for each $v \in \mathcal{T}_1(\Omega)$.

From (1.4.12)

$$I \iint_\Omega \partial_t^2 \boldsymbol{\psi} \cdot \boldsymbol{\phi}\, dA + \iint_\Omega \text{tr}(M\Phi)\, dA + \iint_\Omega \boldsymbol{Q} \cdot \boldsymbol{\phi}\, dA = 0, \tag{1.4.14}$$

for each $\boldsymbol{\phi} \in \mathcal{T}_2(\Omega)$.

We define a bilinear form b_B by

$$\begin{aligned} b_B(\boldsymbol{\psi},\boldsymbol{\phi}) &= \iint_\Omega \operatorname{tr}(M\Phi)\, dA \\ &= \frac{1}{\beta(1-\nu^2)} \iint_\Omega \Big((\partial_1\psi_1 + \nu\partial_2\psi_2)\partial_1\phi_1 + (\partial_2\psi_2 + \nu\partial_1\psi_1)\partial_2\phi_2 \Big)\, dA \\ &\quad + \frac{1}{2\beta(1+\nu)} \iint_\Omega (\partial_1\psi_2 + \partial_2\psi_1)(\partial_1\phi_2 + \partial_2\phi_1)\, dA, \end{aligned}$$

for each $\boldsymbol{\psi}, \boldsymbol{\phi}$ in $H^1(\Omega)^2$. It follows that

$$I \iint_\Omega \partial_t^2 \boldsymbol{\psi} \cdot \boldsymbol{\phi}\, dA + b_B(\boldsymbol{\psi},\boldsymbol{\phi}) + h \iint_\Omega (\nabla w + \boldsymbol{\psi}) \cdot \boldsymbol{\phi}\, dA = 0, \tag{1.4.15}$$

for each $\boldsymbol{\phi} \in \mathcal{T}_2(\Omega)$. Equation (1.4.7) is used for the definition of Q.

Problem RMV

Find w and $\boldsymbol{\psi}$ such that, for $t > 0$, $w(\cdot,t) \in \mathcal{T}_1(\Omega)$, $\boldsymbol{\psi}(\cdot,t) \in \mathcal{T}_2(\Omega)$ and Equations (1.4.13) and (1.4.15) hold for each $v \in \mathcal{T}_1(\Omega)$ and each $\boldsymbol{\phi} \in \mathcal{T}_2(\Omega)$. The initial conditions are

$$w(\cdot,0) = w_0, \quad \boldsymbol{\psi}(\cdot,0) = \boldsymbol{\psi}_0,$$

$$\partial_t w(\cdot,0) = w_1 \quad \text{and} \quad \partial_t \boldsymbol{\psi}(\cdot,0) = \boldsymbol{\psi}_1.$$

Chapter 2

Existence

2.1 Weak variational form

It is desirable to determine whether a given problem is well posed before considering a numerical approximation for the solution. In this chapter we consider the existence of solutions and various publications on the topic. In Section 2.3 we investigate the results of the paper [VV02] and in Section 2.4 we consider alternative existence results.

In this section we use the simple yet nontrivial examples in [VV02] to illustrate the weak variational form of a problem, necessary for existence theory. We consider the small one dimensional transverse vibrations of a cantilever beam to illustrate the three different types of damping that is considered in the general linear vibration problem defined in Section 2.2.1. The beam has length ℓ, cross sectional area A, density ρ, cross sectional area moment of inertia I, and Young's modulus E. All of these values are constants in this case. The reference configuration for the beam is the interval $[0, \ell]$ and the displacement of x at time t is denoted by $u(x, t)$.

2.1.1 Equations of motion and boundary conditions

The equations of motion for the deflection u of the beam is given by

$$\rho A \partial_t^2 u = \partial_x V + Q, \tag{2.1.1}$$

$$\rho I \partial_t^2 \partial_x u = V + \partial_x M. \tag{2.1.2}$$

In these equations M denotes the moment, V the shear force and Q a transverse force density.

The constitutive equation is given by

$$M \quad = \quad EI\partial_x^2 u. \tag{2.1.3}$$

Combining (2.1.1), (2.1.2) and (2.1.3) yields the partial differential equation:

$$\rho A\partial_t^2 u = \partial_x(\rho I\partial_t^2\partial_x u) - \partial_x^2(EI\partial_x^2 u) + Q.$$

Example 1 *Viscous damping*

Here $Q = q - k\partial_t u$ for a given $k > 0$ and q is the load. Equation (2.1.1) takes the form

$$\rho A\partial_t^2 u + k\partial_t u = \partial_x V + q \ \text{ in } \ (0,\ell) \ \text{ for } \ t \geq 0.$$

Example 2 *Kelvin-Voigt damping*

The constitutive equation (2.1.3) is replaced by

$$M = EI\partial_x^2 u + \mu\partial_t\partial_x^2 u.$$

The transverse force Q in (2.1.1) is now the load q.

For Examples 1 and 2 we consider a cantilever beam, so the boundary conditions are given by

$$\begin{aligned} u(0,t) &= \partial_x u(0,t) = 0, \\ V(\ell,t) &= M(\ell,t) = 0. \end{aligned}$$

Example 3 *Boundary damping*

Suppose $k_1 > 0$ and $k_0 > 0$ are given. Consider the undamped problem where $Q = q$ and the constitutive equation is given by (2.1.3). We now have damping at the boundary $x = \ell$:

$$\begin{aligned} V(\ell,t) &= -k_0\partial_t u(\ell,t), \\ M(\ell,t) &= -k_1\partial_t\partial_x u(\ell,t). \end{aligned}$$

2.1.2 Variational form

To obtain the variational form, multiply the equations of motion (2.1.1) and (2.1.2) by arbitrary functions v and w and integrate. Using integration by parts we have

$$\begin{aligned} \int_0^\ell \rho A \partial_t^2 u\, v &= \int_0^\ell (\partial_x V + Q)\, v \\ &= -\int_0^\ell V\, v' + V(\ell)v(\ell) - V(0)v(0) + \int_0^\ell Q\, v \end{aligned} \tag{2.1.4}$$

and

$$\begin{aligned} \int_0^\ell \rho I \partial_t^2 \partial_x u\, w &= \int_0^\ell (V + \partial_x M)\, w \\ &= \int_0^\ell V\, w - \int_0^\ell M\, w' + M(\ell)w(\ell) - M(0)w(0). \end{aligned} \tag{2.1.5}$$

Now replace w by v' in equation (2.1.5) and add the resulting equation to equation (2.1.4). The result is

$$\begin{aligned} &\int_0^\ell \rho A \partial_t^2 u\, v + \int_0^\ell \rho I \partial_t^2 \partial_x u\, v' \\ &\quad = -\int_0^\ell M\, v'' + V(\ell)v(\ell) - V(0)v(0) + M(\ell)v'(\ell) - M(0)v'(0) + \int_0^\ell Q\, v. \end{aligned}$$

Test functions

Define the space of test functions by

$$\mathcal{T}[0,\ell] = \{v \in C^2[0,\ell] : v(0) = v'(0) = 0\}.$$

Problem EV

Using suitable notation, all three the examples may be written in a general variational form. The bilinear forms b and c for all three examples are given by

$$b(u,v) = EI \int_0^\ell u''\, v'' \quad \text{and} \quad c(u,v) = \rho A \int_0^\ell uv + \rho I \int_0^\ell u'v' \quad \text{for} \quad u,v \in \mathcal{T}[0,\ell].$$

The bilinear form a is different for each example.

Example 1: $a(u,v) = k \int_0^\ell u\, v$,

Example 2: $a(u,v) = \mu \int_0^\ell u''\, v''$, Example 3: $a(u, v) = k_1 u'(\ell)v'(\ell) + k_0 u(\ell)v(\ell)$.

The variational equation for each case is then

$$c(\partial_t^2 u(\cdot, t), v) + a(\partial_t u(\cdot, t), v) + b(u(\cdot, t), v) = (q(\cdot, t), v) \tag{2.1.6}$$

for all $v \in \mathcal{T}[0, \ell]$.

2.1.3 The weak variational form

It is important to note that the bilinear form c is defined for functions in $H^1(0, \ell)$ and b is defined for functions $u \in H^2(0, \ell)$ (see Appendix A). But the partial derivatives $\partial_t u$ and $\partial_t^2 u$ do not make sense when a function may be changed arbitrarily on a set of measure zero. If we define a function w by $w(t) = u(\cdot, t)$, then $w'(t)$ may be defined with respect to the norm of $\mathcal{L}^2(0, \ell)$, $H^1(0, \ell)$ or $H^2(0, \ell)$.

In general, let J be a bounded or unbounded interval of real numbers containing zero. J is either an open interval containing zero or it is of the form $[0, T)$ or $[0, \infty)$. Let Y be any Banach space and consider a function u on J with values in Y.

Definition 2.1.1. *Derivative*
Let t be any interior point of J. Suppose there exists a $v \in Y$ such that

$$\lim_{h \to 0} \left\| h^{-1}\Big(u(t + h) - u(t)\Big) - v \right\|_Y = 0,$$

then v is the derivative of u at t. We write $u'(t)$ for the derivative and $u'(t) \in Y$ to show that the derivative exists with respect to the norm of Y. The derivative (function) u' is defined in the usual way as $u'(t)$ for every $t \in J$, with u'' defined by $(u')'$.

Notation for $C^k\Big(J, Y\Big)$

$u \in C(J, Y)$ if u is continuous on J with respect to the norm of Y.

$u \in C^k(J, Y)$ if $u^{(k)} \in C^k(J, Y)$.

Equation (2.1.6) may now be rewritten:

$$c(w''(t), v) + a(w'(t), v) + b(w(t), v) = (q(\cdot, t), v) \tag{2.1.7}$$

for each test function v.

It is generally accepted that Equation (2.1.7) is a generalization of Equation (2.1.6) but it is not easy to prove. Consider the following quote in [Sho77]: "... the construction of a representative $u(\cdot, t)$ of $w(t)$ is non-trivial."

Theorem 2.1.2 ([Sho77], Theorem 7.A, p104)**.** *Let $I = [a, b]$, a closed interval in $\mathbb{R}$ and G be an open (or measurable) set in $\mathbb{R}^n$.*

(a) If $u \in C(I, \mathcal{L}^2(G))$, then there is a measurable function $w : G \times I \to \mathbb{R}$ such that

$$u(t) = w(\cdot, t), \quad t \in I. \tag{2.1.8}$$

(b) If $u \in C^1(I, \mathcal{L}^2(G))$, w and v are measurable real-valued functions on $G \times I$ for which (2.1.8) holds for a.e. $t \in I$ and

$$u'(t) = v(\cdot, t), \quad a.e.\, t \in I,$$

then $v = \partial_t u$ in $\mathcal{D}^*(G \times I)$.

Let $V(0, \ell)$ be the closure of the test functions in $H^2(0, \ell)$, then Equation (2.1.7) holds for all $v \in V(0, \ell)$. It is now possible to formulate the weak variational form of Problem EV. Let $f(t) = q(\cdot, t)$ for each t.

Problem EW

Find a function w with $w'(t) \in V(0, \ell)$ and $w''(t) \in H^1(0, \ell)$ such that

$$c(w''(t), v) + a(w'(t), v) + b(w(t), v) = (f(t), v)$$

for each $v \in V(0, \ell)$, while $w(0) = u_0$ and $w'(0) = u_1$.

Definition 2.1.3. A solution of Problem EW is a weak solution of Problem EV.

2.2 Existence of weak solutions

2.2.1 General linear hyperbolic problem

In this section we consider the general linear hyperbolic problem, also referred to as the general linear vibration problem. We consider the approach in [VV02], because the theorems are given in variational form. Alternatives are considered in Section 2.4.

For the general linear vibration problem let X, V and W be Hilbert spaces such that $V \subset W \subset X$ with inner products and norms in the table below.

Space	Inner product	Norm
X	$(\cdot,\cdot)_X$	$\\|\cdot\\|_X$
W	$c(\cdot,\cdot)$	$\\|\cdot\\|_W$
V	$b(\cdot,\cdot)$	$\\|\cdot\\|_V$

Also, let a be a bilinear form defined on V.

Problem G

Given a function $f : J \to W$, find a function $u \in C(J, V)$ such that u' is continuous at 0 and for each $t \in J$, $u(t) \in V$, $u'(t) \in V$, $u''(t) \in W$ and

$$c\left(u''(t), v\right) + a\left(u'(t), v\right) + b\left(u(t), v\right) = \left(f(t), v\right)_X \quad \text{for each } v \in V, \tag{2.2.1}$$

$$\text{while} \quad u(0) = u_0\,, \quad u'(0) = u_1\,.$$

2.2.2 Existence

Assumptions

The following assumptions are made for the existence results.

E1 V is dense in W and W is dense in X.

E2 There exists a positive constant κ_1 such that $\|v\|_W \leq \kappa_1 \|v\|_V$ for each $v \in V$.

E3 There exists a positive constant κ_2 such that $\|w\|_X \leq \kappa_2 \|w\|_W$ for each $w \in W$.

E4 The bilinear form a is non-negative, symmetric and bounded on V, i.e. there exists a positive constant C_a such that for $v, w \in V$, $|a(u, v)| \leq C_a \|u\|_V \|v\|_V$.

Remark Problem EW in the previous section is a special case of Problem G.

The following three theorems are from [VV02].

Theorem 2.2.1. *Suppose Assumptions **E1**, **E2**, **E3** and **E4** hold. If for $u_0 \in V$ and $u_1 \in W$, and there exists some $y \in W$ such that*

$$b(u_0, v) + a(u_1, v) = c(y, v) \quad \text{for each } v \in V, \tag{2.2.2}$$

then for each $f \in C^1([0,T), X)$ there exists a unique solution

$$u \in C([0,T), V) \cap C^1([0,T), W) \cap C^1((0,T), V) \cap C^2((0,T), W)$$

for Problem G. If $f = 0$ then $u \in C^1([0,\infty), V) \cap C^2([0,\infty), W)$.

Proof. See Section 2.3. □

Definition 2.2.2.

$$E_b = \{\, x \in V \;\big|\; \text{there exists a } y \in W \text{ such that } c(y,v) = b(x,v) \text{ for all } v \in V\}.$$

Assumption E4W *Weak damping*

The bilinear form a is non-negative, symmetric and bounded on W, i.e. there exists a positive constant C_W such that for $v, w \in W$, $|a(u,v)| \leq C_W \|u\|_W \|v\|_W$.

Theorem 2.2.3. *Weak damping*

*Suppose Assumptions **E1**, **E2**, **E3** and **E4W** hold. Let J be an interval containing zero, then there exists a unique solution*

$$u \in C^1(J, V) \cap C^2(J, W)$$

for Problem G for each $u_0 \in E_b$, $u_1 \in V$ and each $f \in C^1(J, X)$. If $f = 0$ then $u \in C^1((-\infty,\infty), V) \cap C^2((-\infty,\infty), W)$.

Proof. See Section 2.3. □

Remark We can still use Theorem 2.2.3 if $a = 0$.

Recall that a is called positive definite on V if there exists a $K > 0$ such that $a(u,u) \geq K\|u\|_V^2$ for any $u \in V$.

Definition 2.2.4. *Strong damping*

When a is positive definite on V the damping is referred to as strong damping.

Assumption E5S *Strong damping*

Assume we have strong damping, i.e. the bilinear form a is positive definite on V.

Theorem 2.2.5. *Strong damping*

*Suppose Assumptions **E1**, **E2**, **E3**, **E4** and **E5S** hold. Let $f : [0,T] \to W$ be locally Lipschitz. Then there exists a unique solution*

$$u \in C([0,T),V) \cap C^1([0,T),W) \cap C^2((0,T),W)$$

for Problem G, for any $u_0 \in V$, $u_1 \in W$. If $f = 0$ then

$$u \in C([0,\infty),V) \cap C^1([0,\infty),W) \cap C^\infty((0,\infty),V).$$

Proof. See Section 2.3. □

Examples

Consider the examples in Section 2.1. In all three examples the space X is $\mathcal{L}^2(0,\ell)$, the space W is the closure of the test functions in $H^1(0,\ell)$ and the space $V = V(0,\ell)$. The three types of damping are present in the examples. In Example 1 we have weak damping and in Example 2 strong damping. The damping in Example 3 is neither weak nor strong.

2.3 Application of semigroup theory

The theorems in the previous section are convenient to use since the assumptions are in terms of the bilinear forms a, b and c and it is not necessary to construct linear operators with suitable properties as in [Sho77], [Kut86] and [AKS96]. The approach in [VV02] is relatively new and therefore we discuss in this section how it is related to semigroup theory.

2.3.1 General damping

Problem G is equivalent to a first-order differential equation in the product space $H = V \times W$ with inner product

$$\left(x,y\right)_H = b(x_1,y_1) + c(x_2,y_2)$$

for all x and $y \in H$. For $x \in H$ denote x and its components by $x = \langle x_1, x_2 \rangle$, where $x_1 \in V$ and $x_2 \in W$. In order to write Problem G as a first-order differential equation, a linear operator $\mathcal{A}$ is constructed in [VV02]. The authors define $\mathcal{A}$ as the inverse of another operator, but provides a characterisation of $\mathcal{A}$ in the following Lemma.

Lemma 2.3.1. *Operator $\mathcal{A}$*
The domain is

$$\mathcal{D}(\mathcal{A}) = \Big\{ x \in H \quad : \quad x_2 \in V \text{ and there exists a } y \in W \text{ such that} \\ b(x_1, v) + a(x_2, v) = c(y, v) \text{ for all } v \in V \Big\}$$

and, $y = \mathcal{A}x$ if $y_1 = x_2$ and $b(x_1, v) + a(x_2, v) = -c(y_2, v)$ for each $v \in V$.

Properties of $\mathcal{A}$ (Proved in [VV02]):

- $\mathcal{A}$ is a densely defined closed linear operator on H. [VV02, Lemma 3]
- For any $\lambda \geq 0$, $\mathcal{R}(\lambda I - \mathcal{A}) = H$. [VV02, Corollary 1]
- $\mathcal{A}$ is dissipative. [VV02, Corollary 3]

Consider the following initial value problem.

Problem IVP-1

Given a function $g : [0, T) \to H$, find $U \in C([0, T), H)$ such that for each $t \in (0, T)$, $U(t) \in \mathcal{D}(A)$, $U'(t) \in H$ and

$$U'(t) = \mathcal{A}U(t) + g(t) \text{ for } t \in (0, T) \tag{2.3.1}$$
$$U(0) = U_0. \tag{2.3.2}$$

We use the same definition for a solution as Pazy [Paz83, p. 105] - see Definition 2.3.2 below.

Definition 2.3.2. A function U is said to be a solution of Problem IVP-1 above if it satisfies (2.3.1) and (2.3.2) and for each $t > 0$, $U(t) \in \mathcal{D}(\mathcal{A})$ and

$$U \in C([0, T), H) \cap C^1((0, T), H).$$

Proof of Theorem 2.2.1

We can now write Problem G as an initial value problem for a first order system. Let $g(t) = \langle 0, f(t)\rangle$. If

$$u \in C\left([0,T),V\right) \cap C^1\left([0,T),W\right)$$

is a solution of Problem G, then $U = \langle u, u'\rangle$ is a solution to Problem IVP with

$$U(0) = U_0 = \langle u_0, u_1\rangle.$$

Also if

$$u \in C\left([0,T),V\right) \cap C^1\left([0,T),W\right) \cap C^1\left((0,T),V\right) \cap C^2\left((0,T),W\right)$$

then

$$U \in C\Big([0,\tau);H\Big) \cap C^1\Big((0,\tau);H\Big).$$

Conversely, if U is a solution to Problem IVP-1 with $U_0 = \langle u_0, u_1\rangle$, then the first component $u = U_1$ of U is a solution of Problem G. If $U \in C^1([0,T),H)$, then

$$u \in C\left([0,T),V\right) \cap C^1\left([0,T),W\right) \cap C^1\left((0,T),V\right) \cap C^2\left((0,\tau),W\right).$$

These results are the result in [VV02, Lemma 7].

From the properties of the operator $\mathcal{A}$, it follows from [Paz83, Theorem 4.3, p.14] that $\mathcal{A}$ is the infinitesimal generator of a C_0 semigroup of contractions. Now, if $g \in C^1([0,T],H)$, then there exists a unique solution $U \in C^1([0,T),H)$ for any $U_0 \in \mathcal{D}(\mathcal{A})$ for Problem IVP-1 [Paz83, Corrolary 2.5, p.107]. If $f \in C^1([0,T],W)$, then $g \in C^1([0,T],H)$. Also, from the definition of $\mathcal{A}$, $U_0 \in \mathcal{D}(\mathcal{A})$ if and only if $u_0 \in V$ and $u_1 \in W$ and there exists some $y \in W$ such that

$$b\left(u_0, v\right) + a\left(u_1, v\right) = c\left(y, v\right) \quad \text{for each } v \in V.$$

It therefore follows from what has been done above that $u = U_1$ is a solution to Problem G with

$$u \in C\left([0,T),V\right) \cap C^1\left([0,T),W\right) \cap C^1\left((0,T),V\right) \cap C^2\left((0,T),W\right).$$

If $g = 0$, then $U \in C^1([0,\infty),H)$ which means that

$$u \in C^1\left([0,\infty),V\right) \cap C^2\left([0,\infty),W\right).$$

□

2.3.2 Weak damping

Let J be an open interval containing 0.

Problem IVP-2

Given a function $g : J \to H$, find $U \in C(J, H)$ such that for each $t \in J$, $U(t) \in \mathcal{D}(A)$, $U'(t) \in H$ and

$$U'(t) = \mathcal{A}U(t) + g(t) \text{ for } t \in J \text{ and } U(0) = U_0. \tag{2.3.3}$$

Suppose $g(t) = \langle 0, f(t)\rangle$ for each $t \in J$. If $u \in C(J, V) \cap C^1(J, W)$ is a solution of Problem G, then $U = \langle u, u'\rangle$ is a solution of Problem IVP-2 with $U_0 = \langle u_0, u_1\rangle$.

If U is a solution of Problem IVP-2 with $U_0 = \langle u_0, u_1\rangle$, then the first component $u = U_1$ of U is a solution of Problem G. If $U \in C^1\big(J, H\big)$ then $u \in C^1\big(J, V\big) \cap C^2\big(J, W\big)$.

If Assumption E4W holds then the operator $\mathcal{A}$ is the infinitesimal generator of a C_0 group in H [VV02, Theorem 4]. This follows from [Paz83, Theorem 6.3, p. 23]. Also note that in this case we have that $\mathcal{D}(\mathcal{A}) = E_b \times V$ (see [VV02, Lemma 8]).

Theorem 2.3.3. *Let $\mathcal{A}$ be the infinitesimal generator of a C_0 group and $g \in C^1(\bar{J}, H)$. Then Problem IVP-2 has a unique solution $U \in C^1(J, H)$ for each $U_0 \in \mathcal{D}(\mathcal{A})$. If $g = 0$, then $U \in C^1((-\infty, \infty), H)$.*

Proof. Suppose that $\bar{J} = [a, b]$ and define a function G on $[0, b - a]$ by $G(t) = g(t + a)$. Since $g \in C^1([a, b], H)$, $G \in C^1([0, b - a], H)$. By [Paz83, Corollary 2.5, p 107 and Theorem 1.3, p 102] there exists a function $y \in C^1((0, b - a), H)$ such that

$$y' = \mathcal{A}y + G \text{ on } (0, b - a).$$

Define a function w on J by $w(t) = y(t - a)$, then $w \in C^1(J, H)$ and $w' = \mathcal{A}w + g$.

Since U_0 and $w(0)$ are in $\mathcal{D}(\mathcal{A})$, the function $T(\cdot)(b - w(0))$ is a solution of the homogeneous differential equation. Consequently $U = T(\cdot)(b - w(0)) + w$ is a solution of the nonhomogeneous differential equation 2.3.1 and since $U(0) = U_0$, U is the solution of Problem IVP-2. □

Proof of Theorem 2.2.3

If $u_0 \in E_b$ and $u_1 \in V$, $\mathcal{D}(\mathcal{A}) = E_b \times V$ yields that $\langle u_0, u_1 \rangle \in D(A)$. For $g(t) = \langle 0, f(t) \rangle$, $f \in C^1(J, W)$ implies that $g \in C^1(J, H)$. By setting $U_0 = \langle u_0, u_1 \rangle$, Theorem 2.3.3 implies that Problem IVP-2 has a unique solution.

It follows that the first component $u = U_1$ of this solution is a solution of Problem G. If $f = 0$, then $g = 0$ and it follows from Theorem 2.3.3 that

$$u \in C^1\Big((-\infty, \infty), V\Big) \cap C^2\Big((-\infty, \infty), W\Big).$$

□

2.3.3 Strong damping

Consider Problem IVP-1. If Assumptions E1, E2 and E3 hold, we know that the operator $\mathcal{A}$ is the infinitesimal generator of a C_0 semigroup and Problem IVP-1 has a unique solution for $U_0 \in \mathcal{D}(\mathcal{A})$ and $g \in C^1([0, T), W)$. In this section we show that both these conditions may be relaxed if a is positive definite on V.

Recall Assumption E5S: The bilinear form a is positive definite on V, i.e. there exists a constant C_s such that for every $u \in V$,

$$a(u, u) \geq C_s \|u\|_V.$$

A real Hilbert space H may be imbedded in a complex Hilbert space

$$\widetilde{H} = \{x + iy \;\big|\; x \in H, y \in H\}.$$

Note that if $\alpha + i\,\beta$ is a complex number, then

$$(\alpha + i\beta)(x + iy) = \alpha x - \beta y + i\beta x + i\alpha y.$$

This construction can be made rigorous, see [Sch71, p. 153].

The inner product for $\widetilde{H}$ is defined as

$$\langle x + iy, u + iv \rangle_{\widetilde{H}} = \langle x, u \rangle_H + \langle y, v \rangle_H + i\langle y, u \rangle_H - i\langle x, v \rangle_H.$$

The linear operator $\mathcal{A}$ may be extended to the space $\widetilde{H}$:

$$\begin{aligned} D(\widetilde{\mathcal{A}}) &= \{x+iy \mid x \in \mathcal{D}(\mathcal{A}), y \in \mathcal{D}(\mathcal{A})\}, \\ \widetilde{\mathcal{A}}(x+iy) &= \mathcal{A}x + i\mathcal{A}y. \end{aligned}$$

Problem IVP-3

Given a function $g : [0,T) \to \widetilde{H}$, find $U \in C([0,T), \widetilde{H})$ such that for each $t \in (0,T)$, $U(t) \in \mathcal{D}(\widetilde{\mathcal{A}})$, $U'(t) \in \widetilde{H}$ and

$$U'(t) = \widetilde{\mathcal{A}}U(t) + g(t) \text{ for } t \in (0,T) \tag{2.3.4}$$

$$U(0) = U_0. \tag{2.3.5}$$

The following results can be proved in [VV02].

- [VV02, Lemma 10] For any $w = x + iy \in D(\widetilde{\mathcal{A}})$

$$\begin{aligned} Re(\widetilde{\mathcal{A}}w, w)_{\widetilde{H}} &= (\mathcal{A}x, x)_H + (\mathcal{A}y, y)_H, \\ Im(\widetilde{\mathcal{A}}w, w)_{\widetilde{H}} &= (\mathcal{A}y, x)_H - (\mathcal{A}x, y)_H. \end{aligned}$$

- There exists a constant C_s such that for any $w = x + iy \in D(\widetilde{\mathcal{A}})$ we have

$$Re(\widetilde{\mathcal{A}}w, w)_{\widetilde{H}} \leq -C_s(\|x_2\|_V^2 + \|y_2\|_V^2).$$

- For any $w = x + iy \in D(\widetilde{\mathcal{A}})$ we have

$$|Im(\widetilde{\mathcal{A}}w, w)|_{\widetilde{H}} \leq \|x_1\|_V^2 + \|x_2\|_V^2 + \|y_1\|_V^2 + \|y_2\|_V^2.$$

- [VV02, Lemma 11] There exists a constant K such that

$$K \; Re((\widetilde{\mathcal{A}} - I)w, w)_{\widetilde{H}} + |Im((\widetilde{\mathcal{A}} - I)w, w)|_{\widetilde{H}} \leq 0 \;\text{ for any } w \in D(\widetilde{\mathcal{A}}).$$

Let $v(t) = e^{-t}U(t)$. Then v is a solution of

$$\begin{cases} v'(t) = \left(\widetilde{\mathcal{A}} - \widetilde{I}\right) v(t) + e^{-t}g(t) \quad \text{for} \quad t > 0, \\ v(0) = U_0, \end{cases}$$

if and only if U is a solution of Problem IVP-3.

Now consider [VV02, Theorem 5]: A linear operator B is the infinitesimal generator of an analytic semigroup in a complex Hilbert space X if $D(B)$ is dense in X, $0 \in \rho(B)$ and there exists a constant $C > 0$ such that

$$C\ Re(Bx, x)_X + |Im(Bx, x)|_X \leq 0 \text{ for any } x \in D(B).$$

It then follows from the properties above ([VV02, Lemma 11]) and [VV02, Theorem 5] that the operator $\widetilde{\mathcal{A}} - \widetilde{I}$ is the infinitesimal generator of an analytic semigroup, and consequently the operator $\widetilde{\mathcal{A}}$ is the infinitesimal generator of an analytic semigroup on $\widetilde{H}$.

Theorem 2.3.4. *If $g : [0, T) \to H$ is locally Lipschitz on $(0, T)$, then Problem IVP has a unique solution U for each $U_0 \in H$. If $g = 0$, then $U \in C^\infty((0, \infty), H)$.*

Proof. The initial value problem $Y' = \widetilde{\mathcal{A}}Y + g$ with $Y(0) = U_0$ has a unique solution and $Y \in C^\infty\left((0, \infty), \widetilde{H}\right)$ if $g = 0$ by [Paz83, Corollary 3.3, p113]. But $Y(t) = U(t) + iW(t)$ with $U(t) \in H$. Since $g(t) \in H$ for each $t \in [0, T)$ we have $U' = AU + g$. Finally $Y(0) = U_0 \in H$ and it follows that $U(0) = U_0$. □

Remark The result in [Paz83] is more general. He only requires that g is locally Hölder continuous (definition on page 112), with exponent α, where $0 < \alpha < 1$.

We can now complete the proof of Theorem 2.2.5.

Proof of Theorem 2.2.5

If $u_0 \in V$ and $u_1 \in W$, then $\langle u_0, u_1 \rangle \in H$ and g is locally Lipschitz with respect to the norm $\|\cdot\|_H$ if f is locally Lipschitz with respect to $\|\cdot\|_W$. Therefore Problem IVP has a unique solution U. It then follows by previous arguments that the first component $u = U_1$ of U is a solution of Problem G, and

$$u \in C([0, T), V) \cap C^1([0, T), W) \cap C^2((0, T), W).$$

If $f = 0$, then $g = 0$ and so

$$u \in C([0, \infty), V) \cap C^1([0, \infty), W) \cap C^\infty((0, \infty), V).$$

□

2.4 Alternatives

2.4.1 Weak solution according to Evans

In this section we present the results in [Eva98]. The notation and formulation of results are as far as possible exactly the same as in the book.

Let L be a strongly elliptic second order differential operator (see Section 1.2). Consider the initial boundary-value problem

$$(1) \qquad \begin{cases} \partial_t^2 w + Lw = f & \text{in } \Omega \text{ for } t > 0 \\ w = 0 & \text{on } \partial\Omega \text{ for } t > 0 \\ w(x,0) = g, \partial_t w(x,0) = h & \text{in } \Omega. \end{cases}$$

Consider the definition of a weak solution by Evans [Eva98, Subsection 7.2.1].

Introduce the function $\widetilde{f} : [0,T] \to \mathcal{L}^2(\Omega)$ by $[\widetilde{f}(t)](x) := f(x,t)$ for $x \in \Omega, 0 \leq t \leq T$, and let $\langle \cdot, \cdot \rangle$ denoting the pairing between $H^{-1}(\Omega)$ and $H_0^1(\Omega)$.

Definition 2.4.1. We say a function

$$u \in \mathcal{L}^2((0,T), H_0^1(\Omega)), u' \in \mathcal{L}^2((0,T), \mathcal{L}^2(\Omega)), u'' \in \mathcal{L}^2((0,T), H^{-1}(\Omega)),$$

is a weak solution of the hyperbolic initial/boundary-value problem (1) provided

1. $\langle u'', v \rangle + b(u,v) = (\widetilde{f}, v)$ for each $v \in H_0^1(\Omega)$ and a.e. $0 \leq t \leq T$, and
2. $u(0) = g$, $u'(0) = h$.

Remarks

1. $u''(t)$ is a distribution and does not belong to a function space.
2. The author only considers homogeneous boundary conditions; $u(t) \in H_0^1(\Omega)$.
3. The author only considers a second order hyperbolic equation which is a generalisation of the multi-dimensional wave equation (without damping). The theory can therefore not be applied to the vibration of plates (see Section 6.3).

In Subsection 7.2.3 Evans proves the existence of a unique weak solution.

Theorem 2.4.2. *[Eva98, Theorem 3, p.384]*
There exists a weak solution of (1).

The method of the proof is based on the proof in [LM72].

A weak solution of the initial boundary-value problem (1) is obtained by first constructing a finite dimensional approximation. Galerkin's method is employed by selecting smooth functions $w_k = w_k(x)(k \geq 1)$ such that

$$\{w_k\}_{k=1}^{\infty} \text{ is an orthogonal basis of } H_0^1(\Omega)$$

and

$$\{w_k\}_{k=1}^{\infty} \text{ is an orthonormal basis of } \mathcal{L}^2(\Omega).$$

For an integer m, let

$$\bar{u}_m(t) := \sum_{k=1}^{m} d_m^k(t) w_k.$$

In [Eva98, Theorem 1, p. 380] it is proved that for each integer $m \geq 1$, $\bar{u}_m$ given by the above is unique. Estimates for $\bar{u}_m$, $\bar{u}_m'$ and $\bar{u}_m''$ are then derived in [Eva98, Theorem 2, p. 381] in terms of $\widetilde{f}$, g and h, in order to eventually let $m \to \infty$. From this, there exists a subsequence (see [Eva98, Theorem 3, p. 384]) $\{\bar{u}_{ml}\}_{l=1}^{\infty} \subset \{\bar{u}_m\}_{m=1}^{\infty}$ and $\bar{u} \in \mathcal{L}^2((0,T), H_0^1(\Omega))$, with $\bar{u}' \in \mathcal{L}^2((0,T), \mathcal{L}^2(\Omega))$, $\bar{u}'' \in \mathcal{L}^2((0,T), H^{-1}(\Omega))$ such that

$$\bar{u}_{ml} \to \bar{u} \text{ weakly in } \mathcal{L}^2((0,T), H_0^1(\Omega)).$$

The limit $\bar{u}$ is a weak solution (according to [Eva98]).

Theorem 2.4.3. *[Eva98, Theorem 4, p.385]*
A weak solution of (1) is unique.

Other publications

Existence results may also be found in other publications e.g. [LM72], [Sho77], [Kut86] and [AKS96]. Their definitions of a weak solution are similar to [LM72], and what will be referred to as a mild solution in the book of [Paz83]. Either way, the existence of a weak solution or a mild solution is not sufficient for the theory of the finite element method as we see in the following chapters.

Remark The model problems in [LM72] does not include damping.

2.4.2 Regularity according to Evans

In Subsection 7.2.3 Evans prove that improved regularity of the weak solution is possible under certain conditions.

Theorem 2.4.4. *[Eva98, Theorem 5, p. 389] Improved regularity*

1. *Assume*
$$g \in H_0^1(\Omega), h \in \mathcal{L}^2(\Omega), f \in \mathcal{L}^2((0,T), \mathcal{L}^2(\Omega)),$$
and suppose also that $u \in \mathcal{L}^2((0,T), H_0^1(\Omega))$*, with* $u' \in \mathcal{L}^2((0,T), \mathcal{L}^2(\Omega))$*,* $u'' \in \mathcal{L}^2((0,T), H^{-1}(\Omega))$*, is the weak solution of the of the problem (1). Then in fact*
$$u \in \mathcal{L}^\infty((0,T), H_0^1(\Omega)), u' \in \mathcal{L}^\infty((0,T), \mathcal{L}^2(\Omega))$$
and we have the estimate
$$\begin{aligned} &\operatorname*{ess\,sup}_{0\le t\le T} \left(\|u(t)\|_{H_0^1(\Omega)} + \|u'(t)\|_{\mathcal{L}^2(\Omega)} \right) \\ &\quad \le C \left(\|f\|_{\mathcal{L}^2((0,T),\mathcal{L}^2(\Omega))} + \|g\|_{H_0^1(\Omega))} + \|h\|_{\mathcal{L}^2(\Omega)} \right). \end{aligned}$$

2. *If, in addition,*
$$g \in H^2(\Omega), h \in H_0^1(\Omega), f' \in \mathcal{L}^2((0,T), \mathcal{L}^2(\Omega)),$$
then
$$\begin{aligned} u &\in \mathcal{L}^\infty((0,T), H^2(\Omega)), & u' &\in \mathcal{L}^\infty((0,T), H_0^1(\Omega)) \\ u'' &\in \mathcal{L}^\infty((0,T), \mathcal{L}^2(\Omega)), & u''' &\in \mathcal{L}^2((0,T), H^{-1}(\Omega)), \end{aligned}$$
with the estimate
$$\begin{aligned} &\operatorname*{ess\,sup}_{0\le t\le T} \left(\|u(t)\|_{H^2(\Omega)} + \|u'(t)\|_{H_0^1(\Omega)} + \|u''(t)\|_{\mathcal{L}^2(\Omega)} \right) + \|u'''(t)\|_{\mathcal{L}^2((0,T),H^{-1}(\Omega))} \\ &\quad \le C \left(\|f\|_{H^1((0,T),\mathcal{L}^2(\Omega))} + \|g\|_{H^2(\Omega))} + \|h\|_{H^1(\Omega)} \right). \end{aligned}$$

Note that now $u''(t) \in \mathcal{L}^2(\Omega)$, i.e. it is actually a function. This result is similar to the result in Theorem 2.2.3 (from Subsection 2.2.2).

2.4.3 Higher regularity

In Subsection 7.2.3 Evans introduce so-called compatibility conditions for the initial values of the weak solution and its derivative which results in the existence of higher order derivatives of the solution.

Theorem 2.4.5. *[Eva98, Theorem 6, p. 391] Higher regularity*
Assume

$$g \in H^{1+m}(\Omega), \quad h \in H^m(\Omega),$$
$$\frac{d^k f}{dt^k} \in \mathcal{L}^2((0,T), H^{m-k}(\Omega)), k = 0, \dots, m.$$

Suppose also the following $m^{th}-$*order compatibility conditions hold:*

$$\begin{aligned} g_0 &:= g \in H_0^1(\Omega), h_1 := h \in H_0^1(\Omega), \dots, \\ g_{2\ell} &:= \frac{d^{2\ell-2} f}{dt^{2\ell-2}}(\cdot,0) - Lg_{2\ell-2} \in H_0^1(\Omega) \quad (\textit{if} m = 2\ell) \\ h_{2\ell} &:= \frac{d^{2\ell-1} f}{dt^{2\ell-1}}(\cdot,0) - Lh_{2\ell-1} \in H_0^1(\Omega) \quad (\textit{if} m = 2\ell+1). \end{aligned}$$

Then

$$\frac{d^k u}{dt^k} \in \mathcal{L}^\infty((0,T), H^{m+1-k}(\Omega)) \quad (k = 0,1,\dots,m+1), \tag{2.4.1}$$

and we have the estimate

$$\begin{aligned} &\operatorname*{ess\,sup}_{0\le t\le T} \sum_{k=0}^{m+1} \left|\left| \frac{d^k u}{dt^k} \right|\right|_{H^{m+1-k}(\Omega)} \\ &\qquad \le C \left(\sum_{k=0}^{m} \left|\left| \frac{d^k f}{dt^k} \right|\right|_{\mathcal{L}^2((0,T),H^{m-k}(\Omega))} + \|g\|_{H^{m+1}(\Omega))} + \|h\|_{H^m(\Omega)} \right). \end{aligned}$$

To determine whether the compatibility conditions hold in an application is problematic.

Recall that it is stated in Chapter 1 that in the article of [Kar11a] a result from the book of [LM72] is cited for the existence of a weak solution, but in proving convergence, assumes more differentiability properties for the solution. In particular, the existence of a weak solution is in the sense of Theorem 2.4.2 or 2.4.4, but for the error estimates it is assumed that

$$u \in C^2(\bar{J}, H^{p+1}(\Omega)), \quad u''' \in C(\bar{J}, \mathcal{L}^2(\Omega)), \quad u^{(4)} \in L^1(\bar{J}, \mathcal{L}^2(\Omega)).$$

These are rather typical assumptions, and can be interpreted in two ways: in the way given in Section 2.1.3 or that $u''(t)$ exists with respect to some other norm, e.g the $\mathcal{L}^2$-norm, but that $u''(t) \in H^{p+1}(\Omega)$. Either way, these are very serious restrictions (see for instance Theorem 2.4.5).

For more on this topic, see Section 6.2. Wu, author of a number of articles [Wu03, Wu04, Wu05, Wu06] in elasto-dynamics, is more careful than other authors when it comes to existence and regularity.

2.5 The multidimensional wave equation with weak damping

2.5.1 Weak variational form

Let $V(\Omega)$ denote the closure of $\mathcal{T}(\Omega)$ in $H^1(\Omega)$. The bilinear form b is defined as in Subsection 1.2.1 (and for heat conduction in Section 1.3.3):

$$b(u,v) \quad = \quad \iiint_\Omega (A\nabla u)\cdot\nabla v \; dV.$$

Since $c_1 \leq \rho \leq c_2$, the bilinear form c is clearly an inner product for $\mathcal{L}^2(\Omega)$ and the corresponding norm is equivalent to the norm of $\mathcal{L}^2(\Omega)$. In this application, W is the space $\mathcal{L}^2(\Omega)$ with inner product c (and norm $\|\cdot\|_W$). Also, $X = \mathcal{L}^2(\Omega)$ and $V = V(\Omega)$.

Let J be an open interval containing zero. Let $\widetilde{f} : t \to f(\cdot, t)$.

Problem MWW

Given $\widetilde{f} \in C(J, \mathcal{L}^2(\Omega))$, find u such that for each $t \in J$, $u(t) \in V(\Omega)$, $u''(t) \in \mathcal{L}^2(\Omega)$ and

$$c(u''(t), v) + a(u'(t), v) + b(u(t), v) = (\widetilde{f}(t), v)_{\mathcal{L}^2(\Omega)} \text{ for each } v \in V(\Omega),$$

while $u(0) = u_0$, and $u'(0) = u_1$.

2.5.2 Existence

Poincaré-Friedrichs inequality

There exists a constant β such that $\|v\|^2 \leq \beta\, b(v,v)$ for each $v \in V(\Omega)$. If $\partial\Omega - \Sigma$ has positive area (positive length for $n = 2$), then the assumption holds, see [Bra01, p.30] or [Eva98, Theorem 2, p.300].

[Bak76] states this without proof or reference. It is not even mentioned in [Dup73].

If the Poincaré-Friedrichs inequality is true, then b is positive definite on $V(\Omega)$ with respect to the norm of $H^1(\Omega)$ and the bilinear form b is an inner product on $V(\Omega)$. (The bilinear form b is symmetric by definition.)

Definition 2.5.1. The norm corresponding to the inner product b is

$$\|v\|_V = \sqrt{b(v,v)} \text{ for any } v \in V(\Omega).$$

The bilinear form b is clearly bounded on $H^1(\Omega)$, hence the norm $\|\cdot\|_V$ is equivalent to the norm of $H^1(\Omega)$ on $V(\Omega)$.

Proposition 2.5.2. *$V(\Omega)$ is dense in $\mathcal{L}^2(\Omega)$.*

Proof. From Appendix A we have that $C_0^\infty(\Omega)$ is dense in $\mathcal{L}^2(\Omega)$. Since $C_0^\infty(\Omega) \subset V(\Omega)$, $V(\Omega)$ is also dense in $\mathcal{L}^2(\Omega)$. □

We have shown that Assumptions E1, E2 and E3 are satisfied. For existence we also need Assumption E4W (weak damping).

Proposition 2.5.3. *The bilinear form a is nonnegative, symmetric and bounded on W, i.e.*

$$|a(u,v)| \leq C_W \|u\|_W \|v\|_W.$$

Proof. Since it was assumed that $0 \leq k \leq c_3$, we have that

$$|a(u,v)| \leq c_3 \iiint_\Omega |uv| \leq c_3\|u\|\|v\|,$$

and the $\mathcal{L}^2$ norm is equivalent to the norm $\|\cdot\|_W$ □

Existence of a unique solution for the weak variational problem depends on the smoothness of the function $\widetilde{f}$.

Recall the definition of E_b in Chapter 2:

$$E_b = \{x \in V \big| \text{ there exists a } y \in W \text{ such that } c(y,v) = b(x,v) \text{ for all } v \in V\}.$$

Theorem 2.5.4. *Suppose $\widetilde{f} \in C^1(J, \mathcal{L}^2(\Omega))$. Problem MWW has a unique solution*

$$u \in C^1(J, V(\Omega)) \cap C^2(J, W)$$

if $u_0 \in E_b$ and $u_1 \in V(\Omega)$.

Proof. Assumptions E1, E2, E3 and E4W in Section 2.2.2 hold. The result follows from Theorem 2.2.3. □

2.5.3 Sufficient conditions for existence

In practice useful sufficient conditions for existence are required. We show that $u_0 \in E_b$ if $u_0 \in H^2(\Omega) \cap V(\Omega)$.

First suppose that $u_0 \in C^2(\bar{\Omega}) \cap \mathcal{T}(\bar{\Omega})$. From the definition of the bilinear form b and using Green's formula (Proposition 1.2.2) we have

$$\begin{aligned} b(u_0, v) &= \iiint_\Omega (A\nabla u_0) \cdot \nabla v \, dV \\ &= \iint_{\partial\Omega} v(A\nabla u_0) \cdot \mathbf{n} \, dA - \iiint_\Omega (\nabla \cdot (A\nabla u_0)) v \, dV \\ &= -\iiint_\Omega (\nabla \cdot (A\nabla u_0)) v \, dV \\ &= c(-\rho^{-1}\nabla \cdot (A\nabla u_0), v). \end{aligned}$$

Since $u_0 \in C^2(\bar{\Omega})$, it follows that $-\rho^{-1}\nabla \cdot (A\nabla u_0) \in C(\bar{\Omega})$.

Now suppose $u_0 \in H^2(\Omega) \cap V(\Omega)$. Then there exists a sequence $\{U_n\}_{n\geq 1} \subset C^2(\bar{\Omega}) \cap \mathcal{T}(\bar{\Omega})$ such that this sequence converges to u_0 with respect to $\|\cdot\|_2$, and so also with respect to $\|\cdot\|_V$. But

$$\|U_n - u_0\|_2^2 = \|U_n - u_0\|_0^2 + |U_n - u_0|_1^2 + |U_n - u_0|_2^2,$$

and so the sequence converges to u_0 with respect to the semi-norm $|\cdot|_2$ as well (The semi-norm $|\cdot|_j$ is defined in Appendix A). This means that all the second order derivatives will also converge in the $\mathcal{L}^2$-norm, and so the sequence $\{-\rho^{-1}\nabla\cdot(A\nabla U_n)\}_{n\geq 1}$ converges to $-\rho^{-1}\nabla\cdot(A\nabla u_0)$ in the $\mathcal{L}^2$-norm. Since the $\mathcal{L}^2$-norm is equivalent to $\|\cdot\|_W$, it follows that there exists a $y \in W$ such that

$$b(u_0, v) = c(y, v) \text{ for all } v \in V,$$

where $y = -\rho^{-1}\nabla\cdot(A\nabla u_0)$. Consequently $u_0 \in E_b$.

2.6 The Dual-Phase-Lag model

2.6.1 Weak variational form

Let $V = V(\Omega)$ denote the closure of $\mathcal{T}(\Omega)$ in $H^1(\Omega)$. The bilinear forms a and b are defined as in Subsection 1.3.6:

$$b(u,v) = \iiint_\Omega A\nabla u\cdot\nabla v\,dV, \text{ and } a(u,v) = \iiint_\Omega \gamma_1 uv + Q\nabla u\cdot\nabla v\,dV$$

where $\partial_i u$ and $\partial_i v$ denote weak derivatives.

Let $u(t) = T(\cdot, t)$.

Problem DPLW

Find u such that for each $t > 0$, $u'(t) \in V(\Omega)$, $u''(t) \in \mathcal{L}^2(\Omega)$ and

$$c(u''(t), v) + a(u'(t), v) + b(u(t), v) = (\widetilde{f}(\cdot, t), v)_\Omega \text{ for each } v \in V(\Omega)$$

while $u(0) = T_0$, and $u'(0) = T_1$.

2.6.2 Existence of a solution

It is assumed that $c_1 \leq \rho \leq c_2$, and therefore $C_* = \tau_q c_1 c_p \leq \gamma_2 \leq \tau_q c_2 c_p = C^*$. From this it follows that the bilinear form c is an inner product for $X = \mathcal{L}^2(\Omega)$ and the corresponding norm is equivalent to the norm of $\mathcal{L}^2(\Omega)$. The space $\mathcal{L}^2(\Omega)$ with inner product c is the space W (with norm $\|\cdot\|_W$).

As in Section 2.5.2 we see that, assuming that the Poincaré-Friedrichs inequality is true, b is positive definite on $V(\Omega)$ with respect to the norm of $H^1(\Omega)$ and the bilinear form b is an inner product on $V(\Omega)$. (The bilinear form b is symmetric by definition.) Recall that the norm corresponding to the inner product b is

$$\|v\|_V = \sqrt{b(v,v)} \text{ for any } v \in V(\Omega).$$

All the properties of the bilinear form b and the space V mentioned in Section 2.5.2 holds here as well. However the bilinear form a does not satisfy Assumption E4W, but rather Assumption E4.

Proposition 2.6.1. *The bilinear form a is nonnegative, symmetric and bounded on V, i.e.*

$$|a(u,v)| \le C_a \|u\|_V \|v\|_V \text{ for any } u,v \in V.$$

Proof. From the definition of a, the fact that $Q = \tau_T A$ and the definition of b we have that

$$\begin{aligned} |a(u,v)| &\le c_2 c_p \|u\| \|v\| + \tau_T \sqrt{b(u,u)}\sqrt{b(v,v)} \\ &\le c_2 c_p \kappa_1^2 \|u\|_V \|v\|_V + \tau_T \|u\|_V \|v\|_V \\ &= C_a \|u\|_V \|v\|_V. \end{aligned}$$

□

Proposition 2.6.2. *The bilinear form a is positive definite on V, i.e.*

$$a(u,u) \ge K\|u\|_V^2 \text{ for any } u \in V.$$

Proof. From the definition of a, the fact that $Q = \tau_T A$ and the definition of b we have that

$$a(u,u) \ge c_1 c_p \|u\|^2 + \tau_T b(u,u) \ge \tau_T \|u\|_V^2 \text{ for any } u \in V.$$

□

Assumption E5S is therefore satisfied and consequently all the assumptions for Theorem 2.2.5 are now satisfied. Existence of a unique solution for the weak variational problem depends on the smoothness of the function $f\colon t \to \tilde{f}(\cdot,t)$.

Theorem 2.6.3. *Suppose $f : [0,T] \to W$ is locally Lipschitz. Then there exists a unique solution*

$$u \in C\left([0,T),V\right) \cap C^1\left([0,T),W\right) \cap C^2\left((0,T),W\right)$$

for Problem DPLW, for any $T_0 \in V$, $T_1 \in W$. If $\widetilde{f} = 0$ then

$$u \in C\left([0,\infty),V\right) \cap C^1\left([0,\infty),W\right) \cap C^\infty\left((0,\infty),V\right).$$

Proof. Assumptions E1, E2, E3, E4 and E5S in Chapter 2, Section 2.2.2 hold. The result follows from Theorem 2.2.5. □

Chapter 3

Error estimates for weak damping

In this chapter we consider the article [BV13]. The article is a generalization of the work done in [Bak76], but also includes weak damping. The results in the article can therefore be applied to any problem of the form given in the general linear vibration problem, Problem G, if the bilinear form a is bounded with respect to the norm of the space W (see Chapter 2). In Chapter 6 we apply the results from [BV13] to the relevant model problems given in Chapter 1, namely the multidimensional wave equation, the hyperbolic heat conduction equation and the Reissner-Mindlin plate model. We also apply the theory to linear elasto-dynamics in Section 6.3 of Chapter 6.

The proofs of the results in the article [BV13] are given in great detail in the article itself, however, some main ideas of the structure of the proofs are provided here for completeness and for comparison with other publications. In particular, the improvements made in [BV13] on [Bak76] are highlighted. We also discuss the significance of the way that the authors split the proofs for the semi-discrete and fully discrete cases and then use both the results to get a final error estimate, unlike other articles where the semi-discrete estimate is not used to derive the error estimate for the fully discrete case.

3.1 Galerkin Approximation

Choose a set of basis functions $\{\phi_1, \phi_2, \ldots, \phi_n\}$ in V. Denote the span of these functions by S^h. We can now formulate the semi-discrete problem for our general linear vibration problem, Problem G. The Galerkin finite element approximation of Problem G is referred to as Problem G^h.

Problem G^h

Given a function $f : [0,T] \to X$, find a function $u_h \in C^2(0,T)$ such that for each $t \in (0,T)$

$$c\left(u_h''(t), v\right) + a\left(u_h'(t), v\right) + b\left(u_h(t), v\right) = \left(f(t), v\right)_X \quad \text{for each } v \in S^h \tag{3.1.1}$$

$$\text{while} \quad u_h(0) = u_0^h\,, \quad u_h'(0) = u_1^h\,.$$

The initial values u_0^h and u_1^h are elements of S^h as close as possible to u_0 and u_1. We discuss the choice of initial values in Section 3.1.2.

3.1.1 Fundamental estimate

The concern here is the difference between the solution u and the Galerkin approximation u_h, i.e. to obtain an estimate for the error $\|u(t) - u_h(t)\|$ for each $t \in [0,T]$. To do this, the projection method is used.

Definition 3.1.1. *Projection operator*
The projection operator P_h is defined by $b(u - P_h u, v) = 0$ for all $v \in S^h$.

If no confusion is possible, we write P for P_h. The projection is used to split the error $e_h(t) = u(t) - u_h(t)$ as follows:

$$e(t) = Pu(t) - u_h(t)\,, \quad e_p(t) = u(t) - Pu(t).$$

Then $e_h(t) = e_p(t) + e(t)$ and $\|u(t) - u_h(t)\|_W \le \|e_p(t)\|_W + \|e(t)\|_W$.

Estimates for the norm of e_p can be obtained from approximation theory. See Appendix B for details on interpolation and see Subsection 3.1.2 for details on the estimate for the norm of e_p. It remains to find an estimate for $e(t)$, the difference between the projection of u and the Galerkin approximation u_h.

The projection P defines a function Pu by $(Pu)(t) = Pu(t)$ for all $t \in [0,T]$.

Lemma 3.1.2. *[BV13, Lemma 3.1]*
If $u \in C^1([0,T], V)$, then $Pu \in C^1[0,T]$ and $(Pu)'(t) = Pu'(t)$.

The proof of this lemma can be found in detail in [BV13]. It relies on the fact that P is a bounded linear operator with norm less than one and that $u \in C^1([0,T],V)$.

Remark It is important to note that in the article [BV13] an assumption is made that the solution u of Problem G has the property that $(Pu) \in C^2([0,T])$. Upon investigation of all the proofs, it was seen that this assumption is not necessary, which was an oversight in [BV13]. It is also worth noting that [Bak76] derived error estimates for the undamped case and **did** use this assumption without mentioning it.

Proposition 3.1.3. *[BV13, Proposition 3.1]*
If u is a solution of Problem G and u_h a solution to Problem G^h, then

$$c\left(e_h''(t),v\right)+a\left(e_h'(t),v\right)+b\left(e(t),v\right)=0 \text{ for each } v\in S^h. \tag{3.1.2}$$

Once again the proof is given in detail in [BV13].

In the proof of the main theorem in [Bak76] (Theorem 3.1, p. 567), he obtains an equation (3.4), which is given here in the notation of this dissertation for comparison:

$$\text{[Bak76, Equation (3.4)]} \qquad c(e''(t),v)+b(e(t),v)=-c(e_p''(t),v). \tag{3.1.3}$$

The proof is almost the same as for Equation (3.1.2). This equation contains the term e_p'', and hence implicitly assumes that $Pu \in C^2[0,T]$, as mentioned earlier. Note that the term does not appear in Equation (3.1.2).

The authors in [BV13] refer to the lemma below as the fundamental estimate. For the proof it is better to use Equation (3.1.2) than Equation (3.1.3).

Lemma 3.1.4. *[BV13, Lemma 4.1]*
For $t\in[0,T]$,

$$\begin{aligned}\|e(t)\|_W \quad &\le \quad \sqrt{2}\bigg(\|e(0)\|_W+3T\|e_h'(0)\|_W+3TC_W\|e_h(0)\|_W \\ &\qquad +3\int_0^T\|e_p'\|_W+3C_W\int_0^T\|e_p\|_W\bigg).\end{aligned} \tag{3.1.4}$$

Proof. The proof of this result is done in detail in [BV13]. The first part of the proof is provided here for completeness and to show the benefit of using Equation (3.1.2) rather than Equation (3.1.3).

Suppose v is an anti-derivative of e. It follows from Equation (3.1.2) that

$$c\left(e_h''(t), v(t)\right) + a\left(e_h'(t), v(t)\right) + b\left(e(t), v(t)\right) = 0 \quad \text{for all } \ t \in [0, T].$$

Now, following from (3.1.2) and the fact that $e'(t) - e_h'(t) = -e_p'(t)$ for $t \in [0, T]$,

$$\begin{aligned}
&\frac{d}{dt}\left[\frac{1}{2}c\,(e,e) - \frac{1}{2}b\,(v,v) - c\,(e_h',v) - a\,(e_h,v)\right] \\
&= c\,(e',e) - b\,(e,v) - c\,(e_h'',v) - c\,(e_h',e) - a\,(e_h',v) - a\,(e_h,e) \\
&= c\,(e' - e_h', e) - a\,(e_h,e) \\
&= -c\left(e_p',e\right) - a\,(e_h,e)\,. \qquad (3.1.5)
\end{aligned}$$

Integrate (3.1.5) over $[0,\tau]$ for some $\tau \in (0,T)$, and choose $v(\tau) = 0$ to obtain

$$\begin{aligned}
&\frac{1}{2}c\,(e(\tau),e(\tau)) - \left[\frac{1}{2}c\,(e(0),e(0)) - \frac{1}{2}b\,(v(0),v(0)) - c\,(e_h'(0),v(0)) - a\,(e_h(0),v(0))\right] \\
&= \int_0^\tau \frac{d}{dt}\left[\frac{1}{2}c\,(e,e) - \frac{1}{2}b\,(v,v) - c\,(e_h',v) - a\,(e_h,v)\right] \\
&= -\int_0^\tau \left(c\left(e_p',e\right) + a\,(e_p,e) + a\,(e,e)\right).
\end{aligned}$$

But $a(e,e) \geq 0$ and $b\,(v(0),v(0)) \geq 0$, so

$$\begin{aligned}
\frac{1}{2}c\,(e(\tau),e(\tau)) \ &\leq\ \frac{1}{2}c\,(e(0),e(0)) - c\,(e_h'(0),v(0)) - a\,(e_h(0),v(0)) \\
&\quad - \int_0^\tau c\left(e_p',e\right) - \int_0^\tau a\,(e_p,e)\,. \qquad (3.1.6)
\end{aligned}$$

But $v'(t) = e(t)$ and since $v(\tau) = 0$ we have

$$-\,c\,(e_h'(0),v(0)) = \int_0^\tau c\,(e_h'(0),e) \ \text{ and } \ -\,a\,(e_h(0),v(0)) = \int_0^\tau a\,(e_h(0),e)\,. \qquad (3.1.7)$$

Now, using (3.1.7) in (3.1.6) we have

$$\begin{aligned}
\|e(\tau)\|_W^2 \ &\leq\ \|e(0)\|_W^2 + 2\left|\int_0^\tau c\,(e_h'(0),e)\right| + 2\left|\int_0^\tau a\,(e_h(0),e)\right| \\
&\quad +2\left|\int_0^\tau c\left(e_p',e\right)\right| + 2\left|\int_0^\tau a\,(e_p,e)\right|. \qquad (3.1.8)
\end{aligned}$$

An estimate for the right hand side of (3.1.8) is needed.

In the article [BV13] it is assumed that Assumptions E1, E2, E3 and E4W hold for the convergence theory. Recall Assumption E4W: The bilinear form a is non-negative, symmetric and bounded on W, i.e. there exists a constant C_W such that for $v, w \in W$, $|a(u,v)| \leq C_W \|u\|_W \|v\|_W$.

Using Assumption E4W, the Cauchy-Schwartz inequality and Young's inequality (see Lemma C.1), the authors obtain estimates for the terms

$$2\left|\int_0^\tau a\left(e_h(0), e\right)\right| \quad \text{and} \quad 2\left|\int_0^\tau a\left(e_p, e\right)\right|.$$

These bounds are substituted into (3.1.8) to obtain:

$$\begin{aligned}\|e(\tau)\|_W^2 \leq \; & \|e(0)\|_W^2 + \frac{4}{9}\max_{t\in[0,T]} \|e(t)\|_W^2 + 9T^2\|e_h'(0)\|_W^2 + 9T^2C_W^2\|e_h(0)\|_W^2 \\ & +9\left(\int_0^\tau \|e_p'\|_W\right)^2 + 9C_W^2\left(\int_0^\tau \|e_p\|_W\right)^2.\end{aligned}$$

But this holds for all $\tau \in (0,T)$, and so $\max_{t\in[0,T]} \|e(t)\|_W^2 \leq \|e(\tau)\|_W^2$, and therefore

$$\begin{aligned}\frac{1}{2}\max_{t\in[0,T]} \|e(t)\|_W^2 \leq \; & \|e(0)\|_W^2 + 9T^2\|e_h'(0)\|_W^2 + 9T^2C_W^2\|e_h(0)\|_W^2 \\ & +9\left(\int_0^\tau \|e_p'\|_W\right)^2 + 9C_W^2\left(\int_0^\tau \|e_p\|_W\right)^2.\end{aligned}$$

Therefore

$$\begin{aligned}\max_{t\in[0,T]} \|e(t)\|_W^2 \leq \; & 2\Bigg(\|e(0)\|_W^2 + 9T^2\|e_h'(0)\|_W^2 + 9T^2C_W^2\|e_h(0)\|_W^2 \\ & +9\int_0^\tau \|e_p'\|_W^2 + 9C_W^2\int_0^\tau \|e_p\|_W^2\Bigg) \\ \leq \; & 2\Bigg(\|e(0)\|_W + 3T\|e_h'(0)\|_W + 3TC_W\|e_h(0)\|_W \\ & +3\int_0^\tau \|e_p'\|_W + 3C_W\int_0^\tau \|e_p\|_W\Bigg)^2.\end{aligned}$$

Taking the square root, the estimate (3.1.4) follows. □

3.1.2 Error estimates

Recall that we have that for $t \in [0, T]$,

$$\|u(t) - u_h(t)\|_W \leq \|e_p(t)\|_W + \|e(t)\|_W.$$

Using Lemma 3.1.4 we therefore have the following result.

Theorem 3.1.5. *[BV13, Theorem 5.1]*
For $t \in [0, T]$,

$$\begin{aligned}\|u(t) - u_h(t)\|_W \quad \leq \quad & \|e_p(t)\|_W + \sqrt{2}\Bigg(\|Pu_0 - u_0\|_W + 3T\|u_1 - u_1^h\|_W \\ & +(1 + 3TC_W)\|u_0 - u_0^h\|_W + 3\int_0^T \|e_p'\|_W + 3C_W \int_0^T \|e_p\|_W\Bigg).\end{aligned}$$

To compare [Bak76] and [BV13], bear in mind that the spaces V and W in [BV13] correspond to the spaces $H_0^1(\Omega)$ and $\mathcal{L}^2(\Omega)$ in [Bak76] respectively. As mentioned in the beginning of this section, Baker chooses the initial values u_0^h and u_1^h to be the $\mathcal{L}^2$-projections of the initial conditions u_0 and u_1 respectively, but does not mention how these projections can be obtained in practice. This is another very important difference between [BV13] and [Bak76]. In [BV13] the initial values are not initially chosen to be the W-projections of the initial conditions u_0 and u_1 respectively. In [BV13] the initial values u_0^h and u_1^h are left in the result, which gives the reader the opportunity to choose the initial values from a variety of options. One of these options is to assume that the initial values are chosen to be the interpolants of u_0 and u_1 respectively, $u_0^h = \Pi u_0$ and $u_1^h = \Pi u_1$. The interpolation operator here is general (see Assumption GI below), but in applications a specific interpolation operator is chosen, see Chapter 6. Consider now the general interpolation assumption as in [BV13].

Assumption GI

There exists a subspace $H(V, k)$ of V, an interpolation operator Π and positive constants C_Π and α (depending on V and k) such that for $u \in H(V, k)$:

$$\|u - \Pi u\|_V \leq C_\Pi h^\alpha \|u\|_{H(V,k)},$$

where $\|\cdot\|_{H(V,k)}$ is a norm or semi-norm associated with $H(V, k)$.

The projection errors in Theorem 3.1.5 can now be estimated. By choosing the initial values to be $u_0^h = \Pi u_0$ and $u_1^h = \Pi u_1$, the following result is obtained.

Theorem 3.1.6. *[BV13, Theorem 5.2]*
Suppose that Assumption GI holds and that $u_0^h = \Pi u_0$ and $u_1^h = \Pi u_1$. If the solution u satisfies $u(t) \in H(V,k)$ and $u'(t) \in H(V,k)$ for $t \in [0,T]$, then

$$\begin{aligned}\|u(t) - u_h(t)\|_W \quad \leq \quad & \kappa_1 C_\Pi h^\alpha \|u(t)\|_{H(V,k)} + \sqrt{2}\kappa_1 C_\Pi h^\alpha \bigg((1 + 3TC_W)\|u_0\|_{H(V,k)} \\ + \quad & 3T\|u_1\|_{H(V,k)} + 3T \max_{t\in[0,T]} \|u'(t)\|_{H(V,k)} + 3C_W T \max_{t\in[0,T]} \|u(t)\|_{H(V,k)} \bigg).\end{aligned}$$

Proof. The result follows directly from Assumption GI and the fact that $\|v\|_W \leq \kappa_1 \|v\|_V$ for all $v \in V$. □

3.2 The fully discrete approximation

3.2.1 A system of ordinary differential equations

The semi-discrete Problem G^h is equivalent to a system of ordinary differential equations. This system is given in Problem ODE below. It is convenient to introduce the following notation. For $\bar{x} \in \mathbb{R}^n$ let

$$T_h \bar{x} = \sum_{i=1}^{n} x_i \, \phi_i \, \in S^h,$$

where S^h is the span of the set of basis functions $\{\phi_1, \phi_2, \ldots, \phi_n\}$. If a function w has values in S^h, then we define a function $\bar{w}$ by

$$\bar{w} = T_h^{-1} w,$$

with values in $\mathbb{R}^n$. (The existence of T_h^{-1} is due to the fact that the basis functions are linearly independent.)

If the matrices M, C, K and the vector F are defined by

$$M_{ij} = c(\phi_j, \phi_i),\ C_{ij} = a(\phi_j, \phi_i),\ K_{ij} = b(\phi_j, \phi_i),\ \text{and}\ F_i(t) = (f(t), \phi_i)_X,$$

and the initial conditions are defined by

$$\bar{d} = \bar{u}_0^h = T_h^{-1} u_0^h \quad \text{and} \quad \bar{v} = \bar{u}_1^h = T_h^{-1} u_1^h,$$

then we have an initial value problem for a system of ordinary differential equations below.

Problem ODE

Determine $\bar{u} \in C^2[0, T]$ such that

$$M\bar{u}'' + C\bar{u}' + K\bar{u} = F(t) \text{ with } \bar{u}(0) = \bar{d} \text{ and } \bar{u}'(0) = \bar{v}.$$

We see in [BV13] that the function u_h is a solution of Problem G^h if and only if the function $\bar{u}$ is a solution of Problem ODE [BV13, Proposition 6.1] and also that if $F \in C[0, T]$, then Problem ODE has a unique solution for each pair of vectors $\bar{d}$ and $\bar{v}$ [BV13, Proposition 6.2].

3.2.2 Time-stepping scheme

A finite difference method is used to approximate the solution of the system in Problem G^h (or Problem ODE). In [BV13] the authors follow [Bak76] with an obvious modification to include the damping term. Note that Baker [Bak76] applies the finite difference scheme to the solution u of Problem G, whereas in [BV13] it is applied to the Galerkin approximation u_h of Problem G^h. To do the time discretization we suppose the interval $[0, T]$ is divided into N steps of length $\tau = \frac{T}{N}$ and denote the approximation of $u_h(t_k)$ by u_k^h.

The aim is to estimate the difference between the solution of Problem G and the fully discrete approximation:

$$u(t_k) - u_k^h = [u(t_k) - u_h(t_k)] + [u_h(t_k) - u_k^h].$$

The approach of estimating this difference in [BV13] is different from the other articles considered, where error estimates are derived for the semi discrete approximation and then for the fully discrete approximation without using the results already obtained (see for example [Bak76], [Dup73], [OR76], [Wu03] [Kar11a], [Kar11b], [Kar12] and [GS09]). Note that an estimate for the error $u(t_k) - u_h(t_k)$ was obtained in Section 3.1.2. Following

[BV13], we now require an estimate for the error $u_h(t_k) - u_k^h$, and then use both the estimates (together with the triangle inequality) to obtain a final estimate. The approach of [BV13] has two advantages. It is not necessary to assume the existence of a third or fourth order derivative for the exact solution and the convergence analysis for the fully discrete approximation is simplified.

Consider the fully discrete problem in variational form. This is also done in [Bak76] and [OR76].

Notation

For any sequence $\{y_k\} \subset \mathbb{R}^n$,

$$\begin{aligned} \delta_t y_k &= \frac{y_{k+1} - y_{k-1}}{\tau}; \\ y_{k+\frac{1}{2}} &= \frac{y_{k+1} + y_k}{2}. \end{aligned}$$

Problem G^h-D

Find a sequence $\{u_k^h\} \subset S^h$ such that for $k = 0, 1, 2, \ldots, N-1$,

$$\delta_t u_k^h = v_{k+\frac{1}{2}}, \tag{3.2.1}$$

$$c(\delta v_k, \varphi) + a(v_{k+\frac{1}{2}}, \varphi) + b(u_{k+\frac{1}{2}}^h, \varphi) = \frac{1}{2}(f(t_k) + f(t_{k+1}), \varphi)_X \tag{3.2.2}$$

for each $\varphi \in S^h$, while $u_0^h = u_h(0) = d^h$ and $u_1^h = u_h'(0) = v^h$.

In [BV13, Proposition 6.3.] we see that Problem G^h-D has a unique solution for any pair of vectors d^h and v^h in S^h. In the proof of this proposition an algorithm is derived, given below in Problem FD.

Problem FD

Find a sequence $\{\bar{u}_k\} \in \mathbb{R}^n$ such that for each k,

$$\begin{aligned} \bar{u}_{k+1} &= \bar{u}_k + \tau \bar{v}_{k+\frac{1}{2}}, \\ \left(M + \frac{\tau}{2}C + \frac{\tau^2}{4}K\right)\bar{v}_{k+1} &= \left(M - \frac{\tau}{2}C - \frac{\tau^2}{4}K\right)\bar{v}_k - \tau K \bar{u}_k + \frac{\tau}{2}\left(F(t_k) + F(t_{k+1})\right) \end{aligned}$$

while $\bar{u}_0 = \bar{d}$ and $\bar{v}_0 = \bar{v}$.

3.2.3 Error estimates

In [BV13] an estimate for $\|u_h(t_k) - u_k^h\|_W$ is now derived. (The direct approach relies on the assumption that the exact solution u has derivatives $u^{(k)} \in \mathcal{L}^2([0,T],V)$ for $k \leq 4$ or $k \leq 3$, which is very restrictive.) For the proof it is required that the Galerkin approximation u_h satisfies $u_h \in C^4[0,T]$, and it is the case if $f \in C^2([0,T],X)$.

Theorem 3.2.1. *[BV13, Theorem 6.1.]*
If $f \in C^2([0,T],X)$, then

$$\begin{aligned} \|u_h(t_k) - u_k^h\|_W \;\; &\leq \;\; 7T^2\tau^2 \max \|u_h^{(4)}\|_W + 7T\tau^2 \max \|u_h'''\|_W \\ &\quad + \sqrt{2C_W}\,\tau^4 \max \|u_h'''\|_W \end{aligned} \tag{3.2.3}$$

for each $t_k \in (0,T)$.

Notation:

- $v_h(t) = u_h'(t)$;
- $\rho_k = \tau^{-1}[v_h(t_{k+1}) - v_h(t_k)] - \frac{1}{2}\,[v_h'(t_{k+1}) + v_h'(t_k)]$;
- $e_k = u_h(t_k) - u_k^h$ and $q_k = u_h'(t_k) - v_k^h$;
- $\sigma_k = \tau^{-1}[u_h(t_{k+1}) - u_h(t_k)] - \frac{1}{2}([v_h(t_{k+1}) + v_h(t_k)]$;
- $\epsilon_n = \frac{\tau}{2}\rho_n + \tau\sum_{k=0}^{n-1} \rho_k + \sigma_n$ for $n = 1, 2, \ldots N-1$.

First, the following stability result is obtained in [BV13].

Lemma 3.2.2. *[BV13, Lemma 6.1.]*

$$\max \|e_n\|_W^2 \leq 8T\tau \sum_{n=0}^{N-1} \|\epsilon_n\|_W^2 + 2\tau^4\|\rho_0\|_W^2 + (8\tau^2 + 2\tau^4 C_W)\|\sigma_0\|_W^2.$$

Next, estimates for the truncation errors are derived. If $u_h \in C^{(4)}[0,T]$ and $v_h = u_h'$, then

$$\begin{aligned} \|\rho_k\|_W^2 \;\; &\leq \;\; \tau^4 \max \|v_h'''\|_W^2; \\ \|\sigma_k\|_W^2 \;\; &\leq \;\; \tau^4 \max \|u_h'''\|_W^2. \end{aligned}$$

Using these estimates an estimate for $\|\epsilon_n\|_W^2$ is obtained:

$$\|\epsilon_n\|_W^2 \leq 5T^2\tau^4 \max \|v_h'''\|_W^2 + 4\tau^4 \max \|u_h'''\|_W^2.$$

Using these estimates together with Lemma 3.2.2, the result in Theorem 3.2.1 is achieved.

3.3 Convergence of the fully discrete approximation

Now error estimates for the fully discrete approximation of the solution of Problem G is obtained by combining the error estimate for the semi-discrete approximation from Section 3.1 with the error estimate obtained in the previous section.

Recall that we assume Assumptions E1, E2, E3, E4W and GI hold for the spaces V, W and X, and we may use the properties of a solution from Theorem 2.2.3.

From Theorem 3.1.6, if $u_0^h = \Pi u_0$ and $u_1^h = \Pi u_1$, we have that for weak damping

$$\begin{aligned} \|u(t) - u_h(t)\|_W \;\leq\; & \kappa_1 C_\Pi h^\alpha \|u(t)\|_{H(V,k)} + \sqrt{2}\kappa_1 C_\Pi h^\alpha \Bigg((1 + 3TC_W)\|u_0\|_{H(V,k)} \\ & + 3T\|u_1\|_{H(V,k)} + 3T \max_{t\in[0,T]} \|u'(t)\|_{H(V,k)} + 3C_W T \max_{t\in[0,T]} \|u(t)\|_{H(V,k)} \Bigg). \end{aligned}$$

for each $t \in [0, T]$. From Theorem 3.2.1, if $f \in C^2([0, T], X)$, then

$$\begin{aligned} \|u_h(t_k) - u_k^h\|_W \;\leq\; & 7T^2\tau^2 \max \|u_h^{(4)}\|_W + 7T\tau^2 \max \|u_h'''\|_W \\ & + \sqrt{2C_W}\, \tau^4 \max \|u_h'''\|_W \end{aligned} \tag{3.3.1}$$

for each $t_k \in (0, T)$.

From the triangle inequality applied to

$$u(t_k) - u_k^h = [u(t_k) - u_h(t_k)] + [u_h(t_k) - u_k^h],$$

the following result is obtained.

Notation

$u^{(k)} \in \mathcal{L}^2\big([0, T]; Y\big)$ if $u^{(k)}(t) \in Y$ for each t and $\int_{[0,T]} \|u^{(k)}\|_Y^2 < \infty$.

Theorem 3.3.1. *Suppose u is the solution to Problem G with $u_0 \in E_b$ and $u_1 \in V$ and the sequence $\{u_k^h\}$ is a solution of Problem G^h-D. Assume*

(a) Assumption GI holds for the space V,

(b) $u_0^h = \Pi u_0$ and $u_1^h = \Pi u_1$,

(c) $u'' \in \mathcal{L}^2([0,T], H(V,k))$,

(d) $f \in C^2([0,T], X)$.

Then,

$$\begin{aligned} \|u(t_k) - u_k^h\|_W &\leq \|u(t_k) - u_h(t_k)\|_W + \|u_h(t_k) - u_k^h\|_W \\ &\leq \kappa_1 C_\Pi h^\alpha \|u(t)\|_{H(V,k)} + \sqrt{2}\kappa_1 C_\Pi h^\alpha \Bigg(3\int_0^T \|u'(t)\|_{H(V,k)} \\ &\quad +3C_W \int_0^T \|u(t)\|_{H(V,k)} + (2 + 3C_W T)\|u_0\|_{H(V,k)} + 3T\|u_1\|_{H(V,k)} \Bigg) \\ &\quad +7T^2\tau^2 \max \|u_h^{(4)}\|_W + 7T\tau^2 \max \|u_h'''\|_W + \sqrt{2C_W}\,\tau^4 \max \|u_h'''\|_W, \end{aligned}$$

for each $t_k \in (0, T)$.

Remarks

1. In [BV13] assumption (c) in the theorem has the following meaning:

$$u \in C^1\left([0,T], V\right), u' \in C^1\left([0,T], W\right) \quad \text{and} \quad u''(t) \in H(V,k) \text{ for each } t.$$

2. Estimates for $\|u_h'''\|_W$ and $\|u_h^{(4)}\|_W$ in terms of the initial data and the forcing function are available, see for instance [Eva98, Theorem 6, p.391]. In this theorem the result is given for the exact solution, and since we are looking at the finite element approximation, similar estimates can be obtained.

Chapter 4

Error estimates for general damping

In this chapter we analyse the article [Kar11a], which is a recent article on the continuous Galerkin finite element method. The aim of the investigation was to study the assumptions that are necessary for convergence and the link with existence results. The proofs of the results are given here in great detail, however some of the proofs could not be done for the cases stated in the article (see Section 4.5).

The proofs from [Kar11a] are short and incomplete. The proofs are therefore done in much greater detail in this dissertation. In doing this, some inconsistencies in the article were discovered.

4.1 Introduction

In the article [Kar11a], the Dual-Phase-Lag model introduced in Section 1.3 is considered and is stated here again for convenience, but in the notation of [Kar11a].

Let Ω be a bounded convex polygonal domain in $\mathbb{R}^m$ with $m = 2, 3$. Let $J = (0, T)$ with $T > 0$. Consider the multidimensional wave equation in the form

$$\gamma_2 \partial_t^2 u + \gamma_1 \partial_t u - \nabla \cdot (Q \nabla \partial_t u) - \nabla \cdot (A \nabla u) = \tilde{f}(x, t); \text{ for } x \in \Omega \text{ and } t \in J, \quad (4.1.1)$$

with mixed boundary conditions

$$\begin{aligned} u(x,t) &= 0 \text{ for } x \in \partial\Omega - \Sigma \text{ and } t \in J \\ A \nabla u \cdot \mathbf{n} &= 0 \text{ for } x \in \Sigma \text{ and } t \in J, \end{aligned}$$

and initial conditions

$$\begin{aligned} u(x,0) &= u^0(x) \text{ for } x \in \Omega \\ \partial_t u(x,0) &= v^0(x) \text{ for } x \in \Omega. \end{aligned}$$

Remarks

1. We refer to the damping in the article as general since weak and strong damping terms occur. However, it is not the most general case as boundary damping is not included.

2. In the article [Kar11a], only homogeneous Dirichlet boundary conditions are specified. This is a significant assumption. To derive the variational form for the DPL model (or the more general version in [Kar11a]), it is either necessary to assume homogeneous Dirichlet boundary conditions or to assume that the matrix Q is a scalar multiple of A.

Recall that $\gamma_k = \gamma_k(x)$, $k = 1, 2$ are nonnegative coefficients given by (1.3.18), and $A = A(x)$ (with $Q = Q(x)$ a scalar multiple of A) is symmetric and nonnegative: for all $x \in \Omega$ and for all $\xi \in \mathbb{R}^m$

$$A = A^T, \qquad \xi^T A \xi \geq 0$$

In the article [Kar11a] it is not assumed that Q is a scalar multiple of A, but this leads to difficulties explained later. The following assumption is made in the article:

$$\gamma_2(x) = 1, \text{ for all } x \in \Omega.$$

It is unnecessary, as can be seen in Section 2.6. There the bilinear form $c(u,v) = \int_\Omega \gamma_2 uv$ is introduced and shown to be an inner product for the space W, which is the space $\mathcal{L}^2(\Omega)$ with norm $\|u\|_W = \sqrt{c(u,u)}$. Since $\|\cdot\|_W$ is equivalent to $\|\cdot\|_0$, the results will be the same.

Weak variational form

Recall the definition of the bilinear forms given in Definition 1.3.1.

$$\begin{aligned} b(u,v) &= \int A\nabla u\cdot\nabla v \text{ for } u,v\in V(\Omega) \\ a_w(u,v) &= \int \gamma_1 uv \text{ for } u,v\in V(\Omega) \\ a_s(u,v) &= \int Q\nabla u\cdot\nabla v \text{ for } u,v\in V(\Omega) \\ a(u,v) &= a_w(u,v)+a_s(u,v). \end{aligned}$$

Recall that the space $V(\Omega)$ is the closure of the space of test functions

$$\mathcal{T}(\Omega)=\{v\in C^1(\bar{\Omega}) : v=0 \text{ on } \partial\Omega-\Sigma\}$$

in the Sobolev space $H^1(\Omega)$.

Remark In [Kar11a] the bilinear form a is given in the form $a(u,v)=\gamma_1(u,v)+(Q\nabla u,\nabla v)$, but γ_1 is not constant according to the article, and hence it can not be taken out of the integral. Consequently, the bilinear form a given here differs slightly from that of [Kar11a].

The weak variational form of the Dual-Phase-Lag model is given in Subsection 2.6.1.

Problem DPLW

Find u such that for each $t\in J$, $u'(t)\in V(\Omega)$, $u''(t)\in\mathcal{L}^2(\Omega)$ and

$$\begin{aligned} (u''(t),v)+a(u'(t),v)+b(u(t),v) &= (f(t),v) \text{ for all } v\in V(\Omega), \qquad (4.1.2)\\ u(0) &= u^0, \\ u'(0) &= v^0. \end{aligned}$$

Remark In the following proposition an estimate for the bilinear for a is derived. This bound is used in the proof of the main result of this chapter (the fully discrete $\mathcal{L}^2$-norm error estimate). In [Kar11a] a different bound is used, but this is discussed in the remark on page 76 in Section 4.4.2.

Proposition 4.1.1. *If $u\in H^{p+1}(\Omega)$ and $v\in V(\Omega)$, there exists a constant $\tilde{K}$ such that the bilinear form a satisfies (for $p\geq 1$)*

$$|a(u,v)|\leq \tilde{K}\|u\|_{p+1}\|v\|_0. \qquad (4.1.3)$$

Proof. Recall that

$$a(u,v) = a_w(u,v) + a_s(u,v)$$

and

$$|a_w(u,v)| \leq C\|u\|_0\|v\|_0$$

for $u, v \in \mathcal{L}^2(\Omega)$. Now we need to estimate $a_s(\cdot,\cdot)$. We use Proposition 1.2.1 in Subsection 1.2.1, with $F = Q\nabla u$. We have for $v \in V(\Omega)$,

$$\begin{aligned} |a_s(u,v)| &= \left|\iiint_\Omega Q\nabla u \cdot \nabla v \, dV\right| \\ &= \left|\iiint_\Omega v\nabla \cdot (Q\nabla u) \; dV\right| \\ &\leq \|\nabla \cdot (Q\nabla u)\|_0\|v\|_0 \\ &\leq K_A\|u\|_2\|v\|_0 \leq K_A\|u\|_{p+1}\|v\|_0. \end{aligned}$$

The last two steps follow from the fact that $a_{ij} \in C(\bar{\Omega}) \cap C^1(\Omega)$. We can now combine the results. Using the fact that $\|\cdot\|_0 \leq \|\cdot\|_{p+1}$ we have

$$|a(u,v)| \leq K\|u\|_0\|v\|_0 + K_A\|u\|_{p+1}\|v\|_0 \leq \tilde{K}\|u\|_{p+1}\|v\|_0.$$

□

Existence

Existence for the problem is proved in Subsection 2.6.2. Following from Theorem 2.2.5, if $f : [0,T] \to \mathcal{L}^2(\Omega)$ is locally Lipschitz, then there exists a unique solution

$$u \in C\left([0,T),V(\Omega)\right) \cap C^1\left([0,T),\mathcal{L}^2(\Omega)\right) \cap C^2\left((0,T),\mathcal{L}^2(\Omega)\right)$$

for the weak variational form Problem DPLW, for any $u^0 \in V(\Omega)$ and $v^0 \in \mathcal{L}^2(\Omega)$. If $f = 0$ then

$$u \in C\left([0,\infty),V(\Omega)\right) \cap C^1\left([0,\infty),\mathcal{L}^2(\Omega)\right) \cap C^\infty\left((0,\infty),V(\Omega)\right).$$

Remarks

1. In [Kar11a], Lions and Magenes [LM72] are referenced for existence and uniqueness of a solution to the problem. However, in [LM72] existence and uniqueness are proved for the undamped wave equation, and since the problem considered

here contains damping terms, the results from [LM72] are not applicable (see Section 2.4).

2. In [Kar11a] the assumption is made that the bilinear form b is positive definite. Note that what we call positive definite in this dissertation, is called coercive in [Kar11a]. This is in fact proved in Section 2.6, assuming the Poincaré-Friedrichs inequality holds. The only property of A mentioned in [Kar11a] is that it is symmetric, and it is therefore implicitly assumed in [Kar11a] that A is positive definite.

3. In the article [Kar11a] it is assumed that the bilinear form $a(\cdot,\cdot)$ is continuous and coercive on $V(\Omega)$. These assumptions are in fact proved in Section 2.6. The fact that a is positive definite is in any event not relevant to the convergence analysis. What is relevant is the fact that the quadratic form of a is non-negative.

4.2 Finite Element Approximation

4.2.1 Semi-discrete Galerkin approximation

The semi-discrete form of the model problem, Problem DPL is a special case of Problem G^h. A finite dimensional subspace S^h of $V(\Omega)$ is constructed using (in two dimensions) piecewise linear basis functions on triangle elements or (in three dimensions) piecewise linear basis functions on tetrahedron elements. Note that the basis functions must satisfy the forced boundary condition: zero on $\partial\Omega - \Sigma$. The semi-discrete Galerkin approximation is then given in the problem below.

Problem DPLG^h
Find a function $u_h \in C(\bar{J}, S^h)$ such that

$$
\begin{aligned}
(u_h''(t), v) + a(u_h(t), v) + b(u_h(t), v) &= (f(t), v) \text{ for all } v \in S^h, \quad t \in J \quad (4.2.1)\\
u_h(0) &= u_0^h,\\
u_h'(0) &= u_1^h,
\end{aligned}
$$

Notation Let P_2 denote the $\mathcal{L}^2$-projection onto S^h.

Remarks

1. In [Kar11a], (as in [Bak76]) the initial conditions u_0^h and u_1^h are chosen as the $\mathcal{L}^2-$projections of $u_h(0)$ and $u_h'(0)$ respectively, i.e. $u_h(0) = P_2u^0$ and $u_h'(0) = P_2v^0$. However, they do not mention how these projections can be obtained in practice. Interpolation theory (see Appendix B) is therefore useful, i.e. choosing the initial conditions $u_h(0)$ and $u_h'(0)$ to be the interpolants of u^0 and v^0 respectively.

2. It is notable that the semi-discrete approximation is only stated in [Kar11a] and a semi-discrete error estimate is not derived (as is sometimes done in articles dealing with convergence). The semi-discrete approximation is stated in order to proceed to the fully discrete approximation.

Projection error estimate

Recall the definition of the elliptic projection P from Chapter 3: $b(u - Pu, v) = 0$ for all $v \in S^h$. An estimate for the projection error $\|u(t) - Pu(t)\|_0 = \|e_p(t)\|_0$ is needed. In [Kar11a] an estimate is given by using "some properties of P". This estimate is, if $u \in H^{p+1}(\Omega)$ (with $p \geq 1$),

$$\|u - Pu\|_0 \leq Ch^{p+1}\|u\|_{p+1}.$$

In Chapter 6 we show how an estimate of this kind can be obtained from interpolation theory.

4.2.2 Fully discrete approximation

For the time approximation of the semi-discrete problem (4.2.1), divide the interval $[0, T]$ into N time steps of length $\tau = \frac{T}{N}$. Denote the approximation by $u_h^n \approx u_h(t_n)$. We consider the following notation - this notation differs from [Kar11a] (see Table 4.1. page 67).

Notation For any sequence $\{y_k\} \subset \mathbb{R}^n$,

$$
\begin{aligned}
\delta_t^2 y_k &= \frac{y_{k+1} - 2y_k + y_{k-1}}{\tau^2} \\
\delta_t^+ y_k &= \frac{y_{k+1} - y_k}{\tau}, \qquad \delta_t^- y_k \frac{y_k - y_{k-1}}{\tau} \\
\delta_{t,\gamma} y_k &= \frac{1}{\tau}\left(\gamma y_{k+1} + (1-2\gamma) y_k + (\gamma - 1) y_{k-1}\right) \\
\delta_{\theta,\gamma} y_k &= \theta y_{k+1} + \left(\frac{1}{2} - 2\theta + \gamma\right) y_k + \left(\frac{1}{2} + \theta - \gamma\right) y_{k-1} \\
y_{k+\frac{1}{2}} &= \frac{1}{2}\left(y_{k+1} + y_k\right) \\
y_{k-\frac{1}{2}} &= \frac{1}{2}\left(y_k + y_{k-1}\right).
\end{aligned}
$$

Note that:

$$
\delta_{t,\frac{1}{2}} y_k = \delta_t y_k = \frac{y_{k+1} - y_{k-1}}{2\tau} \quad \text{and} \quad \delta_{\theta,\frac{1}{2}} y_k = \delta_\theta y_k = \theta y_{k+1} + (1 - 2\theta)\, y_k + \theta y_{k-1}.
$$

As mentioned in the articles [Kar11a] and [Kar12], a general time discretisation method which is well-know in engineering literature is given by the Newmark method [New59].

The Newmark method applied to (4.2.1) yields

$$
(\delta_t^2 u_h^n, v) + a(\delta_{t,\gamma} u_h^n, v) + b(\delta_{\theta,\gamma} u_h^n, v) = (\delta_{\theta,\gamma} f(t_n), v) \text{ for all } v \in S^h. \tag{4.2.2}
$$

The Newmark method is second-order accurate for all θ if $\gamma = \frac{1}{2}$, while it is first-order accurate if $\gamma \neq \frac{1}{2}$. In [Kar11a] it is said that "by Taylor expansions, it is easy to show" that this is true. In this article the author proves this fact for $\gamma = \frac{1}{2}$ only, but not in full generality (see Lemma 4.4.4). If $\gamma = \frac{1}{2}$, the Newmark scheme reduces to

$$
(\delta_t^2 u_h^n, v) + a(\delta_t u_h^n, v) + b(\delta_\theta u_h^n, v) = (\delta_\theta f(t_n), v) \text{ for all } v \in S^h. \tag{4.2.3}
$$

It can be easily verified that (4.2.3) is implicit if $\theta \neq 0$ and it reduces to the Central difference scheme

$$
(\delta_t^2 u_h^n, v) + a(\delta_t u_h^n, v) + b(u_h^n, v) = (f(t_n), v) \text{ for all } v \in S^h \tag{4.2.4}
$$

when $\theta = 0$.

Remark [Kar11a] calls the central difference scheme the Leapfrog scheme.

Some other interesting and notable schemes can be obtained from the General Newmark scheme (4.2.2). For example, by choosing $\theta = \frac{1}{4}$ and $\gamma = \frac{1}{2}$ we obtain the Average Acceleration Method (see [Zie77]):

$$(\delta_t^2 u_h^n, v) + a(\delta_t u_h^n, v) + b(\delta_{\frac{1}{4},\frac{1}{2}} u_h^n, v) = (\delta_{\frac{1}{4},\frac{1}{2}} f(t_n), v) \text{ for all } v \in S^h, \tag{4.2.5}$$

where

$$\begin{aligned} \delta_{\frac{1}{4},\frac{1}{2}} u_h^n &= \frac{1}{4}\left(u_h^{n+1} + 2u_h^n + u_h^{n-1}\right) \\ \delta_{\frac{1}{4},\frac{1}{2}} f(t_n) &= \frac{1}{4}\left(f(t_{n+1}) + 2f(t_n) + f(t_{n-1})\right). \end{aligned}$$

The aim of Karaa's article is to investigate what he calls the stability of the general scheme (4.2.2) and derive optimal error estimates. Appropriate initial conditions are required. Consider $u_h^0 = P_2 u^0$ and define $u_h^1 \in S^h$ by requiring that

$$\begin{aligned} &(u_h^1 - u_h^0, v) + \tau^2 \theta b(u_h^1 - u_h^0, v) \\ &\quad = \tau(v^0, v) + \frac{\tau^2}{2}(\widetilde{u}_h^0, v) + \theta\tau^2(f^1 - f^0, v) \text{ for all } v \in S^h, \end{aligned} \tag{4.2.6}$$

where $\widetilde{u}_h^0 \in S^h$ is the solution of the elliptic problem

$$(\widetilde{u}_h^0, v) = (f^0, v) - b(u^0, v) - a(v^0, v) \text{ for all } v \in S^h. \tag{4.2.7}$$

The choice of initial conditions follows from the derivation of the Newmark scheme. This derivation can be seen in [Kar12].

Proposition 4.2.1. *[Kar11a, Proposition 1]*
For any $\theta, \gamma \geq 0$, the fully discrete approximations $\{u_h^n\}_{n=0}^N$ are uniquely defined in S^h by (4.2.2) and (4.2.6).

Proof. Consider Equations (4.2.6) and (4.2.7). From Riesz's Theorem [Kre78, Theorem 3.8-1, p. 188] it follows that $\widetilde{u}_h^0$ in (4.2.7) is uniquely defined in S^h. The initial condition u_h^1 is then also uniquely defined in S^h for any θ following from (4.2.6). The initial condition u_h^0 is also unique.

Define $\mathcal{Q}^n = \delta_t^2 u_h^n$ for $n \geq 1$. Karaa [Kar11a] now states that

$$\gamma\tau\mathcal{Q}^n + \delta_t^- u_h^n = \delta_{t,\gamma} u_h^n. \tag{4.2.8}$$

To see this note that

$$\begin{aligned}
\mathcal{Q}^n = \delta_t^2 u_h^n &= \frac{u_h^{n+1} - 2u_h^n + u_h^{n-1}}{\tau^2} \\
&= \frac{1}{\gamma\tau}\left(\frac{\gamma u_h^{n+1} - 2\gamma u_h^n + \gamma u_h^{n-1}}{\tau}\right) \\
&= \frac{1}{\gamma\tau}\left(\frac{\gamma u_h^{n+1} + (1-2\gamma)u_h^n + (\gamma-1)u_h^{n-1}}{\tau} - \frac{u_h^n - u_h^{n-1}}{\tau}\right).
\end{aligned}$$

The next statement in the proof is

$$\theta\tau^2\mathcal{Q}^n + \delta_{0,\gamma}u_h^n = \delta_{\theta,\gamma}u_h^n. \tag{4.2.9}$$

To see that this is true, note that

$$\begin{aligned}
\mathcal{Q}^n = \delta_t^2 u_h^n &= \frac{u_h^{n+1} - 2u_h^n + u_h^{n-1}}{\tau^2} \\
&= \frac{1}{\theta\tau^2}\left(\theta u_h^{n+1} - 2\theta u_h^n + \theta u_h^{n-1}\right) \\
&= \frac{1}{\theta\tau^2}\left(\theta u_h^{n+1} + \left(\frac{1}{2} - 2\theta + \gamma\right)u_h^n + \left(\frac{1}{2} + \theta - \gamma\right)u_h^{n-1}\right) \\
&\quad -\frac{1}{\theta\tau^2}\left(\left(\frac{1}{2} + \gamma\right)u_h^n + \left(\frac{1}{2} - \gamma\right)u_h^{n-1}\right).
\end{aligned}$$

Now define a bilinear form $\tilde{c}(\cdot,\cdot)$ by

$$\tilde{c}(u,v) = (u,v) + \gamma\tau a(u,v) + \theta\tau^2 b(u,v) \text{ for all } u,v \in S^h$$

and a linear functional ℓ^n by

$$\ell^n(v) = (\delta_{\theta,\gamma}f(t_n), v) - a(\delta_t^- u_h^n, v) - b(\delta_{0,\gamma}u_h^n, v) \text{ for all } v \in S^h.$$

Substituting (4.2.8) and (4.2.9) into (4.2.2) we find that for $n \geq 1$, $\mathcal{Q}^n$ satisfies

$$\tilde{c}(\mathcal{Q}^n, v) = \ell^n(v) \text{ for all } v \in S^h.$$

The linear functional ℓ^n is bounded on S^h (independent of the dimension n), since the bilinear forms a and b are bounded on S^h. Since the bilinear form b is positive definite on S^h and a is non-negative, it follows that the bilinear form $\tilde{c}$ is an inner product for a space with norm equivalent to $\|\cdot\|_1$. It therefore follows from Riesz's Theorem [Kre78, Theorem 3.8-1, p. 188] that $\mathcal{Q}^n$ is uniquely defined in S^h, and this implies that u_h^{n+1} is uniquely defined in S^h for $n \geq 1$. □

4.3 Energy Analysis

In this section, the stability of the fully discrete scheme (4.2.2) (with the absence of forcing) is investigated. It is not explicitly made clear what the definition of the stability is here.

Theorem 4.3.1. *[Kar11a, Theorem 1] The fully discrete scheme (4.2.2) is stable if*

$$\gamma \geq \frac{1}{2} \text{ and } (\tau)^2 \left(\frac{\gamma}{2} - \theta\right) \sup_{v \in S^h \setminus \{0\}} \frac{b(v,v)}{c(v,v)} \leq 1. \tag{4.3.1}$$

The scheme is unconditionally stable when

$$2\theta \geq \gamma \geq \frac{1}{2}.$$

Proof. First note that

$$\delta_{t,\gamma} u_h^n = \tau \left(\gamma - \frac{1}{2}\right) \delta_t^2 u_h^n + \delta_t u_h^n.$$

This can readily be seen from the definitions of the operators δ_t^2 and δ_t. Also note that

$$\delta_{\gamma,\theta} u_h^n = \tau^2 \left(\gamma - \frac{1}{2}\right) \delta_t^2 u_h^n + \gamma \left(u_h^{n+\frac{1}{2}} + u_h^{n-\frac{1}{2}}\right) + (1-2\gamma)\, u_h^{n-\frac{1}{2}}.$$

This follows immediately from the definitions of δ_t^2, $u_h^{n+\frac{1}{2}}$ and $u_h^{n-\frac{1}{2}}$. We can now rewrite the variational form of the general Newmark method (4.2.2) by using the above two equations. Note that we only investigate the case where $f = 0$. The result is

$$\begin{aligned}(\delta_t^2 u_h^n, v) + \tau \left(\gamma - \frac{1}{2}\right) a(\delta_t^2 u_h^n, v) + a(\delta_t u_h^n, v) + \tau^2 \left(\gamma - \frac{1}{2}\right) b(\delta_t^2 u_h^n, v) \\ + \gamma b\left(u_h^{n+\frac{1}{2}} + u_h^{n-\frac{1}{2}}, v\right) + (1-2\gamma)\, b\left(u_h^{n-\frac{1}{2}}, v\right) = 0 \text{ for all } v \in S^h.\end{aligned}$$

Now choose $v = \delta_t u_h^n$ in the above equation. We therefore have that

$$\begin{aligned}\left((\delta_t^2 u_h^n, \delta_t u_h^n\right) + \tau \left(\gamma - \frac{1}{2}\right) a\left(\delta_t^2 u_h^n, \delta_t u_h^n\right) + a\left(\delta_t u_h^n, \delta_t u_h^n\right) + \tau^2 \left(\gamma - \frac{1}{2}\right) b\left(\delta_t^2 u_h^n, \delta_t u_h^n\right) \\ + \gamma b\left(u_h^{n+\frac{1}{2}} + u_h^{n-\frac{1}{2}}, \delta_t u_h^n\right) + (1-2\gamma)\, b\left(u_h^{n-\frac{1}{2}}, \delta_t u_h^n\right) = 0.\end{aligned} \tag{4.3.2}$$

Now we can rewrite some of these terms. The first term can be rewritten as

$$\begin{aligned}\left(\delta_t^2 u_h^n, \delta_t u_h^n\right) &= \frac{1}{2\tau}\left(\frac{1}{\tau}\left(\left(u_h^{n+1}-u_h^n\right)-\left(u_h^n-u_h^{n-1}\right)\right), \frac{1}{\tau}\left(u_h^{n+1}-u_h^n+u_h^n-u_h^{n-1}\right)\right)\\ &= \frac{1}{2\tau}\left(\delta_t^+ u_h^n - \delta_t^- u_h^n, \delta_t^+ u_h^n + \delta_t^- u_h^n\right)\\ &= \frac{1}{2\tau}\left(\left(\delta_t^+ u_h^n, \delta_t^+ u_h^n\right)-\left(\delta_t^- u_h^n, \delta_t^- u_h^n\right)\right).\end{aligned}$$

In exactly the same way we have that

$$a\left(\delta_t^2 u_h^n, \delta_t u_h^n\right) = \frac{1}{2\tau}\left(a\left(\delta_t^+ u_h^n, \delta_t^+ u_h^n\right) - a\left(\delta_t^- u_h^n, \delta_t^- u_h^n\right)\right)$$

and

$$b\left(\delta_t^2 u_h^n, \delta_t u_h^n\right) = \frac{1}{2\tau}\left(b\left(\delta_t^+ u_h^n, \delta_t^+ u_h^n\right) - b\left(\delta_t^- u_h^n, \delta_t^- u_h^n\right)\right).$$

Notice that the we can write

$$\delta_t u_h^n = \frac{u_h^{n+\frac{1}{2}} - u_h^{n-\frac{1}{2}}}{\tau}.$$

Now using this fact we can rewrite the term $b\left(u_h^{n+\frac{1}{2}} + u_h^{n-\frac{1}{2}}, \delta_t u_h^n\right)$ as follows:

$$\begin{aligned}b\left(u_h^{n+\frac{1}{2}} + u_h^{n-\frac{1}{2}}, \delta_t u_h^n\right) &= b\left(u_h^{n+\frac{1}{2}} + u_h^{n-\frac{1}{2}}, \frac{u_h^{n+\frac{1}{2}} - u_h^{n-\frac{1}{2}}}{\tau}\right)\\ &= \frac{1}{\tau}\left(b\left(u_h^{n+\frac{1}{2}}, u_h^{n+\frac{1}{2}}\right) - b\left(u_h^{n-\frac{1}{2}}, u_h^{n-\frac{1}{2}}\right)\right).\end{aligned}$$

Now consider the inequality in Lemma C.2

$$(x-y, y) \leq \frac{1}{2}(x,x) - \frac{1}{2}(y,y),$$

which is a direct consequence of Young's inequality (Lemma C.1) and the Cauchy-Schwartz inequality. Using this we see that

$$\begin{aligned}b\left(u_h^{n-\frac{1}{2}}, \delta_t u_h^n\right) &= \frac{1}{\tau} b\left(u_h^{n-\frac{1}{2}}, u_h^{n+\frac{1}{2}} - u_h^{n-\frac{1}{2}}\right)\\ &\leq \frac{1}{2\tau} b\left(u_h^{n+\frac{1}{2}}, u_h^{n+\frac{1}{2}}\right) - \frac{1}{2\tau} b\left(u_h^{n-\frac{1}{2}}, u_h^{n-\frac{1}{2}}\right).\end{aligned}$$

Now we define the discrete energy with $\gamma \geq \frac{1}{2}$ as

$$\begin{aligned} E_h^{n+\frac{1}{2}} &= \frac{1}{2}\Bigg(c\left(\delta_t^+ u_h^n, \delta_t^+ u_h^n\right) + \tau\left(\gamma - \frac{1}{2}\right) a\left(\delta_t^+ u_h^n, \delta_t^+ u_h^n\right) \\ &\qquad + \tau^2\left(\theta - \frac{\gamma}{2}\right) b\left(\delta_t^+ u_h^n, \delta_t^+ u_h^n\right) + b(u_h^{n+\frac{1}{2}}, u_h^{n+\frac{1}{2}})\Bigg). \end{aligned} \tag{4.3.3}$$

We therefore see that (4.3.2) is equivalent to

$$\frac{1}{\tau}\left(E_h^{n+\frac{1}{2}} - E_h^{n-\frac{1}{2}}\right) + a\left(\delta_t^+ u_h^n, \delta_t^+ u_h^n\right) = 0. \tag{4.3.4}$$

From (4.3.4) and the fact that a is positive definite on V it follows that the energy $E_h^{n+\frac{1}{2}}$ does not increase with n. Also, if the second condition in (4.3.1) hold, then we see that, from the definition of $E_h^{n+\frac{1}{2}}$ in (4.3.3), we have

$$\begin{aligned} E_h^{n+\frac{1}{2}} &\geq \frac{1}{2}\left(c\left(\delta_t^+ u_h^n, \delta_t^+ u_h^n\right) + \tau^2\left(\theta - \frac{\gamma}{2}\right)\frac{c\left(\delta_t^+ u_h^n, \delta_t^+ u_h^n\right)}{\tau^2\left(\theta - \frac{\gamma}{2}\right)}\right) \\ &\geq c\left(\delta_t^+ u_h^n, \delta_t^+ u_h^n\right). \end{aligned}$$

Therefore $E_h^{n+\frac{1}{2}}$ is positive semi-definite.

According to [Kar11a] it now follows that the scheme (4.2.2) is stable. He does not however give the definition of stability.

Unconditional stability

Considering again the definition of $E_h^{n+\frac{1}{2}}$ in (4.3.3) and assuming that $\gamma \leq 2\theta$ we see that $E_h^{n+\frac{1}{2}}$ is again positive semi-definite.

□

From the inverse property (see for instance [OR76, pages 340 - 341]), we have

$$b(v, v) \leq M h^{-2}\|v\|_0^2 \text{ for all } v \in S^h, \tag{4.3.5}$$

with M a constant independent of h or v. We therefore have the CFL condition

$$\frac{\tau^2}{h^2}\left(\frac{\gamma}{2} - \theta\right) < \frac{1}{M}. \tag{4.3.6}$$

4.4 Convergence Analysis

In this section our attention is devoted to an error analysis of the fully discrete scheme (4.2.2) and (4.2.6). It is assumed that $\gamma \geq 1/2$ and that the mesh size h and the time step τ satisfy the CFL condition (4.3.6).

Theorem 4.4.1. *[Kar11a, Theorem 2] Let the solution u of the wave problem satisfy the regularity properties*

$$u \in C^2(\bar{J}, H^{p+1}(\Omega)), \quad u''' \in C(\bar{J}, \mathcal{L}^2(\Omega)), \quad u^{(4)} \in L^1(\bar{J}, \mathcal{L}^2(\Omega)), \tag{4.4.1}$$

and let the discrete finite element approximations $\{u_h^n\}_{n=0}^N$ be defined by (4.2.2) and (4.2.6). Assume that the CFL condition (4.3.6) is satisfied. Then the following a priori error estimate holds:

$$\max_{n=0,\ldots,N} \|u(t_n) - u_h^n\|_0 \leq \tilde{C}(h^{p+1} + \tau^{q(\gamma)}),$$

where $q(\gamma) = 1$ if $\gamma \neq 1/2$, $q(1/2) = 2$ and $\tilde{C} > 0$ is a constant independent of the mesh size and the time step.

Note that in the article [Kar11a], the proof of Theorem 4.4.1 is limited to the case where $\gamma = 1/2$, “for the sake of conciseness” and that “the general case can be proved without major difficulties”.

Notation

$$e_n^h = Pu(t_n) - u_h^n, \quad e_p(t_n) = u(t_n) - Pu(t_n), \quad e^n = u(t_n) - u_h^n$$

The error e^n can be split in the following manner:

$$e^n = e_n^h + e_p(t_n).$$

The way [Kar11a] and [GS09] prove convergence, is to first define $r^n \in S^h$ by

$$(r^n, v) = (\delta_t^2 Pu(t_n) - \delta_{\theta,\gamma} u''(t_n), v) + a(\delta_t Pu(t_n) - \delta_{\theta,\gamma} u'(t_n), v) \tag{4.4.2}$$

Karaa notation	Our notation
$a(\cdot,\cdot)$	$b(\cdot,\cdot)$
$b(\cdot,\cdot)$	$a(\cdot,\cdot)$
U^n	u_h^n
$\bar{\partial}_t$	δ_t
$\bar{\partial}_{tt}$	δ_t^2
Δt	τ
P_h	P_2
Π_h	P
ω^n	$Pu(t_n)$
ϕ^n	e_n^h
η^n	$e_p(t_n)$

TABLE 4.1: Some differences in notation from [Kar11a]

for all $v \in S^h, n \geq 1$, where

$$\begin{aligned}\delta_{\theta,\gamma}u''(t_n) &= \theta u''(t_{n+1}) + \left(\frac{1}{2} - 2\theta + \gamma\right)u''(t_n) + \left(\frac{1}{2} - \theta + \gamma\right)u''(t_{n-1})\\ \delta_{\theta,\gamma}u'(t_n) &= \theta u'(t_{n+1}) + \left(\frac{1}{2} - 2\theta + \gamma\right)u'(t_n) + \left(\frac{1}{2} - \theta + \gamma\right)u'(t_{n-1})\end{aligned}$$

and for $n = 0$,

$$(r^0, v) = \tau^{-2}(e_1^h - e_0^h, v) + \theta b(e_1^h - e_0^h, v) \text{ for all } v \in S^h. \tag{4.4.3}$$

Next the author defines R^n by $R^n = \tau \sum_{m=0}^{n} r^m$.

The proof of the convergence result (Theorem 4.4.1) is done by first proving three results (Proposition 4.4.2, Lemma 4.4.3 and Lemma 4.4.4). In Proposition 4.4.2 it is shown that the error is bounded by R^n and and the projection error e_p. The truncations errors r^n are estimated in Lemma 4.4.3 and Lemma 4.4.4, in order to then obtain an estimate for $\|R^n\|_0$. These results are then combined to prove the main result. The proofs of these results were found to be not detailed enough, particularly if one wanted to reproduce the results. For this reason the proofs are given in a more detailed manner, i.e. there are at least twice as many steps than in the articles examined in this section.

4.4.1 Stability

The first result is proved in Proposition 4.4.2 and is what we will call the stability result (as in [BV13]), since this result limits the growth of errors. In the article [Kar11a] it is stated, regarding the proof of Proposition 4.4.2 that: "We omit the proof since the result can be obtained by a slight modification of the arguments presented in [Kar11b]." Upon consideration of the article [Kar11b], it was seen that the proof of this proposition was only done for the case when there is no damping, i.e. $a = 0$. We attempted to prove the result for the case when $a \neq 0$, but the "slight modification" mentioned in [Kar11a] is still an obstacle, as the modification is far from trivial. We therefore only give the proof for the case when $a = 0$ here. In Subsection 4.5.1 this problem that arises with the damping term is discussed.

Remark The proof below is given for the general case, i.e. when $\gamma \geq \dfrac{1}{2}$.

Proposition 4.4.2. *[Kar11a, Proposition 2]*
Assume that the CFL condition (4.3.6) holds. Then we have

$$\max_{1\leq n\leq N} \|e^n\|_0 \leq C^* \left(\|e^0\|_0 + \max_{1\leq n\leq N} \|e_p(t_n)\|_0 + \tau \sum_{n=0}^{N-1} \|R^n\|_0 \right), \tag{4.4.4}$$

with a constant $C^* > 0$ independent of h, τ and T.

Proof. By the definition of e^n,

$$\max_{0\leq n\leq N} \|e^n\|_0 \leq \max_{0\leq n\leq N} \|e_n^h\|_0 + \max_{0\leq n\leq N} \|e_p(t_n)\|_0 \tag{4.4.5}$$

and so we need to bound $\max_{0\leq n\leq N} \|e_n^h\|_0$.

Consider the weak variational form (4.1.2) at t_{n+1}, t_n and t_{n-1} (for $n = 1, 2, \ldots, N-1$). Then we have have that for every $v \in S^h$

$$\begin{aligned}
(\theta u''(t_{n+1}), v) + b(\theta u(t_{n+1}), v) &= (\theta f(t_{n+1}), v), \\
\left(\left(\frac{1}{2} - 2\theta + \gamma\right) u''(t_n), v\right) + b\left(\left(\frac{1}{2} - 2\theta + \gamma\right) u(t_n), v\right) &= \left(\left(\frac{1}{2} - 2\theta + \gamma\right) f(t_n), v\right), \\
\left(\left(\frac{1}{2} + \theta - \gamma\right) u''(t_{n-1}), v\right) + b\left(\left(\frac{1}{2} + \theta - \gamma\right) u(t_{n-1}), v\right) &= \left(\left(\frac{1}{2} + \theta - \gamma\right) f(t_{n-1}), v\right),
\end{aligned}$$

and using these three resulting equations we obtain for every $v \in S^h$

$$(\delta_{\theta,\gamma} u''(t_n), v) + b(\delta_{\theta,\gamma} u(t_n), v) = (\delta_{\theta,\gamma} f(t_n), v). \tag{4.4.6}$$

Consider (4.2.2) again

$$(\delta_t^2 u_h^n, v) + b(\delta_{\theta,\gamma} u_h^n, v) = (\delta_{\theta,\gamma} f(t_n), v) \text{ for all } v \in S^h$$

and subtract this from (4.4.6) to obtain

$$(\delta_{\theta,\gamma}u''(t_n) - \delta_t^2 u_h^n, v) + b(\delta_{\theta,\gamma}u(t_n) - \delta_{\theta,\gamma}u_h^n, v) = (\delta_{\theta,\gamma}f(t_n) - \delta_{\theta,\gamma}f(t_n), v) \qquad (4.4.7)$$

for every $v \in S^h$ and $n = 1, 2, \ldots, N-1$.

Now add and subtract the terms $\delta_t^2 Pu(t_n)$ and $\delta_{\theta,\gamma}(Pu(t_n))$ in (4.4.7):

$$\begin{aligned}(\delta_{\theta,\gamma}u''(t_n) - \delta_t^2 Pu(t_n) + \delta_t^2 Pu(t_n) - \delta_t^2 u_h^n, v)& \\ +b(\delta_{\theta,\gamma}u(t_n) - \delta_{\theta,\gamma}(Pu(t_n)) + \delta_{\theta,\gamma}(Pu(t_n)) - \delta_{\theta,\gamma}u_h^n, v) &= 0. \qquad (4.4.8)\end{aligned}$$

Using the definition of the projection P and the definition of r^n we have that (4.4.8) becomes

$$\begin{aligned}(\delta_t^2 e_n^h, v) + b(\delta_{\theta,\gamma}e_n^h, v) &= (\delta_t^2 Pu(t_n) - \delta_{\theta,\gamma}u''(t_n), v) \\ &= (r^n, v) \qquad (4.4.9)\end{aligned}$$

for every $v \in S^h$ and $n = 1, 2, \ldots, N-1$, where

$$\begin{aligned}\delta_{\theta,\gamma}e_n^h &= \theta(Pu(t_{n+1}) - u_h^{n+1}) + \left(\frac{1}{2} - 2\theta + \gamma\right)(Pu(t_n) - u_h^n) \\ &\quad + \left(\frac{1}{2} - \theta + \gamma\right)(Pu(t_{n-1}) - u_h^{n-1}).\end{aligned}$$

We can rearrange (4.4.9) as follows:

$$(r^n, v) = (\delta_t^2 e_n^h, v) + \tau^2\left(\theta - \frac{\gamma}{2}\right) b(\delta_t^2 e_n^h, v) + \gamma b(e_{n+\frac{1}{2}}^h, v) + (1-\gamma) b(e_{n-\frac{1}{2}}^h, v) \quad (4.4.10)$$

for every $v \in S^h$ and $n = 1, 2, \ldots, N-1$. Now multiply (4.4.10) by τ and then sum over $n = 1$ to $n = m$. Some terms will cancel and we are left with the remainder:

$$\begin{aligned}\tau\sum_{n=1}^{m}(r^n, v) &= \left(\frac{1}{\tau}\left(e_{m+1}^h - e_m^h\right), v\right) - \left(\frac{1}{\tau}\left(e_1^h - e_0^h\right), v\right) \\ &\quad +\tau^2\left(\theta - \frac{\gamma}{2}\right) b\left(\frac{1}{\tau}\left(e_{m+1}^h - e_m^h\right), v\right) - \tau^2\left(\theta - \frac{\gamma}{2}\right) b\left(\frac{1}{\tau}\left(e_1^h - e_0^h\right), v\right) \\ &\quad + \sum_{n=1}^{m}\left(\gamma\tau b\left(e_{n+\frac{1}{2}}^h, v\right) + (1-\gamma)\tau b\left(e_{n-\frac{1}{2}}^h, v\right)\right). \qquad (4.4.11)\end{aligned}$$

[Kar11a] now defines a quantity Φ as

$$\Phi^0 = -\gamma e_0^h, \qquad \Phi^m = -\gamma e_0^h + \sum_{n=0}^{m-1} e_{n+\frac{1}{2}}^h.$$

We now have to note

$$\sum_{n=1}^{m} e^h_{n-\frac{1}{2}} = \sum_{n=0}^{m-1} e^h_{n+\frac{1}{2}} \qquad \text{and} \qquad e^h_{0+\frac{1}{2}} = \frac{1}{2}\left(e^h_1 + e^h_0\right).$$

We then have that

$$\begin{aligned}
\gamma\Phi^{m+1} + (1-\gamma)\Phi^m &= \gamma\left(-\gamma e^h_0 + \sum_{n=0}^{m} e^h_{n+\frac{1}{2}}\right) + (1-\gamma)\left(-\gamma e^h_0 + \sum_{n=0}^{m-1} e^h_{n+\frac{1}{2}}\right) \\
&= -\frac{\gamma}{2}e^h_0 + \frac{\gamma}{2}\left(e^h_1 + e^h_0\right) + \gamma\sum_{n=1}^{m} e^h_{n+\frac{1}{2}} - \gamma e^h_0 + \frac{\gamma}{2}e^h_0 \\
&\quad + (1-\gamma)\sum_{n=0}^{m-1} e^h_{n+\frac{1}{2}} \\
&= \frac{\gamma}{2}\left(e^h_1 - e^h_0\right) + \sum_{n=1}^{m}\left(\gamma e^h_{n+\frac{1}{2}} + (1-\gamma)e^h_{n-\frac{1}{2}}\right) \qquad (4.4.12)
\end{aligned}$$

Now add $\tau(r^0, v)$ to both sides of (4.4.11) and use (4.4.12) in the resulting equation. Noting that the terms $\left(\frac{1}{\tau}\left(e^h_1 - e^h_0\right), v\right)$ and $\tau\theta b\left(e^h_1 - e^h_0, v\right)$ cancel we have:

$$\begin{aligned}
(R^m, v) = \tau\sum_{n=0}^{m}(r^n, v) &= \left(\frac{1}{\tau}\left(e^h_{m+1} - e^h_m\right), v\right) + \tau^2\left(\theta - \frac{\gamma}{2}\right) b\left(\frac{1}{\tau}\left(e^h_{m+1} - e^h_m\right), v\right) \\
&\quad + \gamma\tau b\left(\Phi^{m+1} + \Phi^m, v\right) + (1-2\gamma)\tau b\left(\Phi^m, v\right) \qquad (4.4.13)
\end{aligned}$$

for $m = 0, 1, 2, \ldots, N-1$.

Choose $v = e^h_{m+1} + e^h_m = 2(\Phi^{m+1} - \Phi^m)$ in the above equation (4.4.13), and multiply by τ. Then we have:

$$\begin{aligned}
\left(e^h_{m+1} - e^h_m, e^h_{m+1} + e^h_m\right) &= \|e^h_{m+1}\|_0^2 - \|e^h_m\|_0^2 \\
b\left(e^h_{m+1} - e^h_m, e^h_{m+1} + e^h_m\right) &= b\left(e^h_{m+1}, e^h_{m+1}\right) - b\left(e^h_m, e^h_m\right) \\
b\left(\Phi^{m+1} + \Phi^m, \Phi^{m+1} - \Phi^m\right) &= b\left(\Phi^{m+1}, \Phi^{m+1}\right) - b\left(\Phi^m, \Phi^m\right) \\
b\left(\Phi^m, \Phi^{m+1} - \Phi^m\right) &= b\left(\Phi^m, \Phi^{m+1}\right) - b\left(\Phi^m, \Phi^m\right).
\end{aligned}$$

Using this, together with the Cauchy-Schwartz inequality and that $(1-2\gamma) \leq 0$, we see that (4.4.13) (when it is chosen that $v = e^h_{m+1} + e^h_m = 2(\Phi^{m+1} - \Phi^m)$ and multiplied by τ) becomes

$$\begin{aligned}
\tau\left(R^m, e^h_{m+1} + e^h_m\right) \;\geq\; & \|e^h_{m+1}\|_0^2 - \|e^h_m\|_0^2 + \tau^2\left(\theta - \frac{\gamma}{2}\right)\left(b\left(e^h_{m+1}, e^h_{m+1}\right) - b\left(e^h_m, e^h_m\right)\right) \\
& + 2\gamma\tau^2\left(b\left(\Phi^{m+1}, \Phi^{m+1}\right) - b\left(\Phi^m, \Phi^m\right)\right) \\
& + (1-2\gamma)\tau^2\left(b\left(\Phi^{m+1}, \Phi^{m+1}\right) - b\left(\Phi^m, \Phi^m\right)\right) \qquad (4.4.14)
\end{aligned}$$

for $m = 0, 1, 2, \ldots, N-1$. Now sum (4.4.14) from $m = 0$ to $m = n-1$ for $n = 1, 2, 3, \ldots, N$. Considering that some terms will cancel, we then have

$$\begin{aligned}
\tau \sum_{m=0}^{n-1} \left(R^m, e^h_{m+1} + e^h_m\right) &\geq \sum_{m=0}^{n-1} \left(\|e^h_{m+1}\|_0^2 - \|e^h_m\|_0^2\right) \\
&\quad + \tau^2 \left(\theta - \frac{\gamma}{2}\right) \sum_{m=0}^{n-1} \left(b\left(e^h_{m+1}, e^h_{m+1}\right) - b\left(e^h_m, e^h_m\right)\right) \\
&\quad + \tau^2 \left(\sum_{m=0}^{n-1} b\left(\Phi^{m+1}, \Phi^{m+1}\right) - b\left(\Phi^m, \Phi^m\right)\right) \\
&\geq \|e^h_n\|_0^2 - \|e^h_0\|_0^2 + \tau^2 \left(\theta - \frac{1}{4}\right) \left(b\left(e^h_n, e^h_n\right) - b\left(e^h_0, e^h_0\right)\right) \\
&\quad + \tau^2 \left(b\left(\Phi^n, \Phi^n\right) - b\left(\Phi^0, \Phi^0\right)\right) \qquad (4.4.15)
\end{aligned}$$

for $n = 1, 2, 3, \ldots, N$. Note that we have the following

$$b\left(\Phi^0, \Phi^0\right) = b\left(-\gamma e^h_0, -\gamma e^h_0\right) = \gamma^2 b\left(e^h_0, e^h_0\right); \qquad (4.4.16)$$

$$b\left(\Phi^n, \Phi^n\right) \geq 0. \qquad (4.4.17)$$

We also have the inverse property (see (4.3.5))

$$b(v, v) \leq M h^{-2} \|v\|_0^2 \text{ for all } v \in S^h.$$

From the inverse property we have that

$$-\tau^2 \left(\frac{\gamma}{2} - \theta\right) b\left(e^h_n, e^h_n\right) \geq -\tau^2 \left(\frac{\gamma}{2} - \theta\right) M h^{-2} \|e^h_n\|_0^2 \qquad (4.4.18)$$

$$\theta b\left(e^h_0, e^h_0\right) \leq \theta M h^{-2} \|e^h_0\|_0^2. \qquad (4.4.19)$$

We can rewrite (4.4.15) as

$$\begin{aligned}
&\tau \sum_{m=0}^{n-1} \left(R^m, e^h_{m+1} + e^h_m\right) + \|e^h_0\|_0^2 + \tau^2 \left(\theta - \frac{\gamma}{2}\right) b\left(e^h_0, e^h_0\right) + \tau^2 b\left(\Phi^0, \Phi^0\right) \\
&\geq \|e^h_n\|_0^2 + \tau^2 \left(\theta - \frac{\gamma}{2}\right) b\left(e^h_n, e^h_n\right) + \tau^2 b\left(\Phi^n, \Phi^n\right) \qquad (4.4.20)
\end{aligned}$$

for $n = 1, 2, 3, \ldots, N$. Using (4.4.16), (4.4.17), (4.4.18) and (4.4.19) in (4.4.20) we have that

$$\begin{aligned}
&\tau \sum_{m=0}^{n-1} \left(R^m, e^h_{m+1} + e^h_m\right) + \left(1 + D^* \frac{\tau^2}{h^2} \theta M\right) \|e^h_0\|_0^2 \\
&\geq \tilde{D} \|e^h_n\|_0^2
\end{aligned}$$

for $n = 1, 2, 3, \ldots, N$, where

$$D^* = \theta + \gamma^2 - \frac{\gamma}{2}$$

and

$$\tilde{D} = 1 - \frac{\tau^2}{h^2}\left(\frac{\gamma}{2} - \theta\right) M.$$

Now if the CFL condition (4.3.6) holds, we have that $\tilde{D} > 0$.

Now if we use the Cauchy-Schwartz inequality, the triangle inequality and the fact that $n = 1, 2, 3, \ldots, N$, we obtain

$$\begin{aligned}
&\sum_{m=0}^{n-1}\left(R^m, e_{m+1}^h + e_m^h\right)\\
&\quad\le \sum_{n=0}^{N-1}\|R^n\|_0\|e_{n+1}^h + e_n^h\|_0\\
&\quad= \|R^0\|_0\left(\|e_1^h\|_0 + \|e_0^h\|_0\right) + \|R^1\|_0\left(\|e_2^h\|_0 + \|e_1^h\|_0\right)\\
&\quad+\|R^2\|_0\left(\|e_3^h\|_0 + \|e_2^h\|_0\right) + \ldots + \|R^{N-2}\|_0\left(\|e_{N-1}^h\|_0 + \|e_{N-2}^h\|_0\right)\\
&\quad+\|R^{N-1}\|_0\left(\|e_N^h\|_0 + \|e_{N-1}^h\|_0\right)\\
&\quad\le \left(\max_{0\le n\le N}\|e_n^h\|_0\right)\left(2\sum_{n=0}^{N-1}\|R^n\|_0\right).
\end{aligned}$$

Following from Young's inequality, Lemma C.1, with

$$a = 2\tau\sum_{n=0}^{N-1}\|R^n\|_0,\ b = \max_{0\le n\le N}\|e_n^h\|_0 \ \text{ and } \ \varepsilon = \tilde{D}$$

we have

$$\begin{aligned}
\tilde{D}\|e_n^h\|_0^2 &\le \left(1 + D^*\frac{\tau^2}{h^2}\theta M\right)\|e_0^h\|_0^2 + \frac{\tilde{D}}{2}\left(\max_{0\le n\le N}\|e_n^h\|_0\right)^2 + \frac{2}{\tilde{D}}\left(\tau\sum_{n=0}^{N-1}\|R^n\|_0\right)^2\\
&= \left(1 + D^*\frac{\tau^2}{h^2}\theta M\right)\|e_0^h\|_0^2 + \frac{\tilde{D}}{2}\max_{0\le n\le N}\|e_n^h\|_0^2 + \frac{2}{\tilde{D}}\left(\tau\sum_{n=0}^{N-1}\|R^n\|_0\right)^2 \quad (4.4.21)
\end{aligned}$$

Since the right hand side of (4.4.21) does not depend on n, we take the maximum over $n = 0$ to $n = N$ to obtain

$$\max_{0\le n\le N}\|e_n^h\|_0^2 \le \frac{2\left(1 + D^*\tau^2h^{-2}\theta M\right)}{\tilde{D}}\|e_0^h\|_0^2 + \frac{4}{\tilde{D}^2}\left(\tau\sum_{n=0}^{N-1}\|R^n\|_0\right)^2$$

and therefore

$$\max_{0\leq n\leq N}\|e_n^h\|_0 \leq \left(\frac{2\left(1+D^*\tau^2h^{-2}\theta M\right)}{\tilde{D}}\right)^{\frac{1}{2}}\|e_0^h\|_0 + \frac{2\tau}{\tilde{D}}\sum_{n=0}^{N-1}\|R^n\|_0. \tag{4.4.22}$$

Now, using (4.4.5), together with $\|e_0^h\|_0 \leq \|e^0\|_0 + \|e_p(t_0)\|_0$ and (4.4.22) we obtain the desired result

$$\max_{1\leq n\leq N}\|e^n\|_0 \leq C^*\left(\|e^0\|_0 + \max_{1\leq n\leq N}\|e_p(t_n)\|_0 + \tau\sum_{n=0}^{N-1}\|R^n\|_0\right), \tag{4.4.23}$$

where

$$C^* = \max\left\{1+\left(\frac{2\left(1+D^*\tau^2h^{-2}\theta M\right)}{\tilde{D}}\right)^{\frac{1}{2}}, \frac{2}{\tilde{D}}\right\}.$$

□

4.4.2 Convergence

Consider Inequality (4.4.4) in Proposition (4.4.2). We have shown that the error $\|e^n\|_0$ is estimated in terms of $\|R^n\|_0$. It is now necessary to estimate $\|R^n\|_0$ and so it is in the first place necessary to obtain an estimate for $\|r^n\|_0$. This is done in the two lemmas below, for the cases $n=0$ and $n\geq 1$.

Lemma 4.4.3. *[Kar11a, Lemma 1]*
There holds

$$\|r^0\|_0 \leq \tau C_1\left(\|u'''\|_{C(\bar{J},\mathcal{L}^2(\Omega))} + \|u''\|_{C(\bar{J},H^{p+1}(\Omega))}\right) + \tau^{-1}h^{p+1}C_1\|u'\|_{C(\bar{J},H^{p+1}(\Omega))},$$

with the constant $C_1 > 0$ independent of h, τ and T.

Proof. We can write the terms in r^0 as follows: (for all $v \in S^h$)

$$\begin{aligned}
(e_1^h - e_0^h, v) &= (Pu(t_1) - u_h^1, v) - (Pu(t_0) - u_h^0, v)\\
&= (Pu(t_1) - u(t_1), v) + (u(t_1) - u_h^1, v)\\
&\quad -(Pu(t_0) - u(t_0), v) - (u(t_0) - u_h^0, v)\\
&= ((P-I)(u(t_1) - u(t_0)), v) + (u(t_1) - u(t_0), v)\\
&\quad -(u_h^1 - u_h^0, v)
\end{aligned} \tag{4.4.24}$$

and

$$\begin{aligned} b(e_1^h - e_0^h, v) &= b(Pu(t_1) - u_h^1, v) - b(Pu(t_0) - u_h^0, v) \\ &= b(Pu(t_1) - u(t_1), v) + b(u(t_1) - u_h^1, v) \\ &\quad - b(Pu(t_0) - u(t_0), v) - b(u(t_0) - u_h^0, v) \\ &= b(u(t_1) - u(t_0), v) - b(u_h^1 - u_h^0, v), \end{aligned} \tag{4.4.25}$$

since $b(Pu(t_n) - u(t_n), v) = 0$ for $0 \leq n \leq N$.

Also, from Taylor's formula with integral remainder,

$$u(t_1) = u(t_0) + \tau u'(t_0) + \frac{\tau^2}{2} u''(t_0) + \frac{1}{2} \int_{t_0}^{t_1} (\tau - s)^2 u'''(s) ds. \tag{4.4.26}$$

Considering the weak variational form (4.1.2) at t_0 and t_1, and subtracting the resulting equations we obtain (for all $v \in S^h$):

$$\begin{aligned} &(u''(t_1) - u''(t_0), v) + a(u'(t_1) - u'(t_0), v) \\ &\qquad + b(u(t_1) - u(t_0), v) = (f(t_1) - f(t_0), v) \end{aligned} \tag{4.4.27}$$

Multiply (4.4.27) by $\theta\tau^2$ and use (4.4.26) in the resulting equation. Then we have

$$\begin{aligned} (u(t_1) - u(t_0), v) + \theta\tau^2 b(u(t_1) - u(t_0), v) &= (\tau u'(t_0), v) + \frac{\tau^2}{2}(u''(t_0), v) \\ + \frac{1}{2}\int_{t_0}^{t_1} (\tau - s)^2 (u'''(s), v) ds &- \theta\tau^2 (u''(t_1) - u''(t_0), v) \\ -\theta\tau^2 a(u'(t_1) - u'(t_0), v) &+ \theta\tau^2 (f(t_1) - f(t_0), v) \end{aligned} \tag{4.4.28}$$

Now subtract (4.2.6) from (4.4.28). The left hand side will become

$$\begin{aligned} &(u(t_1) - u(t_0), v) + \theta\tau^2 b(u(t_1) - u(t_0), v) - (u_h^1 - u_h^0, v) - \tau^2\theta b(u_h^1 - u_h^0, v) \\ &= \tau^2 (r^0, v) - ((P - I)(u(t_1) - u(t_0), v) \text{ for all } v \in S^h. \end{aligned} \tag{4.4.29}$$

The right hand side will be

$$\begin{aligned} &\tau(u'(t_0), v) + \frac{\tau^2}{2}(u''(t_0), v) + \frac{1}{2}\int_{t_0}^{t_1} (\tau - s)^2 (u'''(s), v) ds \\ &\quad -\theta\tau^2 (u''(t_1) - u''(t_0), v) - \theta\tau^2 a(u'(t_1) - u'(t_0), v) + \theta\tau^2 (f(t_1) - f(t_0), v) \\ &\qquad -\tau(u'(t_0), v) - \frac{\tau^2}{2}(\tilde{u}_h^0, v) - \theta\tau^2 (f(t_1) - f(t_0), v). \end{aligned} \tag{4.4.30}$$

Taking into account the definition of $\widetilde{u}_h^0$, i.e. (4.2.7), (4.4.30) becomes

$$\frac{\tau^2}{2}(u''(t_0),v) - \theta\tau^2(u''(t_1)-u''(t_0),v) - \theta\tau^2 a(u'(t_1)-u'(t_0),v) - a(u'(t_0),v) \\ +\frac{1}{2}\int_{t_0}^{t_1}(\tau-s)^2(u'''(s),v)ds - \frac{\tau^2}{2}\left((f(t_0),v)-b(u(t_0),v)\right). \quad (4.4.31)$$

Considering the weak variational form (4.1.2) at $t=t_0$, and using this in (4.4.31) (together with (4.4.29)), we obtain

$$\begin{aligned}(r^0,v) &= \frac{1}{2\tau^2}\int_{t_0}^{t_1}(\tau-s)^2(u'''(s),v)ds - \theta(u''(t_1)-u''(t_0),v) \\ &\quad -\theta a(u'(t_1)-u'(t_0),v) + \tau^{-2}((P-I)(u(t_1)-u(t_0),v)\end{aligned} \quad (4.4.32)$$

for all $v\in S^h$. The terms on the right of (4.4.32) can be bounded as follows. The first three results follow from the Cauchy-Schwartz inequality and the regularity properties. For all $v\in S^h$,

$$\begin{aligned}|(u''(t_1)-u''(t_0),v)| &= \left|\left(\int_{t_0}^{t_1}u'''(s)ds,v\right)\right| \\ &\le \int_{t_0}^{t_1}|(u'''(s),v)|ds \\ &\le \int_{t_0}^{t_1}\|u'''(s)\|_0\|v\|_0 ds \\ &= \tau\|u'''\|_{C(\bar{J},\mathcal{L}^2(\Omega))}\|v\|_0,\end{aligned} \quad (4.4.33)$$

$$\begin{aligned}\left|\int_{t_0}^{t_1}(\tau-s)^2(u'''(s),v)ds\right| &\le \tau^2\int_{t_0}^{t_1}|(u'''(s),v)|ds \\ &\le \tau^3\|u'''\|_{C(\bar{J},\mathcal{L}^2(\Omega))}\|v\|_0,\end{aligned} \quad (4.4.34)$$

and

$$\begin{aligned}|((P-I)(u(t_1)-u(t_0),v)| &= \left|\int_{t_0}^{t_1}((P-I)u'(s),v)ds\right| \\ &\le \int_{t_0}^{t_1}|((P-I)u'(s),v)|ds \\ &\le \int_{t_0}^{t_1}\|((P-I)u'(s)\|_0\|v\|_0 \\ &\le \tau C h^{p+1}\|u'\|_{C(\bar{J},H^{p+1}(\Omega))}\|v\|_0.\end{aligned} \quad (4.4.35)$$

The last bound follows from Proposition 4.1.3,

$$\begin{aligned}
|a(u'(t_1)-u'(t_0),v)| &= \left|a\left(\int_{t_0}^{t_1} u''(s)ds, v\right)\right| \\
&\leq \int_{t_0}^{t_1} |a(u''(s),v)|ds \\
&\leq \int_{t_0}^{t_1} a(u''(s),u''(s))^{\frac{1}{2}} a(v,v)^{\frac{1}{2}} ds \\
&\leq \beta_1 \tau \|u''\|_{C(\bar{J},H^{p+1}(\Omega))}\|v\|_0. \qquad (4.4.36)
\end{aligned}$$

Combining the bounds (4.4.33) to (4.4.36) with equation (4.4.32), we obtain (for all $v \in S^h$),

$$\begin{aligned}
|(r^0,v)| &\leq \frac{1}{2\tau^2}\left|\int_{t_0}^{t_1}(\tau-s)^2(u'''(s),v)ds\right| + \theta|(u''(t_1)-u''(t_0),v)| \\
&\quad +\theta|a(u'(t_1)-u'(t_0),v)| + \tau^{-2}\left|((P-I)(u(t_1)-u(t_0),v)\right| \\
&\leq \tau\theta\|u'''\|_{C(\bar{J},\mathcal{L}^2(\Omega))}\|v\|_0 + \tau\theta\beta_1\|u''\|_{C(\bar{J},H^{p+1}(\Omega))}\|v\|_0 \\
&\quad +\frac{\tau}{2}\|u'''\|_{C(\bar{J},\mathcal{L}^2(\Omega))}\|v\|_0 + C\tau^{-1}h^{p+1}\|u'\|_{C(\bar{J},H^{p+1}(\Omega))}\|v\|_0.
\end{aligned} \qquad (4.4.37)$$

Since $r^0 \in S^h$, we have that

$$\begin{aligned}
\|r^0\|_0^2 \leq \tau\left(\theta+\frac{1}{2}\right)\|u_h'''\|_{C(\bar{J},\mathcal{L}^2(\Omega))}\|r^0\|_0 + \tau\theta\beta_1\|u_h''\|_{C(\bar{J},H^{p+1}(\Omega))}\|r^0\|_0 \\
+C\tau^{-1}h^{p+1}\|u'\|_{C(\bar{J},H^{p+1}(\Omega))}\|r^0\|_0.
\end{aligned}$$

The result follows with $C_1 = \max\{C, \theta+\frac{1}{2}, \theta\beta_1\}$. □

Remark A very important remark needs to be made here with regards to the proof of the following lemma. The proof of this lemma in the article [Kar11a] requires that we must have for any $u, v \in V(\Omega)$,

$$|a(u,v)| \leq \beta_1\|u\|_0\|v\|_0.$$

This can be seen in going from Equation (4.4.38) to Equation (4.4.39) in the proof of the lemma. If a is merely continuous on the space $V(\Omega)$, then this estimate is not valid. In fact, a is positive definite with respect to the norm $\|\cdot\|_1$. Since the model problems in the other articles of Karaa, [Kar11b] and [Kar12] do not include damping, no help could be gained from them. However, the lemma can be proved with the same results when we are dealing with weak damping, i.e. the bilinear form a is continuous on the space

$\mathcal{L}^2(\Omega)$. This would mean that in the model problem given in the article [Kar11a], the term $\nabla \cdot (Q\nabla\partial_t u)$ should not be taken into account.

Lemma 4.4.4. *[Kar11a, Lemma 2]*
If $a_s = 0$, there holds for $1 \leq n \leq N-1$,

$$\begin{aligned} \|r^n\|_0 &\leq C_2 h^{p+1}\tau^{-1}\left(\int_{t_{n-1}}^{t_{n+1}} \|u''(s)\|_{p+1} ds + \int_{t_{n-1}}^{t_{n+1}} \|u'(s)\|_{p+1} ds\right) \\ &\quad + C_2\tau\left(\int_{t_{n-1}}^{t_{n+1}} \|u'''(s)\|_0 ds + \int_{t_{n-1}}^{t_{n+1}} \|u^{(4)}(s)\|_0 ds\right) \end{aligned}$$

with the constant $C_2 > 0$ independent of h, τ and T.

Proof. Recall that

$$(r^n, v) = (\delta_t^2 Pu(t_n) - \delta_\theta u''(t_n), v) + a(\delta_t Pu(t_n) - \delta_\theta u'(t_n), v)$$

for all $v \in S^h, n \geq 1$. Since $r^n \in S^h$, we can choose $r^n = v$ and so

$$\|r^n\|_0^2 = (\delta_t^2 Pu(t_n) - \delta_\theta u''(t_n), r^n) + a(\delta_t Pu(t_n) - \delta_\theta u'(t_n), r^n). \tag{4.4.38}$$

Now use the Cauchy-Schwartz inequality and the continuity of the bilinear form a on the space $\mathcal{L}^2(\Omega)$ to get

$$\|r^n\|_0 \leq \|\delta_t^2 Pu(t_n) - \delta_\theta u''(t_n)\|_0 + \beta_1\|\delta_t Pu(t_n) - \delta_\theta u'(t_n)\|_0. \tag{4.4.39}$$

Now adding and subtracting the terms $\delta_t^2 u(t_n)$ and $\beta_1\delta_t u(t_n)$ and using the triangle inequality we obtain

$$\begin{aligned} \|r^n\|_0 &\leq \|\delta_t^2 Pu(t_n) - \delta_t^2 u(t_n)\|_0 + \|\delta_t^2 u(t_n) - \delta_\theta u''(t_n)\|_0 \\ &\quad + \beta_1\|\delta_t Pu(t_n) - \delta_t u(t_n)\|_0 + \beta_1\|\delta_t u(t_n) - \delta_\theta u'(t_n)\|_0 \\ &= \|\delta_t^2 (P-I)u(t_n)\|_0 + \|\delta_t^2 u(t_n) - \delta_\theta u''(t_n)\|_0 \\ &\quad + \beta_1\|\delta_t (P-I)u(t_n)\|_0 + \beta_1\|\delta_t u(t_n) - \delta_\theta u'(t_n)\|_0. \end{aligned} \tag{4.4.40}$$

We now want to estimate the terms on the right of (4.4.40). First consider the general ($m \geq 0$) Taylor expansions of $u(t_{n+1})$ and $u(t_{n-1})$ about $u(t_n)$.

Taylor expansion of $u(t_{n+1})$ about $u(t_n)$:

$$\begin{aligned} u(t_{n+1}) &= u(t_n) + (t_{n+1} - t_n)u'(t_n) + \frac{(t_{n+1} - t_n)^2}{2!}u''(t_n) + \dots \\ &\quad + \frac{1}{(m-1)!}\int_{t_n}^{t_{n+1}} (t_{n+1} - s)^{m-1} u^{(m)}(s)ds \\ &= u(t_n) + \tau u'(t_n) + \frac{\tau^2}{2!}u''(t_n) + \dots \\ &\quad + \frac{1}{(m-1)!}\int_0^{\tau} (\tau - |s|)^{m-1} u^{(m)}(t_n + s)ds. \end{aligned} \tag{4.4.41}$$

(The last step follows since $0 \leq s \leq \tau$.)

Taylor expansion of $u(t_{n-1})$ about $u(t_n)$:

$$\begin{aligned} u(t_{n-1}) &= u(t_n) + (t_{n+1} - t_n)u'(t_n) + \frac{(t_{n+1} - t_n)^2}{2!}u''(t_n) + \dots \\ &\quad + \frac{1}{(m-1)!}\int_{t_n}^{t_{n-1}} (t_{n-1} - s)^{m-1} u^{(m)}(s)ds \\ &= u(t_n) - \tau u'(t_n) + \frac{\tau^2}{2!}u''(t_n) + \dots \\ &\quad + \frac{1}{(m-1)!}\int_0^{-\tau} (-1)^{m-1}(\tau - |s|)^{m-1} u^{(m)}(t_n + s)ds. \end{aligned} \tag{4.4.42}$$

(The last step follows since $-\tau \leq s \leq 0$ and so $-|s| = s$.)

Now consider (4.4.41) and (4.4.42) with $m = 1$. Then we have

$$\begin{aligned} \delta_t(P-I)u(t_n) &= \frac{1}{2\tau}\left((P-I)u(t_{n+1}) - (P-I)u(t_{n-1})\right) \\ &= \frac{1}{2\tau}\Bigg((P-I)u(t_n) + \int_0^{\tau}((P-I)u)'(t_n+s)ds \\ &\quad - \left((P-I)u(t_n) + \int_0^{-\tau}((P-I)u)'(t_n+s)ds\right)\Bigg) \\ &= \frac{1}{2\tau}\int_{-\tau}^{\tau}((P-I)u)'(t_n+s)ds \end{aligned} \tag{4.4.43}$$

and so

$$\begin{aligned} \|\delta_t(P-I)u(t_n)\|_0 &\leq \frac{1}{2\tau}\int_{-\tau}^{\tau}\|(P-I)u'(t_n+s)\|_0 ds \\ &\leq \frac{Ch^{p+1}}{2\tau}\int_{-\tau}^{\tau}\|u'(t_n+s)\|_{p+1} ds. \end{aligned} \tag{4.4.44}$$

Now consider (4.4.41) and (4.4.42) with $m = 2$. Then we have

$$\begin{aligned}
\delta_t^2(P-I)u(t_n) &= \frac{1}{\tau^2}\left((P-I)u(t_{n+1}) - 2(P-I)u(t_n) + (P-I)u(t_{n-1})\right) \\
&= \frac{1}{\tau^2}\Bigg((P-I)u(t_n) + \tau((P-I)u)'(t_n) \\
&\quad + \int_0^{\tau}(\tau-|s|)((P-I)u)''(t_n+s)ds - 2(P-I)u(t_n) \\
&\quad + (P-I)u(t_n) - \tau((P-I)u)'(t_n) \\
&\quad - \int_0^{-\tau}(\tau-|s|)((P-I)u)''(t_n+s)ds\Bigg) \\
&= \frac{1}{\tau^2}\int_{-\tau}^{\tau}(\tau-|s|)((P-I)u)''(t_n+s)ds \qquad (4.4.45)
\end{aligned}$$

and so (since $\tau - |s| \leq \tau$ because $s \in [-\tau, \tau]$)

$$\begin{aligned}
\|\delta_t^2(P-I)u(t_n)\|_0 &\leq \frac{1}{\tau^2}\int_{-\tau}^{\tau}|\tau-|s||\|(P-I)u''(t_n+s)\|_0 ds \\
&\leq \frac{Ch^{p+1}}{\tau}\int_{-\tau}^{\tau}\|u''(t_n+s)\|_{p+1}ds. \qquad (4.4.46)
\end{aligned}$$

Consider (4.4.41) and (4.4.42) with $m = 3$. Then

$$\begin{aligned}
\delta_t u(t_n) &= \frac{1}{2\tau}(u(t_{n+1}) - u(t_{n-1})) \\
&= \frac{1}{2\tau}\Bigg(u(t_n) + \tau u'(t_n) + \frac{\tau^2}{2!}u''(t_n) + \frac{1}{2!}\int_0^{\tau}(\tau-|s|)^2 u'''(t_n+s)ds \\
&\quad - \left(u(t_n) - \tau u'(t_n) + \frac{\tau^2}{2!}u''(t_n) + \frac{1}{2!}\int_0^{-\tau}(\tau-|s|)^2 u'''(t_n+s)ds\right)\Bigg) \\
&= u'(t_n) + \frac{1}{4\tau}\int_{-\tau}^{\tau}(\tau-|s|)^2 u'''(t_n+s)ds \qquad (4.4.47)
\end{aligned}$$

Also apply (4.4.41) and (4.4.42) to u' with $m = 2$ to get:

$$u'(t_{n+1}) = u'(t_n) + \tau u''(t_n) + \int_0^{-\tau}(\tau-|s|)u'''(t_n+s)ds \qquad (4.4.48)$$

and

$$u'(t_{n-1}) = u'(t_n) - \tau u''(t_n) + \int_{-\tau}^{0}(\tau-|s|)u'''(t_n+s)ds. \qquad (4.4.49)$$

Then from the Taylor expansions (4.4.48) and (4.4.49) we can write

$$\begin{aligned}
\delta_\theta u'(t_n) &= \theta u'(t_{n+1}) + (1-2\theta)\, u'(t_n) + \theta u'(t_{n-1}) \\
&= \theta\left(u'(t_n) + \tau u''(t_n) + \int_0^\tau (\tau - |s|) u'''(t_n+s)ds\right) + (1-2\theta)\, u'(t_n) \\
&\quad +\theta\left(u'(t_n) - \tau u''(t_n) + \int_{-\tau}^0 (\tau - |s|) u'''(t_n+s)ds\right) \\
&= u'(t_n) + \theta \int_{-\tau}^\tau (\tau - |s|) u'''(t_n+s)ds. \qquad (4.4.50)
\end{aligned}$$

Now subtract (4.4.50) from (4.4.47) to obtain

$$\delta_t u(t_n) - \delta_\theta u'(t_n) = \frac{1}{4\tau}\int_{-\tau}^\tau (\tau-|s|)^2 u'''(t_n+s)ds - \theta\int_{-\tau}^\tau (\tau - |s|) u'''(t_n+s)ds, \qquad (4.4.51)$$

and taking the norm on $\mathcal{L}^2(\Omega)$ and using the triangle inequality we have

$$\begin{aligned}
&\|\delta_t u(t_n) - \delta_\theta u'(t_n)\|_0 \\
&\leq \left\|\frac{1}{4\tau}\int_{-\tau}^\tau (\tau-|s|)^2 u'''(t_n+s)ds\right\|_0 + \left\|\theta\int_{-\tau}^\tau (\tau-|s|) u'''(t_n+s)ds\right\|_0 \\
&\leq \frac{1}{4\tau}\int_{-\tau}^\tau (\tau-|s|)^2 \|u'''(t_n+s)\|_0 ds + \theta\int_{-\tau}^\tau (\tau-|s|)\|u'''(t_n+s)\|_0 ds \\
&\leq \left(\frac{1}{4}+\theta\right)\tau\int_{-\tau}^\tau \|u'''(t_n+s)\|_0 ds. \qquad (4.4.52)
\end{aligned}$$

Lastly consider (4.4.41) and (4.4.42) with $m = 4$:

$$\begin{aligned}
u(t_{n+1}) &= u(t_n) + \tau u'(t_n) + \frac{\tau^2}{2}u''(t_n) + \frac{\tau^3}{6}u'''(t_n) \\
&\quad + \frac{1}{6}\int_0^\tau (\tau-|s|)^3 u^{(4)}(t_n+s)ds, \qquad (4.4.53)
\end{aligned}$$

and

$$\begin{aligned}
u(t_{n-1}) &= u(t_n) - \tau u'(t_n) + \frac{\tau^2}{2}u''(t_n) - \frac{\tau^3}{6}u'''(t_n) \\
&\quad - \frac{1}{6}\int_0^{-\tau} (\tau-|s|)^3 u^{(4)}(t_n+s)ds. \qquad (4.4.54)
\end{aligned}$$

Also apply (4.4.41) and (4.4.42) to u'' with $m = 2$:

$$u''(t_{n+1}) = u''(t_n) + \tau u'''(t_n) + \int_0^\tau (\tau-|s|) u^{(4)}(t_n+s)ds, \qquad (4.4.55)$$

$$u''(t_{n-1}) = u''(t_n) - \tau u'''(t_n) - \int_0^{-\tau} (\tau - |s|) u^{(4)}(t_n + s) ds. \quad (4.4.56)$$

Now we have (following from (4.4.53) and (4.4.54)) that

$$\begin{aligned} \delta_t^2 u(t_n) &= \frac{1}{\tau^2}(u(t_{n+1}) - 2u(t_n) + u(t_{n-1})) \\ &= u''(t_n) + \frac{1}{6\tau^2}\int_{-\tau}^{\tau} (\tau - |s|)^3 u^{(4)}(t_n + s) ds, \end{aligned} \quad (4.4.57)$$

and (following from (4.4.55) and (4.4.56))

$$\delta_\theta u''(t_n) = u''(t_n) + \theta \int_{-\tau}^{\tau} (\tau - |s|) u^{(4)}(t_n + s) ds. \quad (4.4.58)$$

Similarly as before, subtract (4.4.58) from (4.4.57) to obtain

$$\begin{aligned} \delta_t^2 u(t_n) - \delta_\theta u''(t_n) &= \frac{1}{6\tau^2}\int_{-\tau}^{\tau} (\tau - |s|)^3 u^{(4)}(t_n + s) ds \\ &\quad -\theta \int_{-\tau}^{\tau} (\tau - |s|) u^{(4)}(t_n + s) ds, \end{aligned}$$

and taking the norm on $\mathcal{L}^2(\Omega)$ we have

$$\|\delta_t^2 u(t_n) - \delta_\theta u''(t_n)\|_0 \leq \left(\frac{1}{6} + \theta\right) \tau \int_{-\tau}^{\tau} \|u^{(4)}(t_n + s)\|_0 ds. \quad (4.4.59)$$

Returning now to equation (4.4.40), and using the bounds (4.4.44), (4.4.46), (4.4.52) and (4.4.59), we have

$$\begin{aligned} \|r^n\|_0 &\leq \|\delta_t^2 (P - I) u(t_n)\|_0 + \|\delta_t^2 u(t_n) - \delta_\theta u''(t_n)\|_0 \\ &\quad +\beta_1 \|\delta_t (P - I) u(t_n) - \delta_\theta u'(t_n)\|_0 + \beta_1 \|\delta_t u(t_n) - \delta_\theta u'(t_n)\|_0 \\ &\leq \frac{Ch^{p+1}}{\tau} \int_{-\tau}^{\tau} \|u''(t_n + s)\|_{p+1} ds + \left(\frac{1}{6} + \theta\right) \tau \int_{-\tau}^{\tau} \|u^{(4)}(t_n + s)\|_0 ds \\ &\quad + \left(\frac{1}{4} + \theta\right) \tau \int_{-\tau}^{\tau} \|u'''(t_n + s)\|_0 ds + \frac{Ch^{p+1}}{2\tau} \int_{-\tau}^{\tau} \|u'(t_n + s)\|_{p+1} ds \\ &\leq C_2 \Bigg(h^{p+1} \tau^{-1} \int_{t_{n-1}}^{t_{n+1}} \|u''(s)\|_{p+1} ds + \tau \int_{t_{n-1}}^{t_{n+1}} \|u^{(4)}(s)\|_0 ds \\ &\quad + \tau \int_{t_{n-1}}^{t_{n+1}} \|u'''(s)\|_0 ds + h^{p+1} \tau^{-1} \int_{t_{n-1}}^{t_{n+1}} \|u'(s)\|_{p+1} ds \Bigg). \end{aligned}$$

where $C_2 = \max\{\frac{C}{2}, C, \frac{1}{6} + \theta, \beta_1 \left(\frac{1}{c} + \theta\right)\}$.

$\square$

Estimate for $\|R^n\|_0$

From the definition of R^n (with $0 \leq n \leq N-1$) we have that

$$\|R^n\|_0 \leq \tau\|r^0\|_0 + \tau\sum_{m=1}^{n}\|r^m\|_0 \leq \tau\|r^0\|_0 + \tau\sum_{m=1}^{N-1}\|r^m\|_0,$$

and then by the bounds derived in Lemma 4.4.3 and Lemma 4.4.4, we have that

$$\begin{aligned}\|R^n\|_0 \leq\ & \tau^2 C_1\left(\|u'''\|_{C(\bar{J},\mathcal{L}^2(\Omega))} + \|u''\|_{C(\bar{J},H^{p+1}(\Omega))}\right) + h^{p+1}C_1\|u'\|_{C(\bar{J},H^{p+1}(\Omega))}\\ & +C_2 h^{p+1}\sum_{m=1}^{N-1}\left(\int_{t_{m-1}}^{t_{m+1}}\|u''(s)\|_{p+1}ds + \int_{t_{m-1}}^{t_{m+1}}\|u'(s)\|_{p+1}ds\right)\\ & +\tau^2 C_2\sum_{m=1}^{N-1}\left(\int_{t_{m-1}}^{t_{m+1}}\|u'''(s)\|_0 ds + \int_{t_{m-1}}^{t_{m+1}}\|u^{(4)}(s)\|_0 ds\right).\end{aligned}$$

But

$$\begin{aligned}&\sum_{m=1}^{N-1}\left(\int_{t_{m-1}}^{t_{m+1}}\|u'''(s)\|_0 ds + \int_{t_{m-1}}^{t_{m+1}}\|u^{(4)}(s)\|_0 ds\right)\\ &\qquad \leq 2\left(\int_{t_0}^{t_N}\|u'''(s)\|_0 ds + \int_{t_0}^{t_N}\|u^{(4)}(s)\|_0 ds\right)\\ &\qquad = 2\left(\|u'''\|_{C(\bar{J},\mathcal{L}^2(\Omega))} + \|u^{(4)}\|_{L^1(\bar{J},\mathcal{L}^2(\Omega))}\right)\end{aligned}$$

and

$$\begin{aligned}&\sum_{m=1}^{N-1}\left(\int_{t_{m-1}}^{t_{m+1}}\|u''(s)\|_{p+1}ds + \int_{t_{m-1}}^{t_{m+1}}\|u'(s)\|_{p+1}ds\right)\\ &\qquad \leq 2\left(\int_{t_0}^{t_N}\|u''(s)\|_{p+1}ds + \int_{t_0}^{t_N}\|u''(s)\|_{p+1}ds\right)\\ &\qquad = 2\left(\|u''\|_{C(\bar{J},H^{p+1}(\Omega))} + \|u'\|_{C(\bar{J},H^{p+1}(\Omega))}\right).\end{aligned}$$

Now

$$\begin{aligned}\|R^n\|_0 \leq\ & \tau^2 C_1\left(\|u'''\|_{C(\bar{J},\mathcal{L}^2(\Omega))} + \|u''\|_{C(\bar{J},H^{p+1}(\Omega))}\right) + h^{p+1}C_1\|u'\|_{C(\bar{J},H^{p+1}(\Omega))}\\ & +2\tau^2 C_2\left(\|u'''(s)\|_{C(\bar{J},\mathcal{L}^2(\Omega))} + \|u^{(4)}(s)\|_{L^1(\bar{J},\mathcal{L}^2(\Omega))}\right)\\ & +2h^{p+1}C_2\left(\|u''\|_{C(\bar{J},H^{p+1}(\Omega))} + \|u'\|_{C(\bar{J},H^{p+1}(\Omega))}\right)\\ \leq\ & C_3\Bigg(\tau^2\left(\|u''\|_{C(\bar{J},H^{p+1}(\Omega))} + \|u'''\|_{C(\bar{J},\mathcal{L}^2(\Omega))} + \|u^{(4)}(s)\|_{L^1(\bar{J},\mathcal{L}^2(\Omega))}\right)\\ & +h^{p+1}\left(\|u'\|_{C(\bar{J},H^{p+1}(\Omega))} + \|u''\|_{C(\bar{J},H^{p+1}(\Omega))}\right)\Bigg),\end{aligned}$$

where $C_3 = \max\{C_1, C_1 + 2C_2, 2C_2\}$. Therefore

$$\sum_{n=0}^{N-1} \|R^n\|_0 \leq NC_3 \Bigg(\tau^2 \left(\|u''\|_{C(\bar{J},H^{p+1}(\Omega))} + \|u'''\|_{C(\bar{J},\mathcal{L}^2(\Omega))} + \|u^{(4)}(s)\|_{L^1(\bar{J},\mathcal{L}^2(\Omega))} \right) + h^{p+1} \left(\|u'\|_{C(\bar{J},H^{p+1}(\Omega))} + \|u''\|_{C(\bar{J},H^{p+1}(\Omega))} \right) \Bigg).$$

It is now possible to complete the **proof of Theorem 4.4.1**.

Consider Inequality (4.4.4):

$$\max_{1\leq n\leq N} \|e^n\|_0 \leq C^* \left(\|e^0\|_0 + \max_{1\leq n\leq N} \|e_p\|_0 + \tau \sum_{n=0}^{N-1} \|R^n\|_0 \right)$$

From Subsection 4.2.1 it follows that

$$\max_{1\leq n\leq N} \|e_p(t_n)\|_0 \leq Ch^{p+1} \max_{1\leq n\leq N} \|u(t_n)\|_{p+1}$$

and

$$\|e^0\|_0 = \|u^0 - P_2 u^0\|_0 \leq Ch^{p+1} \|u^0\|_{p+1}.$$

From the estimate for $\sum_{n=0}^{N-1} \|R^n\|_0$, the two results above and since

$$\|u(t_n) - u_h^n\|_0 = \|e^n\|_0 \leq \max_{1\leq n\leq N} \|e^n\|_0 = \max_{1\leq n\leq N} \|u(t_n) - u_h^n\|_0$$

we have the result. □

Remark Note that the constant $\tilde{C}$ depends on the length of the time interval T. This is not mentioned in [Kar11a].

4.5 Conclusion

In this section the problems that arose in proving the results in [Kar11a] are discussed, and a conclusion is also given.

4.5.1 Stability with strong damping

As mentioned, for the proof of Proposition 4.4.2 [Kar11a, Proposition 2], the following is stated: "We omit the proof since the result can be obtained by a slight modification of the arguments presented in [Kar11b]." The proof in [Kar11b] does not make use of the damping term in [Kar11a]. Also, the damping term provides an obstacle that is not so trivial to overcome, as discussed below. We give the modifications that have to be made to the proof of Proposition 4.4.2 to include a damping term.

Consider the weak variational form (4.1.2) at t_{n+1}, t_n and t_{n-1} (for $n = 1, 2, \ldots, N-1$). By using appropriate weights, we have that for every $v \in S^h$

$$(\delta_{\theta,\gamma} u''(t_n), v) + a(\delta_{\theta,\gamma} u'(t_n), v) + b(\delta_{\theta,\gamma} u(t_n), v) = (\delta_{\theta,\gamma} f(t_n), v). \tag{4.5.1}$$

Consider (4.2.2) again

$$(\delta_t^2 u_h^n, v) + a(\delta_{t,\gamma} u_h^n, v) + b(\delta_{\theta,\gamma} u_h^n, v) = (\delta_{\theta,\gamma} f(t_n), v) \text{ for all } v \in S^h$$

and subtract this from (4.5.1) to obtain

$$\begin{aligned}(\delta_{\theta,\gamma} u''(t_n) - \delta_t^2 u_h^n, v) &+ a(\delta_{\theta,\gamma} u'(t_n) - \delta_{t,\gamma} u_h^n, v) \\ &+ b(\delta_{\theta,\gamma} u(t_n) - \delta_{\theta,\gamma} u_h^n, v) = (\delta_{\theta,\gamma} f(t_n) - \delta_{\theta,\gamma} f(t_n), v)\end{aligned} \tag{4.5.2}$$

for every $v \in S^h$ and $n = 1, 2, \ldots, N-1$.

Now add and subtract the terms $\delta_t^2 Pu(t_n)$, $\delta_{t,\gamma} Pu(t_n)$ and $\delta_{\theta,\gamma}(Pu(t_n))$ in (4.5.2):

$$\begin{aligned}&(\delta_{\theta,\gamma} u''(t_n) - \delta_t^2 Pu(t_n) + \delta_t^2 Pu(t_n) - \delta_t^2 u_h^n, v) \\ &\quad + a(\delta_{\theta,\gamma} u'(t_n) - \delta_{t,\gamma} Pu(t_n) + \delta_{t,\gamma} Pu(t_n) - \delta_{t,\gamma} u_h^n, v) \\ &\quad\quad + b(\delta_{\theta,\gamma} u(t_n) - \delta_{\theta,\gamma}(Pu(t_n)) + \delta_{\theta,\gamma}(Pu(t_n)) - \delta_{\theta,\gamma} u_h^n, v) = 0.\end{aligned} \tag{4.5.3}$$

Using the definition of the projection P and the definition of r^n we have that (4.5.3) becomes

$$\begin{aligned}(\delta_t^2 e_n^h, v) + a(\delta_{t,\gamma} e_n^h, v) + b(\delta_{\theta,\gamma} e_n^h, v) &= (\delta_t^2 Pu(t_n) - \delta_{\theta,\gamma} u''(t_n), v) \\ &\quad + a(\delta_{t,\gamma} Pu(t_n) - \delta_{\theta,\gamma} u'(t_n), v) \\ &= (r^n, v) \end{aligned} \tag{4.5.4}$$

for every $v \in S^h$ and $n = 1, 2, \ldots, N-1$, where

$$\begin{aligned}\delta_{\theta,\gamma} e_n^h &= \theta(Pu(t_{n+1}) - u_h^{n+1}) + \left(\frac{1}{2} - 2\theta + \gamma\right)(Pu(t_n) - u_h^n) \\ &\quad + \left(\frac{1}{2} - \theta + \gamma\right)(Pu(t_{n-1}) - u_h^{n-1})\end{aligned}$$

and

$$\begin{aligned}&\delta_{t,\gamma} e_n^h \\ &\quad = \frac{1}{\tau}\Big(\gamma(Pu(t_{n+1}) - u_h^{n+1}) + (1-2\gamma)(Pu(t_n) - u_h^n) + (\gamma - 1)(Pu(t_{n-1}) - u_h^{n-1})\Big).\end{aligned}$$

We can rearrange (4.5.4) as follows:

$$\begin{aligned}(r^n, v) &= (\delta_t^2 e_n^h, v) + \tau\gamma a(\delta_t^2 e_n^h, v) + a(\delta_t^- e_n^h, v) \\ &\quad + \tau^2\left(\theta - \frac{\gamma}{2}\right) b(\delta_t^2 e_n^h, v) + \gamma b(e_{n+\frac{1}{2}}^h, v) + (1-\gamma) b(e_{n-\frac{1}{2}}^h, v)\end{aligned} \tag{4.5.5}$$

for every $v \in S^h$ and $n = 1, 2, \ldots, N-1$. Now multiply (4.5.5) by τ and then sum over $n = 1$ to $n = m$. Some terms will cancel and we are left with the remainder:

$$\begin{aligned}\tau\sum_{n=1}^{m}(r^n, v) &= \left(\frac{1}{\tau}\left(e_{m+1}^h - e_m^h\right), v\right) - \left(\frac{1}{\tau}\left(e_1^h - e_0^h\right), v\right) \\ &\quad + \tau\gamma a\left(\frac{1}{\tau}\left(e_{m+1}^h - e_m^h\right), v\right) - \tau\gamma a\left(\frac{1}{\tau}\left(e_1^h - e_0^h\right), v\right) + a\left(e_m^h - e_0^h, v\right) \\ &\quad + \tau^2\left(\theta - \frac{\gamma}{2}\right) b\left(\frac{1}{\tau}\left(e_{m+1}^h - e_m^h\right), v\right) - \tau^2\left(\theta - \frac{\gamma}{2}\right) b\left(\frac{1}{\tau}\left(e_1^h - e_0^h\right), v\right) \\ &\quad + \sum_{n=1}^{m}\left(\gamma\tau b\left(e_{n+\frac{1}{2}}^h, v\right) + (1-\gamma)\tau b\left(e_{n-\frac{1}{2}}^h, v\right)\right).\end{aligned} \tag{4.5.6}$$

The terms $\tau\gamma a\left(\frac{1}{\tau}\left(e_{m+1}^h - e_m^h\right), v\right) - \tau\gamma a\left(\frac{1}{\tau}\left(e_1^h - e_0^h\right), v\right)$ can be bounded in the same way as the corresponding terms for the bilinear form b, as in the proof of Proposition 4.4.2. The term $a\left(e_m^h - e_0^h, v\right)$ also needs to be bounded. The way that v is chosen, that is $v = e^h_{m+1} + e^h_m = 2(\Phi^{m+1} - \Phi^m)$, results in

$$a\left(e^h_m - e^h_0, e^h_{m+1} + e^h_m\right) = a\left(e^h_m, e^h_{m+1}\right) - a\left(e^h_0, e^h_{m+1}\right) + a\left(e^h_m, e^h_m\right) - a\left(e^h_0, e^h_m\right).$$

For the proof to be completed, we need to find an upper bound (with respect to the $\mathcal{L}^2$-norm) for $a\left(e^h_m - e^h_0, e^h_{m+1} + e^h_m\right)$. The damping that is present in the problem is strong damping: the bilinear form a is positive definite on $V(\Omega)$, i.e. there exists a constant C such that $a(u,u) \geq C\|u\|^2_{V(\Omega)}$. We also have that a is bounded on $V(\Omega)$, i.e. there exists a constant C_a such that $|a(u,v)| \geq C_a\|u\|_{V(\Omega)}\|v\|_{V(\Omega)}$.

An upper bound for $a\left(e^h_m - e^h_0, e^h_{m+1} + e^h_m\right)$ can therefore not be found.

4.5.2 Truncation error with general γ

As mentioned before, it is said in [Kar11a] that the proof of the main result Theorem 4.4.1 is done with $\gamma = \dfrac{1}{2}$ “for the sake of conciseness” and that “the general case can be proved without major difficulties”. In [Kar12] the general scheme is used to prove the main result, but in that article no damping is present. When [Kar12] was consulted to aid in the proofs for the general case $\left(\gamma \geq \dfrac{1}{2}\right)$, some irregularities were found in the proofs.

In this section we consider Lemma 4.4.4, the estimate for the terms r^n. In the articles [Kar11a] and [Kar11b], the proof is only done for the case when $\gamma = \frac{1}{2}$. The article of [Kar12] was therefore consulted, even though no damping is present. It was found that in the proof of Lemma 4.4.4 we need the following result: (see [Kar12, p. 697])

$$\delta_{\theta,\gamma}u''(t_n) = u''(t_n) + \theta\int_{-\tau}^{\tau}(\tau - |s|)u^{(4)}(t_n+s)ds + \left(\gamma - \frac{1}{2}\right)\int_{-\tau}^{\tau}u'''(t_n+s)ds. \quad (4.5.7)$$

It is said in [Kar12] that this “ …can be easily obtained from Taylor expansions of $u''(t_{n+1})$ and $u''(t_{n-1})$ about $u''(t_n)$ …”. Let's consider this. These Taylor expansions are given by

$$u''(t_{n+1}) = u''(t_n) + \tau u'''(t_n) + \int_0^{\tau}(\tau - |s|)u^{(4)}(t_n+s)ds,$$

and

$$u''(t_{n-1}) = u''(t_n) - \tau u'''(t_n) - \int_0^{-\tau}(\tau - |s|)u^{(4)}(t_n+s)ds.$$

From definition we have that

$$\delta_{\theta,\gamma}u''(t_n) = \theta u''(t_{n+1}) + \left(\frac{1}{2} - 2\theta + \gamma\right) u''(t_n) + \left(\frac{1}{2} + \theta - \gamma\right) u''(t_{n-1}),$$

and combining this with the Taylor expansions above we will have (after cancellations)

$$\begin{aligned}
\delta_{\theta,\gamma}u''(t_n) &= \theta\left(u''(t_n) + \tau u'''(t_n) + \int_0^\tau (\tau - |s|)u^{(4)}(t_n+s)ds\right) + \left(\frac{1}{2} - 2\theta + \gamma\right) u''(t_n) \\
&\quad + \left(\frac{1}{2} + \theta - \gamma\right)\left(u''(t_n) - \tau u'''(t_n) - \int_0^{-\tau} (\tau - |s|)u^{(4)}(t_n+s)ds\right) \\
&= u''(t_n) + \theta \int_{-\tau}^{\tau} (\tau - |s|)u^{(4)}(t_n+s)ds + \left(\gamma - \frac{1}{2}\right)\tau u'''(t_n) \\
&\quad + \left(\frac{1}{2} - \gamma\right)\int_{-\tau}^{0} (\tau - |s|)u^{(4)}(t_n+s)ds. \qquad (4.5.8)
\end{aligned}$$

The next step will be to subtract (4.5.8) from $\delta_t^2 u(t_n)$. This will give

$$\begin{aligned}
\delta_t^2 u(t_n) - \delta_{\theta,\gamma}u''(t_n) &= \frac{1}{6\tau^2}\int_{-\tau}^{\tau} (\tau - |s|)^3 u^{(4)}(t_n+s)ds - \theta\int_{-\tau}^{\tau} (\tau - |s|)u^{(4)}(t_n+s)ds \\
&\quad - \left(\gamma - \frac{1}{2}\right)\tau u'''(t_n) - \left(\frac{1}{2} - \gamma\right)\int_{-\tau}^{0} (\tau - |s|)u^{(4)}(t_n+s)ds.
\end{aligned}$$

Taking the $\mathcal{L}^2$-norm we therefore have

$$\begin{aligned}
\|\delta_t^2 u(t_n) - \delta_{\theta,\gamma}u''(t_n)\|_0 &\leq \frac{\tau^3}{6\tau^2}\int_{-\tau}^{\tau} \|u^{(4)}(t_n+s)\|_0 ds + \theta\tau\int_{-\tau}^{\tau} \|u^{(4)}(t_n+s)\|_0 ds \\
&\quad + \left|\gamma - \frac{1}{2}\right|\tau\|u'''(t_n)\|_0 + \tau\left|\frac{1}{2} - \gamma\right|\int_{-\tau}^{0} \|u^{(4)}(t_n+s)\|_0 ds \\
&= \left(\frac{1}{6} + \theta\right)\tau\int_{-\tau}^{\tau} \|u^{(4)}(t_n+s)\|_0 ds \\
&\quad + \left|\gamma - \frac{1}{2}\right|\tau\|u'''(t_n)\|_0 + \tau\left|\frac{1}{2} - \gamma\right|\int_{-\tau}^{0} \|u^{(4)}(t_n+s)\|_0 ds. \qquad (4.5.9)
\end{aligned}$$

Recall that we want estimates for the four terms on the right hand side of the inequality below:

$$\begin{aligned}
\|r^n\|_0 &\leq \|\delta_t^2 Pu(t_n) - \delta_t^2 u(t_n)\|_0 + \|\delta_t^2 u(t_n) - \delta_{\theta,\gamma}u''(t_n)\|_0 \\
&\quad + \beta_1\|\delta_{t,\gamma}Pu(t_n) - \delta_{t,\gamma}u(t_n)\|_0 + \beta_1\|\delta_{t,\gamma}u(t_n) - \delta_{\theta,\gamma}u'(t_n)\|_0 \\
&= \|\delta_t^2(P-I)u(t_n)\|_0 + \|\delta_t^2 u(t_n) - \delta_{\theta,\gamma}u''(t_n)\|_0 \\
&\quad + \beta_1\|\delta_{t,\gamma}(P-I)u(t_n)\|_0 + \beta_1\|\delta_{t,\gamma}u(t_n) - \delta_{\theta,\gamma}u'(t_n)\|_0. \qquad (4.5.10)
\end{aligned}$$

The first term is estimated in Lemma 4.4.4, and the third term can be estimated in a similar way as in Lemma 4.4.4. We have:

$$\begin{aligned}&\|\delta_{t,\gamma}(P-I)u(t_n)\|_0\\ &\quad\le \frac{Ch^{p+1}}{\tau}\left(\gamma\int_{-\tau}^{\tau}\|u'(t_n+s)\|_{p+1}ds+\int_{-\tau}^{0}\|u'(t_n+s)\|_{p+1}ds\right). \end{aligned} \tag{4.5.11}$$

The second term was estimated above in Equation (4.5.9).

A similar result can be obtained for the fourth term on the right hand side of (4.5.10). After applying (4.4.41) and (4.4.42) to u' with $m=2$ and cancelling terms:

$$\begin{aligned}\delta_{\theta,\gamma}u'(t_n) &= u'(t_n)+\theta\int_{-\tau}^{\tau}(\tau-|s|)u'''(t_n+s)ds+\left(\gamma-\frac{1}{2}\right)\tau u''(t_n)\\ &\quad+\left(\frac{1}{2}-\gamma\right)\int_{-\tau}^{0}(\tau-|s|)u'''(t_n+s)ds. \end{aligned} \tag{4.5.12}$$

Also, after applying (4.4.41) and (4.4.42) with $m=3$ and cancelling terms:

$$\begin{aligned}\delta_{t,\gamma}u(t_n) &= u'(t_n)-\frac{\tau}{2}u''(t_n)+\frac{\gamma}{2\tau}\int_{0}^{\tau}(\tau-|s|)^2u'''(t_n+s)ds\\ &\quad+\frac{(\gamma-1)}{2\tau}\int_{0}^{-\tau}(\tau-|s|)^2u'''(t_n+s)ds. \end{aligned} \tag{4.5.13}$$

Subtracting (4.5.12) from (4.5.13) and taking the $\mathcal{L}^2$-norm we obtain:

$$\begin{aligned}&\|\delta_{t,\gamma}u(t_n)-\delta_{\theta,\gamma}u'(t_n)\|_0\\ &\le \tau\gamma\|u''(t_n)\|_0+\frac{\gamma\tau}{2}\int_{0}^{\tau}\|u'''(t_n+s)\|_0ds+\frac{|\gamma-1|\,\tau}{2}\int_{-\tau}^{0}\|u'''(t_n+s)\|_0ds\\ &\quad+\theta\tau\int_{-\tau}^{\tau}\|u'''(t_n+s)\|_0ds+\left|\frac{1}{2}-\gamma\right|\tau\int_{-\tau}^{0}\|u'''(t_n+s)\|_0ds.\end{aligned}$$

Although the result above is not exactly the same as the estimate in [Kar12], it is essentially the same. It still provides a proof for Lemma 4.4.4 for the case $\gamma>\dfrac{1}{2}$.

4.5.3 Concluding remarks

As mentioned in the beginning of this chapter, additional work was required to supplement what is provided in the article [Kar11a] specifically with regards to certain proofs which the author claims are easy extensions. Certain errors were discovered, for instance in [Kar11a] there is said that the proof of Proposition 4.4.2 “ ... can be obtained by a slight modification of the arguments presented in [Kar11b].” The modification turns out to be a lot more comprehensive than expected. The results that could be proved and those that could not have been summarised in Table 4.2.

		Stability (Proposition 4.4.2)		$\|r^n\|_0$ (Lemma 4.4.4)	
		$\gamma = \frac{1}{2}$	$\gamma > \frac{1}{2}$	$\gamma = \frac{1}{2}$	$\gamma > \frac{1}{2}$
Damping	None	Yes	Yes	No	No
	Weak	No	No	Yes	Yes
	Strong	No	No	No	No

Table 4.2: Summary

The assumptions that are made on the regularity of the solution u of Problem DPL are very restrictive. It is not mentioned that compatibility conditions must be imposed on the boundary and initial data to yield these higher regularity properties (see Section 2.4). These compatibility conditions required to obtain higher regularity place serious restrictions (often not realistic) on the initial and boundary data.

It is also mentioned in the conclusion in [Kar11a] that the analysis in the article extends to discontinuous Galerkin finite element methods, and that the convergence results hold true, as long as the corresponding discontinuous Galerkin bilinear forms are symmetric, continuous, positive definite and consistent. In the articles [Kar11b] and [Kar12], this convergence analysis has been extended to the interior penalty symmetric discontinuous Galerkin finite element method in [GSS06] and [GS09]. However, this is only done for the wave equation with no damping. Also, the results that are obtained rely on the same higher regularity assumptions being maintained on the solution u as in [Kar11a].

Chapter 5

The discontinuous Galerkin Finite Element method for the wave equation

5.1 Introduction

In this chapter we investigate the work done in articles [GSS06] and [GS09]. In [GSS06] semi-discrete error estimates are derived for the symmetric interior penalty discontinuous Galerkin finite element method for the wave equation (without damping) in both the energy norm and the $\mathcal{L}^2$-norm. In [GS09] the fully discrete error estimate in the $\mathcal{L}^2$-norm is derived.

We read in [GSS06]: "... continuous Galerkin methods impose significant restrictions on the underlying mesh and discretization; in particular, they do not accommodate hanging nodes."

"To avoid these difficulties, we consider instead discontinuous Galerkin (DG) methods. Based on discontinuous finite element spaces, these methods easily handle elements of various types and shapes, irregular nonmatching grids, and even locally varying polynomial order; thus, they are ideally suited for hp-adaptivity. Here continuity is weakly enforced across mesh interfaces by adding suitable bilinear forms, so-called numerical fluxes, to standard variational formulations. These fluxes are easily included within an existing conforming finite element code."

It is clear that the discontinuous Galerkin approximation has certain advantages. However, the aim of the investigation of DG methods in this dissertation is to examine if the regularity assumptions needed on the solution for optimal error estimates are less restrictive than those made to obtain optimal error estimates for the continuous Galerkin finite element method. A glance at Theorem 4.2 in the article [GSS06] and Theorem 3.4 in [GS09] reveals that this is not the case; in [GS09, Theorem 3.4] it is assumed that the solution has a fourth order time derivative. Nevertheless, the possibility exists that suboptimal error estimates were obtained by [GSS06] and therefore it was decided to study the articles in detail.

The investigation of [GSS06] was a substantial undertaking. The proofs done here are given in greater detail making them much more accessible. In some cases additional explanations were provided.

5.1.1 Model problem

In [GSS06], the model under consideration is a special case of the multidimensional ($\mathbb{R}^d$ with $d = 1, 2$ or 3) wave equation (Problem MW) with $k = 0$, i.e. there is no damping present. The problem is given below partly in the notation of [GSS06] for convenience.

Problem W (Special case):

Given functions f, u_0 and v_0, find u defined on $\bar{\Omega} \times \bar{J}$ such that

$$\begin{aligned} \partial_t^2 u - \nabla \cdot (c \nabla u) &= f \quad \text{in} \quad \Omega \times J, \\ u &= 0 \quad \text{on} \quad \partial\Omega - \Sigma, \\ (c\ \nabla u) \cdot \mathbf{n} &= 0 \quad \text{on} \quad \Sigma, \end{aligned}$$

while $u(\cdot, 0) = u_0$ and $\partial_t u(\cdot, 0) = v_0$. Note that $J = (0, T)$ in [GSS06].

Note that the matrix A is cI in this case. From the properties of the matrix A assumed for Problem MW, we have that there are real numbers c_* and c^* such that (in the notation of [GSS06])

$$0 < c_* \leq c(x) \leq c^* < \infty \text{ for } x \in \bar{\Omega}. \tag{5.1.1}$$

Remark For the problem in [GSS06], $u = 0$ on the entire boundary $\partial\Omega$.

5.1.2 Weak variational form and existence

The weak variational form of Problem W was derived in Section 2.5.1 and is given below in Problem WW for completeness and comparison with that of the weak variational form of Problem W given in [GSS06]. The assumptions made for existence in [GSS06] is compared to that of [VV02].

Note that the space $V(\Omega)$ is the closure of the space of test functions

$$\mathcal{T}(\Omega) := \{v \in C^1(\bar{\Omega}) : v = 0 \text{ on } \partial\Omega - \Sigma\}$$

in $H^1(\Omega)$. In [GSS06], $V(\Omega) = H_0^1(\Omega)$, since $u = 0$ on the entire boundary $\partial\Omega$.

Weak variational form in [GSS06]:

Find $u \in \mathcal{L}^2(J, V(\Omega))$, $u' \in \mathcal{L}^2(J, \mathcal{L}^2(\Omega))$ and $u'' \in \mathcal{L}^2(J, H^{-1}(\Omega))$ such that $u(0) = u_0$ and $u'(0) = v_0$, and

$$\langle u'', v\rangle + b(u, v) = (f, v) \text{ for all } v \in V(\Omega),$$

where $\langle \cdot, \cdot\rangle$ is the duality pairing between $H^{-1}(\Omega)$ and $V(\Omega)$.

The bilinear form b is given by

$$b(u, v) = (c\nabla u, \nabla v).$$

Remark From the formulation of Problem GSS06W we see that u'' is assumed to be a distribution, not a function. This assumption, although it is a weak assumption for the existence of a solution, serves no purpose for the outcome of the article [GSS06], i.e. to obtain convergence of the semi-discrete discontinuous Galerkin method. Higher regularity is assumed for the convergence of the semi-discrete discontinuous Galerkin method regardless of the assumptions made for existence. Assuming $u'' \in \mathcal{L}^2(J, \mathcal{L}^2(\Omega))$ would be more constructive.

For the existence of a solution it is assumed in [GSS06] that $f \in \mathcal{L}^2(J, \mathcal{L}^2(\Omega))$, $u_0 \in V(\Omega)$ and $v_0 \in \mathcal{L}^2(\Omega)$, and then [LM72] is cited; that a solution of the problem exists. As mentioned, this serves no purpose.

As is shown in Chapter 2 (Section 2.5.1), the weak variational form of Problem W (undamped case) is given by:

Problem WW (Special case):

Find u such that for each $t \in J$, $u(t) \in V(\Omega)$, $u''(t) \in \mathcal{L}^2(\Omega)$ and

$$(u''(t), v) + b(u(t), v) = (f(\cdot, t), v)_\Omega \text{ for each } v \in V(\Omega)$$

while $u(0) = u_0$, and $u'(0) = v_0$.

Existence of a unique solution of Problem WW is proved in Section 2.5. In particular, sufficient conditions for this result is given and we have

$$u \in C^1(J, V(\Omega)) \cap C^2\left(J, \mathcal{L}^2(\Omega)\right).$$

In Section 5.2 Problem WW is discretised in space by using the symmetric interior penalty discontinuous Galerkin finite element approximation (SIPDG) from [GSS06]. The main results concerning the semi-discrete approximation of the article are stated in Section 5.3. The two main results, Theorem 5.3.2 and Theorem 5.3.3 are proved in Sections 5.4 and 5.5 respectively. In Section 5.6 the fully discrete error estimates from [GS09] are given and proved. Some numerical experiments done in [GSS06] are discussed in Section 5.7.

5.2 Discontinuous Galerkin discretisation

In this section Problem WW is discretised in space by using the SIPDG method from [GSS06]. We first provide the definitions needed for the SIPDG method.

5.2.1 Preliminaries

Consider shape-regular meshes $\mathcal{M}_h$ that partition the domain Ω into disjoint elements $\{E\}$. Let h_E denote the diameter of element E, and let the mesh size h be given by $h = \max_{E \in \mathcal{M}_h} h_E$. It is assumed that the partition is aligned with the discontinuities of $\sqrt{c}$. Generally, irregular meshes with hanging nodes are allowed, however, in this article it is assumed that the local mesh sizes are of bounded variation: there exists a constant $\kappa > 0$ (depending only on the shape-regularity of the mesh) such that

$$\kappa h_E \leq h_{E'} \leq \kappa^{-1} h_E \tag{5.2.1}$$

for all **neighboring** elements E and E'.

Note The term "shape-regular mesh" refers here to the following definition: A family of partitions Ω_h is called shape regular provided that there exists a number $k > 0$ such that every $E \in \Omega_h$ contains a circle of radius ρ_E with $\rho_E \geq \frac{h_E}{k}$, where h_E is half the diameter of E.

Definition 5.2.1. *Interior face of $\mathcal{M}_h$*
The nonempty interior of $\partial E \cap \partial E'$, where E and E' are two adjacent elements of $\mathcal{M}_h$, is called an interior face and denoted by F.

Definition 5.2.2. *Boundary face of $\mathcal{M}_h$*
The nonempty interior of $\partial E \cap \partial \Omega$ which consists of entire faces of ∂E, is called a boundary face and also denoted by F.

Definition 5.2.3. *The set $\mathcal{F}_h$*
Let $\mathcal{F}_h^{\mathcal{I}}$: The set of all interior faces of $\mathcal{M}_h$;
$\mathcal{F}_h^{\mathcal{B}}$: The set of all boundary faces of $\mathcal{M}_h$.
Define $\mathcal{F}_h = \mathcal{F}_h^{\mathcal{I}} \cup \mathcal{F}_h^{\mathcal{B}}$.

Remark Any element of $\mathcal{F}_h$ is generally referred to as a "face" in both two and three dimensions, although in reality a "face" in two dimensions is an edge.

Definition 5.2.4. *Jump and average for piecewise smooth functions*
Suppose v is any piecewise smooth function. Let $F_I \in \mathcal{F}_h^{\mathcal{I}}$ be an interior face shared by two neighboring elements E^+ and E^- and let $x \in F_I$; let $\mathbf{n}^{\pm}$ denote the unit outward normal vectors on the boundaries $\partial E^{\pm}$. Let $v^{\pm}$ denote the trace (Section A.4) of v taken from within $E^{\pm}$, and then define the jump and average of v at $x \in F_I$ by

$$[\![v]\!] := v^+\mathbf{n}^+ + v^-\mathbf{n}^-, \qquad \{\!\{v\}\!\} := \frac{1}{2}\left(v^+ + v^-\right)$$

respectively.

On every boundary face $F_B \in \mathcal{F}_h^{\mathcal{B}}$ let $[\![v]\!] := v\mathbf{n}$ and $\{\!\{v\}\!\} := v$, with $\mathbf{n}$ the unit outward normal vector on $\partial\Omega$.

Definition 5.2.5. *Average for piecewise smooth vector-valued functions*
Suppose $\mathbf{q}$ is any piecewise smooth vector-valued function. Define the average across interior faces and boundary faces by

$$\{\!\{\mathbf{q}\}\!\} := \frac{1}{2}\left(\mathbf{q}^+ + \mathbf{q}^-\right) \quad \text{and} \quad \{\!\{\mathbf{q}\}\!\} := \mathbf{q},$$

respectively.

5.2.2 Discretization in space

Define the finite element space by

$$V^h := \{v \in \mathcal{L}^2(\Omega) : v|_E \in \mathcal{S}^j(E) \text{ for all } E \in \mathcal{M}_h\},$$

where $\mathcal{S}^j(E)$ is the space $\mathcal{P}^j(E)$ of polynomials of total degree at most j on E if E is a triangle or a tetrahedra, or the space $Q^j(E)$ of polynomials of total degree at most j on E if E is a parallelogram or a parallelepiped.

Definition 5.2.6. *Bilinear form b_h*
The discrete bilinear form b_h on $V^h \times V^h$ is given by

$$\begin{aligned} b_h(u,v) \quad := \quad & \sum_{E\in\mathcal{M}_h} \int_E c\nabla u \cdot \nabla v dx - \sum_{F\in\mathcal{F}_h} \int_F [\![u]\!] \cdot \{\!\{c\nabla v\}\!\} ds \\ & - \sum_{F\in\mathcal{F}_h} \int_F [\![v]\!] \cdot \{\!\{c\nabla u\}\!\} ds + \sum_{F\in\mathcal{F}_h} \int_F \boldsymbol{\eta}\, [\![v]\!] \cdot [\![u]\!] ds, \end{aligned} \qquad (5.2.2)$$

where $\boldsymbol{\eta}|_F := \beta\, \bar{\boldsymbol{c}}\, \bar{\boldsymbol{h}}^{-1}$ on each $F \in \mathcal{F}_h$ (referred to as the interior penalty stabilization function) with

$$\bar{\boldsymbol{h}}|_F = \begin{cases} \min\{h_E, h_{E'}\}, & F \in \mathcal{F}_h^{\mathcal{I}}, F = \partial E \cap \partial E' \\ h_E, & F \in \mathcal{F}_h^{\mathcal{B}}, F = \partial E \cap \partial \Omega \end{cases}$$

and for $x \in F$

$$\bar{\boldsymbol{c}}|_F = \begin{cases} \max\{c|_E(x), c|_{E'}(x)\}, & F \in \mathcal{F}_h^{\mathcal{I}}, F = \partial E \cap \partial E' \\ c|_E(x), & F \in \mathcal{F}_h^{\mathcal{B}}, F = \partial E \cap \partial \Omega, \end{cases}$$

and $\beta > 0$ a parameter independent of the local mesh sizes and the coefficient c, where E and E' are neighbouring elements.

Remark The last three terms in (5.2.2) correspond to jump and flux terms at element boundaries. These terms vanish when $u, v \in V(\Omega) \cap H^{1+m}(\Omega)$ for $m > \dfrac{1}{2}$. The semi-discrete DG approximation given below in Problem WDG is therefore consistent with Problem WW (Section 2.5.1).

Problem WDG

Given a function $f : J \to \mathcal{L}^2(\Omega)$ and a partition $\mathcal{M}_h$ of Ω, find a function $u_h \in C(\bar{J}, V^h)$ such that

$$\begin{aligned} (u_h''(t), v) + b_h(u_h(t), v) &= (f(t), v) \text{ for all } v \in V^h \\ u_h(0) = u_0^h, &\qquad u_h'(0) = u_1^h. \end{aligned} \tag{5.2.3}$$

Notation Let P_2 denote the $\mathcal{L}^2$-projection onto V^h.

Remark It was mentioned previously (see Section 3.1 and Section 3.1.2) that most publications choose the initial conditions u_0^h and u_1^h to be the $\mathcal{L}^2$-projections of the initial conditions u_0 and v_0 respectively, that is, $u_h(0) = P_2 u^0$ and $u_h'(0) = P_2 v^0$. [GSS06] makes the same choice.

Now consider the following result, referred to as the stability result for the DG form b_h in [GSS06]. (Recall that the dimension is denoted by d).

Notation To avoid confusion, note that $\|\cdot\|_{0,F}$ is the norm for $\mathcal{L}^2(F)$ and $\|\cdot\|_{0,E}$ is the norm for $\mathcal{L}^2(E)^d$

Lemma 5.2.7. *[GSS06, Lemma 3.1]*
There exists a threshold value $\beta_{\min} > 0$ which depends only on the shape-regularity of the mesh, the approximation order p , the dimension d, and the bounds in (5.1.1), such that for $\beta \geq \beta_{\min}$ and $v \in V^h$

$$b_h(v, v) \geq C_{pd} \left(\sum_{E \in \mathcal{M}_h} \|\sqrt{c}\nabla v\|_{0,E}^2 + \sum_{F \in \mathcal{F}_h} \|\sqrt{\boldsymbol{\eta}}[[v]]\|_{0,F}^2 \right)$$

where the constant C_{pd} is independent of c and h.

The proof of this lemma can be found in [ABCM02]. A slightly more general stability result is proven in [GSS06, Lemma 4.4] (Lemma 5.3.7 in this dissertation).

5.3 Main results

Optimal error estimates are derived in [GSS06] for the symmetric interior penalty DG finite element method. Two main results are obtained: an error estimate with respect to the DG energy norm and an error estimate with respect to the $\mathcal{L}^2$-norm. We follow the order of [GSS06], where these two main results are stated in this section and the proofs postponed to Section 5.4 and Section 5.5.

We now introduce what the article [GSS06] refers to as the DG energy space, together with its norm.

Definition 5.3.1.

$$V(h) = V(\Omega) + V^h.$$

On $V(h)$ define the DG energy norm

$$\|v\|_{V(h)}^2 := \sum_{E\in\mathcal{M}_h} \|c^{\frac{1}{2}}\nabla v\|_{0,E}^2 + \sum_{F\in\mathcal{F}_h} \|\boldsymbol{\eta}^{\frac{1}{2}}[\![v]\!]\|_{0,F}^2.$$

Note that in the rest of this chapter, $J = (0, T)$.

Theorem 5.3.2. *[GSS06, Theorem 4.1]*
Let the solution u of Problem W satisfy

$$u \in C(J, H^{1+m}(\Omega)), \quad u' \in C(J, H^{1+m}(\Omega)), \quad u'' \in \mathcal{L}^1(J, H^m(\Omega))$$

for $m > \frac{1}{2}$, and let u_h be the semi-discrete DG approximation in Problem WDG, with $\beta \geq \beta_{\min}$. Then the error $e_h = u - u_h$ satisfies

$$\|e_h'\|_{C(J,\mathcal{L}^2(\Omega))} + \|e_h\|_{C(J,V(h))} \leq \mathcal{C}_1 \left(\|e_h'(0)\|_0 + \|e_h(0)\|_{V(h)}\right)$$
$$+\mathcal{C}_2 h^{\min\{m,j\}} \left(\|u\|_{V(h)} + T\|u'\|_{V(h)} + \|u''\|_{\mathcal{L}^1(J,H^m(\Omega))}\right),$$

with constants $\mathcal{C}_1$ and $\mathcal{C}_2$ independent of T and h.

Remarks

1. In order to obtain an **optimal** error estimate with respect to the $\mathcal{L}^2$-norm, elliptic regularity is assumed; i.e. there is a stability constant C_S such that for any $f \in \mathcal{L}^2(\Omega)$ the solution u of the problem

$$b(u, v) = (f, v) \quad \text{for all} \quad v \in V(\Omega) \tag{5.3.1}$$

belongs to $H^2(\Omega)$ and satisfies the bound

$$\|u\|_2 \leq C_S\|f\|_0. \tag{5.3.2}$$

2. Elliptic regularity as stated above is assumed to obtain an **optimal** error estimate with respect to the $\mathcal{L}^2$-norm. This is relevant in the proof of Lemma 5.5.2 in Section 5.5, where a higher order error estimate in the $\mathcal{L}^2$-norm is obtained. This is where we differ from [GSS06]: we also give an error estimate that is not necessarily **optimal**, but it is of practical importance.

Theorem 5.3.3. *[GSS06, Theorem 4.2]*
Let the solution u satisfy

$$u \in C(J, H^{1+m}(\Omega)), \quad u' \in C(J, H^{1+m}(\Omega)), \quad u'' \in \mathcal{L}^1(J, H^m(\Omega))$$

for $m > \frac{1}{2}$, and let u_h be the semi-discrete DG approximation in Problem WDG, with $\beta \geq \beta_{\min}$.

1. *The error $e_h = u - u_h$ satisfies*

$$\begin{aligned} \|e_h\|_{C(J,\mathcal{L}^2(\Omega))} \leq\ & \mathcal{C}_3 h^{\min\{m,j\}} \left(\|u_0\|_{1+m} + \|u\|_{C(J,H^{1+m}(\Omega))} + \sqrt{T}\|u'\|_{C(J,H^{1+m}(\Omega))}\right) \\ & + \mathcal{C}_4 \left(\|e_h(0)\|_0 + T\|e_h'(0)\|_0\right). \end{aligned} \tag{5.3.3}$$

with a constant C that is independent of T and h.

2. *If elliptic regularity as in (5.3.1) and (5.3.2) is assumed, then the error $e_h = u - u_h$ satisfies*

$$\begin{aligned} \|e_h\|_{C(J,\mathcal{L}^2(\Omega))} \leq\ & \mathcal{C}_5 h^{\min\{m,j\}+1} \left(\|u_0\|_{1+m} + \|u\|_{C(J,H^{1+m}(\Omega))} + \sqrt{T}\|u'\|_{C(J,H^{1+m}(\Omega))}\right) \\ & + \mathcal{C}_4 \left(\|e_h(0)\|_0 + T\|e_h'(0)\|_0\right). \end{aligned} \tag{5.3.4}$$

with constants $\mathcal{C}_i$ $(i = 3, 4, 5)$ that is independent of T and h.

We have now stated the main results of the article [GSS06]. In Subsections 5.3.1 to 5.3.3 some preliminary results are obtained in order to prove these main results: Theorem 5.3.2 is proved in Section 5.4 and Theorem 5.3.3 is proved in Section 5.5.

Remark We remark here that the proof of Theorem 5.3.3 ([GSS06, Theorem 4.2]) in the article [GSS06] follows an argument in [Bak76] for conforming finite element approximations. However, the proof of this theorem in this dissertation in Section 5.5 follows the work already done in Chapter 3 on the article [BV13].

For convenience we include Table 5.1.

Notation	Refer to
DG bilinear form b_h	Definition 5.2.6, page 95
Space $V(h)$ and DG energy norm $\|\cdot\|_{V(h)}$	Page 97
Auxiliary DG form $\tilde{b}_h$	Equation (5.3.9), page 102
Lifted function $\ell(v)$	Definition 5.3.4, page 99
The form r_h	Equation (5.3.13), page 107
Operator $\mathcal{P}_h$	Equation (5.5.1), page 116
Broken norm $\|\cdot\|_*$	Equation (5.6.2), page 123
Truncation: r^n	Equations (5.6.6) and (5.6.7), page 126

TABLE 5.1: Some notation in Chapter 5

5.3.1 Extension of the DG form b_h

The DG form b_h is defined on the space $V^h \times V^h$. However, it does not extend in a standard way to a continuous form on the larger space $V(h) \times V(h)$, since in general the average $\{\!\{c\nabla v\}\!\}$ on a face $F \in \mathcal{F}_h$ is not well defined for $v \in H^1(\Omega)$. To get around this difficulty, [GSS06] extends the form b_h to the space $V(h) \times V(h)$ by making use of the lifting operators defined in [ABCM02] and the approach in [PS02].

Definition 5.3.4. *Lifted function*
The lifted function $\ell(v) \in (V^h)^d (d = 2, 3)$ is defined for $v \in V(h)$ by

$$\int_\Omega \ell(v) \cdot \mathbf{w} dx = \sum_{F \in \mathcal{F}_h} \int_F [\![v]\!] \cdot \{\!\{c\mathbf{w}\}\!\} dA \quad \text{for all} \quad \mathbf{w} \in (V^h)^d.$$

Remarks

1. The notation for the lifting operator differs here from [GSS06], where it is denoted by $\mathcal{L}_c$. In [ABCM02] it is denoted by r (on $\Gamma = \cup_{E \in \mathcal{M}_h} \partial E$) and l (on $\Gamma^0 = \Gamma \backslash \partial\Omega$).

2. The following definition for the norm of a lifted element is not explicitly given in [GSS06].

3. The existence of $\ell(v)$ is guaranteed from Riesz's Theorem [Kre78, Theorem 3.8-1, p. 188].

Definition 5.3.5. *Norm of a lifted element*

For $v \in V(h)$ we define

$$\|\ell(v)\|_0 := \max_{\mathbf{w}\in(V^h)^d} \frac{\sum_{F\in\mathcal{F}_h}\int_F [\![v]\!]\cdot\{\!\{c\mathbf{w}\}\!\}dA}{\|\mathbf{w}\|_0}$$

Lemma 5.3.6. *[GSS06, Lemma 4.3]*

There exists a constant C_{inv} which depends only on the shape-regularity of the mesh, the approximation order j, and the dimension d such that

$$\|\ell(v)\|_0^2 \le \beta^{-1} c^* C_{inv}^2 \sum_{F\in\mathcal{F}_h} \|\boldsymbol{\eta}^{\frac{1}{2}}[\![v]\!]\|_{0,F}^2$$

for any $v \in V(h)$. Moreover, if $c^{\frac{1}{2}}$ is piecewise constant, with discontinuities aligned with the finite element mesh $\mathcal{M}_h$, then

$$\|c^{-\frac{1}{2}}\ell(v)\|_0^2 \le \beta^{-1} C_{inv}^2 \sum_{F\in\mathcal{F}_h} \|\boldsymbol{\eta}^{\frac{1}{2}}[\![v]\!]\|_{0,F}^2.$$

Proof. **1.** First estimate

By making use of the Cauchy-Schwartz inequality, we have

$$\begin{aligned}
\|\ell(v)\|_0 &= \max_{\mathbf{w}\in(V^h)^d} \frac{\sum_{F\in\mathcal{F}_h}\int_F [\![v]\!]\cdot\{\!\{c\mathbf{w}\}\!\}dA}{\|\mathbf{w}\|_0} \\
&= \max_{\mathbf{w}\in(V^h)^d} \frac{\sum_{F\in\mathcal{F}_h}\int_F \boldsymbol{\eta}^{\frac{1}{2}}\boldsymbol{\eta}^{-\frac{1}{2}}[\![v]\!]\cdot\{\!\{c\mathbf{w}\}\!\}dA}{\|\mathbf{w}\|_0} \\
&\le \max_{\mathbf{w}\in(V^h)^d} \frac{\left(\sum_{F\in\mathcal{F}_h}\int_F \boldsymbol{\eta}|[\![v]\!]|^2\,dA\right)^{\frac{1}{2}} \left(\sum_{F\in\mathcal{F}_h}\int_F \boldsymbol{\eta}^{-1}|\{\!\{c\mathbf{w}\}\!\}|^2\,dA\right)^{\frac{1}{2}}}{\|\mathbf{w}\|_0}.
\end{aligned}$$

Following from the definition of $\boldsymbol{\eta}$ and the bound (5.1.1) on c, we have

$$\begin{aligned}
\|\ell(v)\|_0 &\le \beta^{-\frac{1}{2}} \max_{\mathbf{w}\in(V^h)^d} \frac{\left(\sum_{F\in\mathcal{F}_h}\int_F \boldsymbol{\eta}|[\![v]\!]|^2\,dA\right)^{\frac{1}{2}} \left(\sum_{F\in\mathcal{F}_h}\int_F \bar{\mathbf{h}}\bar{\mathbf{c}}^{-1}|\{\!\{c\mathbf{w}\}\!\}|^2\,dA\right)^{\frac{1}{2}}}{\|\mathbf{w}\|_0} \\
&\le \beta^{-\frac{1}{2}}(c^*)^{\frac{1}{2}} \max_{\mathbf{w}\in(V^h)^d} \frac{\left(\sum_{F\in\mathcal{F}_h}\int_F \boldsymbol{\eta}|[\![v]\!]|^2\,dA\right)^{\frac{1}{2}} \left(\sum_{E\in\mathcal{M}_h} h_E\int_{\partial E} |\mathbf{w}|^2\,dA\right)^{\frac{1}{2}}}{\|\mathbf{w}\|_0}
\end{aligned} \tag{5.3.5}$$

The following inequality (what [GSS06] calls the inverse inequality) is stated without proof in the article:

$$\|\mathbf{w}\|_{0,\partial E}^2 \le C_{inv}^2 h_E^{-1}\|\mathbf{w}\|_{0,E}^2 \text{ for all } \mathbf{w}\in(S^j(E))^d, \tag{5.3.6}$$

with a constant C_{inv} that depends only on the shape regularity of the mesh, the approximation order j and the dimension d. (We discuss this inequality at the end of Section 5.5 on page 121). This inequality implies the following:

$$\left(\sum_{E \in \mathcal{M}_h} h_E \int_{\partial E} |\mathbf{w}|^2 \, dA \right)^{\frac{1}{2}} \leq C_{inv} \|\mathbf{w}\|_0 \text{ for all } \mathbf{w} \in (S^j(E))^d,$$

Substituting this into (5.3.5) we obtain

$$\begin{aligned}
\|\ell(v)\|_0 &\leq \beta^{-\frac{1}{2}} (c^*)^{\frac{1}{2}} \max_{\mathbf{w} \in (V^h)^d} \frac{\left(\sum_{F \in \mathcal{F}_h} \int_F \boldsymbol{\eta} |[[v]]|^2 \, dA \right)^{\frac{1}{2}} C_{inv} \|\mathbf{w}\|_0}{\|\mathbf{w}\|_0} \\
&= \beta^{-\frac{1}{2}} (c^*)^{\frac{1}{2}} C_{inv} \left(\sum_{F \in \mathcal{F}_h} \int_F \boldsymbol{\eta} |[[v]]|^2 \, dA \right)^{\frac{1}{2}} \\
&= \beta^{-\frac{1}{2}} (c^*)^{\frac{1}{2}} C_{inv} \left(\sum_{F \in \mathcal{F}_h} \|\boldsymbol{\eta}^{\frac{1}{2}} [[v]]\|_{0,F}^2 \right)^{\frac{1}{2}}.
\end{aligned}$$

2. Second estimate

We have that for a piecewise constant c, for any element $z \in (V^h)^d$, $c^{-\frac{1}{2}} z \in (V^h)^d$ and so

$$\|c^{-\frac{1}{2}} \ell(v)\|_0 = \max_{\mathbf{w} \in (V^h)^d} \frac{\sum_{F \in \mathcal{F}_h} \int_F [[v]] \cdot \{\{c^{\frac{1}{2}} \mathbf{w}\}\} dA}{\|\mathbf{w}\|_0}.$$

Now we apply the same steps as we did for the first estimate to obtain

$$\begin{aligned}
\|c^{-\frac{1}{2}} \ell(v)\|_0 &\leq \beta^{-\frac{1}{2}} \max_{\mathbf{w} \in (V^h)^d} \frac{\left(\sum_{F \in \mathcal{F}_h} \int_F \boldsymbol{\eta} |[[v]]|^2 \, dA \right)^{\frac{1}{2}} \left(\sum_{F \in \mathcal{F}_h} \int_F \bar{\mathbf{h}} \bar{\mathbf{c}}^{-1} |\{\{c\mathbf{w}\}\}|^2 \, dA \right)^{\frac{1}{2}}}{\|\mathbf{w}\|_0} \\
&\leq \beta^{-\frac{1}{2}} \max_{\mathbf{w} \in (V^h)^d} \frac{\left(\sum_{F \in \mathcal{F}_h} \int_F \boldsymbol{\eta} |[[v]]|^2 \, dA \right)^{\frac{1}{2}} \left(\sum_{E \in \mathcal{M}_h} h_E \int_{\partial E} |w|^2 \, dA \right)^{\frac{1}{2}}}{\|\mathbf{w}\|_0} \\
&\leq \beta^{-\frac{1}{2}} \max_{\mathbf{w} \in (V^h)^d} \frac{\left(\sum_{F \in \mathcal{F}_h} \int_F \boldsymbol{\eta} |[[v]]|^2 \, dA \right)^{\frac{1}{2}} C_{inv} \|\mathbf{w}\|_0}{\|\mathbf{w}\|_0} \qquad (5.3.7) \\
&= \beta^{-\frac{1}{2}} C_{inv} \left(\sum_{F \in \mathcal{F}_h} \int_F \boldsymbol{\eta} |[[v]]|^2 \, dA \right)^{\frac{1}{2}} \\
&= \beta^{-\frac{1}{2}} C_{inv} \left(\sum_{F \in \mathcal{F}_h} \|\boldsymbol{\eta}^{\frac{1}{2}} [[v]]\|_{0,F}^2 \right)^{\frac{1}{2}}. \qquad (5.3.8)
\end{aligned}$$

□

The auxillary bilinear form $\tilde{b}_h$ defined on $V(h) \times V(h)$ is now introduced.

$$\begin{aligned}\tilde{b}_h(u,v) \quad := \quad & \sum_{E\in\mathcal{M}_h} \int_E c\nabla u \cdot \nabla v dx - \sum_{E\in\mathcal{M}_h} \int_E \ell(u) \cdot \nabla v dx \\ & - \sum_{E\in\mathcal{M}_h} \int_E \ell(v) \cdot \nabla u dx + \sum_{F\in\mathcal{F}_h} \int_F \boldsymbol{\eta} [\![v]\!] \cdot [\![u]\!] ds, \end{aligned} \tag{5.3.9}$$

We have that $V^h \subset V(h)$ and $V(\Omega) \subset V(h)$, and so from the definitions of b and b_h it follows that:

$$\tilde{b}_h = b_h \text{ on } V^h \times V^h, \qquad \tilde{b}_h = b \text{ on } H_0^1(\Omega) \times H_0^1(\Omega). \tag{5.3.10}$$

This suggests that the bilinear form $\tilde{b}_h$ can be viewed as an extension of the two bilinear forms b_h and b to the space $V(h) \times V(h)$.

In the following lemma it is proved that $\tilde{b}_h$ is continuous and positive definite on the entire space $V(h) \times V(h)$.

Lemma 5.3.7. *[GSS06, Lemma 4.4]*
Set

$$\beta_{\min} = 4c_*^{-1}c^*C_{inv}^2$$

for a general piecewise smooth c, and

$$\beta_{\min} = 4C_{inv}^2$$

for a piecewise constant c, with discontinuities aligned with the finite element mesh $\mathcal{M}_h$. C_{inv} is the constant from Lemma 5.3.6. Setting $C_{pd} = \frac{1}{2}$, we have for $\beta \geq \beta_{\min}$

$$\begin{aligned} |\tilde{b}_h(u,v)| \quad &\leq \quad \|u\|_{V(h)}\|v\|_{V(h)}, \quad u,v \in V(h), \\ \tilde{b}_h(u,u) \quad &\geq \quad C_{pd}\|u\|_{V(h)}^2, \quad u \in V(h). \end{aligned}$$

In particular, the positive definiteness bound implies the result in Lemma 5.2.7.

Proof. **1.** First we prove that $\tilde{b}_h$ is bounded.

From the definition of $\tilde{b}_h$ and the Cauchy-Schwartz inequality we have for every $u, v \in V(h)$ that

$$\begin{aligned}
|\tilde{b}_h(u,v)| &= \Big| \sum_{E\in\mathcal{M}_h} \int_E c\nabla u \cdot \nabla v dx - \sum_{E\in\mathcal{M}_h} \int_E \ell(u)\cdot \nabla v dx \\
&\quad - \sum_{E\in\mathcal{M}_h} \int_E \ell(v)\cdot\nabla u dx + \sum_{F\in\mathcal{F}_h} \int_F \boldsymbol{\eta} [\![v]\!] \cdot [\![u]\!] ds \Big| \\
&\leq \sum_{E\in\mathcal{M}_h} \|c^{\frac{1}{2}}\nabla u\|_{0,E} \|c^{\frac{1}{2}}\nabla v\|_{0,E} + \sum_{E\in\mathcal{M}_h} \|\ell(u)\|_{0,E}\|\nabla v\|_{0,E} \\
&\quad + \sum_{E\in\mathcal{M}_h} \|\ell(v)\|_{0,E}\|\nabla u\|_{0,E} + \sum_{F\in\mathcal{F}_h} \|\boldsymbol{\eta}^{\frac{1}{2}}[\![v]\!]\|_{0,F}\|\boldsymbol{\eta}^{\frac{1}{2}}[\![u]\!]\|_{0,F}.
\end{aligned}$$

Since $\|\cdot\| \geq 0$,

$$\begin{aligned}
|\tilde{b}_h(u,v)| &\leq \sum_{E\in\mathcal{M}_h} \|c^{\frac{1}{2}}\nabla u\|_{0,E} \sum_{E\in\mathcal{M}_h} \|c^{\frac{1}{2}}\nabla v\|_{0,E} + \sum_{E\in\mathcal{M}_h} \|\ell(u)\|_{0,E} \sum_{E\in\mathcal{M}_h} \|\nabla v\|_{0,E} \\
&\quad + \sum_{E\in\mathcal{M}_h} \|\ell(v)\|_{0,E} \sum_{E\in\mathcal{M}_h} \|\nabla u\|_{0,E} + \sum_{F\in\mathcal{F}_h} \|\boldsymbol{\eta}^{\frac{1}{2}}[\![v]\!]\|_{0,F} \sum_{F\in\mathcal{F}_h} \|\boldsymbol{\eta}^{\frac{1}{2}}[\![u]\!]\|_{0,F} \\
&\leq \sum_{E\in\mathcal{M}_h} \|c^{\frac{1}{2}}\nabla u\|_{0,E} \sum_{E\in\mathcal{M}_h} \|c^{\frac{1}{2}}\nabla v\|_{0,E} + \|\ell(u)\|_0 \sum_{E\in\mathcal{M}_h} \|\nabla v\|_{0,E} \\
&\quad + \|\ell(v)\|_0 \sum_{E\in\mathcal{M}_h} \|\nabla u\|_{0,E} + \sum_{F\in\mathcal{F}_h} \|\boldsymbol{\eta}^{\frac{1}{2}}[\![v]\!]\|_{0,F} \sum_{F\in\mathcal{F}_h} \|\boldsymbol{\eta}^{\frac{1}{2}}[\![u]\!]\|_{0,F}.
\end{aligned}$$

Since $\|\nabla u\|_{0,E} \leq c_*^{-\frac{1}{2}} \|c^{\frac{1}{2}}\nabla u\|_{0,E}$ and from the first estimate in Lemma 5.3.6 we have

$$\begin{aligned}
|\tilde{b}_h(u,v)| &\leq \sum_{E\in\mathcal{M}_h} \|c^{\frac{1}{2}}\nabla u\|_{0,E} \sum_{E\in\mathcal{M}_h} \|c^{\frac{1}{2}}\nabla v\|_{0,E} \\
&\quad + \beta^{-\frac{1}{2}}(c^*)^{\frac{1}{2}} C_{inv} c_*^{-\frac{1}{2}} \sqrt{\sum_{F\in\mathcal{F}_h} \|\boldsymbol{\eta}^{\frac{1}{2}}[\![u]\!]\|_{0,F}^2} \sum_{E\in\mathcal{M}_h} \|c^{\frac{1}{2}}\nabla v\|_{0,E} \\
&\quad + \beta^{-\frac{1}{2}}(c^*)^{\frac{1}{2}} C_{inv} c_*^{-\frac{1}{2}} \sqrt{\sum_{F\in\mathcal{F}_h} \|\boldsymbol{\eta}^{\frac{1}{2}}[\![v]\!]\|_{0,F}^2} \sum_{E\in\mathcal{M}_h} \|c^{\frac{1}{2}}\nabla u\|_{0,E} \\
&\quad + \sum_{F\in\mathcal{F}_h} \|\boldsymbol{\eta}^{\frac{1}{2}}[\![v]\!]\|_{0,F} \sum_{F\in\mathcal{F}_h} \|\boldsymbol{\eta}^{\frac{1}{2}}[\![u]\!]\|_{0,F} \\
&\leq \sum_{E\in\mathcal{M}_h} \|c^{\frac{1}{2}}\nabla u\|_{0,E} \sum_{E\in\mathcal{M}_h} \|c^{\frac{1}{2}}\nabla v\|_{0,E} \\
&\quad + \beta^{-\frac{1}{2}}(c^*)^{\frac{1}{2}} C_{inv} c_*^{-\frac{1}{2}} \sum_{F\in\mathcal{F}_h} \|\boldsymbol{\eta}^{\frac{1}{2}}[\![u]\!]\|_{0,F} \sum_{E\in\mathcal{M}_h} \|c^{\frac{1}{2}}\nabla v\|_{0,E} \\
&\quad + \beta^{-\frac{1}{2}}(c^*)^{\frac{1}{2}} C_{inv} c_*^{-\frac{1}{2}} \sum_{F\in\mathcal{F}_h} \|\boldsymbol{\eta}^{\frac{1}{2}}[\![v]\!]\|_{0,F}^2 \sum_{E\in\mathcal{M}_h} \|c^{\frac{1}{2}}\nabla u\|_{0,E} \\
&\quad + \sum_{F\in\mathcal{F}_h} \|\boldsymbol{\eta}^{\frac{1}{2}}[\![v]\!]\|_{0,F} \sum_{F\in\mathcal{F}_h} \|\boldsymbol{\eta}^{\frac{1}{2}}[\![u]\!]\|_{0,F}.
\end{aligned}$$

It then follows that

$$\begin{aligned}|\tilde{b}_h(u,v)| &\leq \max\{1, \beta^{-\frac{1}{2}}(c^*)^{\frac{1}{2}}C_{inv}c_*^{-\frac{1}{2}}\}\Bigg(\sum_{E\in\mathcal{M}_h}\|c^{\frac{1}{2}}\nabla u\|_{0,E}\sum_{E\in\mathcal{M}_h}\|c^{\frac{1}{2}}\nabla v\|_{0,E} \\ &\quad + \sum_{F\in\mathcal{F}_h}\|\boldsymbol{\eta}^{\frac{1}{2}}[[u]]\|_{0,F}\sum_{E\in\mathcal{M}_h}\|c^{\frac{1}{2}}\nabla v\|_{0,E} + \sum_{F\in\mathcal{F}_h}\|\boldsymbol{\eta}^{\frac{1}{2}}[[v]]\|^2_{0,F}\sum_{E\in\mathcal{M}_h}\|c^{\frac{1}{2}}\nabla u\|_{0,E} \\ &\quad + \sum_{F\in\mathcal{F}_h}\|\boldsymbol{\eta}^{\frac{1}{2}}[[v]]\|_{0,F}\sum_{F\in\mathcal{F}_h}\|\boldsymbol{\eta}^{\frac{1}{2}}[[u]]\|_{0,F}\Bigg),\end{aligned}$$

and so finally (since for any real a and b, $a+b \leq \sqrt{a^2+b^2}$),

$$\begin{aligned}|\tilde{b}_h(u,v)| &\leq K^*\left(\sum_{E\in\mathcal{M}_h}\|c^{\frac{1}{2}}\nabla u\|_{0,E} + \sum_{F\in\mathcal{F}_h}\|\boldsymbol{\eta}^{\frac{1}{2}}[[u]]\|_{0,F}\right) \\ &\qquad \times\left(\sum_{E\in\mathcal{M}_h}\|c^{\frac{1}{2}}\nabla v\|_{0,E} + \sum_{F\in\mathcal{F}_h}\|\boldsymbol{\eta}^{\frac{1}{2}}[[v]]\|_{0,F}\right) \\ &\leq K^*\sqrt{\sum_{E\in\mathcal{M}_h}\|c^{\frac{1}{2}}\nabla u\|^2_{0,E} + \sum_{F\in\mathcal{F}_h}\|\boldsymbol{\eta}^{\frac{1}{2}}[[u]]\|^2_{0,F}} \\ &\qquad \times\sqrt{\sum_{E\in\mathcal{M}_h}\|c^{\frac{1}{2}}\nabla v\|^2_{0,E} + \sum_{F\in\mathcal{F}_h}\|\boldsymbol{\eta}^{\frac{1}{2}}[[v]]\|^2_{0,F}} \\ &= K^*\|u\|_{V(h)}\|v\|_{V(h)}.\end{aligned}$$

where $K^* = \max\{1, \beta^{-\frac{1}{2}}(c^*)^{\frac{1}{2}}C_{inv}c_*^{-\frac{1}{2}}\}$. Since we choose $\beta_{\min} = 4c_*^{-1}c^*C_{inv}^2$, we have that $K^* = 1$.

The case of piecewise constant c follows analogously, using the second estimate in Lemma 5.3.6.

2. Now we prove that $\tilde{b}_h$ is positive definite.

We have that

$$\begin{aligned}\tilde{b}_h(u,u) &= \sum_{E\in\mathcal{M}_h}\int_E c\nabla u\cdot\nabla u dx - \sum_{E\in\mathcal{M}_h}\int_E \ell(u)\cdot\nabla u dx \\ &\quad - \sum_{E\in\mathcal{M}_h}\int_E \ell(u)\cdot\nabla u dx + \sum_{F\in\mathcal{F}_h}\int_F \boldsymbol{\eta}\,[[u]]\cdot[[u]]ds \\ &= \sum_{E\in\mathcal{M}_h}\|c^{\frac{1}{2}}\nabla u\|^2_{0,E} - 2\sum_{E\in\mathcal{M}_h}\int_E \ell(u)\cdot\nabla u dx + \sum_{F\in\mathcal{F}_h}\|\boldsymbol{\eta}^{\frac{1}{2}}\,[[u]]\|^2_{0,F}.\end{aligned} \tag{5.3.11}$$

Now consider the second term on the right hand side of (5.3.11). Using the Cauchy-Schwartz inequality and Young's inequality in Lemma C.1 we obtain

$$\begin{aligned} 2\left|\sum_{E\in\mathcal{M}_h}\int_E \ell(u)\cdot\nabla u dx\right| &= 2\sum_{E\in\mathcal{M}_h}\int_E c^{-\frac{1}{2}}\ell(u)\cdot c^{\frac{1}{2}}\nabla u dx \\ &\leq 2\sum_{E\in\mathcal{M}_h}\|c^{-\frac{1}{2}}\ell(u)\|_{0,E}\|c^{\frac{1}{2}}\nabla u\|_{0,E} \\ &\leq \varepsilon\sum_{E\in\mathcal{M}_h}\|c^{\frac{1}{2}}\nabla u\|_{0,E}^2+\varepsilon^{-1}{c_*}^{-1}\sum_{E\in\mathcal{M}_h}\|\ell(u)\|_{0,E}^2 \end{aligned}$$

Now, for a general piecewise smooth c, we use the first estimate in Lemma 5.3.6.

$$2\sum_{E\in\mathcal{M}_h}\int_E \ell(u)\cdot\nabla u dx \leq \varepsilon\sum_{E\in\mathcal{M}_h}\|c^{\frac{1}{2}}\nabla u\|_{0,E}^2+\varepsilon^{-1}{c_*}^{-1}\beta^{-1}c^*C_{inv}^2\sum_{F\in\mathcal{F}_h}\|\boldsymbol{\eta}^{\frac{1}{2}}[[u]]\|_{0,F}^2.$$

Combining this with equation (5.3.11) we have

$$\tilde{b}_h(u,u)\geq(1-\varepsilon)\sum_{E\in\mathcal{M}_h}\|c^{\frac{1}{2}}\nabla u\|_{0,E}^2+\left(1-\varepsilon^{-1}{c_*}^{-1}\beta^{-1}c^*C_{inv}^2\right)\sum_{F\in\mathcal{F}_h}\|\boldsymbol{\eta}^{\frac{1}{2}}[[u]]\|_{0,F}^2.$$

Now, if we choose $\varepsilon=\frac{1}{2}$ and $\beta_{\min}=4c_*^{-1}c^*C_{inv}^2$ where $\beta\geq\beta_{\min}$, we obtain the positive definiteness bound as was desired;

$$\begin{aligned} \tilde{b}_h(u,u) &\geq (1-\frac{1}{2})\sum_{E\in\mathcal{M}_h}\|c^{\frac{1}{2}}\nabla u\|_{0,E}^2+\left(1-2{c_*}^{-1}{\beta_{\min}}^{-1}c^*C_{inv}^2\right)\sum_{F\in\mathcal{F}_h}\|\boldsymbol{\eta}^{\frac{1}{2}}[[u]]\|_{0,F}^2 \\ &\geq \frac{1}{2}\sum_{E\in\mathcal{M}_h}\|c^{\frac{1}{2}}\nabla u\|_{0,E}^2+\left(1-2{c_*}^{-1}(4c_*^{-1}c^*C_{inv}^2)^{-1}c^*C_{inv}^2\right)\sum_{F\in\mathcal{F}_h}\|\boldsymbol{\eta}^{\frac{1}{2}}[[u]]\|_{0,F}^2 \\ &= \frac{1}{2}\sum_{E\in\mathcal{M}_h}\|c^{\frac{1}{2}}\nabla u\|_{0,E}^2+\frac{1}{2}\sum_{F\in\mathcal{F}_h}\|\boldsymbol{\eta}^{\frac{1}{2}}[[u]]\|_{0,F}^2 \\ &= \frac{1}{2}\|u\|_{V(h)}^2=C_{pd}\|u\|_{V(h)}^2. \end{aligned}$$

Now, for a general piecewise constant c, we use the second estimate in Lemma 5.3.6. Similarly as above we will obtain

$$\tilde{b}_h(u,u)\geq(1-\varepsilon)\sum_{E\in\mathcal{M}_h}\|c^{\frac{1}{2}}\nabla u\|_{0,E}^2+\left(1-\varepsilon^{-1}\beta^{-1}C_{inv}^2\right)\sum_{F\in\mathcal{F}_h}\|\boldsymbol{\eta}^{\frac{1}{2}}[[u]]\|_{0,F}^2.$$

Again, if we choose $\varepsilon = \frac{1}{2}$ and $\beta_{\min} = 4C_{inv}^2$ where $\beta \geq \beta_{\min}$, we obtain the coercivity bound as was desired.

$$\begin{aligned}
\tilde{b}_h(u,u) &\geq (1-\frac{1}{2})\sum_{E\in\mathcal{M}_h}\|c^{\frac{1}{2}}\nabla u\|_{0,E}^2 + \left(1-2\beta_{\min}{}^{-1}C_{inv}^2\right)\sum_{F\in\mathcal{F}_h}\|\boldsymbol{\eta}^{\frac{1}{2}}[[u]]\|_{0,F}^2 \\
&\geq \frac{1}{2}\sum_{E\in\mathcal{M}_h}\|c^{\frac{1}{2}}\nabla u\|_{0,E}^2 + \left(1-2(4C_{inv}^2)^{-1}C_{inv}^2\right)\sum_{F\in\mathcal{F}_h}\|\boldsymbol{\eta}^{\frac{1}{2}}[[u]]\|_{0,F}^2 \\
&= \frac{1}{2}\sum_{E\in\mathcal{M}_h}\|c^{\frac{1}{2}}\nabla u\|_{0,E}^2 + \frac{1}{2}\sum_{F\in\mathcal{F}_h}\|\boldsymbol{\eta}^{\frac{1}{2}}[[u]]\|_{0,F}^2 \\
&= \frac{1}{2}\|u\|_{V(h)}^2 = C_{pd}\|u\|_{V(h)}^2.
\end{aligned}$$

□

Remark The above proof looks rather messy. It has been done in great detail, especially proving continuity of $\tilde{b}_h$, since the proof given in [GSS06] only states the following: "By taking into account the bounds in (2.5) and Lemma 4.3, application of the Cauchy-Schwartz inequality readily gives in the general case ... " From this the estimate obtained was not obvious requiring us to fill in the details in this work. Note that the estimates obtained here differs from those of [GSS06], but the results are the same.

5.3.2 Error equation

In this subsection the so-called error equation is derived. This equation is used to prove the main results (Theorem 5.3.2 and Theorem 5.3.3).

Since $\tilde{b}_h = b_h$ on $V^h \times V^h$ (see Equation (5.3.10) on page 102), Problem WDG (see Subsection 5.2.2) is equivalent to the following problem.

Problem $\widetilde{\textbf{WDG}}$

For a given partition $\mathcal{M}_h$ of Ω, find a function $u_h \in C(J, V^h)$ such that

$$\begin{aligned}
(u_h''(t), v) + \tilde{b}_h(u_h(t), v) &= (f(t), v) \text{ for all } v \in V^h \\
u_h(0) &= u_0^h \\
u_h'(0) &= u_1^h.
\end{aligned} \tag{5.3.12}$$

To derive an error equation, [GSS06] defines, for $u \in H^{1+m}(\Omega)$ with $m > \frac{1}{2}$

$$r_h(u, v) = \sum_{F \in \mathcal{F}_h} \int_F [\![v]\!] \cdot \{\!\{c\nabla u - cP_2(\nabla u)\}\!\}\, dA \tag{5.3.13}$$

for all $v \in V(h)$.

Remark The definition above is the reason the bilinear form $\tilde{b}_h$ is introduced.

Lemma 5.3.8. *Let the solution u of Problem W satisfy*

$$u \in \mathcal{L}^\infty(J, H^{1+m}(\Omega)), \quad u'' \in \mathcal{L}^1(J, \mathcal{L}^2(\Omega)),$$

with $m \geq \frac{1}{2}$. Let u_h be the semi-discrete DG approximation obtained by (5.3.12). Then the error $e_h = u - u_h$ satisfies

$$(e_h'', v) + \tilde{b}_h(e_h, v) = r_h(u, v) \text{ for all } v \in V^h \text{ a.e. in } J, \tag{5.3.14}$$

with r_h given as in (5.3.13).

Proof. Subtracting u in equation (5.3.12), we have that

$$(e_h''(t), v) + \tilde{b}_h(e_h(t), v) = (u''(t), v) + \tilde{b}_h(u(t), v) - (f(t), v) \text{ for all } v \in V^h \text{ and all } t \in J.$$

Since we have that $u'' \in \mathcal{L}^1(J, \mathcal{L}^2(\Omega))$, we know that $\ell(u) = 0$ and also that $[\![u]\!] = 0$ on all faces. Now, using the definition of the lifted element $\ell(v)$ and the properties of the $\mathcal{L}^2$-projection P_2 we have

$$\int_\Omega \ell(v) \cdot \nabla u dx = \sum_{F \in \mathcal{F}_h} \int_F [\![v]\!] \cdot \{\!\{c\nabla u\}\!\} dA = \sum_{F \in \mathcal{F}_h} \int_F [\![v]\!] \cdot \{\!\{cP_2(\nabla u)\}\!\} dA.$$

Using these facts in the definition of $\tilde{b}_h$ we have

$$\tilde{b}_h(u, v) = \sum_{E \in \mathcal{M}_h} \int_E c\nabla u \cdot \nabla v\, dx - \sum_{F \in \mathcal{F}_h} \int_F [\![v]\!] \cdot \{\!\{cP_2 \nabla u\}\!\}\, dA. \tag{5.3.15}$$

Now note that $\nabla \cdot (c\nabla u) \in \mathcal{L}^2(\Omega)$ almost everywhere in J, since $u'' \in C(J, \mathcal{L}^2(\Omega))$ and $f \in C^1(J, \mathcal{L}^2(\Omega))$. This therefore implies that $c\nabla u$ has continuous normal components across all interior faces. Now we integrate by parts elementwise (and taking jumps and averages into account) to get

$$\sum_{E \in \mathcal{M}_h} \int_E c\nabla u \cdot \nabla v dx = - \sum_{E \in \mathcal{M}_h} \int_E \nabla \cdot (c\nabla u) v\, dx + \sum_{F \in \mathcal{F}_h} \int_F [\![v]\!] \cdot \{\!\{c\nabla u\}\!\}\, dA.$$

Substituting this back into Equation (5.3.15), and using the definition of r_h we have that the following holds:

$$\tilde{b}_h(u,v) = -\sum_{E\in\mathcal{M}_h}\int_E \nabla\cdot(c\nabla u)v dx + r_h(u,v).$$

Now $(u'',v)+\tilde{b}_h(u,v) = (u''-\nabla\cdot(c\nabla u),v)+r_h(u,v)$ and from Problem WDG

$$(e_h'',v)+\tilde{b}_h(e_h,v) = (u''-\nabla\cdot(c\nabla u)-f,v)+r_h(u,v).$$

Therefore, since $(u''-\nabla\cdot(c\nabla u)-f,v)=0$ for all $v\in V^h$, we have the desired result. □

Remark It is worth noting here that this error equation (5.3.14) is similar to the equation obtained in Lemma 3.1.3 (Chapter 3, Section 3.1.1).

5.3.3 Approximation properties

The proofs of the main results (Theorem 5.3.2 and Theorem 5.3.3) rely on similar techniques as used in [Bak76] and [BV13]. A projection is added and subtracted and then from known theory the projection error estimates can be obtained. These approximation properties is given in this subsection. The error estimates for the initial conditions also rely on these approximation properties.

The approximation properties in Lemma 5.3.9 below are from [Cia78]. Recall that P_2 denotes the $\mathcal{L}^2$-projection onto V^h.

Lemma 5.3.9. *[GSS06, Lemma 4.6]*
Let $E\in\mathcal{M}_h$. Then the following holds:

1. *For $v\in H^m(E)$, $m\geq 0$, we have*

$$\|v-P_2v\|_{0,E}\leq Ch_E^{\min\{m,j+1\}}\|v\|_{m,E},$$

with a constant $C>0$ that is independent of the local mesh size h_E and depends only on the shape regularity of the mesh, the approximation order j, the dimension d, and the regularity exponent m.

2. *For $v \in H^{1+m}(E)$, $m > \frac{1}{2}$, we have*

$$\begin{aligned} \|\nabla v - \nabla(P_2 v)\|_{0,E} &\leq Ch_E^{\min\{m,j\}} \|v\|_{1+m,E}, \\ \|v - P_2 v\|_{0,\partial E} &\leq Ch_E^{\min\{m,j\}+\frac{1}{2}} \|v\|_{1+m,E}, \\ \|\nabla v - P_2(\nabla v)\|_{0,\partial E} &\leq Ch_E^{\min\{m,j+1\}-\frac{1}{2}} \|v\|_{1+m,E}, \end{aligned}$$

with a constant $C > 0$ that is independent of the local mesh size h_E and depends only on the shape regularity of the mesh, the approximation order j, the dimension d, and the regularity exponent m.

Lemma 5.3.10. *[GSS06, Lemma 4.7]*
Let $u \in H^{1+m}(E)$, $m \geq 0$. Then the following holds:

1. *We have*

$$\|u - P_2 u\|_{V(h)} \leq C_A h^{\min\{m,j\}} \|u\|_{1+m},$$

with a constant $C_A > 0$ that is independent of local mesh size and depends only on β, the constant κ in (5.2.1), the bounds in (5.1.1) and the constants in Lemma 5.3.9.

2. *For $v \in V(h)$, the form $r_h(u,v)$ in (5.3.13), can be bounded by*

$$|r_h(u,v)| \leq C_R h^{\min\{m,j\}} \left(\sum_{F \in \mathcal{F}_h} \|\boldsymbol{\eta}^{\frac{1}{2}} [[v]]\|_{0,F}^2 \right)^{\frac{1}{2}} \|u\|_{1+m} \tag{5.3.16}$$

with a constant $C_R > 0$ that is independent of h and depends only on β, the bounds in (5.1.1) and the constants in Lemma 5.3.9.

Proof.

1. From the definition of the norm $\|\cdot\|_{V(h)}$ and Lemma 5.3.9 we have

$$\begin{aligned} \|u - P_2 u\|_{V(h)}^2 &= \sum_{E \in \mathcal{M}_h} \|c^{\frac{1}{2}} \nabla(u - P_2 u)\|_{0,E}^2 + \sum_{F \in \mathcal{F}_h} \|\boldsymbol{\eta}^{\frac{1}{2}} [[u - P_2 u]]\|_{0,F}^2 \\ &\leq \sum_{E \in \mathcal{M}_h} c^* C^2 h_E^{2\min\{m,j\}} \|v\|_{1+m,E}^2 + \sum_{F \in \mathcal{F}_h} \|\boldsymbol{\eta}^{\frac{1}{2}} [[u - P_2 u]]\|_{0,F}^2. \end{aligned}$$

It follows from the definition of η and Lemma 5.3.9 that

$$\begin{aligned} \|u - P_2 u\|_{V(h)} &\leq C_A \left(h^{\min\{m,j\}} \|u\|_{1+m} + h^{\min\{m,j\}+1} \|u\|_{1+m} \right) \\ &\leq C_A h^{\min\{m,j\}} \|u\|_{1+m}, \end{aligned}$$

where C_A depends on independent of the local mesh size h_E and depends only on the shape regularity of the mesh, the approximation order j, the dimension d, and the regularity exponent m.

2. We have from definition and the Cauchy-Schwartz inequality that

$$\begin{aligned}
r_h(u,v) &= \sum_{F\in\mathcal{F}_h}\int_F [\![v]\!]\cdot\{\!\{c\nabla u - cP_2(\nabla u\}\!\}\,dA \\
&\le \sum_{F\in\mathcal{F}_h}\left(\int_F \boldsymbol{\eta}|[\![v]\!]|^2\,ds\right)^{\frac{1}{2}}\left(\int_F \boldsymbol{\eta}^{-1}|\{\!\{c\nabla u - cP_2(\nabla u)\}\!\}|^2\,ds\right)^{\frac{1}{2}} \\
&\le \left(\sum_{F\in\mathcal{F}_h}\int_F \boldsymbol{\eta}|[\![v]\!]|^2\,ds\right)^{\frac{1}{2}}\left(\sum_{F\in\mathcal{F}_h}\int_F \boldsymbol{\eta}^{-1}|\{\!\{c\nabla u - cP_2(\nabla u)\}\!\}|^2\,ds\right)^{\frac{1}{2}} \\
&\le \beta^{\frac{1}{2}}{c_*}^{\frac{1}{2}}c^*\left(\sum_{F\in\mathcal{F}_h}\|\boldsymbol{\eta}[\![v]\!]\|_{0,F}^2\right)^{\frac{1}{2}}\left(\sum_{E\in\mathcal{M}_h} h_E\int_{\partial E}|\nabla u - P_2(\nabla u)|^2\,ds\right)^{\frac{1}{2}} \\
&= \beta^{\frac{1}{2}}{c_*}^{\frac{1}{2}}c^*\left(\sum_{F\in\mathcal{F}_h}\|\boldsymbol{\eta}[\![v]\!]\|_{0,F}^2\right)^{\frac{1}{2}}\left(\sum_{E\in\mathcal{M}_h} h_E\|\nabla u - P_2(\nabla u)\|_{0,\partial E}^2\,ds\right)^{\frac{1}{2}}.
\end{aligned}$$

Now using the result from Lemma 5.3.9 we have the result.

$$\begin{aligned}
r_h(u,v) &\le \beta^{\frac{1}{2}}{c_*}^{\frac{1}{2}}c^*\left(\sum_{F\in\mathcal{F}_h}\|\boldsymbol{\eta}[\![v]\!]\|_{0,F}^2\right)^{\frac{1}{2}}\left(\sum_{E\in\mathcal{M}_h} Ch_E h_E^{\min\{m,j\}-\frac{1}{2}}\|u\|_{0,E}\right)^{\frac{1}{2}} \\
&= C_R h_E^{\min\{m,j\}}\left(\sum_{F\in\mathcal{F}_h}\|\boldsymbol{\eta}[\![v]\!]\|_{0,F}^2\right)^{\frac{1}{2}}\|u\|_{0,E},
\end{aligned}$$

where $C_R = \beta^{\frac{1}{2}}{c_*}^{\frac{1}{2}}c^*$.

□

Remark [GSS06] states that in Theorem 5.3.2 (the DG energy norm error estimate) "it is implicitly assumed that $u^0 \in H^{1+m}(\Omega)$ and $v^0 \in H^m(\Omega)$". We refer back to Chapter 2, Subsection 2.5.3. It was found that sufficient conditions (on the initial conditions) for existence is $u^0 \in H^{1+m}(\Omega)\cap V(\Omega)$ and $v^0 \in V(\Omega)$. It therefore follows from Lemma 5.3.10 that

$$\begin{aligned}
\|e_h(0)\|_{V(h)} &= \|u_0 - P_2 u_0\|_{V(h)} \le Ch^{\min\{m,j\}}\|u_0\|_{1+m}, \\
\|e_h'(0)\|_{V(h)} &= \|v_0 - P_2 v_0\|_{V(h)} \le Ch^{\min\{m,j+1\}}\|v_0\|_{m}.
\end{aligned}$$

We now have the building blocks to prove the main error estimates.

5.4 Proof of the $\|\cdot\|_{V(h)}$-norm error estimate

The result in the lemma below is used in the proof of Theorem 5.3.2 (the DG energy norm error estimate).

Lemma 5.4.1. *[GSS06, Lemma 4.8]*
Let the analytical solution u of the wave equation satisfy

$$u \in C(J, H^{1+m}(\Omega)), \quad u' \in C(J, H^{1+m}(\Omega))$$

for $m > \frac{1}{2}$. Let $v \in C(J, V(h))$ and $v' \in \mathcal{L}^1(J, V(h))$. Then we have

$$\begin{aligned}
&\int_J |r_h(u, v')|\, dt \\
&\qquad \le C_R h^{\min\{m,j\}} \|v\|_{C(J,V(h))} \left(2\|u\|_{C(J,H^{1+m}(\Omega))} + T\|u'\|_{C(J,H^{1+m}(\Omega))}\right),
\end{aligned}$$

where C_R is the constant in Lemma 5.3.10.

Proof. From the definition of r_h and integration by parts, we obtain

$$\begin{aligned}
\int_J r_h(u, v')\, dt &= \int_J \sum_{F\in\mathcal{F}_h} \int_F [\![v']\!] \cdot \{\!\{c\nabla u - cP_2(\nabla u)\}\!\}\, dA\, dt \\
&= -\int_J \sum_{F\in\mathcal{F}_h} \int_F [\![v]\!] \cdot \{\!\{c\nabla u' - cP_2(\nabla u')\}\!\}\, dA\, dt \\
&\qquad + \left[\sum_{F\in\mathcal{F}_h} \int_F [\![v]\!] \cdot \{\!\{c\nabla u - cP_2(\nabla u)\}\!\}\, dA\right]_{t=0}^{t=T} \\
&= -\int_J r_h(u', v)\, dt + [r_h(u, v)]_{t=0}^{t=T}
\end{aligned}$$

Now, from the estimates in Lemma 5.3.10,

$$\begin{aligned}
\left|\int_J r_h(u', v)\, dt\right| &\le \int_J |r_h(u', v)|\, dt \\
&\le \int_J |r_h(u', v)|\ dt \\
&\le \int_J C_R h^{\min\{m,j\}} \left(\sum_{F\in\mathcal{F}_h} \|\boldsymbol{\eta}^{\frac{1}{2}} [\![v(t)]\!]\|_{0,F}^2\right)^{\frac{1}{2}} \|u'\|_{1+m}\, dt \\
&\le \int_J C_R h^{\min\{m,j\}} \left(\sum_{E\in\mathcal{M}_h} \|c^{\frac{1}{2}}\nabla v\|_{0,E}^2 + \sum_{F\in\mathcal{F}_h} \|\boldsymbol{\eta}^{\frac{1}{2}} [\![v]\!]\|_{0,F}^2\right)^{\frac{1}{2}} \|u'\|_{1+m}\, dt \\
&\le C_R h^{\min\{m,j\}} T \|v\|_{C(J,V(h))} \|u'\|_{C(J,H^{1+m}(\Omega))}.
\end{aligned}$$

We also have from Lemma 5.3.10 that

$$\begin{aligned}\left|[r_h(u,v)]_{t=0}^{t=T}\right| &= |r_h(u(T),v(T)) - r_h(u(0),v(0))| \\ &\leq 2C_R h^{\min\{m,j\}}\|v\|_{C(J,H^{1+m}(\Omega))}\|u\|_{C(J,H^{1+m}(\Omega))}.\end{aligned}$$

Adding, we therefore obtain the desired result:

$$\begin{aligned}\int_J |r_h(u,v')|\ dt &\leq C_R h^{\min\{m,j\}} T\|v\|_{C(J,V(h))}\|u'\|_{C(J,H^{1+m}(\Omega))} \\ &\quad +2C_R h^{\min\{m,j\}}\|v\|_{C(J,H^{1+m}(\Omega))}\|u\|_{C(J,H^{1+m}(\Omega))} \\ &\leq C_R h^{\min\{m,j\}}\|v\|_{C(J,H^{1+m}(\Omega))}\left(2\|u\|_{C(J,H^{1+m}(\Omega))} + T\|u'\|_{C(J,H^{1+m}(\Omega))}\right).\end{aligned}$$

□

The following lemma is not given in [GSS06]. It is a trivial matter to prove, but is important nonetheless, since it is used (but not mentioned) in the proof of the DG norm error estimate.

Lemma 5.4.2. *If $u \in C^k(J,\mathcal{L}^2(\Omega))$, then $P_2u \in C^k(J)$ and $(P_2u)^{(k)}(t) = P_2u^{(k)}(t)$, for $k = 1,2$.*

Proof. First consider the case when $k=1$. Since $\|u\|_0^2 = \|P_2u\|_0^2 + \|u - P_2u\|_0^2$, we have that $\|P_2u\|_0 \leq \|u\|_0$ for all $u \in C^1(J,\mathcal{L}^2(\Omega))$. If $u \in C^1(J,\mathcal{L}^2(\Omega))$, it follows that:

$$\begin{aligned}&\|(\delta t)^{-1}\Big(P_2u(t+\delta t) - P_2u(t)\Big) - P_2u'(t)\|_0 \\ =\ &\|(\delta t)^{-1}P_2\left(u(t+\delta t) - u(t) - u'(t)\right)\|_0 \\ \leq\ &\|(\delta t)^{-1}\left(u(t+\delta t) - u(t)\right) - u'(t)\|_0\end{aligned}$$

But $\|(\delta t)^{-1}\left(u(t+\delta t) - u(t)\right) - u'(t)\|_0$ converges to 0 as $\delta t \to 0$, and so

$$P_2u \in C^1(J,\mathcal{L}^2(\Omega))$$

and $(P_2u)'(t) = P_2u'(t)$.

Now it follows that if $u' \in C^1(J,\mathcal{L}^2(\Omega))$, then $P_2u' \in C^1(J,\mathcal{L}^2(\Omega))$ and so

$$(P_2u')'(t) = P_2u''(t).$$

□

Proof of Theorem 5.3.2:

We can now prove Theorem 5.3.2. Recall that P_2 denotes the $\mathcal{L}^2$-projection onto V^h. Since the bilinear form $\tilde{b}_h$ is symmetric, and adding and subtracting the term $(P_2u)'$ we have

$$\begin{aligned} \frac{1}{2}\frac{d}{dt}\left(\|e_h'(t)\|_0^2 + \tilde{b}_h(e_h(t), e_h(t))\right) &= (e_h''(t), e_h'(t)) + \tilde{b}_h(e_h(t), e_h'(t)) \\ &= (e_h''(t), (u - P_2u)'(t)) + \tilde{b}_h(e_h(t), (u - P_2u)'(t)) \\ &\quad + r_h(u(t), (P_2u - u_h)'(t)) \end{aligned} \tag{5.4.1}$$

Fix $s \in J$ and integrate (5.4.1) from 0 to s to obtain:

$$\begin{aligned} &\frac{1}{2}\|e_h'(s)\|_0^2 + \frac{1}{2}\tilde{b}_h(e_h(s), e_h(s)) \\ &\quad = \frac{1}{2}\|e_h'(0)\|_0^2 + \frac{1}{2}\tilde{b}_h(e_h(0), e_h(0)) + \int_0^s (e_h''(t), (u - P_2u)'(t))\, dt \\ &\qquad + \int_0^s \tilde{b}_h(e_h(t), (u - P_2u)'(t))\, dt + \int_0^s r_h(u(t), (P_2u - u_h)'(t))\, dt \end{aligned} \tag{5.4.2}$$

In the next step we use Lemma 5.3.7, integration by parts on the third term on the right hand side of (5.4.2), and the Cauchy-Schwartz inequality on the inner product of $\mathcal{L}^2$. We then have that

$$\begin{aligned} &\frac{1}{2}\|e_h'(s)\|_0^2 + \frac{1}{2}C_{pd}\|e_h(s)\|_{V(h)}^2 \\ &\quad \leq \Bigg| \frac{1}{2}\|e_h'(0)\|_0^2 + \frac{1}{2}\|e_h(0)\|_{V(h)}^2 - \int_0^s (e_h'(t), (u - P_2u)''(t))\, dt \\ &\qquad + (e_h'(t), (u - P_2u)'(t))\Big|_{t=0}^{t=s} + \int_0^s \|e_h(t)\|_{V(h)}\|(u - P_2u)'(t)\|_{V(h)}\, dt \\ &\qquad + \int_0^s r_h(u(t), (P_2u - u_h)'(t))\, dt \Bigg| \\ &\quad \leq \frac{1}{2}\|e_h'(0)\|_0^2 + \frac{1}{2}\|e_h(0)\|_{V(h)}^2 + \int_0^s \|e_h'(t)\|_0\|(u - P_2u)''(t)\|_0\, dt \\ &\qquad + |(e_h'(s), (u - P_2u)'(s))| + |(e_h'(0), (u - P_2u)'(0))| \\ &\qquad + \int_0^s \|e_h(t)\|_{V(h)}\|(u - P_2u)'(t)\|_{V(h)}\, dt + \left|\int_0^s r_h(u(t), (P_2u - u_h)'(t))\, dt\right|. \end{aligned} \tag{5.4.3}$$

Bounding the third term on the right hand side of (5.4.3) we obtain:

$$\begin{aligned}\int_0^s \|e_h'(t)\|_0 \|(u-P_2u)''(t)\|_0\,dt &\leq \max_{t\in(0,s)} \|e_h'(t)\|_0 \int_0^s \|(u-P_2u)''(t)\|_0\,dt \\ &\leq \max_{t\in[0,T]} \|e_h'(t)\|_0 \int_0^T \|(u-P_2u)''(t)\|_0\,dt.\end{aligned}$$

Bounding the fourth and fifth terms on the right hand side of (5.4.3) we obtain:

$$\begin{aligned}&|(e_h'(s),(u-P_2u)'(s))| + |(e_h'(0),(u-P_2u)'(0))| \\ &\quad\leq \|e_h'(s)\|_0 \|(u-P_2u)'(s)\|_0 + \|e_h'(0)\|_0 \|(u-P_2u)'(0)\|)_0 \\ &\quad\leq 2\max_{t\in[0,T]} \|e_h'(t)\|_0 \max_{t\in[0,T]} \|(u-P_2u)'(t)\|_0.\end{aligned}$$

Also, bounding the sixth term on the right hand side of (5.4.3) we obtain:

$$\begin{aligned}&\int_0^s \|e_h(t)\|_{V(h)} \|(u-P_2u)'(t)\|_{V(h)}\,dt \\ &\qquad\leq T\max_{t\in[0,T]} \|e_h'(t)\|_{V(h)} \max_{t\in[0,T]} \|(u-P_2u)'(t)\|_{V(h)}.\end{aligned}$$

Therefore, since (5.4.3) holds for all $s\in J$, we have that

$$\max_{t\in[0,T]} \|e_h'(t)\|_0^2 + C_{pd}\max_{t\in[0,T]} \|e_h(t)\|_{V(h)}^2 \leq \|e_h'(0)\|_0^2 + \|e_h(0)\|_{V(h)}^2 + T_1 + T_2 + T_3$$

where

$$\begin{aligned}T_1 &= 2\max_{t\in[0,T]} \|e_h'(t)\|_0 \left(\int_0^s \|(u-P_2u)''(t)\|_0\,dt + 2\max_{t\in[0,T]} \|(u-P_2u)'(t)\|_0\right) \\ T_2 &= 2T\max_{t\in[0,T]} \|e_h'(t)\|_{V(h)} \max_{t\in[0,T]} \|(u-P_2u)'(t)\|_{V(h)} \\ T_3 &= 2\int_0^s |r_h(u(t),(P_2u-u_h)'(t))|\,dt\end{aligned}$$

To bound the terms T_1 and T_2 we use Lemma C.1 (Young's inequality). For T_1 we use $\varepsilon = 1$.

$$\begin{aligned} T_1 &\leq \frac{1}{2}\max_{t\in[0,T]}\|e_h'(t)\|_0^2 + 2\left(\int_0^s \|(u-P_2u)''(t)\|_0\,dt + 2\max_{t\in[0,T]}\|(u-P_2u)'(t)\|_0\right)^2 \\ &\leq \frac{1}{2}\max_{t\in[0,T]}\|e_h'(t)\|_0^2 + 2\left(\int_0^s \|(u-P_2u)''(t)\|_0\,dt\right)^2 + 8\max_{t\in[0,T]}\|(u-P_2u)'(t)\|_0^2 \\ &+ 8\left(\int_0^s \|(u-P_2u)''(t)\|_0\,dt\right)\max_{t\in[0,T]}\|(u-P_2u)'(t)\|_0 \\ &\leq \frac{1}{2}\max_{t\in[0,T]}\|e_h'(t)\|_0^2 + 4\left(\int_0^s \|(u-P_2u)''(t)\|_0\,dt\right)^2 + 16\max_{t\in[0,T]}\|(u-P_2u)'(t)\|_0^2 \\ &\leq \frac{1}{2}\max_{t\in[0,T]}\|e_h'(t)\|_0^2 + 4\left(\int_0^s \|(u-P_2u)''(t)\|_0\,dt\right)^2 + 16\max_{t\in[0,T]}\|(u-P_2u)'(t)\|_0^2. \end{aligned}$$

From the approximation results in Lemma 5.3.9 we therefore have

$$T_1 \leq \frac{1}{2}\max_{t\in[0,T]}\|e_h'(t)\|_0^2 + 16Ch^{2\min\{m,j\}}\left(\left(\int_0^s \|u''(t)\|_0\,dt\right)^2 + h^2\max_{t\in[0,T]}\|u'(t)\|_0^2\right).$$

And using $\varepsilon = \frac{2}{C_{pd}}$ and the approximation results in Lemma 5.3.9 we have,

$$\begin{aligned} T_2 &\leq \frac{1}{4}C_{pd}\max_{t\in[0,T]}\|e_h'(t)\|_{V(h)}^2 + \frac{1}{C_{pd}}T^2\max_{t\in[0,T]}\|(u-P_2u)'(t)\|_{V(h)}^2 \\ &\leq \frac{1}{4}C_{pd}\max_{t\in[0,T]}\|e_h'(t)\|_{V(h)}^2 + \frac{1}{C_{pd}}T^2\max_{t\in[0,T]}\|(u-P_2u)'(t)\|_{V(h)}^2 \\ &\leq \frac{1}{4}C_{pd}\max_{t\in[0,T]}\|e_h'(t)\|_{V(h)}^2 + \frac{C}{C_{pd}}T^2\max_{t\in[0,T]}\|u'(t)\|_{1+m}^2. \end{aligned}$$

To bound T_3 we use Lemma 5.4.1. Clearly $P_2u - u_h$ satisfy the properties in Lemma 5.4.1 and so we have:

$$\begin{aligned} T_3 &\leq 2\int_0^s |r_h(u(t),(P_2u-u_h)'(t))|\,dt \\ &\leq C_R h^{\min\{m,j\}}\max_{t\in[0,T]}\|(P_2u-u_h)(t)\|_{V(h)}\left(2\max_{t\in[0,T]}\|u(t)\|_{1+m} + T\max_{t\in[0,T]}\|u'(t)\|_{1+m}\right). \end{aligned}$$

By using the triangle inequality, Young's inequality (Lemma C.1) and Lemma 5.3.10 we then have

$$\begin{aligned} T_3 &\leq 2C_R\mathcal{K}(T)h^{\min\{m,j\}}\left(\max_{t\in[0,T]}\|e_h(t)\|_{V(h)} + \max_{t\in[0,T]}\|(u-P_2u)(t)\|_{V(h)}\right) \\ &\leq 2C_R\mathcal{K}(T)h^{\min\{m,j\}}\left(\max_{t\in[0,T]}\|e_h(t)\|_{V(h)} + C_A h^{\min\{m,j\}}\max_{t\in[0,T]}\|u(t)\|_{1+m}\right) \\ &\leq \frac{1}{4}C_{pd}\max_{t\in[0,T]}\|e_h(t)\|_{V(h)}^2 + \mathcal{C}h^{2\min\{m,j\}}\left(\max_{t\in[0,T]}\|u(t)\|_{1+m}^2 + \mathcal{K}^2(T)\right). \end{aligned}$$

where

$$\mathcal{K}(T) = 2 \max_{t\in[0,T]} \|u(t)\|_{1+m} + T \max_{t\in[0,T]} \|u'(t)\|_{1+m}$$

and $\mathcal{C} = \max\{16C_R^2, C_R C_A\}$.

From the work done above, it now follows that for every $t \in J$

$$\|e_h'(t)\|_0 + \|e_h(t)\|_{V(h)} \leq \mathcal{C}_1 \left(\|e_h'(0)\|_0 + \|e_h(0)\|_{V(h)}\right)$$
$$+\mathcal{C}_2 h^{\min\{m,j\}} \left(\max_{t\in[0,T]} \|u(t)\|_{1+m} + T \max_{t\in[0,T]} \|u'(t)\|_{1+m} + \int_0^s \|u''(t)\|_0 \, dt\right),$$

for constants $\mathcal{C}_1$ and $\mathcal{C}_2$ independent of T and h. □

The error estimate in the DG energy norm has been proved.

5.5 Proof of the $\mathcal{L}^2$-norm error estimate

As mentioned, the proof of Theorem 5.3.3 ([GSS06, Theorem 4.2], the $\mathcal{L}^2$-norm error estimate) is based on the proof for the error estimate with respect to the $\mathcal{L}^2$-norm in [Bak76] for conforming finite element approximations. [GSS06] defines a mapping that is used in the same way as the Galerkin projection in [Bak76] and [BV13]. The mapping is defined below.

The mapping $\mathcal{P}_h$ is defined for $u \in H^{1+m}(\Omega)$ with $m > \frac{1}{2}$ by

$$\tilde{b}_h(\mathcal{P}_h u, v) = \tilde{b}_h(u, v) - r_h(u, v) \text{ for all } v \in V^h. \tag{5.5.1}$$

This defines a function $\mathcal{P}_h u$ by $(\mathcal{P}_h u)(t) = \mathcal{P}_h u(t)$ for $t \in [0, T]$.

Now consider for $t \in [0, T]$,

$$\|e_h(t)\|_0^2 = \|u(t) - u_h(t)\|_0^2 \leq \|u(t) - \mathcal{P}_h u(t)\|_0^2 + \|\mathcal{P}_h u(t) - u_h(t)\|_0^2. \tag{5.5.2}$$

The first term on the right hand side of (5.5.2) can be estimated from the bounds in Lemma 5.5.1 below, while an estimate for the second term on the right hand side of (5.5.2) have to be estimated.

Remark Since [GSS06] follows the proof for the $\mathcal{L}^2$-norm error estimate in [Bak76], the same steps are followed in the proof and hence the assumptions made in [GSS06] should be the same as in [Bak76]. As was mentioned in Chapter 3 of this dissertation, the assumption that $(P_h u)''$ exists is used without mention or discussion in [Bak76], and the same happens in [GSS06] for $(\mathcal{P}_h u)''$. We therefore use the results of [BV13] in the proof of Theorem 5.3.3, since the results in [BV13] (and Chapter 3 in this dissertation) do not require that $(\mathcal{P}_h u)''$ exist. Note that the results are trivially true for $a = 0$.

Recall the error equation obtained in Lemma 5.3.8,

$$(e_h'', v) + \tilde{b}_h(e_h, v) = r_h(u, v) \text{ for all } v \in V^h \text{ a.e. in } J,$$

where $e_h = u - u_h$. We then have that

$$(e_h'', v) + \tilde{b}_h(u - \mathcal{P}_h u, v) + \tilde{b}_h(\mathcal{P}_h u - u_h, v) = r_h(u, v) \text{ for all } v \in V^h \text{ a.e. in } J,$$

and from the definition of $\mathcal{P}_h$, we obtain

$$(e_h'', v) + \tilde{b}_h(\mathcal{P}_h u - u_h, v) = r_h(u, v) - r_h(u, v) = 0 \text{ for all } v \in V^h \text{ a.e. in } J.$$

We can now use the results for the semi-discrete approximation error estimate from [BV13], given in Chapter 3. Recall the result in Lemma 3.1.4, adapted for the case where the bilinear form $a = 0$: for $t \in [0, T]$,

$$\begin{aligned}
&\|(u_h - \mathcal{P}_h u)(t)\|_0 \\
&\qquad \leq \sqrt{2}\|(u_h - \mathcal{P}_h u)(0)\|_0 + 2T\|(u - u_h)'(0)\|_0 + 4\sqrt{T} \max_{t\in[0,T]} \|(u - \mathcal{P}_h)'(t)\|_0 \\
&\qquad \leq \sqrt{2}\left(\|u_0^h - u_0\|_0 + \|u_0 - \mathcal{P}_h u_0\|_0\right) + 2T\|v_0 - u_1^h\|_0 \\
&\qquad\qquad + 4\sqrt{T} \max_{t\in[0,T]} \|(u - \mathcal{P}_h)'(t)\|_0. \qquad (5.5.3)
\end{aligned}$$

The following two lemmas are from [GSS06] and gives estimates for some of the terms on the right hand side of Equation (5.5.3).

Lemma 5.5.1. *[GSS06, Lemma 4.9]*
Let the mapping $\mathcal{P}_h$ be defined by (5.5.1). Then we have

$$\|u - \mathcal{P}_h u\|_{V(h)} \leq C_E h^{\min\{m,j\}} \|u\|_{1+m}$$

with a constant C_E that is independent of h and depends only on C_{pd} in Lemma 5.3.7 and C_A, C_R in Lemma 5.3.10.

Moreover, if the elliptic regularity defined in (5.3.1) and (5.3.2) holds, we have

$$\|u - \mathcal{P}_h u\|_0 \leq C_L h^{\min\{m,j\}+1} \|u\|_{1+m}$$

with a constant C_L that is independent of h and depends only on the stability constant C_S in (5.3.2), C_{pd} in Lemma 5.3.7 and C_A, C_R in Lemma 5.3.10.

Proof.

The $\|\cdot\|_{V(h)}$ bound

Using the triangle inequality we have

$$\|u - \mathcal{P}_h u\|_{V(h)} \leq \|u(t) - P_2 u\|_{V(h)} + \|P_2 u(t) - \mathcal{P}_h u\|_{V(h)}.$$

From Lemma 5.3.10 we know that $\|u - P_2 u\|_{V(h)} \leq C_A h^{\min\{m,j\}} \|u\|_{1+m}$. We therefore need to bound the term

$$\|P_2 u - \mathcal{P}_h u\|_{V(h)}.$$

Note that from Lemma 5.3.7 (the continuity and positive definiteness of $\tilde{b}_h$) and the definition of $\mathcal{P}_h$ we have

$$\begin{aligned} C_{pd}\|P_2 u - \mathcal{P}_h u\|^2_{V(h)} &\leq \tilde{b}_h(P_2 u - \mathcal{P}_h u, P_2 u - \mathcal{P}_h u) \\ &= \tilde{b}_h(P_2 u - u, P_2 u \mathcal{P}_h u \\ &\qquad + \tilde{b}_h(u - \mathcal{P}_h u, P_2 u - \mathcal{P}_h u) \\ &= \tilde{b}_h(P_2 u - u, P_2 u - \mathcal{P}_h u) \\ &\qquad + r_h(u, P_2 u - \mathcal{P}_h u) \\ &\leq \|P_2 u - u\|_{V(h)} \|P_2 u - \mathcal{P}_h u\|_{V(h)} \\ &\qquad + r_h(u, P_2 u - \mathcal{P}_h u). \end{aligned}$$

[GSS06] now uses Lemma 5.3.10 to obtain an estimate for the above result. However, to use Lemma 5.3.10, it must hold that $P_2 u - \mathcal{P}_h u \in V(h)$, that is, $\mathcal{P}_h u \in V(h)$. This is not obviously true, since $\mathcal{P}_h$ is defined to be in V^h. If this is true though, then use Lemma 5.3.10 to obtain

$$\begin{aligned} C_{pd}\|P_2 u(t) - \mathcal{P}_h u(t)\|^2_{V(h)} &\leq C_A h^{\min\{m,j\}} \|u(t)\|_{1+m} \|P_2 u(t) - \mathcal{P}_h u(t)\|_{V(h)} \\ &\qquad + C_R h^{\min\{m,j\}} \|u(t)\|_{1+m} \|P_2 u(t) - \mathcal{P}_h u(t)\|_{V(h)}. \end{aligned}$$

Therefore,

$$\|P_2 u - \mathcal{P}_h u\|_{V(h)} \leq \left(\frac{C_A + C_R}{C_{pd}}\right) h^{\min\{m,j\}} \|u\|_{1+m}.$$

The $\mathcal{L}^2$-bound

This bound is proved in a similar way as the Aubin-Nitsche trick and relies on the estimate $\|\cdot\|_{V(h)}$ obtained above.

□

It follows from Lemma 5.5.1 that the following estimates hold.

Lemma 5.5.2. *[GSS06, Lemma 4.10] Let $\mathcal{P}_h u$ be defined by (5.5.1). Under the regularity assumptions of Theorem 5.3.3, we have*

$$\begin{aligned} \|(u - \mathcal{P}_h u)'\|_{C(J,V(h))} &\leq C_E h^{\min\{m,j\}} \|u'\|_{C(J,H^{1+m}(\Omega))} \\ \|(u - \mathcal{P}_h u)(0)\|_{V(h)} &\leq C_E h^{\min\{m,j\}} \|u_0\|_{1+m}. \end{aligned}$$

Moreover, if elliptic regularity as defined in (5.3.1) and (5.3.2) holds, we have the $\mathcal{L}^2$-bounds

$$\begin{aligned} \|(u - \mathcal{P}_h u)'\|_{C(J,\mathcal{L}^2(\Omega))} &\leq C_L h^{\min\{m,j\}+1} \|u'\|_{C(J,H^{1+m}(\Omega))} \\ \|(u - \mathcal{P}_h u)(0)\|_0 &\leq C_E h^{\min\{m,j\}+1} \|u_0\|_{1+m}. \end{aligned}$$

Proof of Theorem 5.3.3:

Recall Equation 5.5.2:

$$\|e_h(t)\|_0 = \|u(t) - u_h(t)\|_0 \leq \|u(t) - \mathcal{P}_h u(t)\|_0 + \|\mathcal{P}_h u(t) - u_h(t)\|_0.$$

We now have an estimate for $\|u(t) - \mathcal{P}_h u(t)\|_0$:

$$\|u(t) - \mathcal{P}_h u(t)\|_0 \leq C_E h^{\min\{m,j\}} \|u(t)\|_{1+m},$$

and a higher order estimate:

$$\|u(t) - \mathcal{P}_h u(t)\|_0 \leq C_L h^{\min\{m,j\}+1} \|u(t)\|_{1+m}.$$

We need estimates for the terms on the right hand side of Equation (5.5.3) (an estimate for $\|\mathcal{P}_h u(t) - u_h(t)\|_0$). [GSS06] proves in Lemma 5.5.2 and Lemma 5.5.2 estimates for the projection error. Since $\|\cdot\|_0 \leq \|\cdot\|_{V(h)}$, we therefore have estimates for two terms on the right hand side of Equation (5.5.3):

$$\begin{aligned}
&\|(u_h - \mathcal{P}_h u)(t)\|_0 \\
&\quad \leq \sqrt{2}\left(\|u_0^h - u_0\|_0 + \|u_0 - \mathcal{P}_h u_0\|_0\right) + 2T\|v_0 - u_1^h\|_0 \\
&\qquad\qquad + 4\sqrt{T}\max_{t\in[0,T]}\|(u - \mathcal{P}_h)'(t)\|_0 \\
&\quad \leq \sqrt{2}\left(\|u_0^h - u_0\|_0 + C_E h^{\min\{m,j\}}\|u_0\|_{1+m}\right) + 2T\|v_0 - u_1^h\|_0 \\
&\qquad\qquad + 4\sqrt{T}\max_{t\in[0,T]} C_E h^{\min\{m,j\}}\|u'(t)\|_{1+m}.
\end{aligned}$$

A result similar to the Aubin-Nitsche trick to get a higher order error estimate in the $\mathcal{L}^2$-norm is also proven in Lemma 5.5.2 and Lemma 5.5.2 in [GSS06]. If elliptic regularity as defined in (5.3.1) and (5.3.2) holds, the higher order $\mathcal{L}^2$-bounds then give:

$$\begin{aligned}
&\|(u_h - \mathcal{P}_h u)(t)\|_0 \\
&\quad \leq \sqrt{2}\left(\|u_0^h - u_0\|_0 + C_L h^{\min\{m,j\}+1}\|u_0\|_{1+m}\right) + 2T\|v_0 - u_1^h\|_0 \\
&\qquad\qquad + 4\sqrt{T}\max_{t\in[0,T]} C_L h^{\min\{m,j\}+1}\|u'(t)\|_{1+m}.
\end{aligned}$$

Therefore:

$$\begin{aligned}
\|e_h(t)\|_0 \quad \leq \quad & C_E h^{\min\{m,j\}}\|u(t)\|_{1+m} + \sqrt{2}\left(\|u_0^h - u_0\|_0 + C_E h^{\min\{m,j\}}\|u_0\|_{1+m}\right) \\
& + 2T\|v_0 - u_1^h\|_0 + 4\sqrt{T}\max_{t\in[0,T]} C_E h^{\min\{m,j\}}\|u'(t)\|_{1+m},
\end{aligned}$$

and a higher order estimate:

$$\begin{aligned}
\|e_h(t)\|_0 \quad \leq \quad & C_L h^{\min\{m,j\}+1}\|u(t)\|_{1+m} + \sqrt{2}\left(\|u_0^h - u_0\|_0 + C_L h^{\min\{m,j\}+1}\|u_0\|_{1+m}\right) \\
& + 2T\|v_0 - u_1^h\|_0 + 4\sqrt{T}\max_{t\in[0,T]} C_L h^{\min\{m,j\}+1}\|u'(t)\|_{1+m}.
\end{aligned}$$

□

Final estimate

The terms $\|u_0^h - u_0\|_0 = \|e_h(0)\|_0$ and $\|v_0 - u_1^h\|_0 = \|e_h'(0)\|_0$ still need to be estimated. If we choose $u_0^h = P_2 u_0$ and $u_1^h = P_2 v_0$ as is done in [GSS06], we can use the approximation properties in Lemma 5.3.9. Recall the remark made on page 110. Therefore

$$\|e_h(0)\|_0 \leq Ch^{\min\{m,j\}}\|u_0\|_{1+m},$$
$$\|e_h'(0)\|_0 \leq Ch^{\min\{m,j+1\}}\|v_0\|_m.$$

Remark The aim of the investigation in this dissertation regarding the discontinuous Galerkin finite element method was to ascertain if it has any impact on the assumptions made on the regularity of the solution u, in order to obtain convergence of the semi-discrete (and fully discrete) approximations. It was found that the assumptions made on the regularity of the solution u by making use of the discontinuous Galerkin finite element method to obtain a semi-discrete error estimate (Theorem 5.3.2 and Theorem 5.3.3) are only fractionally better than those for the the continuous Galerkin finite element method. In particular, it is assumed that $u''(t) \in H^{1+m}(\Omega)$ (with $m > \frac{1}{2}$) for the DG approximation, while it is assumed that $u''(t) \in H^2(\Omega)$ for the continuous Galerkin approximation, for $t \in J$. However, the discontinuous Galerkin finite element method has the advantage of being able to use non-conforming finite element meshes, as stated in [GSS06]: "Based on discontinuous finite element spaces, the proposed DG method easily handles elements of various types and shapes, irregular nonmatching grids, and even locally varying polynomial order."

Proof of the so called inverse inequality

Recall the following so called inverse inequality (5.3.6):

$$\|\mathbf{w}\|^2_{0,\partial E} \leq C^2_{inv}h_E^{-1}\|\mathbf{w}\|^2_{0,E} \text{ for all } \mathbf{w} \in (S^j(E))^d,$$

with a constant C_{inv} that depends only on the shape regularity of the mesh, the approximation order j and the dimension d.

Consider the inverse property in [OR76, p.341] (due to Babuška and Aziz): there exists real number $m > 0$ such that for every real number $s \leq m$,

$$\|u\|_{H^m(\Omega)} \leq \mathcal{C}h^{-(m-s)}\|u\|_{H^s(\Omega)}, \tag{5.5.4}$$

where $\mathcal{C}$ does not depend on u or h. Note that it is not well known that this result is holds for real numbers and is not only true for integers.

The so-called inverse inequality (5.3.6) therefore follows from the estimates for the trace operator (Section A.4) and the inverse property (5.5.4) above:

$$\|u\|_{0,\partial E} \leq C\|u\|_{H^{\frac{1}{2}}(E)} \leq C\mathcal{C}h_E^{-\frac{1}{2}}\|u\|_{0,E}.$$

5.6 Fully discrete approximation

In this section the article [GS09] is investigated. In it the error analysis done in [GSS06] is extended to the fully discrete numerical scheme. A centered second-order finite difference approximation (called the "leap-frog" scheme in [GS09]) is used for the time discretisation. Optimal error estimates are obtained for the fully discrete approximation in the $\mathcal{L}^2$-norm. This error analysis has also been done in [Kar12], but the time discretisation scheme used in [Kar12] is the general Newmark method in Chapter 4. It is worth noting again that the semi-discrete approximation error estimate and the fully discrete error estimate are kept separately here.

The semi-discrete approximation is given by Problem WDG (Section 5.2.2).

For the fully discrete approximation, divide the interval $[0, T]$ into N time steps of length $\tau = \frac{T}{N}$. Denote the approximation by $u_h^n \approx u_h(t_n)$, where $u_h(t_n)$ denotes the DG approximation. Recall that we defined for any sequence $\{y_k\} \subset \mathbb{R}^n$,

$$\delta_t^2 y_k = \frac{y_{k+1} - 2y_k + y_{k-1}}{\tau^2} \text{ for } k = 1, 2, \ldots, N-1.$$

Recall the definition of the discontinuous Galerkin finite element space V^h in Chapter 5:

$$V^h := \{v \in \mathcal{L}^2(\Omega) : v|_E \in \mathcal{S}^j(E) \text{ for all } E \in \mathcal{M}_h\}.$$

We now have the fully discrete numerical approximation to the wave equation.

Problem WFD

Find a sequence $\{u_h^n\} \subset V^h$ such that for $n = 1, 2, \ldots, N-1$,

$$\begin{aligned} \left(\delta_t^2 u_h^n, v\right) + b_h(u_h^n, v) &= (f^n, v) \text{ for all } v \in V^h, \\ u_h^0 &= P_2 u^0, \\ u_h^1 &= u_h^0 + \tau P_2 v^0 + \frac{\tau^2}{2}\widetilde{u}_h^0, \end{aligned} \tag{5.6.1}$$

where $\widetilde{u}_h^0 \in V^h$ is defined by

$$(\widetilde{u}_h^0, v) = (f^0, v) - b_h(u^0, v) \text{ for all } v \in V^h.$$

In (5.6.1), every time step involves the inversion of the DG mass matrix M. Since it is symmetric positive definite, the new approximations u_h^{n+1} are well-defined for $n \geq 1$. Therefore the fully discrete DG approximations $\{u_h^n\}_{n=0}^{n=N}$ are uniquely defined by (5.6.1) and the initial conditions in Problem WFD.

The following remark is made in [GSS06] with regards to the DG mass matrix M: "It can be inverted at very low computational cost, and the scheme is essentially fully explicit. In fact, if the basis functions are chosen mutually orthogonal, M reduces to the identity."

5.6.1 Properties of the bilinear form b_h

In order to show the key properties of the bilinear form b_h [GS09] introduces the broken norm below, where D^2u denotes the matrix of second derivatives of the solution u.

$$\|u\|_*^2 := \sum_{E\in\mathcal{M}_h} \|\nabla u\|_{0,E}^2 + \sum_{E\in\mathcal{M}_h} h_E^2 \|D^2 u\|_{0,E}^2 + \sum_{F\in\mathcal{F}_h} h_F^{-1} \|[[u]]\|_{0,F}^2. \tag{5.6.2}$$

Remark Here [GS09] differs from [GSS06]. As was seen in the beginning of this chapter, a bilinear form $\tilde{b}_h$ had to be introduced to obtain an error estimate for the semi-discrete problem. However, here this is not necessary, but the broken norm above is needed for the derivation of the error estimate for the fully discrete case. As stated in [GS09, Remark 3.1], "the norm $\|\cdot\|_*$ is the natural one for obtaining continuity of the bilinear form b_h on $H^2(\Omega) + V^h$, while the weaker DG norm $\|\cdot\|_{V(h)}$ is enough for obtaining coercivity." Also see Appendix A.5 for more detail on broken Sobolev spaces.

The following results are special cases of the results obtained in [ABCM02, Sections 4.1 and 4.2]. The results are similar to the results in Lemma 5.3.7.

Lemma 5.6.1. *[GS09, Lemma 3.2]*
There exists a threshold value $\beta_{\min} > 0$, independent of the mesh size, such that for $\beta \geq \beta_{\min}$ there holds

$$\begin{aligned} b_h(u,u) &\geq C_{pd}\|u\|_*^2, \quad u \in V^h, \\ |b_h(u,v)| &\leq C_E(c^*)^2 \max\{1,\beta\}\|u\|_*\|v\|_*, \quad u,v \in H^2(\Omega) + V^h, \end{aligned}$$

with a positive definiteness constant $C_{pd} > 0$ that is independent of the mesh size, and a continuity constant $C_E > 0$ that is independent of the mesh size, c and β.

Lemma 5.6.2. *[GS09, Lemma 3.3]*
For quasi-uniform meshes $\mathcal{M}_h$, there holds

$$b_h(u,u) \leq C_S(c^*)^2 \max\{1,\beta\} h^{-2}\|u\|_0^2 \text{ for } u \in V^h,$$

with a stability constant $C_S > 0$ that is independent of the mesh size, c and β.

The proof of this lemma is given in detail in [GS09]. It relies on the continuity of b_h in Lemma 5.6.1, the definition of the norm $\|\cdot\|_*$, and the inverse inequalities

$$\|\nabla u\|_{0,E} \leq Ch^{-1}\|u\|_{0,E} \quad \text{and} \quad \|D^2 u\|_{0,E} \leq Ch^{-2}\|u\|_{0,E}.$$

Finally, suppose that the mesh size h and the time step τ satisfy the CFL condition

$$\frac{\tau}{h} < \frac{2}{c^*\sqrt{C_S \max\{1,\beta\}}}. \tag{5.6.3}$$

5.6.2 Convergence

Following [GS09], we first state the main theorem below (the fully discrete error estimate), and then prove it in the sections that follow. In this subsection the Galerkin projection is defined and estimates are obtained. In Subsection 5.6.3 a stability result is obtained (Proposition 5.6.7) and estimates for some of the terms in this stability result are given (Lemmas 5.6.8 and 5.6.9). Theorem 5.6.3 is then proved by combining these results.

Theorem 5.6.3. *[GS09, Theorem 3.4]*
Let the solution u of the wave problem satisfy the regularity properties (p

$$u \in C^2(J, H^{p+1}(\Omega)), \quad u''' \in C(J, \mathcal{L}^2(\Omega)), \quad u^{(4)} \in L^1(J, \mathcal{L}^2(\Omega)), \tag{5.6.4}$$

and let the discrete finite element approximations $\{u_h^n\}_{n=0}^N$ be defined by (5.6.1) together with the initial conditions in Problem WFD. Assume that the CFL condition (5.6.3) is satisfied. Then the following a priori error estimate holds:

$$\max_{n=0,\ldots,N} \|u(t_n) - u_h^n\|_0 \leq \tilde{C}(h^{p+1} + \tau^2),$$

where $\tilde{C} > 0$ is a constant independent of the mesh size and the time step.

Galerkin projection

The steps in the proof of Theorem 5.6.3 are the same as that of Theorem 4.4.1 in Section 4.4. However, we are now dealing with the DG method, and therefore error estimates for the Galerkin projection $P_h u \in V^h$ defined below are required.

Definition 5.6.4. Let $u \in H^2(\Omega)$ and define the Galerkin projection $P_h u \in V^h$ of u by

$$b_h(P_h u, v) = b_h(u, v) \text{ for all } v \in V^h. \tag{5.6.5}$$

Lemma 5.6.5. *[GS09, Lemma 4.1]*
If, additionally, $u \in H^{p+1}(\Omega)$ for $p \geq 1$, then

$$\begin{aligned} \|u - P_h u\|_* &\leq Ch^p \|u\|_{p+1}, \\ \|u - P_h u\|_0 &\leq Ch^{p+1} \|u\|_{p+1}, \end{aligned}$$

with a constant $C > 0$ that is independent of the mesh size.

The proof of this lemma uses Lemma 5.6.1 and interpolation error estimates, not properly cited in [GS09]. The estimate with respect to the $\mathcal{L}^2$-norm is a higher order estimate.

A function $P_h u(t) \in V^h$ is defined by $(P_h u)(t) = P_h u(t)$ for all $t \in [0, T]$. Then the following lemma holds.

Lemma 5.6.6. *[GS09, Lemma 4.2]*
Let u satisfy the regularity properties in Theorem 5.6.3 and let $P_h u$ be defined by Definition 5.6.4. Then

$$\|(u - P_h u)^{(k)}(t)\|_0 \leq Ch^{p+1} \|u^{(k)}(t)\|_{p+1}, \ \textit{for } k = 0, 1, 2, \ \textit{and } t \in [0, T]$$

with a constant $C > 0$ that is independent of the mesh size.

Proof. This follows immediately from Lemma 5.6.5. □

5.6.3 Proof of Theorem 5.6.3

The steps in the proof of Theorem 5.6.3 are the same as those in the proof of Theorem 4.4.1 in Section 4.4. However, in [GS09] the undamped wave equation is considered and the central difference scheme (the Newmark method with $\theta = 0$ and $\gamma = \frac{1}{2}$) is used for the time discretisation. In the proofs of the different steps the similarities and differences will become clear.

Recall the following notation from Section 4.4 (adapted for the undamped case and using the central difference method).

Notation

$$e_n^h = P_h u(t_n) - u_h^n, \quad e_p(t_n) = u(t_n) - P_h u(t_n), \quad e^n = u(t_n) - u_h^n$$

The error e^n is split in the following manner:

$$e^n = e_n^h + e_p(t_n).$$

Define $r^n \in V^h$ for $1 \leq n \leq N-1$, by

$$(r^n, v) = (\delta_t^2 P_h u(t_n) - u''(t_n), v), \tag{5.6.6}$$

and for $n = 0$,

$$(r^0, v) = \tau^{-2}(e_1^h - e_0^h, v) \tag{5.6.7}$$

for all $v \in V^h$.

Set

$$R^n = \tau \sum_{m=0}^{n} r^m. \tag{5.6.8}$$

Proposition 5.6.7. *[GS09, Proposition 3.4]*
Assume that the CFL condition (4.3.6) *holds. Then we have*

$$\max_{1\leq n\leq N} \|e^n\|_0 \leq C^* \left(\|e^0\|_0 + \max_{1\leq n\leq N} \|e_p(t_n)\|_0 + \tau \sum_{n=0}^{N-1} \|R^n\|_0 \right), \tag{5.6.9}$$

with a constant $C > 0$ independent of h, τ and T.

Proof. We follow the exact same steps as in the proof of Proposition 4.4.2 in Section 4.4 with $\theta = 0$ and $\gamma = \frac{1}{2}$, and since the bilinear form b_h is symmetric we can use it in the same way as we used the bilinear form b in Proposition 4.4.2. Note that the steps of the proof of [GS09, Proposition 4.3] differ slightly from the proof of Proposition 4.4.2 in Section 4.4 ([Kar11a, Proposition 2]), but we get to the following step with either.

We have that for $1 \leq n \leq N$,

$$\|e_n^h\|_0^2 - \frac{\tau^2}{4} b_h\left(e_n^h, e_n^h\right) \leq \|e_0^h\|_0^2 + \tau \sum_{m=0}^{n-1} \left(R^m, e_{m+1}^h + e_m^h\right). \tag{5.6.10}$$

Recall the estimate from Lemma 5.6.2:

$$b_h(e_n^h, e_n^h) \leq C_S (c^*)^2 \max\{1, \beta\} h^{-2} \|e_n^h\|_0^2.$$

Now, if the CFL condition (5.6.3) holds, we then have for $1 \leq n \leq N$,

$$D_* \|e_n^h\|_0^2 \leq \|e_0^h\|_0^2 + \tau \sum_{m=0}^{n-1} \left(R^m, e_{m+1}^h + e_m^h\right), \tag{5.6.11}$$

where

$$D_* = 1 - \frac{\tau^2}{4} C_S (c^*)^2 \max\{1, \beta\} h^{-2} > 0.$$

Recalling from the proof of Proposition 4.4.2, we know that

$$\sum_{m=0}^{n-1} \left(R^m, e_{m+1}^h + e_m^h\right) \leq 2 \left(\max_{0 \leq n \leq N} \|e_n^h\|_0\right) \left(\sum_{n=0}^{N-1} \|R^n\|_0\right). \tag{5.6.12}$$

Following from Young's inequality, Lemma C.1, with

$$a = \max_{0 \leq n \leq N} \|e_n^h\|_0, \qquad b = 2\tau \sum_{n=0}^{N-1} \|R^n\|_0, \qquad \varepsilon = \frac{1}{D_*},$$

we have

$$D_* \|e_n^h\|_0^2 \leq \|e_0^h\|_0^2 + \frac{D_*}{2} \max_{0 \leq n \leq N} \|e_n^h\|_0^2 + \frac{2}{D_*} \left(\sum_{n=0}^{N-1} \|R^n\|_0\right)^2. \tag{5.6.13}$$

Since the right hand side of (5.6.13) does not depend on n, we take the maximum over $n = 0$ to $n = N$ to obtain

$$\frac{D_*}{2} \max_{0 \leq n \leq N} \|e_n^h\|_0^2 \leq \|e_0^h\|_0^2 + \frac{2}{D_*} \left(\sum_{n=0}^{N-1} \|R^n\|_0\right)^2$$

and therefore

$$\max_{0\leq n\leq N}\|e_n^h\|_0 \leq \sqrt{\frac{2}{D_*}}\|e_0^h\|_0 + \frac{\sqrt{2}}{D_*}\left(\sum_{n=0}^{N-1}\|R^n\|_0\right).$$

Using this together with

$$\max_{0\leq n\leq N}\|e^n\|_0 \leq \max_{0\leq n\leq N}\|e_n^h\|_0 + \max_{0\leq n\leq N}\|e_p(t_n)\|_0,$$

we have the desired result, with $C^* = \min\{1, \sqrt{\frac{2}{D_*}}, \frac{\sqrt{2}}{D_*}\}$. □

Lemma 5.6.8. *[GS09, Lemma 4.4] There holds*

$$\|r^0\|_0 \leq \tau C_1\left(\|u'''\|_{C(J,\mathcal{L}^2(\Omega))} + \tau^{-1}h^{p+1}\|u'\|_{C(J,H^{p+1}(\Omega))}\right),$$

with the constant $C_1 > 0$ independent of h, τ and T.

Proof. Once again, the steps in this proof follows the exact same way as the proof of Lemma 4.4.3 in Section 4.4 with $\theta = 0$ and $\gamma = \frac{1}{2}$. Also note that we have no damping term a as in Lemma 4.4.3.

Note that the proof relies on Lemma 5.6.6. □

Lemma 5.6.9. *[GS09, Lemma 4.5] There holds for $1 \leq n \leq N-1$,*

$$\|r^n\|_0 \quad \leq \quad C_2\left(h^{p+1}\tau^{-1}\int_{t_{n-1}}^{t_{n+1}}\|u''(s)\|_{p+1}ds + \tau\int_{t_{n-1}}^{t_{n+1}}\|u^{(4)}(s)\|_0 ds\right)$$

with the constant $C_2 > 0$ independent of h, τ and T.

Proof. Once again, the steps in this proof follows the format of the proof of Lemma 4.4.4 in Section 4.4 with $\theta = 0$ and $\gamma = \frac{1}{2}$. Also note that we have no damping term a as in Lemma 4.4.4.

Note that the proof relies on Lemma 5.6.6. □

Proposition 5.6.10. *[GS09, Proposition 4.6] For $0 \leq n \leq N-1$, there holds*

$$\begin{aligned}\|R^n\|_0 \quad \leq \quad & C_3h^{p+1}\left(\|u'\|_{C(J,H^{p+1}(\Omega))} + \|u''\|_{C(J,H^{p+1}(\Omega))}\right) \\ & + C_3\tau^2\left(\|u'''\|_{C(J,\mathcal{L}^2(\Omega))} + \|u^{(4)}\|_{\mathcal{L}^1(J,\mathcal{L}^2(\Omega))}\right),\end{aligned}$$

with the constant $C_3 > 0$ independent of h, τ and T.

Proof. From the definition of R^n (with $0 \leq n \leq N-1$) and using the triangle inequality we have that

$$\begin{aligned}\|R^n\|_0 = (R^n, R^n)_0 &= (\tau \sum_{m=1}^{n} r^m, \tau \sum_{m=1}^{n} r^m)_0 \\ &\leq \tau\|r^0\|_0 + \tau \sum_{m=1}^{n} \|r^m\|_0 \\ &\leq \tau\|r^0\|_0 + \tau \sum_{m=1}^{N-1} \|r^m\|_0.\end{aligned}$$

Then, by the bounds derived in Lemma 5.6.8 and Lemma 5.6.9, we have that

$$\begin{aligned}\|R^n\|_0 &\leq C_1 \left(\tau^2\|u'''\|_{C(J,\mathcal{L}^2(\Omega))} + h^{p+1}\|u'\|_{C(J,H^{p+1}(\Omega))}\right) \\ &\quad + C_2 \sum_{m=1}^{N-1} \left(h^{p+1}\int_{t_{m-1}}^{t_{m+1}} \|u''(s)\|_{p+1} ds + \tau^2 \int_{t_{m-1}}^{t_{m+1}} \|u^{(4)}(s)\|_0 ds\right) \\ &\leq h^{p+1}\left(C_1\|u'\|_{C(J,H^{p+1}(\Omega))} + 2C_2\|u''\|_{C(J,H^{p+1}(\Omega))}\right) \\ &\quad +\tau^2\left(C_1\|u'''\|_{C(J,\mathcal{L}^2(\Omega))} + 2C_2\|u^{(4)}\|_{\mathcal{L}^1(J,\mathcal{L}^2(\Omega))}\right) \\ &\leq C_3 h^{p+1}\left(\|u'\|_{C(J,H^{p+1}(\Omega))} + \|u''\|_{C(J,H^{p+1}(\Omega))}\right) \\ &\quad + C_3\tau^2\left(\|u'''\|_{C(J,\mathcal{L}^2(\Omega))} + \|u^{(4)}\|_{\mathcal{L}^1(J,\mathcal{L}^2(\Omega))}\right),\end{aligned}$$

where $C_3 = \min\{C_1, 2C_2\}$. □

Proof of Theorem 5.6.3

Now note that

$$\tau \sum_{n=0}^{N-1} \|R^n\|_0 \leq \tau N \max_{n=0}^{N-1} \|R^n\|_0 = T \max_{n=0}^{N-1} \|R^n\|_0,$$

and use this in Proposition 5.6.7 to obtain:

$$\max_{n=1}^{N} \|e^n\|_0 \leq C^* \left(\|e^0\|_0 + \max_{n=1}^{N} \|e_p(t_n)\|_0 + T \max_{n=0}^{N-1} \|R^n\|_0\right).$$

Using Lemma 5.6.6 we have

$$\max_{n=1}^{N} \|e_p(t_n)\|_0 \leq Ch^{p+1}\|u\|_{C(J,H^{p+1}(\Omega))}.$$

Also note that since we chose that $u_h^0 = P_2 u^0$, we have

$$\begin{aligned} \|e^0\|_0 = \|u^0 - u_h^0\|_0 &= \|u^0 - P_2 u^0\|_0 \\ &\leq Ch^{p+1}\|u^0\|_{p+1} \\ &\leq Ch^{p+1}\|u\|_{C(J,H^{p+1}(\Omega))}. \end{aligned}$$

Now, since the right hand side of the estimate in Proposition 5.6.10 does not depend on n we have

$$\begin{aligned} \max_{n=0}^{N-1} \|R^n\|_0 &\leq C_3 h^{p+1}\left(\|u'\|_{C(J,H^{p+1}(\Omega))} + \|u''\|_{C(J,H^{p+1}(\Omega))}\right) \\ &\quad + C_3\tau^2\left(\|u'''\|_{C(J,\mathcal{L}^2(\Omega))} + \|u^{(4)}\|_{\mathcal{L}^1(J,\mathcal{L}^2(\Omega))}\right). \end{aligned}$$

Finally we have

$$\begin{aligned} \max_{n=1}^{N} \|e^n\|_0 &\leq C^*\Bigg(Ch^{p+1}\|u\|_{C(J,H^{p+1}(\Omega))} + Ch^{p+1}\|u\|_{C(J,H^{p+1}(\Omega))} \\ &\quad + TC_3 h^{p+1}\left(\|u'\|_{C(J,H^{p+1}(\Omega))} + \|u''\|_{C(J,H^{p+1}(\Omega))}\right) \\ &\quad + TC_3\tau^2\left(\|u'''\|_{C(J,\mathcal{L}^2(\Omega))} + \|u^{(4)}\|_{\mathcal{L}^1(J,\mathcal{L}^2(\Omega))}\right)\Bigg) \\ &\leq C_T h^{p+1}\left(\|u\|_{C(J,H^{p+1}(\Omega))} + \|u'\|_{C(J,H^{p+1}(\Omega))} + \|u''\|_{C(J,H^{p+1}(\Omega))}\right) \\ &\quad + C_T\tau^2\left(\|u'''\|_{C(J,\mathcal{L}^2(\Omega))} + \|u^{(4)}\|_{\mathcal{L}^1(J,\mathcal{L}^2(\Omega))}\right), \end{aligned}$$

where the constant C_T depends on T, and so the constant $\tilde{C}$ in Theorem 5.6.3 grows linearly with T.

Since $\|u(t_n) - u_h^n\|_0 = \|e^n\|_0 \leq \max_{1\leq n\leq N} \|e^n\|_0 = \max_{1\leq n\leq N} \|u(t_n) - u_h^n\|_0$ we have the result. □

5.7 Numerical experiments

In the article [GSS06] three numerical examples are given. In each of the first two examples, an explicit (analytical) solution is available. The first example is to "confirm" the theoretical estimates. In the second example the solution has a "spatial singularity" at a boundary point and "the $\mathcal{L}^2$-error rates" are less than for Example 1. The authors claim that these results establish the "sharpness" of the regularity assumptions. The last example is not related to rates of convergence, but serves to illustrate the flexibility of the DG method.

Consider Example 1. It is stated that: "This solution is arbitrarily smooth so that all our theoretical regularity assumptions are satisfied." The authors then describe numerical experiments and calculations. "In Figure 5.2, the relative errors for the fully discrete approximation of (5.6) show convergence rates of order h in the energy norm and order h^2 in the $\mathcal{L}^2$-norm, thereby confirming the theoretical estimates of Theorems 4.1 and 4.2."

In Example 2 the authors consider the wave equation on a two dimensional L-shaped domain. Again an explicit solution is available. They claim that the solution is in $C^\infty(J, H^{\frac{5}{3}}(\Omega))$ without proof. Then they state "... regularity assumptions of Theorem 4.1 hold with $\sigma = \dfrac{2}{3}$. Thus, Theorem 4.1 predicts numerical convergence rates of $\dfrac{2}{3}$ in the energy norm, as confirmed by our numerical results in Table 5.1." Then they proceed to argue: "As the elliptic regularity assumptions (4.1) - (4.2) from Theorem 4.2 are violated, we do not expect $\mathcal{L}^2$-error rates of the order $1+\sigma$ for this problem. Indeed, in Table 5.1 we observe convergence rates close to $\dfrac{4}{3}$."

"To explain this behaviour, let us consider the following weaker **elliptic regularity assumption**: for any $f \in \mathcal{L}^2(\Omega)$ we assume that the solution of the problem

$$\begin{aligned} -\nabla \cdot (c\nabla z) &= f \quad \text{in} \quad \Omega, \\ z &= 0 \quad \text{on} \quad \partial\Omega, \end{aligned}$$

belongs to $H^{1+s}(\Omega)$ for a parameter $s \in \left(\frac{1}{2}, 1\right]$ and satisfies the following bound:
$\|z\|_{1+s} \leq C_S \|\lambda\|_0$."

No motivation or reference is given for this assumption. We conclude that not all the observed results are adequately explained by the theory and further research is indicated.

Chapter 6

Applications

In this chapter we apply the theory in Chapter 3 to some of the model problems introduced in Chapter 1.

6.1 The multidimensional wave equation with weak damping

6.1.1 Error estimate for the semi-discrete approximation

Recall from Section 2.5 the spaces:

$V = V(\Omega)$ is the closure of the test functions in $H^1(\Omega)$;

W is the space $\mathcal{L}^2(\Omega)$ with inner product c;

$X = \mathcal{L}^2(\Omega)$.

The semi-discrete form of the model problems, Problem MW and Problem HHCE are special cases of Problem G^h. A finite dimensional subspace S^h of V is constructed using (in two dimensions) piecewise linear basis functions on triangle elements and (in three dimensions) piecewise linear basis functions on tetrahedron elements. Note that the basis functions must satisfy the forced boundary condition: zero on $\partial\Omega - \Sigma$. We can therefore give the Galerkin finite element approximation of the weak variational forms of the model problems.

Recall that $\widetilde{f} : t \to f(\cdot, t)$.

Problem MW^h

Given a function $\widetilde{f} \in C([0,T], \mathcal{L}^2(\Omega))$, find a function $u_h \in C^2((0,T), S^h)$ such that u_h' is continuous at 0 and for each $t \in (0,T)$

$$c\left(u_h''(t), v\right) + a\left(u_h'(t), v\right) + b\left(u_h(t), v\right) = \left(\widetilde{f}(t), v\right)_{\mathcal{L}^2(\Omega)} \quad \text{for each } v \in S^h \tag{6.1.1}$$

$$\text{while} \quad u_h(0) = u_0^h\,, \quad u_h'(0) = u_1^h\,.$$

We now consider the semi-discrete error estimate in Theorem 3.1.5 and apply it to the model problems. We need to investigate whether the assumptions used in this theorem hold for the problems.

Piecewise linear basis functions on triangle elements

We consider the interpolation theory discussed in Appendix B. A general interpolation assumption (Assumption GI) is used in Section 3.1.2. For this situation we have that

$$H(V,k) = H^k(\Omega) \cap V(\Omega).$$

Instead of Assumption GI we have a specific estimate which depends on a concrete subspace S^h and a specific interpolation operator. For simplicity, first consider the two dimensional case on a rectangle with triangle elements. In Section B.2 we denote the interpolation operator for piecewise linear basis functions on triangle elements by $\Pi_\triangle$. If $k \geq 2$, then there exists a constant $\widehat{C}_\triangle$ such that for any $u \in H^k(\Omega)$

$$|\Pi_\triangle u - u|_{m,\Omega} \leq \widehat{C}_\triangle h^{2-m} |u|_{k,\Omega} \text{ for } m = 0,\ 1,\ 2.$$

Since the space $V(\Omega) \subset H^1(\Omega)$, we have that $\|\Pi_\triangle u(t) - u(t)\|_V \leq \widehat{C}_\triangle h |u(t)|_{k,\Omega}$, and so

$$\|e_p(t)\|_W \leq \|u(t) - \Pi_\triangle u(t)\|_W \leq \kappa_1 \|u(t) - \Pi_\triangle u(t)\|_V \leq \kappa_1 \widehat{C}_\triangle h |u(t)|_{k,\Omega},$$

for $t \in [0,T]$ and $k \geq 2$. Similarly,

$$\|e_p'(t)\|_W \leq \kappa_1 \widehat{C}_\triangle h |u'(t)|_{k,\Omega},$$

for $t \in [0,T]$ and $k \geq 2$.

Error estimate

Suppose that $u_0^h = \Pi_\triangle u_0$ and $u_1^h = \Pi_\triangle u_1$. If the solution u of Problem MW (and Problem HHCE) satisfies $u(t) \in H^2(\Omega)$ and $u'(t) \in H^2(\Omega)$, then it follows that for $t \in [0,T]$,

$$\begin{aligned}
\|u(t)-u_h(t)\|_W &\leq \|e_p(t)\|_W + \sqrt{2}\Bigg(\|Pu_0 - u_0\|_W + 3T\|u_1 - u_1^h\|_W \\
&\qquad + (1+3TC_W)\|u_0 - u_0^h\|_W + 3\int_0^T \|e_p'\|_W + 3C_W \int_0^T \|e_p\|_W \Bigg) \\
&\leq \kappa_1 \widehat{C}_\triangle h |u(t)|_{2,\Omega} + \kappa_1 \widehat{C}_\triangle \sqrt{2} h \Bigg(|u_0|_{2,\Omega} + 3T|u_1|_{2,\Omega} \\
&\qquad + (1+3TC_W)|u_0|_{2,\Omega} + 3\int_0^T |u'(t)|_{2,\Omega} + 3C_W \int_0^T |u(t)|_{2,\Omega} \Bigg) \\
&\leq \kappa_1 \widehat{C}_\triangle h |u(t)|_{2,\Omega} + \kappa_1 \widehat{C}_\triangle \sqrt{2} h \Bigg(3T|u_1|_{2,\Omega} + (2+3TC_W)|u_0|_{2,\Omega} \\
&\qquad + 3 \max_{t\in[0,T]} |u'(t)|_{2,\Omega} + 3C_W \max_{t\in[0,T]} |u(t)|_{2,\Omega} \Bigg).
\end{aligned}$$

Piecewise linear basis functions on tetrahedron elements

We now consider the three dimensional case on a 3-interval (“brick”) with tetrahedron elements. From Section B.4 we have that for piecewise linear basis functions on tetrahedron elements the interpolation operator is denoted by Π_t. If $k \geq 2$, then there exists a constant $\widehat{C}_t$ such that for any $u \in H^k(\Omega)$ we have

$$|\Pi_t u - u|_{m,\Omega} \leq \widehat{C}_t h^{2-m} |u|_{k,\Omega} \text{ for } m = 0,\ 1,\ 2.$$

The error estimate follows in the exact same way as for piecewise linear basis functions on triangle elements as done above.

6.1.2 Fully discrete error estimate

We can apply the error estimates for the fully discrete approximation of Problem G to our model problems. Consider the scheme from Section 3.2.

Problem MW^h-D

Find a sequence $\{u_k^h\} \subset S^h$ such that for $k = 0, 1, 2, \dots, N-1$,

$$\delta_t u_k^h = v_{k+\frac{1}{2}}, \tag{6.1.2}$$

$$c(\delta v_k, \varphi) + a(v_{k+\frac{1}{2}}, \varphi) + b(u_{k+\frac{1}{2}}^h, \varphi) = \frac{1}{2}(f(t_k) + f(t_{k+1}), \varphi)_X \tag{6.1.3}$$

for each $\varphi \in S^h$, while $u_0^h = u_h(0) = d^h$ and $u_1^h = u_h'(0) = v^h$.

For the result below, we either have that $\Pi = \Pi_\triangle$ or $\Pi = \Pi_t$. Recall the notation from Section 3.3: $u^{(k)} \in \mathcal{L}^2\Big([0,T]; Y\Big)$ if $u^{(k)}(t) \in Y$ for each t and $\int_{[0,T]} \|u^{(k)}\|_Y^2 < \infty$.

Suppose

(a) $u_0^h = \Pi u_0$ and $u_1^h = \Pi u_1$,

(b) $u'' \in \mathcal{L}^2([0,T], H^2(\Omega) \cap V(\Omega))$,

(c) $\tilde{f} \in C^2([0,T], \mathcal{L}^2(\Omega))$ and

(d) the sequence $\{u_k^h\}$ is a solution of Problem MW^h-D.

Then

$$\begin{aligned} \|u(t_k) - u_k^h\|_W \leq\ & \kappa_1 \widehat{C} h |u(t)|_{2,\Omega} + \kappa_1 \widehat{C}\sqrt{2} h \Bigg(3T|u_1|_{2,\Omega} + (2 + 3TC_W)|u_0|_{2,\Omega} \\ & + 3 \max_{t\in[0,T]} |u'(t)|_{2,\Omega} + 3C_W \max_{t\in[0,T]} |u(t)|_{2,\Omega} \Bigg) \\ & + 7T^2\tau^2 \max \|u_h^{(4)}\|_0 + 7T\tau^2 \max \|u_h'''\|_0 + \sqrt{2C_W}\,\tau^4 \max \|u_h'''\|_0, \end{aligned}$$

for each $t_k \in (0, T)$.

6.1.3 Convergence with less restrictive regularity assumptions

The assumption that $u''(t) \in H^2(\Omega)$ is necessary to obtain the error estimate of order h in Subsection 6.1.2. However, under less restrictive regularity assumptions, the solution u_h of Problem MW^h converges to the solution u of Problem MW as h tends to zero. (This is not mentioned in [BV13].) The convergence can be obtained without assuming anything more than the result of Theorem 2.5.4 in Section 2.5. Note that $H^2(\Omega)$ is dense in $H^1(\Omega)$, and since $V \subset H^1(\Omega)$, it follows that $H^2(\Omega) \cap V$ is dense in V (see Appendix A).

Lemma 6.1.1. *For any $\varepsilon > 0$ and any $u \in V(\Omega)$, there exists a $\delta > 0$, such that*

$$\|u - \Pi u\|_V < \varepsilon \quad \textit{for} \quad h < \delta$$

and

$$\|u - Pu\|_V < \varepsilon \quad \textit{for} \quad h < \delta.$$

Proof. Since Π is bounded (Appendix B),

$$\begin{aligned} \|u - \Pi u\|_V &\leq \|u - y\|_V + \|y - \Pi y\|_V + \|\Pi y - \Pi u\|_V \\ &\leq (1 + \|\Pi\|_V)\,\|u - y\|_V + \|y - \Pi y\|_V. \end{aligned}$$

Since $H^2(\Omega) \cap V(\Omega)$ is dense in $V(\Omega)$, and $u \in V(\Omega)$, there exists a $y \in H^2(\Omega) \cap V(\Omega)$ such that $\|u - y\|_V < (1 + \|\Pi\|_V)^{-1}\,\dfrac{\varepsilon}{2}$. It follows that

$$\begin{aligned} \|u - \Pi u\|_V &\leq (1 + \|\Pi\|_V)\,\|u - y\|_V + \|y - \Pi y\|_V \\ &\leq \frac{\varepsilon}{2} + Ch\|y\|_{H^2(\Omega)}. \end{aligned}$$

There exists a δ such that for $h < \delta$, $Ch\|y\|_{H^2(\Omega)} \leq \dfrac{\varepsilon}{2}$, and so

$$\|u - \Pi u\|_V \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

This proves the first estimate.

The second estimate follows from the fact that P is a projection. □

Theorem 6.1.2. *Suppose u is the solution of Problem MW and u_h is the solution of Problem MW^h. If the initial conditions are chosen such that $u_0^h = \Pi u_0$ and $u_1^h = \Pi u_1$, then*

$$\lim_{h \to 0} \|u(t) - u_h(t)\|_W = 0 \ \textit{for}\ t \in [0, T].$$

Proof. Following from Theorem 3.1.5 we have

$$\begin{aligned}
\|u(t) - u_h(t)\|_W \quad &\leq \quad \|u(t) - Pu(t)\|_W + \sqrt{2}\Bigg(\|Pu_0 - u_0^h\|_W + 3T\|u_1 - u_1^h\|_W \\
&\qquad + 3TC_W\|u_0 - u_0^h\|_W + 3\int_0^T \|u' - Pu'\|_W \\
&\qquad + 3C_W \int_0^T \|u - Pu\|_W \Bigg) \\
&\leq \quad \kappa_1 \Bigg(\|u(t) - Pu(t)\|_V + \sqrt{2}\Bigg(\|Pu_0 - u_0^h\|_V + 3T\|u_1 - u_1^h\|_V \\
&\qquad + 3TC_W\|u_0 - u_0^h\|_V + 3\int_0^T \|u' - Pu'\|_V \\
&\qquad + 3C_W \int_0^T \|u - Pu\|_V \Bigg).
\end{aligned}$$

Note that $u(t), u'(t) \in V$ for all $t \in (0, T)$ (Section 2.5) and so $Pu(t), Pu'(t) \in V$ for all $t \in (0, T)$. We can therefore use Lemma 6.1.1 to estimate all the terms on the right hand side of the above equation. Consider for example:

$$\|Pu_0 - u_0^h\|_V \leq \|Pu_0 - u_0\|_V + \|u_0 - \Pi u_0\|_V \leq 2\varepsilon$$

for $h < \delta$.

Also,

$$\|u'(t) - Pu'(t)\|_V \leq \varepsilon,$$

and so

$$\int_0^T \|u'(t) - Pu'(t)\|_V \, dt \leq T\varepsilon,$$

for $h < \delta$.

The rest of the terms are estimated in the same way, and so $\|u(t) - u_h(t)\|_W \to 0$ as h tends to zero. $\square$

Fully discrete error estimate

A fully discrete error estimate can be obtained by using the result above together with Equation (3.3.1) in Section 3.3. Recall that

$$\begin{aligned}\|u_h(t_k) - u_k^h\|_W &\leq 7T^2\tau^2 \max \|u_h^{(4)}\|_W + 7T\tau^2 \max \|u_h'''\|_W \\ &\quad + \sqrt{2C_W}\,\tau^4 \max \|u_h'''\|_W.\end{aligned}$$

Corollary 6.1.3. *For h and τ sufficiently small, we have*

$$\|u(t_k) - u_k^h\|_W \leq \|u(t_k) - u_h(t_k)\|_W + \|u_h(t_k) - u_k^h\|_W \leq 2\varepsilon.$$

6.2 The Reissner-Mindlin plate

We can also apply the theory discussed in Chapter 2 and Chapter 3 ([VV02] and [BV13]) to the Reissner-Mindlin plate model introduced in Section 1.4.

6.2.1 Weak variational form

To obtain the weak variational form of the problem, we consider the variational equations in Section 1.4.5. We show that this problem is of the same form as Problem G. Add equations (1.4.13) and (1.4.15):

$$\begin{aligned}h &\iint_\Omega \partial_t^2 w v \, dA + I \iint_\Omega \partial_t^2 \boldsymbol{\psi} \cdot \boldsymbol{\phi} \, dA \\ &+ b_B(\boldsymbol{\psi}, \boldsymbol{\phi}) + h \iint_\Omega (\nabla w + \boldsymbol{\psi})(\nabla v + \boldsymbol{\phi}) \, dA = \iint_\Omega qv \, dA. \qquad (6.2.1)\end{aligned}$$

To make the formulation precise, we introduce the following product spaces.

Product spaces

$$\begin{aligned}X &= \mathcal{L}^2(\Omega) \times \mathcal{L}^2(\Omega)^2 \\ H^k &= H^k(\Omega) \times H^k(\Omega)^2 \\ \mathcal{T} &= \mathcal{T}_1(\Omega) \times \mathcal{T}_2(\Omega)\end{aligned}$$

We define the space V as the closure of $\mathcal{T}$ in H^1.

Bilinear forms For u and v in V, define

$$\begin{aligned} c(u,v) &= h(u_1,v_1)_\Omega + I(u_2,v_2)^\Omega_{0,2} \text{ and} \\ b(u,v) &= b_B(u_2,v_2) + h\Big(\nabla u_1 + u_2, \nabla v_1 + v_2\Big)^\Omega_{0,2}. \end{aligned}$$

A natural inner product for X is

$$(x,y)_X = (x_1,y_1)_0 + (x_2,y_2)_{0,2},$$

where $(\cdot,\cdot)_0$ and $(\cdot,\cdot)_{0,2}$ denote the inner products for $\mathcal{L}^2(\Omega)$ and $\mathcal{L}^2(\Omega)^2$ respectively. (See Appendix A, Section A.2.3 on Sobolev spaces of vector valued functions.) Denote the corresponding norm by $\|\cdot\|_X$.

Proposition 6.2.1. *Let u and v be any elements of X. Then there exist positive constants K_i such that*

$$K_1\|u\|_X \leq c(u,u) \leq K_2\|u\|^2_X.$$

Proof. The result follows from the fact that $c(u,u) = h\|u_1\|^2_\Omega + I(\|u_2\|^\Omega_{0,2})^2$.

□

Proposition 6.2.2. *The bilinear form c is an inner product for the space X.*

Proof. The bilinear form c is a symmetric bilinear form and $c(u,u) \geq K_1\|u\|^2_X$ by Proposition 6.2.1. □

Definition 6.2.3. *Inertia space*

The norm $\|\cdot\|_W$ is defined by $\|u\|_W = \sqrt{c(u,u)}$. We refer to the vector space X equipped with the norm $\|\cdot\|_W$ as the space W.

Proposition 6.2.4. *The norms $\|\cdot\|_W$ and $\|\cdot\|_X$ are equivalent.*

Proof. It follows directly from Proposition 6.2.1. □

Let $f(t) = \langle q(\cdot,t), \mathbf{0}\rangle$ and J an open interval containing zero. We can now formulate the weak variational form of the Reissner-Mindlin plate model.

Problem RMW

Find u such that for each $t \in J$, $u(t) \in V$, $u'(t) \in V$, $u''(t) \in W$ and

$$c(u''(t), v) + b(u(t), v) = (f(t), v)_X \text{ for each } v \in V,$$

while $u(0) = u_0 = \langle w_0, \boldsymbol{\psi}_0 \rangle$ and $u'(0) = u_d = \langle w_1, \boldsymbol{\psi}_1 \rangle$.

Before we discuss existence and convergence of the problem, we need the following results.

Proposition 6.2.5. *V is a dense subset of W.*

Proof. $\mathcal{T}_1(\Omega)$ is dense in $\mathcal{L}^2(\Omega)$ and $\mathcal{T}_2(\Omega)$ is dense in $\mathcal{L}^2(\Omega)^2$. Consequently $\mathcal{T}$ is dense in X. Since $\mathcal{T} \subset V \subset X$, the result follow from Proposition 6.2.4. □

Theorem 6.2.6. *There exist a constant K_b such that*

$$|b(u, v)| \leq K_b \|u\|_{H^1} \|v\|_{H^1},$$

for each $u \in V$.

Proof. The terms in $b_B(u_2, v_2)$ are all of the form

$$\iint_\Omega \partial_i u_{2,i} \, \partial_j v_{2,j} \, dA.$$

Applying the Cauchy-Schwartz inequality to each term and adding we obtain

$$|b_B(u_2, v_2)| \leq K \|u_2\|_{1,2} \|v_2\|_{1,2}.$$

Consider the other term in b. By the Cauchy-Schwartz inequality,

$$\begin{aligned} |(\nabla u_1 + u_2, \nabla v_1 + v_2)_{0,2}| &\leq \|\nabla u_1 + u_2\|_{0,2} \|\nabla v_1 + v_2\|_{0,2} \\ &\leq 2\|\nabla u_1 + u_2\|_{0,2}^2 + 2\|\nabla v_1 + v_2\|_{0,2}^2. \end{aligned}$$

Now, $\|\nabla u_1 + u_2\|_{0,2}^2 = |u_1|_1^2 + \|u_2\|_{0,2}^2 \leq \|u_1\|_1^2 + \|u_2\|_{0,2}^2$, and hence

$$\|\nabla u_1 + u_2\|_{0,2}^2 \leq 2\|u_1\|_1^2 + 2\|u_2\|_{0,2}^2 \leq 2\|u\|_{H^1}^2.$$

□

Theorem 6.2.7. *Korn's lemma*
There exists a constant C_b such that

$$b_B(u,u) \quad \geq \quad C_b\|u\|_{1,2}^2 \text{ for each } u \in V.$$

Proof. See [Bra01, p.289] or the references in [Wu06]. □

Corollary 6.2.8. *The bilinear form b is an inner product for V.*

Definition 6.2.9. *Energy space*
The space V equipped with the inner product b is referred to as the energy space. The norm $\|\cdot\|_V$ is defined by $\|u\|_V = \sqrt{b(u,u)}$.

From Theorems 6.2.6 and 6.2.7 we have the next result.

Corollary 6.2.10. *The norms $\|\cdot\|_V$ and $\|\cdot\|_{H^1}$ are equivalent on V.*

We have shown that Assumptions E1, E2 and E3 hold. Since the vibration is undamped, E4 is automatically satisfied. Let $\tilde{q} : t \to q(\cdot,t)$. Recall the definition of the space E_b from Section 2.2.2. For this application, E_b is defined to be

$$E_b = \{\ x \in V \ \Big|\ \text{there exists a } y \in \mathcal{L}^2(\Omega) \text{ such that } (y,v)_{\mathcal{L}^2(\Omega)} = b(x,v) \text{ for all } v \in V\}.$$

Again in a similar manner as in Subsection 2.5.3, a sufficient condition for existence of a solution will be when $u_0 \in H^2 \cap V$.

Theorem 6.2.11. *Suppose $\tilde{q} \in C^1(J,\mathcal{L}^2(\Omega))$, then there exists a unique solution*

$$u \in C^1(J,V) \cap C^2(J,W),$$

for Problem RMW for each $u_0 \in E_b$ and $u_d \in V$.

Proof. Since $\tilde{q} \in C^1(J,\mathcal{L}^2(\Omega))$, $f \in C^1(J,X)$ and the result follows from Theorem 2.2.3. □

Remark In [Wu04] existence and uniqueness for solutions of the Reissner-Mindlin plate model (analogous to that of [LM72]) is proved, as well as some regularity results similar to those in [Eva98]. This paper has not been investigated in detail.

6.2.2 Finite element approximation

The finite element approximation is considered in [Wu05]. From the introduction to the article [Wu05], we conclude that this is the first result on convergence of the finite element method for the vibration of a Reissner-Mindlin plate. Wu formulates it as follows: "The static analysis is the foundation of the dynamic analysis. On the other hand, the a priori estimates for hyperbolic problems are developed in [Dup73] and [Bak76]. The method is extended to elasto-dynamic problems by explicit finite elements" [Wu03].

Wu uses the quadrilateral four-node Bath-Dvorkin (B-D) element, which employs piecewise bilinear shape functions. He observed that the convergence rates were optimal for the deflection and rotation, but deterioration in convergence rate and locking are clearly observed for the velocities.

Problem RMW^h

Find a function $u_h \in C^2(0,T)$ such that u_h' is continuous at 0 and for each $t > 0$, $u_h(t) \in S^h$ and

$$c({u_h}''(t), v) + b(u_h(t), v) = (q(\cdot, t), v_1) \quad \text{for each } v \in S^h \tag{6.2.2}$$

while $u_h(0) = u_0^h = \langle w_0^h, \boldsymbol{\psi}_0^h \rangle$ and $u_h'(0) = u_d^h = \langle w_1^h, \boldsymbol{\psi}_1^h \rangle$.

The existence of a unique solution follows from [BV13, Theorem 3.1] (see Chapter 3) provided that $\tilde{q}$ is continuous w.r.t. the norm of $\mathcal{L}^2(\Omega)$.

Interpolation

In [Wu05] the domain Ω is a rectangle. We define an interpolation operator on $H^k = H^k(\Omega) \times H^k(\Omega)^2$. For piecewise bilinear basis functions on rectangular elements Πu is defined as follows. If $u \in H^k$, then

$$\Pi u = \langle \Pi_b u_1, \Pi_B u_2 \rangle,$$

where Π_b and Π_B are defined in Appendix B.

Proposition 6.2.12. *There exists a $\widehat{C} > 0$ such that*

$$\|\Pi u - u\|_{H^1} \leq \widehat{C} h \|u\|_2 \textit{ for each } u \in H^2.$$

Proof. From Appendix B we have that

$$\|\Pi_b u - u\|_{1,\Omega} \leq \widehat{C}h|u|_{2,\Omega} \text{ for } u \in H^2(\Omega)$$

and

$$\|\Pi_B u - u\|_{1,2} \leq \widehat{C}h|u|_{2,2} \text{ for } u \in H^2(\Omega)^2.$$

From the definition of Π, we have

$$\|\Pi u - u\|_{H^1}^2 = \|\Pi_b u_1 - u_1\|_{1,\Omega}^2 + \|\Pi_B u_2 - u_2\|_{1,2}^2.$$

The result follows. □

Using the result above and equivalence of the norms $\|\cdot\|_V$ and $\|\cdot\|_{H^1}$ (Corollary 6.2.10), we obtain the following estimate: For $u \in H^2$,

$$\|\Pi u - u\|_V \leq \widehat{C}h\|u\|_2. \tag{6.2.3}$$

6.2.3 Inertia norm error estimates

Since no damping is present, it follows that the results for weak damping in Chapter 3 (based on [BV13]) are valid. For this problem we have that $H = H^k \cap V$.

We first apply the result from Theorem 3.1.6.

Suppose that we choose $u_0^h = \Pi u_0$ and $u_1^h = \Pi u_1$. If the solution u of Problem RMW satisfies $u' \in \mathcal{L}^2([0,T], H^2 \cap V)$, then

$$\begin{aligned}\|u(t) - u_h(t)\|_W \quad \leq \quad & \kappa_1 \widehat{C}h\|u(t)\|_2 + \sqrt{2}\kappa_1\widehat{C}h\Bigg(3T\|u_1\|_2 + (2 + 3TC_W)\|u_0\|_2 \\ & + 3\int_0^T \|u'(\cdot)\|_W + 3C_W\int_0^T \|u(\cdot)\|_W\Bigg).\end{aligned}$$

for each $t \in [0,T]$.

Remark Theoretically convergence follows from the discussion above, but in practice locking may be a problem as observed in the numerical experiments in [Wu05]. However, with piecewise bicubic basis functions satisfactory results are obtained (see [LVV09]).

Remark Alternatively, we can let S_i^h denote a subspace of $H^2(\Omega)$ consisting of piecewise Hermite bicubic functions that vanish on the boundary of Ω. Defining

$$S^h = S_1^h \times S_2^h \times S_3^h,$$

then S^h is a finite dimensional subspace of V. These elements were used in [LVV09] with success.

6.2.4 Fully discrete

Since we used the method of [BV13] to split the semi-discrete and fully discrete error estimates, one is now in a fortunate position to decide what algorithm to use for the fully discrete approximation. Suppose one were to choose to use the algorithm in Problem G^h-D (Section 3.2).

Problem RMW^h-D
Find a sequence $\{u_k^h\} \subset S^h$ such that for $k = 0, 1, 2, \ldots, N-1$,

$$\delta_t u_k^h = v_{k+\frac{1}{2}},$$
$$c(\delta_t v_k, \psi) + b(u_{k+\frac{1}{2}}^h, \psi) = \frac{1}{2}([f(t_k) + f(t_{k+1})], \psi)_X$$

for each $\psi \in S^h$, while $u_0^h = u_h(0) = d^h$ and $v_0 = u_h'(0) = v^h$.

Recall that we know from Theorem 3.2.1 that if $\tilde{q} \in C^2([0,T], \mathcal{L}^2(\Omega))$, then

$$\begin{aligned} \|u_h(t_k) - u_k^h\|_W \quad \leq \quad & 7T^2\tau^2 \max \|u_h^{(4)}\|_W + 7T\tau^2 \max \|u_h'''\|_W \\ & + \sqrt{2C_W}\, \tau^4 \max \|u_h'''\|_W \end{aligned}$$

for each $t_k \in (0, T)$.

Now we can use this result together with the result in the previous subsection (Section 6.2.3) to obtain the main result.

Suppose $u_0^h = \Pi u_0$ and $u_d^h = \Pi u_d$. If $u'(t) \in \mathcal{L}^2([0,T], H^2 \cap V)$, $\tilde{q} \in C^2([0,T], \mathcal{L}^2(\Omega))$ and the sequence $\{u_k^h\}$ is a solution of Problem RMW^h-D, then

$$
\begin{aligned}
\|u(t_k) - u_k^h\|_W \quad \leq \quad & \kappa_1 \widehat{C} h \|u(t)\|_2 + \sqrt{2}\kappa_1 \widehat{C} h \Bigg(3T\|u_1\|_2 + (2 + 3TC_W)\|u_0\|_2 \\
& + 3\int_0^T \|u'(\cdot)\|_W + 3C_W \int_0^T \|u(\cdot)\|_W \Bigg) + 7T^2\tau^2 \max \|u_h^{(4)}\|_W \\
& + 7T\tau^2 \max \|u_h'''\|_W + \sqrt{2C_W}\,\tau^4 \max \|u_h'''\|_W.
\end{aligned}
$$

for each $t_k \in (0,T)$, where $\alpha = 1$ for $k = 3$ and $\alpha = 2$ for $k \geq 4$.

Recall that an estimate for $\|u_h^{(4)}\|_W$ in terms of the data can be obtained.

6.2.5 The approach taken by Wu

In Section 2 of [Wu05] the plate vibration problem (Problem RM) is formulated and the weak variational form is stated (but not derived). This is Equation (2.6). Estimates and the existence theorem with references to other articles are given.

The fully discrete problem is formulated in Section 3 and the explicit scheme explained. The fully discrete method is the same as in [Wu03], see Section 6.3 of this dissertation.

In Section 4 the projection is defined and the decomposition of the errors introduced as in Section 2.3 of this dissertation. The notation in [Wu05] is the author's own and differs from [Dup73] and [Bak76]. (The fact that all these papers use different notation makes comparison a time consuming task.)

6.3 Linear elasto-dynamics

In this section we briefly consider the work done in [Wu03]. In the introduction we read. This paper is devoted to the evaluation of the accuracy and convergence of the explicit finite element method for linear structural dynamics.

6.3.1 Equation of motion

Consider an elastic body with density ρ. The displacement of a point $\mathbf{x}$ in the reference configuration at time t is $\mathbf{u}(\mathbf{x}, t)$ and the velocity is $\mathbf{v} = \partial_t \mathbf{u}$.

From the conservation law for momentum, we have the **equation of motion** (see [Fun65, Sec 5.5, 5.7]))

$$\rho \partial_t^2 \mathbf{u} = \text{div } \mathbf{T} + \mathbf{f},$$

where $\mathbf{T}$ is the first Piola stress tensor and $\mathbf{f}$ an external body force (density force). In the case of small local displacements, we may assume that $\mathbf{T}$ is the Cauchy stress tensor (which is symmetric).

In the matrix representation of $\mathbf{T}$ the stress components are denoted by σ_{ij} and div $\mathbf{T}$ is a vector with components

$$[\text{div } \mathbf{T}]_i = \partial_1 \sigma_{i1} + \partial_2 \sigma_{i2} + \partial_3 \sigma_{i3} \quad \text{for } i = 1,\ 2,\ 3.$$

The strain tensor $\mathcal{E}$ is defined by

$$\varepsilon_{ij} = \frac{1}{2} \left(\partial_i u_j + \partial_j u_i \right).$$

It is possible to derive a system of partial differential equations using Hooke's law

$$\mathbf{T} = \frac{E}{1+\nu} \mathcal{E} + \frac{\nu E}{(1+\nu)(1-2\nu)} \text{tr}(\mathcal{E}) I,$$

where E is Young's modulus and ν Poisson's ratio. Note that here the symbol "tr" denotes the trace of the tensor (matrix).

Problem Wu03

Suppose $\Omega \subset E_3$ is the reference configuration for a solid executing small vibrations. The boundary of Ω consists of two parts Σ and Γ. The problem is to find $\mathbf{u}$ such that

the equation of motion is satisfied in Ω;

Hooke's law is satisfied in Ω;

the specified displacement for $\mathbf{u} = \mathbf{U}$ is satisfied on Σ;

the specified traction $\mathbf{Tn} = \mathbf{g}$ is satisfied on Γ;

the initial conditions $\mathbf{u}(\mathbf{x}, 0) = \mathbf{u_0}$ and $\partial_t \mathbf{u}(\mathbf{x}, 0) = \mathbf{u_1}$ is satisfied in Ω.

Remark The problem in [Wu03] is the system of equations labelled (2.1). Initially [Wu03] assumes an arbitrary $\mathbf{U}$, but when giving the variational form assumes that $\mathbf{U} = \mathbf{0}$, "without loss of generality."

6.3.2 Variational form

It follows from the divergence theorem that

$$\iiint_\Omega \text{div } (\mathbf{Tv}) \, dV = \iint_{\partial\Omega} \mathbf{T}\, \mathbf{v} \cdot \mathbf{n} \, dS,$$

where $\mathbf{n}$ is the unit outward normal on $\partial\Omega$. Also, from the properties of the divergence we have

$$\text{div } (\mathbf{Tv}) = \text{div } \mathbf{T} \cdot \mathbf{v} + \text{tr}(\mathbf{TV}),$$

where $\mathbf{V}$ is the Jacobi matrix for $\mathbf{v}$. Since $\mathbf{T}$ is symmetric, it follows that

$$\iiint_\Omega \text{div } \mathbf{T} \cdot \mathbf{v} \, dV = -\iiint_\Omega \text{tr}(\mathbf{TV}) \, dV + \iint_{\partial\Omega} \mathbf{T}\, \mathbf{n} \cdot \mathbf{v} \, dS.$$

It therefore follows that the variational form of Problem Wu03 is to find a displacement $\mathbf{u}$ such that the boundary condition on Σ is satisfied and

$$\int_\Omega \rho \partial_t^2 \mathbf{u} \cdot \mathbf{v} \, dV \quad = \quad \int_\Omega c \, \text{tr} \left(\mathcal{E} \, \mathbf{V} \right) + k \, \text{tr} \left(\mathcal{E} \right) \text{tr} \left(\mathbf{V} \right) \, dV + \int_\Omega \mathbf{f} \cdot \mathbf{v} \, dV + \int_\Gamma \mathbf{g} \cdot \mathbf{v} \, dS,$$

for each $\mathbf{v} \in \mathcal{T}(\bar{\Omega}) = \{\mathbf{v} \in C^1(\bar{\Omega})^3 : \mathbf{v} = 0 \;\text{ on }\; \Sigma\}$ and where

$$c = \frac{E}{1+\nu}, \qquad k = \frac{\nu E}{(1+\nu)(1-2\nu)}.$$

If the bilinear forms is defined by

$$b(\mathbf{u}, \mathbf{v}) = \int_\Omega c \, \text{tr} \left(\mathcal{E} \, \mathbf{V} \right) + k \, \text{tr} \left(\mathcal{E} \right) \text{tr} \left(\mathbf{V} \right) \, dV,$$

and

$$c(\mathbf{u}, \mathbf{v}) = \int_\Omega \rho \partial_t^2 \mathbf{u} \cdot \mathbf{v} \, dV,$$

then the variational form of Problem Wu03 is similar to that for the wave equation.

6.3.3 Weak variational form and existence

The two-dimensional vibration problem is similar to the Reissner-Mindlin plate model in Section 6.2, Problem RMW. The three-dimensional case differs slightly. Note that the undamped case is considered. The abstract formulation is the same and the theory depends on Korn's inequality (see [Bra01, p.289] and Theorem 6.2.7), which gives the positive definiteness of the bilinear form b. The spaces under consideration are:

- $X = \mathcal{L}^2(\Omega)^3$;
- W is the space $\mathcal{L}^2(\Omega)^3$ with norm $\|\cdot\|_W = \sqrt{c(\cdot,\cdot)}$;
- V is the closure of the space of test functions in $H^1(\Omega)^3$.

Problem Wu03W

Find u such that for each $t > 0$, $u(t) \in V$, $u'(t) \in V$, $u''(t) \in W$ and

$$c(u''(t), v) + c(u(t), v) = (f(t), v)_\Omega + (g(t), \gamma v)_\Gamma \text{ for each } v \in V,$$

while $u(0) = u_0$ and $u'(0) = u_d$.

If $g = 0$, we find that the problem is a special case of the general linear vibration problem introduced and discussed in Chapter 2 (with no damping). We can therefore use the results from the article [VV02], discussed in Chapter 2, in the same way as for Problem RMW. Note that the existence result cited by Wu are also not applicable if $g \neq 0$.

Let J be an interval containing zero, and suppose $f \in C^1(J, \mathcal{L}^2(\Omega))$. Then there exists a unique solution

$$u \in C^1\big(J, V\big) \cap C^2\big(J, W\big),$$

for Problem Wu03W for each $u_0 \in E_b$ and $u_d \in V$.

Remark In [Wu03], Lions and Magenes [LM72] is cited for existence. Also, in [Wu03] it is mentioned that "…the regularity of a solution is discussed in [Eva98, Chapter 7], for the system of one function with homogeneous Dirichlet type displacement boundary conditions, similar to Eq (2.1′). Here, we suppose the argument can be extended to the

2-D and 3-D elasto-dynamics system, without proof." However, it should be noted that Problem Wu03W is similar to Problem RMW and the proofs are given in [Wu04], see the remark on page 141 at the end of Subsection 6.2.1.

6.3.4 Semi-discrete approximation

The finite element approximation for this problem is considered in [Wu03] and [Wu06]. However, no error estimate for the semi-discrete case is derived. We have a special case of Problem G^h in Chapter 3 and apart from the obvious differences between two-dimensional and three-dimensional elements, we can apply [BV13] (Theorem 3.1.6 in Chapter 3 in this dissertation). The results are not stated here since it is analogous to that of the Reissner-Mindlin plate model, Sections 6.2.2 and 6.2.3. Recall that S^h is a finite dimensional subspace of V.

Problem Wu03G$^{\mathbf{h}}$

Find $u_h \in C^2(0,T)$ such that u_h' is continuous at 0 and for each $t > 0$, $u_h(t) \in S^h$ and

$$c(u_h''(t), v) + c(u_h(t), v) = (f(t), v)_\Omega + (g(t), v)_\Gamma \text{ for each } v \in S^h,$$

while $u_h(0) = u_0^h$ and $u_h'(0) = u_d^h$.

6.3.5 Fully discrete approximation

The fully discrete scheme in [BV13] (Chapter 3 in this dissertation), which is implicit, can theoretically be applied to this three-dimensional elasto-dynamics problem.

In the introduction of [Wu03] we read: "Among many numerical schemes implemented in commercial software, the explicit finite element method has been successfully employed to solve transient, large deformation, dynamics problems which are subject to impact loading." Also, in the introduction to [Wu06] the importance of the explicit method is stressed. "The explicit finite element method has been extensively developed for the transient dynamic analysis to meet the increasing demand of engineering application."

We also read in the abstract of [Wu06] that " with smooth solutions, it is shown that by using diagonal mass matrix or consistent mass matrix, the displacement, velocity, and the energies have the same convergence rates." In the introduction of [Wu06], it is also remarked that " the diagonal mass matrix is one of the important features making the explicit method sufficient and practical."

From this it is clear that according to Wu, the explicit method is preferable to an implicit method, due to high computational costs. Any finite element method tends to be implicit due to the presence of the mass matrix M. Therefore, to make a finite element method explicit, the mass matrix M is changed to a diagonal matrix by so called mass lumping. We therefore briefly discuss the explicit scheme used in [Wu03].

The semi-discrete problem can be discretised into a system of ordinary differential equations as in Section 3.2. Recall the following notation. For $\bar{x} \in \mathbb{R}^n$ let

$$T_h\bar{x} = \Sigma_{i=1}^{n} x_i\,\phi_i \in S^h,$$

where S^h is the span of the set of basis functions $\{\phi_1, \phi_2, \dots, \phi_n\}$. If a function w has values in S^h, then we define a function $\bar{w}$ by

$$\bar{w}(t) = T_h^{-1}w(t),$$

with values in $\mathbb{R}^n$.

Problem Wu03G^h can now be written as a system of ordinary differential equations:

$$M\bar{u}'' + K\bar{u} = F(t) \quad \text{with} \quad \bar{u}(0) = \bar{u}_0^h \quad \text{and} \quad \bar{u}'(0) = \bar{u}_d^h,$$

with the matrices M and K given in Section 3.2, and $\bar{u}_0^h = T_h^{-1}u_0^h$ and $\bar{u}_d^h = T_h^{-1}u_d^h$.

Recall that the time interval $[0, T]$ is divided into N steps with a step length $\tau = \frac{T}{N}$ and that we denote the approximation of $u_h(t_k)$ by u_k^h. [Wu03] uses the central difference method for time discretisation. In variational form:

$$\begin{aligned}
(\rho\delta_t^2 u_k^h, v) + b(u_k^h, v) &= (f(t_k), v)\,, \\
(u_0^h, v) &= (P_2 u_0, v)\,, \\
(u_1^h, v) &= (P_2 u_d, v)\,,
\end{aligned}$$

where

$$\begin{aligned}
\delta_t u_k^h &= \frac{u_{k+1}^h - u_k^h}{\tau} \quad \text{and} \\
\delta_t^2 u_k^h &= \frac{\delta_t u_k^h - \delta_t u_{k-1}^h}{\tau} = \frac{u_{k+1}^h - 2u_k^h + u_{k-1}^h}{\tau^2}.
\end{aligned}$$

Implementing this scheme for the system of ordinary differential equations we obtain

$$\tau^{-2}M\left(\bar{u}_{k+1}-2\bar{u}_k+\bar{u}_{k-1}\right)+K\bar{u}_k=F(t_k),$$

with

$$\bar{u}(0)=\bar{u}_0^h \quad \text{and} \quad \bar{u}'(0)=\bar{u}_d^h.$$

Using mass lumping the matrix M is replaced by the diagonal matrix D to become

$$\tau^{-2}D\left(\bar{u}_{k+1}-2\bar{u}_k+\bar{u}_{k-1}\right)+K\bar{u}_k=F(t_k).$$

The system is now essentially explicit since it is a trivial matter to compute D^{-1}.

The time integration procedure followed by Wu is explained in [Wu03]. The following explanation follows the article [Wu03] precisely.

Let the velocity be denoted by $\bar{v}_k=\delta_t\bar{u}_k$.

(1) Move one step

$$\bar{u}_k=\bar{u}_{k-1}+\tau\bar{v}_{k-\frac{1}{2}},$$

(2) calculate forces

$$\mathbf{F}_k=K\bar{u}_k+F(t_k),$$

(3) calculate acceleration

$$\bar{a}_k=D^{-1}\mathbf{F}_k,$$

(4) update the velocity

$$\bar{v}_{k+\frac{1}{2}}=\delta_t\bar{u}_{k-\frac{1}{2}}+\tau\bar{a}_k,$$

(5) go back to step (1).

Wu ([Wu03]) does not explain how the first step works. For the above to work, the first step is obtained by setting

$$\bar{u}_d^h=\bar{v}_0=\frac{\bar{v}_{\frac{1}{2}}+\bar{v}_{-\frac{1}{2}}}{2}$$

and using

$$\bar{a}_0=\frac{\bar{v}_{\frac{1}{2}}-\bar{v}_{-\frac{1}{2}}}{\tau}$$

to get

$$\bar{u}_1=\bar{u}_0+\tau\bar{v}_{\frac{1}{2}}=\bar{u}_0^h+\tau\left(\frac{\tau}{2}\bar{a}_0+\bar{v}_0\right).$$

6.3.6 Numerical experiments

In the article [Wu03], the author provides numerical experiments to compare to the theoretical results. The first two examples concern the longitudinal vibration of a rod. In both cases explicit solutions (analytical solutions) are available.

In Example 1 [Wu03] the solution has a fourth order time derivative, hence the solution satisfies the assumption of the theory. Wu then states "it is observed that the numerical results match the theoretical a priori estimates, ... Both displacement and velocity achieve the optimal rates."

In Example 2 Wu examined a second case with possibly lower regularity. However, it is a solution with a third order time derivative so the predictions of the theory differs slightly from Example 1. Wu noted: "it is observed that the convergence rates of displacement in both $\mathcal{L}^2$ and H^1 norm match the theoretical results, which are still the optimal. However, the convergence rates of velocity and energies are better than the theoretical results."

In Example 3 Wu considers a two-dimensional problem: in-plane vibration of a rectangular plate. He assumes that the theory in [Eva98] (Section 2.4 of this dissertation) is applicable and thus the solution has a fourth order time derivative. The author found "... that the convergence rates in both the $\mathcal{L}^2$-norm and H^1 semi-norm match the theoretical results ... which are still optimal. However, the convergence rates of velocity and energies are better than the theoretical results."

In Example 3 the initial displacement is zero. If the theory in [Eva98] is applicable the solution is smoother than a classical solution. If the conditions of Theorem 2.2.3 are met the solution will be less smooth than in Example 3, but almost a classical solution. In our view that would have been a more useful experiment.

We conclude that the theory does not explain all the observed numerical results.

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Chapter 7

Conclusion

7.1 Overview

As mentioned in the introduction, the general aim of the research is to investigate the disparity that was noticed in the theory between the existence of solutions and the regularity assumed on these solutions for convergence of the finite element method. In almost all of the articles that were considered, an existence result (in [LM72]) for the existence of a weak solution is quoted, but in proving convergence of the Galerkin approximation, substantially more differentiability properties for the solution are assumed. These articles are [Bak76], [GSS06], [Kar11a], [Kar11b], [Kar12] and [Dup73]. In [GS09] and [Wu06] existence is not mentioned. The assumptions used for convergence theory are very restrictive; the solution is required to be smoother than even a classical solution. The detail of this is given in Section 2.4.

We first consider the theory of existence of a solution to a general linear vibration problem, called Problem G, in Chapter 2. We investigated the article of Van Rensburg and Van der Merwe [VV02] published in 2002. The theory in [VV02] is convenient to use in this dissertation since it is given in terms of bilinear forms - a requirement for the finite element method. To compare we examined alternative theories on existence of solutions to hyperbolic partial differential equations, such as those in [Eva98] and [LM72]. We also presented the improved regularity and higher regularity results of a solution to the multi-dimensional wave equation without damping (in [Eva98]). To obtain this, compatibility conditions are required on the initial and boundary data. The limitations of these results are that no damping is considered and only homogeneous Dirichlet boundary conditions are considered.

In conducting the literature study, the article of [BV13] published in 2013 came under consideration (see Chapter 3). In the article convergence is proved, but with weaker assumptions than the other articles considered. However, it is still necessary to assume higher regularity of the solution. In the article semi-discrete and fully discrete error estimates for the Galerkin approximation of a general linear second order hyperbolic problem are derived. Viscous type damping is also incorporated and as such the results in [BV13] could be applied to problems like the multi-dimensional wave equation with weak damping, the hyperbolic heat conduction equation and the Reissner-Mindlin plate model. The results and proofs in the article [BV13] are mostly given in sufficient detail in the article, hence the focus in this dissertation was to compare the article to other research results, and to highlight significant parts of proofs.

A recent article on the continuous Galerkin method, [Kar11a], is analysed in detail in Chapter 4. The problem that Karaa [Kar11a] considers is the general Dual-Phase-Lag model introduced in Subsection 1.3.4. A fully discrete error estimate in the $\mathcal{L}^2$-norm is derived. The approximation method in time in [Kar11a] is the general Newmark method. Special cases of this scheme include the central difference scheme and the average acceleration method. The Dual-Phase-Lag model that is considered includes strong damping (see Section 2.6), where other articles usually only include weak damping or no damping at all. The proofs in [Kar11a] posed a challenge to follow and it was necessary to provide more steps and reasons for greater readability. Possible oversights or omissions in the proofs were discovered and either rectified or reported on.

It was decided to include the discontinuous Galerkin finite element method in the investigation, to see if it has any impact on the assumptions on the regularity of solutions required for convergence. To be specific, we wanted to determine whether the assumptions made on the regularity of solutions for convergence of the solution are less restrictive than those made for the continuous Galerkin method. In Chapter 5 the work done on the discontinuous Galerkin (DG) method by Grote, Schneebeli and Schötzau [GSS06] is investigated. This article deals with the symmetric interior penalty DG method (SIPDG) for the spatial discretization of the second-order scalar wave equation. In the article error estimates for the semi-discrete DG method are derived, in both the energy norm and the weaker $\mathcal{L}^2$-norm. These two estimates are proven in greater detail in this dissertation, and some results that aid in the proofs of these estimates are also provided in more detail. The $\mathcal{L}^2$-norm error estimate is based on the proof of the error estimate in [Bak76] for the continuous Galerkin finite element method, but in this dissertation we showed that the proof could be streamlined following [BV13] (see Section 3.1.1). We also report on the article [GS09] where a fully discrete error estimate in the $\mathcal{L}^2$-norm is derived for the

SIPDG method. The authors use the semi-discrete formulation from [GSS06] and the central difference scheme in time.

We apply the general theory from Chapters 2 and 3 to some of the model problems, in Chapter 6. To be precise, a section each is devoted to the multidimensional wave equation with weak damping, the Reissner-Mindlin plate model and linear elasto-dynamics.

7.2 Results

To start, consider Theorem 3.1.5 in Subsection 3.1.2. We have the following error estimate for the semi-discrete approximation: for $t \in [0, T]$,

$$\begin{aligned} \|u(t) - u_h(t)\|_W \;\; &\leq \;\; \|e_p(t)\|_W + \sqrt{2}\Bigg(\|Pu_0 - u_0\|_W + 3T\|u_1 - u_1^h\|_W \\ &\quad + (1 + 3TC_W)\|u_0 - u_0^h\|_W + 3\int_0^T \|e_p'\|_W + 3C_W \int_0^T \|e_p\|_W \Bigg). \end{aligned}$$

For this result no additional assumptions are necessary; the properties of the solution guaranteed by Theorem 2.2.3 are sufficient. To proceed, estimates are necessary for the projection errors

$$\|u(t) - Pu(t)\|_W \quad \text{and} \quad \|u'(t) - Pu'(t)\|_W, \tag{7.2.1}$$

and for the errors in approximating the initial states

$$\|u_1 - u_1^h\|_W \quad \text{and} \quad \|u_0 - u_0^h\|_W. \tag{7.2.2}$$

As mentioned, most articles assume that the initial conditions u_0^h and u_1^h are the projections with respect to the weaker space W of the initial conditions u_0 and u_1 respectively. In none of the articles is it mentioned how to obtain these projections.

Now consider the multi-dimensional wave equation with weak damping. To obtain optimal estimates for the projection errors (7.2.1), the **regularity assumption** $u'(t) \in H^2(\Omega)$ for $t \in [0, T]$ is necessary (see Subsection 6.1.1).

Also, if it is assumed that the initial conditions u_0^h and u_1^h are the $\mathcal{L}^2$-projections of the initial conditions u_0 and u_1 respectively, it should be **assumed** that $u_0, u_1 \in H^2(\Omega)$, in order to obtain optimal order estimates for (7.2.2).

There is clearly a lack of convergence results for classical solutions or weak solutions. The situation for other model problems, e.g. the vibration of a plate is the same. The assumptions made on the regularity of the solution u by making use of the discontinuous Galerkin finite element method to obtain a semi-discrete error estimate (Theorem 5.3.2 and Theorem 5.3.3) are only fractionally better than those for the continuous Galerkin finite element method.

In the article of Basson and Van Rensburg [BV13], a fully discrete error estimate is obtained by first deriving estimates for

$$\|u(t_k) - u_h(t_k)\|_W \quad \text{and} \quad \|u_h(t_k) - u_k^h\|_W,$$

and then combining these two estimates with the use of the triangle inequality to obtain a fully discrete error estimate. Consider Theorem 3.2.1 in Subsection 3.2.3. We have the following estimate for $\|u_h(t_k) - u_k^h\|_W$:

$$\begin{aligned}\|u_h(t_k) - u_k^h\|_W \;\leq\; & 7T^2\tau^2 \max \|u_h^{(4)}\|_W + 7T\tau^2 \max \|u_h'''\|_W \\ & + \sqrt{2C_W}\,\tau^4 \max \|u_h'''\|_W.\end{aligned}$$

For this estimate to hold, it is assumed that $u_h \in C^4[0,T]$ and this is the case if $f \in C^2([0,T],X)$, which is realistic. However, note that nothing is mentioned in [BV13] about estimates for $\|u_h^{(4)}\|_W$ and $\|u_h'''\|_W$. It is not necessary to assume the existence of a third or fourth order derivative on the exact solution u as is done in all the other articles considered. Strict compatibility conditions must be imposed on the initial data and regularity conditions on the forcing function to yield these higher regularity properties (see Section 2.4).

Again the result is that convergence results for classical solutions and weak solutions are not available.

In Section 6.1 we proved that the results obtained from existence theory in Section 2.2.1 are sufficient for convergence of the semi-discrete and fully discrete approximations, but no result on the order of convergence could be obtained.

In Section 5.7 and Subsection 6.3.6 numerical experiments in [GSS06] and [Wu03] are discussed briefly. Interesting phenomena are observed and to a large degree the theory and experiments agree. However, there are indications that the order of convergence may in some cases be better than predicted by the theory.

7.3 Further research

The investigation has confirmed that there is a problem when applying theoretical results to real world problems. Further research is required to either improve on the regularity theory for hyperbolic type partial differential equations or improve the convergence theory for finite element approximations.

Numerical experiments suggest that it should be possible to obtain error estimates with less restrictive regularity assumptions.

One possibility is to do further research to extend the work done in [BV13]:

- obtain estimates for $u_h^{(4)}$
- investigate if weaker regularity assumptions can be made for the case when strong damping is present;
- apply the general method on structures consisting of linked elastic bodies.

Further research is possible on the time-stepping schemes in [Kar11a], [Kar11b] and [Kar12]. For example, one could investigate the application of methods to the multi-dimensional wave equation with general damping.

The discontinuous Galerkin method looks promising. Further investigation of the proofs in [GSS06] may lead to results where less restrictive spatial regularity is required.

Fundamentally the approach to prove optimal order convergence in all articles considered is related to the approach of Baker [Bak76]. A challenge would be to find an alternative method to obtain an estimate for the difference between the elliptic projection error and the Galerkin approximation. Hopefully an alternative method may yield convergence results with less restrictive regularity assumptions. In this context mixed finite elements may provide an alternative.

Another consideration would be to study how the compatibility conditions placed on the initial and boundary data in the books of Evans [Eva98] and Wloka [Wlo87] can be interpreted when applying it to systems of linked elastic bodies.

Model problems with "almost" classical solutions are important in practice. For example the solution u of the multi-dimensional wave equation has the properties $u(t) \in H^2(\Omega)$, $u'(t) \in H^1(\Omega)$ and $u''(t) \in \mathcal{L}^2(\Omega)$. The conjecture is that $\|u(t) - u_h(t)\|_{\mathcal{L}^2(\Omega)} \leq Kh^\alpha$

where $\alpha > 0$. Further research should start with either doing numerical experiments or search the literature for numerical experiments that can be used to "test" this conjecture.

Appendix A

Sobolev Spaces

A.1 The space $\mathcal{L}^2(\Omega)$

Consider an open subset Ω of $\mathbb{R}^n$ and denote its closure by $\bar{\Omega}$. The space $\mathcal{L}^2(\Omega)$ consists of functions f such that f^2 is Lebesgue integrable on Ω.

Theorem A.1.1. *The space $\mathcal{L}^2(\Omega)$ is a Hilbert space with inner product*

$$(u, v) = \int_\Omega uv = \int_\Omega uv\, d\mu,$$

where μ is the n-dimensional Lebesgue measure.

Proof. See [Rud76, Theorem 3.11, p. 69]. □

Notation Unless otherwise stated, the norm on $\mathcal{L}^2(\Omega)$ is denoted by $\|\cdot\|$ or $\|\cdot\|_0$.

Definition A.1.2. Set $A_f := \{x \in \Omega : f(x) \neq 0\}$ and S_f to be the closure of A_f in Ω, where Ω is open. Then define $C_0^\infty(\Omega) := \{f \in C^\infty(\Omega) : S_f \subset \Omega\}$.

If $f \in C_0^\infty(\Omega)$, then the distance between S_f and the boundary of Ω is positive.

Theorem A.1.3. *The space $C_0^\infty(\Omega)$ is dense in $\mathcal{L}^2(\Omega)$.*

Proof. See [Ada75, Theorem 2.13, p.28]. □

A.2 Sobolev Spaces

A.2.1 The one dimensional case

Let the open interval (a, b) be denoted by I and the closed interval $[a, b]$ by $\bar{I}$. The Sobolev space $H^m(I)$ is the subspace of functions in $\mathcal{L}^2(I)$ with weak derivatives up to order m.

Definition A.2.1. Weak Derivative of order m
Suppose $u \in \mathcal{L}^2(I)$ and there exists a $v \in \mathcal{L}^2(I)$ such that

$$(u, \phi^{(m)}) = (-1)^m (v, \phi) \quad \forall \quad \phi \in C_0^\infty(I),$$

then v is called the weak derivative of order m of u and is denoted by $D^m u$.

In this dissertation no distinction will be made between weak and ordinary derivatives as far as notation is concerned, i.e. Du is denoted by u'.

Definition A.2.2. Inner product on $H^m(I)$
For u and v in $H^m(I)$ we define

$$(u, v)_m = \sum_{k=0}^{m} (u^{(k)}, v^{(k)}) \text{ for } m = 0, 1, 2, \ldots.$$

Definition A.2.3. Semi-norm on $H^m(I)$
For any $m \geq 1$ and any function $u \in H^m(I)$, we define

$$|u|_m = \sqrt{(u^{(m)}, v^{(m)})} = \|u^{(m)}\|.$$

Definition A.2.4. Norm on $H^m(I)$
For any function $u \in H^m(I)$ we define

$$\|u\|_m = \sqrt{(u, u)_m} \text{ for } m = 0, 1, 2, \ldots.$$

A.2.2 The higher dimensional case

Let Ω be an open subset of $\mathbb{R}^n$. The Sobolev spaces $H^m(\Omega)$ is the subspace of functions in $\mathcal{L}^2(\Omega)$ with weak partial derivatives up to order m in $\mathcal{L}^2(\Omega)$.

Notation Let $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)$, then $D^\alpha = D^{\alpha_1} D^{\alpha_2} \cdots D^{\alpha_n}$ and $|\alpha| = \alpha_1 + \alpha_2 + \ldots + \alpha_n$.

Definition A.2.5. Weak partial derivative of order m
Suppose $u \in \mathcal{L}^2(\Omega)$ and there exists a $v \in \mathcal{L}^2(\Omega)$ such that

$$(u, D^\alpha \phi) = (-1)^{|\alpha|}(v, \phi) \quad \forall \;\; \phi \in C_0^\infty(\Omega),$$

then v is called the weak derivative of order $|\alpha|$ of u and is denoted by $D^m u$.

Remark In this dissertation no distinction will be made between weak and ordinary derivatives as far as notation is concerned.

Definition A.2.6. Inner product on $H^m(\Omega)$
For u and v in $H^m(\Omega)$ we define

$$(u, v)_m = \sum_{|\alpha| \leq m} (D^\alpha u, D^\alpha v) \text{ for } m = 0, 1, 2, \ldots.$$

The bilinear form $(u, v)_m$ has all the properties of an inner product.

Definition A.2.7. Semi-norm on $H^m(\Omega)$
For any $m \geq 1$ and any function $u \in H^m(\Omega)$ we define

$$|u|_m = \sqrt{\sum_{|\alpha| = m} (D^\alpha u, D^\alpha u)}.$$

Definition A.2.8. Norm on $H^m(\Omega)$
For any function $u \in H^m(\Omega)$ we define

$$\|u\|_m = \sqrt{(u, u)_m} \text{ for } m = 0, 1, 2, \ldots.$$

Definition A.2.9. The space $H_0^1(\Omega)$
The space $H_0^1(\Omega)$ is defined to be the closure of the space $C_0^\infty(\Omega)$.

Notation If we need to distinguish between different domains, we will denote the norm and semi-norm by $\| \cdot \|_{m,\Omega}$ and $| \cdot |_{m,\Omega}$.

A.2.3 Sobolev spaces of vector valued functions

Definition A.2.10. For f, g in $H^m(\Omega)$,

$$[f,g]_m = (f^{(m)}, g^{(m)}) \quad \text{and}$$

$$|f|_m = \sqrt{[f,f]_m} \quad \text{for} \quad m = 0,\, 1,\, \ldots$$

The function $|\cdot|_m$ is a semi-norm for $m \geq 1$.

Definition A.2.11. $u \in \mathcal{L}^2(\Omega)^2$ if $u_i \in \mathcal{L}^2(\Omega)$ for $i = 1,\, 2.$

$u \in \mathcal{L}^2(\Gamma)^2$ if $u_i \in \mathcal{L}^2(\Gamma)$ for $i = 1,\, 2.$

$u \in H^k(\Omega)^2$ if $u_i \in H^k(\Omega)$ for $i = 1,\, 2.$

$[u,v]_{m,2} = [u_1, v_1]_m + [u_2, v_2]_m$ for $u \in H^m(\Omega)^2$ and $v \in H^m(\Omega)^2$.

$|u|_{m,2} = \sqrt{[u,u]_{m,2}}$ for $u \in \mathcal{L}^2(\Omega)^2$.

The function $|\cdot|_{m,2}$ is a semi-norm for $m \geq 1$.

When we need to distinguish between domains, we will use superscripts Ω and Γ in the cases of a double subscript, e.g. $\|\cdot\|^{\Omega}_{m,2}$ and $\|\cdot\|^{\Gamma}_{m,2}$.

Definition A.2.12. The inner product for $H^m(\Omega)^2$ is defined by

$$(f,g)_{m,2} = \sum_{k=0}^{m} [f,g]_{k,2} \quad \text{for} \quad m = 0,\, 1,\, \ldots$$

Definition A.2.13. The norm for $H^m(\Omega)^2$ is defined by

$$\|f\|_{m,2} = \sqrt{(f,g)_{m,2}} \quad \text{for} \quad m = 0,\, 1,\, \ldots$$

Notation $H^0(\Omega) = \mathcal{L}^2(\Omega)$ and $H^0(\Omega)^2 = \mathcal{L}^2(\Omega)^2$.

A.3 Fundamental properties of Sobolev Spaces

Suppose Ω is a bounded open interval or a bounded open convex subset of $\mathbb{R}^n$.

It is not necessary to require that Ω be convex, but it is sufficient for our purpose. In the theory it is usually assumed that Ω is star shaped or has the cone property.

Remark $\mathcal{L}^2(\Omega) = H^0(\Omega)$.

Theorem A.3.1. *The space $H^m(\Omega)$ is complete.*

Proof. See [Ada75, Theorem 3.2, p.45] □

Theorem A.3.2. *$C^m(\bar{\Omega})$ is dense in $H^m(\Omega)$ with respect to the norm of $H^m(\Omega)$.*

Proof. See [OR76, Theorem 2.10, p. 53]. □

Remark A function $v \in H^m(\Omega)$ can be approximated by a function in $C^m(\bar{\Omega})$: if $v \in H^m(\Omega)$, then for any $\varepsilon > 0$ there exists a $\varphi \in C^m(\bar{\Omega})$ such that $\|v - \varphi\|_m < \varepsilon$.

Theorem A.3.3. *Sobolev's Lemma*
Let m be any non-negative integer. If $u \in H^p(\Omega)$ where $p > m + \frac{n}{2}$, then $u \in C^m(\bar{\Omega})$ and

$$\|D^\alpha u\|_{\sup} \leq \|u\|_p \text{ for } |\alpha| \leq m.$$

Proof. See [OR76, Theorem 3.10, p. 80] □

A.4 Trace

Lemma A.4.1. *Let $u \in H^s(\Omega)$, $s > \dfrac{1}{2}$. Then there exists a trace γ_0 of the function u on $\partial\Omega$ and $C > 0$ such that*

$$\|\gamma_0 u\|_{H^{s-\frac{1}{2}}(\partial\Omega)} \leq C\|u\|_{H^s(\Omega)}.$$

Further, if $s > \dfrac{3}{2}$, there exists a trace $\dfrac{\partial u}{\partial n}$ on $\partial\Omega$ and

$$\left\|\frac{\partial u}{\partial n}\right\|_{H^{s-\frac{3}{2}}(\partial\Omega)} \leq C\|u\|_{H^s(\Omega)}.$$

This lemma is a special case of the Trace Theorem [OR76, Theorem 4.18, p.143].

A.5 Broken Sobolev Spaces

This section is included for convenience and is from [Riv08].

Broken Sobolev spaces are natural spaces to work with the discontinuous Galerkin methods. Let Ω be a subspace of $\mathbb{R}^n$, and subdivide it into elements E. Here an element E can be a triangle or quadrilateral when $n = 2$ and a tetrahedron or hexahedron when $n = 3$. It is assumed that the intersection of any two elements is either empty, a vertex, an edge, or a face. Such a subdivision is called a conforming mesh, and is denoted by $\mathcal{M}_h$, with $h = \max_{E \in \mathcal{M}_h} h_E$, where h_E is the diameter of element E.

Assumption BSs1

The subdivision $\mathcal{M}_h$ is regular, i.e. there is a constant $\rho > 0$ such that for every $E \in \mathcal{M}_h$,

$$\frac{h_E}{\rho_E} \leq \rho,$$

where ρ_E the maximum diameter of a ball inscribed in E.

Definition A.5.1. Broken Sobolev space of order m

$$H^m(\mathcal{M}_h) := \{v \in \mathcal{L}^2(\Omega) : \text{ for every } E \in \mathcal{M}_h, v|_E \in H^m(E)\}.$$

Definition A.5.2. Broken norm and seminorm
The broken Sobolev space $H^m(\mathcal{M}_h)$ is equipped with the norm

$$|||v|||_{H^m(\mathcal{M}_h)} = \left(\sum_{E \in \mathcal{M}_h} \|v\|^2_{m,E} \right)^{\frac{1}{2}}$$

and the broken seminorm

$$|||\nabla v|||_{H^0(\mathcal{M}_h)} = \left(\sum_{E \in \mathcal{M}_h} \|\nabla v\|^2_{0,E} \right)^{\frac{1}{2}}.$$

We now have

$$H^m(\Omega) \subset H^m(\mathcal{M}_h) \text{ and } H^{m+1}(\mathcal{M}_h) \subset H^m(\mathcal{M}_h).$$

A.6 Elliptic regularity

Theorem A.6.1. *[Eva98, Theorem 4, p. 317] (Boundary H^2-regularity)*
Assume that

(a). $a^{ij} \in C^1(\Omega)$, $b^i, c \in C(\Omega)(i,j = 1,\ldots,n)$ *and*

(b). $f \in \mathcal{L}^2(\Omega)$.

Suppose that $u \in H_0^1(\Omega)$ *is a weak solution of the elliptic boundary value problem*

$$\begin{cases} Lu = f & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega. \end{cases} \tag{A.6.1}$$

Assume finally $\partial\Omega \in C^2$. *Then* $u \in H^2(\Omega)$ *and we have the estimate*

$$\|u\|_2 \leq C\left(\|f\|_0 + \|u\|_0\right), \tag{A.6.2}$$

with the constant C *depending only on* Ω *and the coefficients of* L.

Appendix B

Interpolation

The results in this appendix are simplified versions of the theory in [OR76, Chapter 6], [OC83, Chapter 4] and [SF73, Chapter 5].

B.1 The one dimensional case

The interpolation error

Theorem B.1.1 below is formulated as a special case of a general result. This result may be found in [SF73, p.144], [OC83, p.76] and [OR76, p.279].

We will use $\widehat{C}$ to denote a generic constant. Also denote the interpolation operator on an element by Π_e and the interpolation operator on the entire domain by Π. (The definitions are given in [SF73], [OR76] and [OC83]).

Theorem B.1.1. *Suppose there exists an integer k such that for each element*

$$s(\Pi_e) + 1 \leq k \leq r(\Pi_e) + 1$$

for the interpolation operator Π. Then there exists a constant $\widehat{C}$ such that for any $u \in H^k(I)$ we have

$$|\Pi u - u|_{m,I} \leq \widehat{C} h^{k-m} |u|_{k,I} \quad \textit{for} \quad m = 0,\ 1,\ \ldots, k.$$

The interpolation operator is denoted by Π_L for piecewise linear basis functions and by Π_c for Hermite cubics.

Corollary B.1.2. *Hermite cubic basis functions.*
There exists a constant $\widehat{C}_c$ *such that if* $u \in H^k(I)$ *for*

a) $2 \leq k \leq 4$, *then*

$$\|u - \Pi_c u\|_m \leq \widehat{C}_c h^{k-m} |u|_k, \quad m = 0, 1, \ldots, k.$$

b) $k > 4$, *then*

$$\|u - \Pi_c u\|_m \leq \widehat{C}_c h^{4-m} |u|_4, \quad m = 0, 1, \ldots, 4.$$

Proof. It is clear that $s(\Pi_c) = 1$ and it can be shown that $r(\Pi_c) = 3$. Consequently Theorem B.1.1 is applicable with $k = 2,\ 3$ or 4. □

Corollary B.1.3. *Piecewise linear basis functions*
There exists a constant $\widehat{C}_L$ *such that if* $u \in H^k(I)$ *for* $k \geq 2$, *then*

$$\|\Pi_L u - u\|_1 \leq \widehat{C}_L h |u|_2$$

Proof. It is clear that $s(\Pi_L) = 1$ and it can be shown that $r(\Pi_L) = 1$. Consequently Theorem B.1.1 is applicable with $k = 2$. □

B.2 The two-dimensional case

Theorem B.2.1 below is formulated as a special case of a general result. As mentioned before, this result may be found in [SF73, p.144], [OC83, p.76] and [OR76, p.279]. In the theorem, $h = \max h_e$, where h_e is the diameter of the element Ω_e.

Theorem B.2.1. *Suppose there exists an integer* k *such that for each element*

$$s(\Pi_e) + 2 \leq k \leq r(\Pi_e) + 1$$

for the interpolation operator Π. *Then there exists a constant* $\widehat{C}$ *such that for any* $u \in H^k(\Omega)$ *we have*

$$|\Pi u - u|_{m,\Omega} \leq \widehat{C} h^{k-m} |u|_{k,\Omega} \quad \textit{for} \quad m = 0,\ 1,\ \ldots, k.$$

Remark The constant $\widehat{C}$ depends on the shape of the elements in the finite element mesh.

Corollary B.2.2. *Piecewise linear basis functions on triangle elements.*
The interpolation operator is denoted by $\Pi_\triangle$. *If* $k \geq 2$, *then there exists a constant* $\widehat{C}_\triangle$ *such that for any* $u \in H^k(\Omega)$ *we have*

$$|\Pi_\triangle u - u|_{m,\Omega} \leq \widehat{C}_\triangle h^{2-m}|u|_{k,\Omega} \quad \textit{for} \quad m = 0,\ 1,\ 2.$$

Corollary B.2.3. *Piecewise bilinear basis functions on rectangle elements.*
The interpolation operator is denoted by Π_b. *If* $k \geq 2$, *then there exists a constant* $\widehat{C}_b$ *such that for any* $u \in H^k(\Omega)$ *we have*

$$|\Pi_b u - u|_{m,\Omega} \leq \widehat{C}_b h^{2-m}|u|_{k,\Omega} \quad \textit{for} \quad m = 0,\ 1,\ 2.$$

Corollary B.2.4. *Hermite cubic basis functions.*
The interpolation operator is denoted by Π_c. *If* $k \geq 2$, *then there exists a constant* $\widehat{C}_c$ *such that for any* $u \in H^k(\Omega)$ *we have*

$$|\Pi_c u - u|_{m,\Omega} \leq \widehat{C}_c h^{2-m}|u|_{k,\Omega} \quad \textit{for} \quad m = 0,\ 1,\ 2.$$

B.3 Vector-valued functions

If an interpolation operator Π is defined on $H^k(\Omega)$ we may define one on $H^k(\Omega)^2$. For $u = \langle u_1,\, u_2 \rangle \in H^k(\Omega)^2$, we define

$$\Pi_2\, u = \langle \Pi\, u_1,\, \Pi\, u_2 \rangle\,.$$

The **seminorm** of order k for $H^k(\Omega)^2$ is denoted by $|\cdot|_{k,2}$ and

$$|u|_{k,2}^2 = |u_1|_k^2 + |u_2|_k^2\,.$$

(See Appendix A.)

Lemma B.3.1. *If* $\|\Pi\ v - v\|_m \leq \hat{C}h^{k-m}|v|_k$ *for* $v \in H^k(\Omega)$, *then*

$$\|u - \Pi_2 u\|_{m,2} \leq \hat{C}h^{k-m}|u|_{k,2} \textit{ for } u \in H^k(\Omega)^2.$$

Proof.

$$|\,u - \Pi_2\, u|_{m,2}^2 = |\,u_1 - \Pi\, u_1|_m^2 + |\,u_2 - \Pi\, u_2|_m^2$$

□

For piecewise bilinear basis functions on rectangles, let $\Pi_B\, u = \langle \Pi_b\, u_1, \Pi_b\, u_2 \rangle$.

Corollary B.3.2. *If $k \geq 2$, then there exists a constant $\widehat{C}$ such that, for all $u \in H^k(\Omega)^2$*

$$|\,u - \Pi_B\, u|\,_{m,2} \leq \widehat{C} h |u|_{2,2} \text{ for } m = 0, 1, 2.$$

Proof. The result follows from Corollary B.2.3 and Lemma B.3.1. □

For piecewise bicubic basis functions on rectangles, let $\Pi_{Bc}\, u = \langle \Pi_{bc}\, u_1, \Pi_{bc}\, u_2 \rangle$.

Corollary B.3.3. *If $k \geq 2$, then there exists a constant $\widehat{C}$ such that, for all $u \in H^k(\Omega)^2$*

$$|\,u - \Pi_{Bc}\, u|\,_{m,2} \leq \widehat{C} h |u|_{2,2} \text{ for } m = 0, 1, 2.$$

B.4 Three-dimensional case

Theorem B.2.1 is also valid for the three-dimensional case.

Corollary B.4.1. *Piecewise linear basis functions on tetrahedron elements.*
The interpolation operator is denoted by Π_t. If $k \geq 2$, then there exists a constant $\widehat{C}_t$ such that for any $u \in H^k(\Omega)$ we have

$$|\Pi_t u - u|_{m,\Omega} \leq \widehat{C}_t h^{2-m} |u|_{2,\Omega} \quad \text{for} \quad m = 0,\ 1,\ 2.$$

B.5 General interpolation assumption

Assumption GI

There exists a subspace $H(V,k)$ of V, and positive constants C_Π and α (depending on V and k) such that for $u \in H(V,k)$:

$$\|u - Pu\|_V \leq C_\Pi h^\alpha \|u\|_{H(V,k)},$$

where $\| \cdot \|_{H(V,k)}$ is a norm or semi-norm associated with $H(V, k)$.

Appendix C

Some inequalities

Lemma C.1. Young's inequality
For any $\varepsilon > 0$ we have

$$ab \leq \frac{a^2}{2\varepsilon} + \frac{\varepsilon b^2}{2}.$$

Proof. We have

$$(\varepsilon a - \varepsilon^{-1} b)^2 \geq 0,$$

which implies that

$$\varepsilon^2 a^2 + \varepsilon^{-2} b^2 \geq 2ab$$

and the result follows. □

Lemma C.2. *We have*

$$(x - y, y) \leq \frac{1}{2}(x, x) - \frac{1}{2}(y, y).$$

Proof. Following from the Cauchy-Schwartz inequality and Young's inequality we have

$$\begin{aligned}
(x - y, y) &= (x, y) - (y, y) \\
&\leq (x, x)^{\frac{1}{2}} (y, y)^{\frac{1}{2}} - (y, y) \\
&\leq \frac{1}{2}(x, x) + \frac{1}{2}(y, y) - (y, y) \\
&= \frac{1}{2}(x, x) - \frac{1}{2}(y, y).
\end{aligned}$$

□

Appendix D

Notation

Nonstandard notation in Chapter 3

$u^{(k)} \in \mathcal{L}^2\Big(J;Y\Big)$ if $u^{(k)}(t) \in Y$ for each t and $\int_J \|u^{(k)}\|_Y^2 < \infty$.

Bibliography

[ABCM02] D N Arnold, F Brezzi, B Cockburn, and L D Marini. Unified analysis of discontinuous galerkin methods for elliptic problems. *SIAM Journal of Numerical Analysis*, 39:1749–1779, 2002.

[Ada75] R A Adams. *Sobolev Spaces*. Series 3. Academic Press, New York, San Francisco, London, 1975.

[AKS96] K T Andrews, K L Kuttler, and M Shillor. Second order evolution equations with dynamic boundary conditions. *J. Math. Anal. Appl.*, 197:781–795, 1996.

[Apo67] T M Apostol. *Calculus*, volume II. Blaisdell Publishing Company, Waltham, Massachusetts, second edition, 1967.

[Bak76] G A Baker. Error estimates for finite element methods for second order hyperbolic equations. *SIAM Journal Numerical Analysis*, pages 564–576, 1976.

[Bra01] D Braess. *Finite elements*. Cambridge university press, Cambridge, ii edition, 2001.

[BSL60] R B Bird, W E Stewart, and E N Lightfoot. *Transport phenomena*. John Wiley and Sons, New York, 1960.

[BV13] M Basson and N F J Van Rensburg. Galerkin finite element approximation of general linear second order hyperbolic equations. *Numerical Functional Analysis and Optimization*, 34(9):976–1000, 2013.

[Cat48] M C Cattaneo. Sulla conduzione de calore (on the conduction of heat). *Atti del Semin. Mat. e Fis. Univ. Modena*, 3(3):83–101, 1948.

[Cia78] P G Ciarlet. *The Finite Element Method for Elliptic Problems*. North-Holland, Amsterdam, 1978.

[Dup73] T Dupont. $\mathcal{L}^2$- estimates for galerkin methods for second order hyperbolic equations. *SIAM Journal on Numerical Analysis*, 10:880–889, 1973.

[Eva98] L C Evans. *Partial differential equations.* American Mathematical Society, 1998.

[Fun65] Y C Fung. *Foundations of solid mechanics.* Prentice-Hall, New Jersey, 1965.

[GS09] M J Grote and D Schötzau. Optimal error estimates for the fully discreet interior penalty discontinuous galerkin method for the wave equation. *Journal of Scientific Computing*, 40:257 – 272, 2009.

[GSS06] 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.

[Inm94] D J Inman. *Engineering vibration.* Prentice-Hall Inc, Englewood Cliffs, New Jersey, 1994.

[Kar11a] S Karaa. Error estimates for finite element approximations of a viscous wave equation. *Numerical Functional Analysis and Optimization*, 32(7):750–767, 2011.

[Kar11b] S Karaa. Finite element *theta*-schemes for the acoustic wave equation. *Advances in Applied Mathematics and Mechanics*, 3:181–203, 2011.

[Kar12] S Karaa. Stability and convergence of fully discrete finite element schemes for the acoustic wave equation. *Journal of Applied Mathematics and Computing*, 2012.

[Kre78] E Kreyzig. *Introductory functional analysis with applications.* John Wiley & Son Inc, 1978.

[Kut86] K L Kuttler. Time-dependent implicit evolution equations. *Nonlinear Analysis:Theory, Methods and Applications*, 10(5):447–463, 1986.

[LM72] J L Lions and E Magenes. *Nonhomogeneous Boundary Value Problems and Applications*, volume 1. Springer-Verlag, 1972.

[LVV09] A Labuschagne, N F J Van Rensburg, and A J Van der Merwe. Vibration of a Reissner-Mindlin-Timoshenko plate-beam system. *Mathematical and Computer Modelling*, 50:1033–1044, 2009.

[Min51] R D Mindlin. Influence of rotatory inertia and shear on flexural motions of isotropic elastic plates. *Journal of Applied Mechanics*, 18(1):31–38, 1951.

[New59] M N Newmark. A method of computation for structural dynamics. *Prc. ASCE J. Engrg. Mech.*, 85:67–94, 1959.

[OC83] J T Oden and F Carey. *Finite elements: Mathematical aspects*. Prentice-Hall International, London, 1983.

[OR76] J T Oden and J N Reddy. *An introduction to the mathematical theory of finite elements*. John Wiley and Sons, 1976.

[Paz83] A Pazy. *Semigroups of linear operators and applications to partial differential equations*. Springer-Verlag, 1983.

[PR05] Y Pinchover and J Rubinstein. *An introduction to partial differential equations*. Cambridge University Press, 2005.

[PS02] I Perugia and D Schötzau. An hp-analysis of the local discontinuous galerkin method for diffusion problems. *Journal of Scientific Computing*, 17:561–571, 2002.

[Rei88] H Reismann. *Elastic Plates: Theory and Application*. John Wiley & Sons, New York, 1988.

[Riv08] B Rivière. *Discontinuous Galerkin methods for solving elliptic and parabolic equations: Theory and Implementation*. SIAM: Frontiers in Applied Mathematics, 2008.

[Rud76] W Rudin. *Principles of Mathematical Analysis*. McGraw-Hill International Editions, third edition edition, 1976.

[Sch71] M Schecter. *Principles of functional analysis*. Academic Press, New York, 1971.

[SF73] G Strang and G J Fix. *An Analysis of the Finite Element Method*. Prentice-Hall, 1973.

[Sho77] R E Showalter. *Hilbert space methods for partial differential equations*. Pitman, London, 1977.

[Tzo95] D Y Tzou. A unified field approach for heat conduction from macro- to micro-scales. *Trans ASME Journal of Heat Transfer*, 117(9):8 – 16, 1995.

[Ver58] P Vernotte. Les paradoxes de la théorie continue de l'équation de la chaleur. *C. R. Acad. Sci. Paris Ser. I Math.*, 246:3154, 1958.

[VV02] N F J Van Rensburg and A J Van der Merwe. Analysis of the solvability of linear vibration models. *Applicable Analysis*, 81(5):1143–1159, 2002.

[Wlo87] J Wloka. *Partial differential equations.* Cambridge University Press, 1987.

[Wu03] S R Wu. A priori error estimates for explicit finite element for linear elastodynamics by Galerkin method and central difference method. *Comput. Methods Appl. Mech. Engrg.*, 192(51-52):5329–5353, 2003.

[Wu04] S R Wu. Reissner-mindlin plate theory for elastodynamics. *Journal of Applied Mathematics*, 3:179 – 189, 2004.

[Wu05] S R Wu. A priori error estimation of a four-node Reissner-Mindlin plate element for elasto-dynamics. *Comput. Methods Appl. Mech. Engrg.*, 194(18-20):2257–2281, 2005.

[Wu06] S R Wu. Lumped mass matrix in explicit finite element method for transient dynamics of elasticity. *Comput. Methods Appl. Mech. Engrg.*, 195(44-47):5983–5994, 2006.

[Zie77] O C Zienkiewicz. *The finite element method.* McGraw-Hill, London, New York, iii edition, 1977.