École Doctorale Carnot-Pasteur

Thèse de Doctorat

présentée par

Olga GORYNINA

pour obtenir le grade de
Docteur en Mathématiques et Applications
de l'Université de Bourgogne Franche-Comté

Eléments finis adaptatifs pour l'équation des ondes instationnaire

Thèse soutenue le 22 février 2018, devant le jury composé de :

Omar LAKKIS	University of Sussex	Rapporteur
Martin VOHRALIK	INRIA	Rapporteur
Emmanuel CREUSÉ	Université Lille 1	Examinateur
François-Xavier ROUX	Université Paris 6	Examinateur
Alexei LOZINSKI	Université Bourgogne - Franche-Comté	Directeur
Marco PICASSO	EPFL	Co-Directeur

UBFC
UNIVERSITÉ
BOURGOGNE FRANCHE-COMTÉ

ECOLE DOCTORALE
CARNOT - PASTEUR

LmB

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my supervisor Alexei Lozinski. I could not have imagined having a better supervisor for my PhD study. Without his bright ideas, guidance and dedicated involvement in every step throughout the process, this work would have never been accomplished. I'm also very grateful to my co-supervisor Marco Picasso for his continuous support and help during my PhD study with motivation and knowledge.

I am grateful to the jury members, Prof. François-Xavier Roux, Dr. Omar Lakkis, Dr. Martin Vohralik, Prof. Emmanuel Creusé, for spending time to correct my thesis and providing their critical reviews on it.

I would also like to thank all of those with whom I have had the pleasure to work during the PhD study and especially to Mustafa Gaja and Samuel Dubuis.

I also thank many friends I met in Laboratoire de Mathématiques de Besançon for their help and accompany during the last three years: Eva, Aude, Charlotte, Lucie, Johann, Guillaume, Colin, Youssef, Thao, Alessandro. Especially I want to thank Clement and Runlian for spending a really good time together; Marine and Isabelle for their support and help.

For many memorable evenings I must thank everyone above as well as Baptiste, Tanya, Yulia, Anya and Masha. Many thanks to Timur and my sister Masha for their especially precious support.

Finally, I must express my very profound gratitude to my whole family for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. Many thanks to my grandmother Geta, grandfather Lesha, grandmother Galya, grandfather Lenya, aunt Tanya, uncle Serega, my little brother and especially to my mother and father. This accomplishment would not have been possible without them.

Contents

Introduction

The error estimates control the difference between the exact solution of some partial differential equation and its approximation. Usually they have the following form

$$\|u - u_h\| \leq C(h, \ldots),$$

where u is the exact solution of the problem, u_h is a computed approximate solution, h is an approximation parameter and $C(h, \ldots)$ is some function of h and other quantities. In *a priori* error estimates, the right-hand side depends on the exact solution u. This is not very useful in practice since generally we don't know the exact solution. Conversely, *a posteriori* error estimates depend on known quantities only, i.e. on the computed solution and thus they can be evaluated in practice.

A posteriori error analysis of finite element approximations for partial differential equations plays an important role in mesh adaptivity techniques. The cases of elliptic and parabolic problems are well studied in the literature [EJ91; LMP14; LPP09]. On the contrary, the *a posteriori* error analysis for hyperbolic equations of second order in time is much less developed. Some *a posteriori* bounds are proposed in [BS05; GLM13] for the wave equation using the Euler discretization in time, which is however known to be too diffusive and thus rarely used for the wave equation. More popular schemes, i.e. the leap-frog and cosine methods, are studied in [Geo+16] but only the error caused by discretization in time is considered. On the other hand, error estimators for the space discretization only are proposed in [Pic10; Adj02]. Goal-oriented error estimation and adaptivity for the wave equation were developed in [BGR10; BR01; BR99].

Thesis contribution

The aim of the thesis is to develop *a posteriori* error analysis in time and space in energy norm for the wave equation discretized with the Newmark scheme in time and with finite elements in space. We present the *a posteriori* error bounds and a corresponding adaptive algorithm in time and in space. The theoretical analysis is validated by various numerical experiments.

Scope of the thesis

In the first chapter we introduce the settings for *a posteriori* error estimates and give a brief overview of the so-called residual type *a posteriori* error estimators for elliptic, parabolic and hyperbolic problems. The main purpose here is to highlight the relevance of the derivation of the *a posteriori* error bounds for the wave equation discretized by the second order scheme in time. We start from the space error bounds and remind the classical results of *a posteriori* error analysis for finite element discretization on isotropic meshes on the example of Laplace problem. We also present the technique for anisotropic finite element developed in [Pic03], which is in turn

based on the anisotropic interpolation estimates developed in [FP01; FP03]. We turn then to the *a posteriori* error bound in time for the heat equation discretized in time with Crank-Nicolson method from [LPP09], that is constructed using continuous, piecewise quadratic polynomial reconstruction in time and some properties of the Crank-Nicolson scheme. The rest of the chapter is dedicated to *a posteriori* error estimates for the wave equation available in the literature. We start from the *a posteriori* error bound for the wave equation discretized with the Euler method in time [BS05]. For the sake of brevity, we explain the technique of deriving the time *a posteriori* bound on the example of ordinary differential equation of second order. The basic technical tool is the piecewise linear in time reconstruction of the discrete solution which leads to the first order in time error bound. We also derive two new alternative error estimators that are sharper then Bernardi-Süli estimator. Next we describe the approach for deriving the *a posteriori* error estimates for general cosine-type second order methods controlling the time discretization error from [Geo+16]. The estimator is based on the rewriting the scheme as the one-step system as in [BS05] and on the appropriate time reconstruction adapted from [AMN06]. Finally, we present the anisotropic estimator proposed in [Pic10] for the error due only to the the finite element discretization of the wave equation.

The aim of the second chapter is to obtain the *a posteriori* error bounds of optimal order in time and space for the linear second-order wave equation discretized by the Newmark scheme in time and by the finite elements in space. The main results of this chapter were announced in [GLP17b]. Error estimate is derived in the L^∞-in-time/energy in space norm. We set the parameters in the Newmark scheme $\beta = 1/4, \gamma = 1/2$ [BW76], since it provides a conservative method with respect to the energy norm. Another interesting feature of this variant of the method, which is in fact essential for our analysis, is the fact that the method can be reinterpreted as the Crank-Nicolson discretization of the reformulation of the governing equation in the first-order system, as in [Bak76]. Therefore we use the techniques stemming from *a posteriori* error analysis for the Crank-Nicolson discretization of the heat equation in [LPP09], based on a piecewise quadratic polynomial in time reconstruction of the numerical solution. The proposed strategy leads to *a posteriori* error estimate in time and also allows us to easily recover the estimates in space. The resulting estimates are referred to as the 3-point estimator since our quadratic reconstruction is drawn through the values of the discrete solution at 3 points in time. The reliability of the 3-point estimator is proved theoretically for general regular meshes in space and non-uniform meshes in time. It is also illustrated by numerical experiments. We discuss through this part also the question of optimality for our error estimate. Although we do not have a lower bound for our error estimators, we prove that the error indicator in time provides the estimate of second order at least on sufficiently smooth solutions and on uniform meshes. This result is obtained under a particular discretization of the initial conditions and the right-hand side function. This is validated by numerical experiments. Finally, we give the anisotropic *a posteriori* error estimate based on the technique from [FP01; FP03].

In the third chapter we propose a cheaper version of the 3-point *a posteriori* error estimator. This is achieved by replacing the second derivatives in space (Laplacian of the discrete solution) in the 3-point error estimate with the fourth derivatives in time. We call the resulting estimate the 5-point estimator since it contains the fourth order finite differences in time and thus involves the discrete solution at 5 points in time at each time step. The main results of this chapter were announced

in [GLP17a]. The new estimator preserves all the properties of the previous one (reliability, optimality on smooth solutions and quasi-uniform meshes) but no longer requires an extra computation of the Laplacian of the discrete solution on each time step. We perform some numerical tests that show the equivalence between the 3-point error estimator and the 5-point error estimator.

Chapter 4 is dedicated to the *a posteriori* error analysis in time for general second order Newmark scheme ($\gamma = 1/2$) in the case of second order ordinary differential equation. This is done by extending the approach for the 3-point error estimator from Chapter 2. Numerical experiments confirm similarity between convergence rate of the time error estimator and that of the true error. We consider explicit ($\beta = 0$) and implicit ($\beta = 1/4$) second order Newmark scheme. We also present numerical comparison between our time error estimator and the staggered grids time error estimator from [Geo+16] for the case of a constant time step.

Our goal in Chapter 5 is to apply the *a posteriori* analysis presented in Chapter 2 to mesh adaptivity in time and space. Numerical studies are reported for several test cases and show that the manner of interpolation of the numerical solution from mesh to mesh plays an important role for optimal behavior of the time error estimator and thus for the whole adaptation algorithm.

Finally, Chapter 6 is dedicated to PANLIM research project realized during CEMRACS 2016 project. We investigate the feasibility of applying the Parareal algorithm [Lio01; MT05] for quasi-static nonlinear structural analysis problems. We describe how this proposal has been realized and present some preliminary numerical results of applying this algorithm to a beam undergoing nonlinear deflection with a contact boundary condition. Further numerical experiments are needed to provide an evidence for the efficiency of the method.

Chapter 1

State of the art

We present in this chapter the main results in *a posteriori* error analysis for elliptic, parabolic and hyperbolic problems. We focus on residual type *a posteriori* error estimators. The purpose of this chapter is to motivate the derivation of *a posteriori* error bounds for the wave equation discretized by a second order scheme in time. Since high aspect ratio finite elements reduce the number of degrees of freedom, anisotropic *a posteriori* error estimates have received much more attention and thus we present the technique for anisotropic finite element developed in [Pic03], which is in turn based on the anisotropic interpolation estimates developed in [FP01; FP03]. This chapter is organized into three sections.

The first part introduces some notations and preliminary results that we will use all along the thesis. Then we give a brief summary of *a posteriori* error analysis for finite element method discretization on isotropic and anisotropic meshes on the example of Laplace problem. We recall the classical result of the residual type *a posteriori* error estimator for Laplace equation in the case of isotropic finite elements and briefly explain the approach to anisotropic finite elements from [Pic03].

The second part of this chapter is devoted to residual-based *a posteriori* error analysis for the heat equation discretized in time with Crank-Nicolson method and with continuous, piecewise linear finite elements in space [LPP09]. The *a posteriori* error bound in time is constructed using continuous, piecewise quadratic polynomial reconstruction in time and some properties of Crank-Nicolson scheme. The anisotropic space error estimator is derived using the same approach as in Laplace problem.

The third section of this chapter presents a review of *a posteriori* error estimates for the wave equation available in the literature. We start from the pioneering result for the wave equation using the Euler discretization in time [BS05]. For the sake of simplicity we explain the technique of deriving the time *a posteriori* bound on the example of ordinary differential equation of second order. The basic technical tool is the piecewise linear in time reconstruction of the discrete solution which leads to the optimal first order in time error bound. Then we present analogous estimate of the global error for fully discretized wave equation with standard residual-based local space estimator in space. Next we turn to *a posteriori* error estimates for general cosine-type second order methods controlling the time discretization error from [Geo+16]. The approach is based on the scheme as the one-step system as in [BS05] and uses an appropriate time reconstruction adapted from [AMN06]. Finally, we

present the anisotropic error estimator for the wave equation for the space discretization only, proposed in [Pic10].

Chapter contents

1.0.1 Preliminary results and notations

Given a polygonal domain $\Omega \subset \mathbb{R}^2$, for any $0 \leq h \leq 1$, let $\mathcal{T}_h$ be a conforming triangulation of Ω into triangles K with diameter h_K less than h. Let $\mathcal{E}_h$ represent the internal edges of the mesh $\mathcal{T}_h$.

Definition 1.0.1. *A family of meshes $\{\mathcal{T}_h\}_{h>0}$ is said to be shape-regular if there exists σ_0 such that*

$$\forall h, \quad \forall K \in \mathcal{T}_h, \qquad \sigma_K = \frac{h_K}{\rho_K} \leq \sigma_0, \tag{1.1}$$

where ρ_K is the diameter of the largest ball that can be inscribed in K.

Let $V_h \subset H_0^1(\Omega)$ be the usual finite element space of continuous, piecewise linear functions on the triangles of $\mathcal{T}_h$, vanishing on $\partial\Omega$:

$$V_h = \left\{ v_h \in C(\bar{\Omega}) : v_h|_K \in \mathbb{P}_1 \; \forall K \in \mathcal{T}_h \text{ and } v_h|_{\partial\Omega} = 0 \right\}. \tag{1.2}$$

Here and in what follows, we consider the usual Hilbertian Sobolev spaces $H^s(\Omega)$ for all non-negative real number s. The norms and semi-norms in Sobolev spaces $H^s(\Omega)$ are denoted, respectively, by $\|\cdot\|_{H^s(\Omega)}$ and $|\cdot|_{H^s(\Omega)}$. $H_0^1(\Omega)$ is the closure in $H^1(\Omega)$ of the space $C_c^\infty(\Omega)$ of infinitely differentiable functions with a compact support in Ω. Its dual space is denoted by $H^{-1}(\Omega)$ and is equipped with the corresponding dual norm $\|\cdot\|_{H^{-1}(\Omega)}$.

Definition 1.0.2. *The L^2-orthogonal projection $P_h : L^2(\Omega) \to V_h$ is defined by*

$$\forall v \in L^2(\Omega) : (P_h v, \varphi_h) = (v, \varphi_h), \quad \forall \varphi_h \in V_h. \tag{1.3}$$

Definition 1.0.3. *The H_0^1-orthogonal projection operator $\Pi_h : H_0^1(\Omega) \to V_h$ is defined by*

$$\forall v \in H_0^1(\Omega) : (\nabla \Pi_h v, \nabla \varphi_h) = (\nabla v, \nabla \varphi_h), \qquad \forall \varphi_h \in V_h. \tag{1.4}$$

We denote by I_h a Scott-Zhang interpolation operator $I_h : H_0^1(\Omega) \to V_h$ which is also a projection, i.e. $I_h V_h = V_h$ [EG04; SZ90].

Let us recall, for future reference, the well known stability and approximation properties of the H_0^1-orthogonal projection operator:

Proposition 1. *For every sufficiently smooth function v the following inequalities hold*

$$|\Pi_h v|_{H^1(\Omega)} \leq |v|_{H^1(\Omega)}, \quad |v - \Pi_h v|_{H^1(\Omega)} \leq Ch|v|_{H^2(\Omega)}, \tag{1.5}$$

with a constant $C > 0$ which depends only on the regularity of the mesh.

Proof. The stability estimate from (1.5) directly follows from the definition of elliptic projection. Indeed,

$$|\Pi_h v|^2_{H^1(\Omega)} = (\nabla \Pi_h v, \nabla \Pi_h v) = (\nabla v, \nabla \Pi_h v) \leq |v|_{H^1(\Omega)} |\Pi_h v|_{H^1(\Omega)}. \tag{1.6}$$

The approximation property (1.5) follows from interpolation properties of the Scott-Zhang interpolation operator [EG04]

$$|v - I_h v|_{H^1(\Omega)} \leq Ch\, |v|_{H^2(\Omega)}. \tag{1.7}$$

Since $\Pi_h v$ is the best approximation in H^1-norm, the desired result follows easily. □

We now recall the approximation properties of the Scott-Zhang interpolation operator:

Proposition 2. *For every sufficiently smooth function v, for all $K \in \mathcal{T}_h$ and $E \in \mathcal{E}_h$ we have*

$$\|v - I_h v\|_{L^2(K)} \leq Ch_K |v|_{H^1(\omega_K)} \text{ and } \|v - I_h v\|_{L^2(E)} \leq Ch_E^{1/2} |v|_{H^1(\omega_E)}. \tag{1.8}$$

Here ω_K (resp. ω_E) represents the set of triangles of $\mathcal{T}_h$ having a common vertex with triangle K (resp. edge E) and the constant $C > 0$ depends only on the regularity of the mesh.

Proof. See [SZ90]. □

1.0.2 Anisotropic finite elements

In this chapter we will work with *a posteriori* error estimates in 2-dimensional space for anisotropic finite elements. In order to describe the mesh anisotropy we introduce some definitions and properties from [FP01; FP03] which are used in *a posteriori* error estimates for the Laplace equation [Pic03]and in space *a posteriori* indicators for the heat equation [LPP09] and for the wave equation [Pic10].

For any triangle K of the mesh $\mathcal{T}_h$, let T_K be the affine transformation mapping the reference triangle $\hat{K}$ into K defined by

$$\boldsymbol{x} = T_K(\hat{\boldsymbol{x}}) = M_K \hat{\boldsymbol{x}} + \boldsymbol{t}_K,$$

where M_K is the Jacobian of T_K. M_K admits a singular value decomposition

$$M_K = R_K^T \Lambda_K P_K,$$

where R_K and P_K are orthogonal matrices and

$$\Lambda_K = \begin{pmatrix} \lambda_{1,K} & 0 \\ 0 & \lambda_{2,K} \end{pmatrix}, \quad \lambda_{1,K} \geq \lambda_{2,K} > 0. \tag{1.9}$$

We set

$$R_K = \begin{pmatrix} \boldsymbol{r}_{1,K}^T \\ \boldsymbol{r}_{2,K}^T \end{pmatrix}, \tag{1.10}$$

where $\boldsymbol{r}_{1,K}^T$, $\boldsymbol{r}_{2,K}^T$ are the unit vectors corresponding to directions of maximum and minimum stretching.

In the framework of anisotropic meshes, the classical minimum angle condition is not required. However, for each vertex, the number of neighboring vertices should be bounded from above, uniformly with respect to the mesh size h. Also, for each triangle K of the mesh, there is a restriction related to the patch Δ_K of elements around K. More precisely, the diameter of the reference patch Δ_K, that is, $\Delta_{\hat{K}} = T_K^{-1}(\Delta_K)$, must be uniformly bounded independently of the mesh geometry (see [Pic03] for illustrations). We suppose in the rest of this work that the family $\mathcal{T}_h$ meets the above mentioned restrictions.

Proposition 3. *Let I_h be the Scott-Zhang interpolation operator. There is a constant C independent of the mesh size and aspect ratio such that, for any $v \in H^1(\Omega)$ and any $K \in \mathcal{T}_h$ we have:*

$$\|v - I_h v\|_{L^2(K)} \leq C\omega_K(v), \tag{1.11}$$

$$\lambda_{2,K}\|\nabla(v - I_h v)\|_{L^2(K)} \leq C\omega_K(v), \tag{1.12}$$

$$\|v - I_h v\|_{L^2(\partial K)} \leq C\frac{1}{\lambda_{2,K}^{1/2}}\omega_K(v). \tag{1.13}$$

Here $\omega_K(v)$ is defined by

$$\omega_K^2(v) = \lambda_{1,K}^2 \left(\boldsymbol{r}_{1,K}^T G_K(v) \boldsymbol{r}_{1,K}\right) + \lambda_{2,K}^2 \left(\boldsymbol{r}_{2,K}^T G_K(v) \boldsymbol{r}_{2,K}\right), \tag{1.14}$$

where $\lambda_{i,K}$ and $\boldsymbol{r}_{i,K}$ are given by (1.9) and (1.10) and $G_K(v)$ is the following 2×2 matrix

$$G_K(v) = \sum_{T \in \Delta_K} \begin{pmatrix} \displaystyle\int_T \left(\frac{\partial v}{\partial x_1}\right)^2 dx & \displaystyle\int_T \frac{\partial v}{\partial x_1}\frac{\partial v}{\partial x_2} dx \\ \displaystyle\int_T \frac{\partial v}{\partial x_1}\frac{\partial v}{\partial x_2} dx & \displaystyle\int_T \left(\frac{\partial v}{\partial x_2}\right)^2 dx \end{pmatrix}. \tag{1.15}$$

We shall use the interpolation estimates of Proposition 3 in order to derive *a posteriori* error estimates at first for the model Laplace equation and after to derive space *a posteriori* error estimates for the heat equation and for the wave equation. More precisely, in all cases the technique is based on the following principles: the error is first related to the equation residual Scott-Zhang interpolation operator is introduced, as in standard textbooks [EG04]. Then, the anisotropic interpolation

estimates presented in Proposition 3 are used, and finally, a Z-Z error estimator is used to approach the error gradient matrix.

We are going to prove now the stability properties for the Scott-Zhang interpolator on anisotropic mesh. We choose a particular implementation of the Scott-Zhang interpolator I_h, defined for any $v \in L^2(\Omega)$ as follows. For any interior mesh node x, consider the patch Δ_x of mesh triangles attached to x. Let K_x be the triangle in Δ_x with maximal area. We set then

$$I_h v(x) = \frac{1}{|K_x|} \int_{K_x} v\psi_x,$$

where ψ_x is the polynomial of degree 1 which is dual to the hat function φ_x (the standard basis function of V_h associated to node x) in the scalar product of $L^2(K_x)$. We mean by this

$$(\psi_x, \varphi_x)_{L^2(K_x)} = 1 \text{ and } (\psi_x, \varphi_y)_{L^2(K_x)} = 0,$$

for all the hat function φ_y associated to mesh nodes y other than x. For the boundary nodes x we should proceed differently in order to preserve homogeneous Dirichlet boundary conditions. We thus denote by E_x a boundary edge attached to any boundary node x and define

$$I_h v(x) = \frac{1}{|E_x|} \int_{E_x} v\psi_x,$$

with ψ_x redefined as the dual function in the scalar product of $L^2(E_x)$. This defines uniquely Scott-Zhang interpolator $I_h : H_0^1(\Omega) \to V_h$.

Proposition 4. *We have for any triangle $K \in \mathcal{T}_h$ and any $v \in H_0^1(\Omega)$*

$$\|I_h v\|_{L^2(K)} \leqslant 3\|v\|_{L^2(\omega_K)}.$$

Proof. Consider any triangle $K_x \in \mathcal{T}_h$ with the vertices, say, x, y, z. Denoting by $\varphi_x, \varphi_y, \varphi_z$ the hat functions associated to these vertices and applying a quadrature rule, we compute easily

$$(\varphi_x, \varphi_x)_{L^2(K_x)} = \frac{|K_x|}{6},\ (\varphi_x, \varphi_y)_{L^2(K_x)} = (\varphi_x, \varphi_z)_{L^2(K_x)} = \frac{|K_x|}{12}, \dots$$

Recalling the definition of the dual function ψ_x, i.e. the polynomial of degree 1 such that

$$(\psi_x, \varphi_x)_{L^2(K_x)} = 1,\ (\psi_x, \varphi_y)_{L^2(K_x)} = (\psi_x, \varphi_z)_{L^2(K_x)} = 0,$$

we can easily compute it as a linear combination of $\varphi_x, \varphi_y, \varphi_z$. This gives

$$\psi_x = \frac{1}{|K_x|}(9\varphi_x - 3\varphi_y - 3\varphi_z),$$

so that

$$\|\psi_x\|_{L^2(K_x)} = \frac{3}{|K_x|^{1/2}}.$$

Take now any triangle $K \in \mathcal{T}_h$ with vertices x_1, x_2, x_3. Suppose for the moment that all these vertices are interior nodes of the mesh. Also denote the midpoints of

the edges of K by m_1, m_2, m_3. Observe, using the quadrature rule exact for polynomials of degree 2,

$$\|I_h v\|^2_{L^2(K)} = \frac{|K|}{3}\sum_{i=1}^{3}|I_h v(m_i)|^2 \leqslant \frac{|K|}{3}\sum_{i=1}^{3}|I_h v(x_i)|^2 \tag{1.16}$$
$$= \frac{|K|}{3}\sum_{i=1}^{3}\left|\int_{K_{x_i}}\psi_{x_i} v\right|^2 \leqslant \frac{|K|}{3}\sum_{i=1}^{3}\frac{9}{|K_{x_i}|}\|v\|^2_{L^2(K_{x_i})} \leqslant 9\|v\|^2_{L^2(\omega_K)}.$$

The last inequality is valid since K_{x_i} is the mesh triangle of maximal area in the patch Δ_{x_i} and $K \in \Delta_{x_i}$ so that $|K| \leq |K_{x_i}|$.

Note, finally, that all the estimates above remain valid if some of the vertices of triangle K lie on the boundary $\partial\Omega$. Indeed, we simply have $I_h v(x_i) = 0$ at such a node x_i so that it can be neglected in the sums over the vertices in (1.16). □

1.1 The Laplace equation

Given $f \in L^2(\Omega)$, we are searching for $u : \Omega \to \mathbb{R}$ such that

$$\begin{aligned} -\Delta u &= f, \quad \text{in } \Omega, \\ u &= 0, \quad \text{on } \partial\Omega. \end{aligned} \tag{1.17}$$

The simplest finite element approximation of (1.17) therefore consists in seeking $u_h \in \mathcal{T}_h$ such that

$$\int_\Omega \nabla u_h \cdot \nabla \varphi_h = \int_\Omega f\varphi_h, \ \forall \varphi_h \in V_h. \tag{1.18}$$

1.1.1 Isotropic *a posteriori* error estimate

In order to remind the basic technique for residual-based *a posteriori* error estimates we start from the following classical isotropic *a posteriori* error estimate [EG04].

Lemma 1. *There is a constant C depending only on the interpolation constants of Proposition 2 such that the following a posteriori error estimate holds*

$$|u - u_h|_{H^1(\Omega)} \leq C\sum_{K\in\mathcal{T}_h}\eta_K(u_h, f), \tag{1.19}$$

where local error indicator is defined by

$$\eta_K(u_h, f) \ = \ \sum_{K\in\mathcal{T}_h}\left(h_K\,\|f + \Delta u_h\|_{L^2(K)} + \frac{1}{2}\sum_{E\in\mathcal{E}_K} h_E^{1/2}\|\,[\boldsymbol{n}\cdot\nabla u_h]\,\|_{L^2(E)}\right). \tag{1.20}$$

Here $[\cdot]$ stands for the jump of the bracketed quantity across an internal edge $E \in \mathcal{E}_h$, $\mathcal{E}_K$ is the set of internal edges of K, and $\boldsymbol{n}$ is the unit edge normal (in arbitrary direction).

Proof. Using (1.18) and the standard idea of integration by parts, we have, for any $v \in H_0^1(\Omega)$

$$\begin{aligned}\int_\Omega f\varphi - \int_\Omega \nabla u_h \cdot \nabla\varphi &= \int_\Omega f(\varphi - I_h\varphi) - \int_\Omega \nabla u_h \cdot \nabla(\varphi - I_h\varphi) \\ &= \sum_{K\in\mathcal{T}_h}\left(\int_K (f+\Delta u_h)(\varphi - I_h\varphi) + \sum_{E\in\partial K}\int_E [\boldsymbol{n}\cdot\nabla u_h]\,(\varphi - I_h\varphi)\right).\end{aligned}$$

Second, using the Cauchy-Shwarz inequality and the fact that $\varphi - I_h\varphi$ vanishes on $\partial\Omega$, we obtain

$$\begin{aligned}\int_\Omega f\varphi - \int_\Omega \nabla u_h \cdot \nabla\varphi \leq \sum_{K\in\mathcal{T}_h}\Bigg(&\|f+\Delta u_h\|_{L^2(K)}\,\|\varphi - I_h\varphi\|_{L^2(K)} \\ &+ \frac{1}{2}\sum_{E\in\mathcal{E}_K}\|[\boldsymbol{n}\cdot\nabla u_h]\|_{L^2(E)}\,\|\varphi - I_h\varphi\|_{L^2(E)}\Bigg).\end{aligned}$$

We now use the standard interpolation estimates of Proposition 2 and obtain

$$\begin{aligned}&\int_\Omega f\varphi - \int_\Omega \nabla u_h \cdot \nabla\varphi \qquad (1.21)\\ &\leq C\left(\sum_{K\in\mathcal{T}_h}\left(h_K\,\|f+\Delta u_h\|_{L^2(K)} + \frac{1}{2}\sum_{E\in\mathcal{E}_K} h_E^{1/2}\|\,[\boldsymbol{n}\cdot\nabla u_h]\,\|_{L^2(E)}\right)\right)|\varphi|_{H^1(\Omega)}.\end{aligned}$$

It then suffices to choose $\varphi = u - u_h$ in the estimate above and use the fact that, by definition of u and u_h we have

$$|u-u_h|^2_{H^1(\Omega)} = \int_\Omega f(u-u_h) - \int_\Omega \nabla u_h\cdot\nabla(u-u_h), \qquad (1.22)$$

to obtain the desired result. □

We now present the lower bound of the error. Following [EG04], [Ver96] and [Ver94], we introduce some notations and present two technical lemmas.

We first introduce an extension operator which maps functions define on an interface to triangles sharing the interface. Let E be an interior mesh face and K be a triangle containing the edge E. Consider the reference triangle $\hat{K}$ and let T_K be the affine transformation mapping from reference triangle $\hat{K}$ into K. Without loss of generality we assume, that $E = T_K(\hat{E})$. We define the extension operator $\hat{P}_{\hat{E}} : \mathbb{P}_1(\hat{E}) \to \mathbb{P}_1(\hat{K})$ as

$$\hat{P}_{\hat{E}}\hat{\phi}(x,y) = \hat{\phi}(x),\ \forall\hat{\phi}\in\mathbb{P}_1(\hat{E}),\ \forall(x,y)\in\hat{K}.$$

The extension operator $P_{E,K} : \mathbb{P}_1(\hat{E}) \to \mathbb{P}_1(\hat{K})$ is defined $\forall\phi\in\mathbb{P}_1(E)$ as

$$P_{E,K}(\phi) = \hat{P}_{\hat{E}}(\phi\circ T_K)\circ T_K^{-1}.$$

Let D_K be the union of the triangles sharing an edge with the triangle K and D_E be the union of two triangles K and K' sharing the edge E. The extension operator

P_E is defined $\forall \phi \in \mathbb{P}_1(E)$ as

$$\begin{cases} P_{E,K}(\phi) & \text{on } E, \\ P_{E,K'}(\phi) & \text{on } E'. \end{cases} \tag{1.23}$$

Lemma 2. *Let $b_K \in \mathbb{P}_3(K)$ be a function such that:*

(i) $0 \leq b_K \leq 1$.

(ii) $\exists D \subset K$ *with* $\mathrm{meas}(D) > 0$ *and* $b_{K|D} \geq 1/2$.

Here $\mathrm{meas}(D)$ *is the Lebesgue measure of* D. *Then, there exists* c_1 *and* c_2 *such that,* $\forall K \in \mathcal{T}_h$, $\forall \phi \in \mathbb{P}_1(K)$,

$$\|b_K \phi\|_{L^2(K)} \leq \|\phi\|_{L^2(K)} \leq c_1 \left\| b_K^{1/2} \phi \right\|_{L^2(K)}, \tag{1.24}$$

$$|b_K \phi|_{H^1(K)} \leq c_2 h_K^{-1} \|\phi\|_{L^2(K)}. \tag{1.25}$$

Proof. See [Ver96] and [Ver94]. □

Lemma 3. *Let $b_E \in \mathbb{P}_3(D_E)$ be a function such that:*

(i) $0 \leq b_E \leq 1$.

(ii) $\exists D \subset D_E$ *with* $\mathrm{meas}(D) > 0$ *and* $b_{E|D} \geq 1/2$.

(iii) $b_{E|E} \in H_0^1(E)$.

(iv) $\exists D' \subset E$ *with* $\mathrm{meas}(D') > 0$ *and* $b_{E|D'} \geq 1/2$.

Then, there exists c_1, c_2, c_3 *and* c_4 *such that,* $\forall E \in \mathcal{E}_h$, $\forall \phi \in \mathbb{P}_1(E)$,

$$\|b_E \phi\|_{L^2(E)} \leq \|\phi\|_{L^2(E)} \leq c_1 \left\| b_E^{1/2} \phi \right\|_{L^2(E)}, \tag{1.26}$$

$$c_2 h_E^{1/2} \|\phi\|_{L^2(E)} \leq \|b_E P_E(\phi)\|_{L^2(D_E)} \leq c_3 h_E^{1/2} \|\phi\|_{L^2(E)}, \tag{1.27}$$

$$|b_E P_E(\phi)|_{H^1(D_E)} \leq c_4 h_E^{-1/2} \|\phi\|_{L^2(E)}. \tag{1.28}$$

Proof. See [Ver96] and [Ver94]. □

The optimality result for the error indicator (1.19) is stated in the following:

Theorem 4. *There exists a constant* c, *depending only on the shape-regularity of the mesh* $\mathcal{T}_h$ *and the reference finite element, such that*

$$\eta_K(u_h, f) \leq c \left(|u - u_h|_{H^1(D_E)} + h_K \inf_{v_h \in V_h} \|f - v_h\|_{L^2(D_E)} \right), \tag{1.29}$$

where local error indicator $\eta_K(u_h, f)$ *is defined by (1.20).*

Proof. The proof is standard and can be found in [EG04]. For brevity, we provide here only the bound for term $\|f + \Delta u_h\|_{L^2(K)}$. For all $v_h \in V_h$ we have

$$\|f + \Delta u_h\|_{L^2(K)} \leq \|f + v_h\|_{L^2(K)} + \|v_h + \Delta u_h\|_{L^2(K)}.$$

From Lemma 2, using the fact that b_K vanishes on the boundary of K, we obtain

$$\begin{aligned} c\|v_h+\Delta u_h\|^2_{L^2(K)} \le \|b_K^{1/2}(v_h+\Delta u_h)\|^2_{L^2(K)} &= \int_K (v_h+\Delta u_h) b_K (v_h+\Delta u_h) \\ \int_K (f+\Delta u_h) b_K(v_h+\Delta u_h) &+ \int_K (v_h-f) b_K(v_h+\Delta u_h) \\ \le \int_K \nabla(u-u_h)\cdot\nabla(b_K(v_h+\Delta u_h)) &+ \|v_h-f\|_{L^2(K)}\|v_h+\Delta u_h\|_{L^2(K)}. \end{aligned}$$

The inverse inequality from Lemma 2 yields

$$\begin{aligned} c\|v_h+\Delta u_h\|^2_{L^2(K)} &\le |u-u_h|_{H^1(K)}|b_K(v_h+\Delta u_h)|_{H^1(K)} + \|v_h-f\|_{L^2(K)}\|v_h+\Delta u_h\|_{L^2(K)} \\ &\le \left(ch_K^{-1}|u-u_h|_{H^1(K)} + \|v_h-f\|_{L^2(K)}\right)\|v_h+\Delta u_h\|_{L^2(K)}. \end{aligned}$$

Thus

$$\|f+\Delta u_h\|_{L^2(K)} \le c\left(h_K^{-1}|u-u_h|_{H^1(K)} + \|v_h-f\|_{L^2(K)}\right).$$

□

1.1.2 Anisotropic *a posteriori* error estimate

Anisotropic mesh adaptation is now widely used in numerical simulations to improve the accuracy of the solutions as well as to capture the behavior of physical phenomena. An anisotropic, adaptive finite elements allow to reduce the number of vertices required to reach a given level of accuracy. Thereby anisotropic *a posteriori* error estimates plays an important role for improving the accuracy of the numerical solution.

Reproducing the proof of Lemma 1 and using the results from Proposition 3 we obtain the following anisotropic error estimate [Pic03].

Lemma 5. *There is a constant $C = C(\hat{K})$ depending only on the interpolation constants of Proposition 3 (thus not on the mesh size or aspect ratio) such that*

$$\begin{aligned} &|u-u_h|^2_{H^1(\Omega)} \\ &\le C\sum_{K\in\mathcal{T}_h}\left(\|f+\Delta u_h\|_{L^2(K)} + \frac{1}{2\lambda_{2,K}^{1/2}}\|\left[\boldsymbol{n}\cdot\nabla u_h\right]\|_{L^2(\partial K)}\right)\omega_K(u-u_h), \end{aligned} \tag{1.30}$$

where ω_K is defined by (1.14).

Proof. The proof is similar to that of Lemma 1. We reproduce it up to equation (1.21). We now use the interpolation estimates of Proposition 3 and obtain

$$\begin{aligned} &\int_\Omega f\varphi - \int_\Omega \nabla u_h\cdot\nabla\varphi \\ &\le C\sum_{K\in\mathcal{T}_h}\left(\|f+\Delta u_h\|_{L^2(K)} + \frac{1}{2\lambda_{2,K}^{1/2}}\|\left[\boldsymbol{n}\cdot\nabla u_h\right]\|_{L^2(\partial K)}\right)\omega_K(\varphi), \end{aligned} \tag{1.31}$$

where $C = C(\hat{K})$. Then, as in Lemma 1, we conclude by choosing $v = u - u_h$ and using (1.22). □

Remark 1. *Note that the constant C in Lemma 1 depends on the triangles aspect ratio, while in the proof of Lemma 5 the classical regularity (or minimum angle) is not used. That is to say the estimate holds for highly anisotropic meshes. Moreover, if we assume that $\lambda_{1,K}/\lambda_{2,K}$ is bounded from above (that said, the triangles aspect ratio is bounded above), then from estimate (1.30) we easily recover the classical isotropic estimate (1.19).*

1.1.3 Zienkiewicz-Zhu recovery

Note that estimate (1.30) is not a standard *a posteriori* error estimate since the exact solution u is still involved in $\omega_K(u-u_h)$ at the right-hand side. In this paragraph we explain the approach to derive an anisotropic error indicators from papers [Pic03; LPP09; Pic10]. The technique to approach the error gradient is based on Z-Z error estimator [ZZ87; Ain+89; ZZ92]. More precisely, following Z-Z recovery technique, we replace the first order space partial derivatives of the exact solution $\dfrac{\partial u}{\partial x_1}$ and $\dfrac{\partial u}{\partial x_2}$ by, respectively, $\tilde{P}_h\left(\dfrac{\partial u_h}{\partial x_1}\right)$ and $\tilde{P}_h\left(\dfrac{\partial u}{\partial x_2}\right)$, where $\tilde{P}_h : L^2(\Omega) \to V_h$ is an approximate $L^2(\Omega)$ projection on V_h defined for any $g \in L^2(\Omega)$ as

$$\int_\Omega r_h\left(\left(\tilde{P}_h g\right) v_h\right) = \int_\Omega g v_h \quad \forall v_h \in V_h.$$

Here r_h denotes the piecewise linear Lagrange interpolant operator. That is to say from constant values of ∇u_h on triangles, $\tilde{P}_h\left(\dfrac{\partial u_h}{\partial x_i}\right)$ is defined by its values at each vertex P as

$$\begin{pmatrix} \tilde{P}_h\left(\dfrac{\partial u_h}{\partial x_1}\right)(P) \\ \tilde{P}_h\left(\dfrac{\partial u_h}{\partial x_2}\right)(P) \end{pmatrix} = \frac{1}{\displaystyle\sum_{\substack{K\in\mathcal{T}_h \\ P\in K}} |K|} \begin{pmatrix} \displaystyle\sum_{\substack{K\in\mathcal{T}_h \\ P\in K}} |K| \left(\frac{\partial u_h}{\partial x_1}\right)|_K \\ \displaystyle\sum_{\substack{K\in\mathcal{T}_h \\ P\in K}} |K| \left(\frac{\partial u_h}{\partial x_2}\right)|_K \end{pmatrix}.$$

Thus the anisotropic error estimator is obtained by replacing the matrix $G_K(u-u_h)$ in (1.30) by the matrix $\tilde{G}_K(u_h)$ defined by

$$\tilde{G}_K(u_h) = \begin{pmatrix} \displaystyle\int_K (\eta_1^{ZZ}(u_h))^2 dx & \displaystyle\int_K \eta_1^{ZZ}(u_h)\eta_2^{ZZ}(u_h) dx \\ \displaystyle\int_K \eta_1^{ZZ}(u_h)\eta_2^{ZZ}(u_h) dx & \displaystyle\int_K (\eta_2^{ZZ}(u_h))^2 dx \end{pmatrix}, \tag{1.32}$$

where

$$\boldsymbol{\eta}^{ZZ}(u_h) = \begin{pmatrix} \eta_1^{ZZ}(u_h) \\ \eta_2^{ZZ}(u_h) \end{pmatrix} = \begin{pmatrix} (I-\tilde{P}_h)\left(\dfrac{\partial u_h}{\partial x_1}\right) \\ (I-\tilde{P}_h)\left(\dfrac{\partial u_h}{\partial x_2}\right) \end{pmatrix}. \tag{1.33}$$

Numerical results show the efficiency of Z-Z post-processing for anisotropic space *a*

posteriori error estimate for elliptic [Pic03], parabolic [LPP09] and hyperbolic problems [Pic10].

1.2 The heat equation

The aim of this subsection is to give the main ideas for deriving time and space *a posteriori* error estimates for the heat equation given in [LPP09]. The approach from [AMN06] is used in order to estimate error from time discretization and the approach to anisotropic finite elements developed in [Pic98; Pic03; Pic10] is used for space error estimators.

1.2.1 *A posteriori* error estimate for a first order ordinary differential equation

Consider a given final time $T > 0$. In order to explain the main idea for *a posteriori* error estimate in time we consider first the following ordinary differential equation

$$\begin{cases} \dfrac{du(t)}{dt} + Au(t) = 0, \quad t \in [0;T] \\ u(0) = u_0, \end{cases} \tag{1.34}$$

with a constant $A > 0$. This problem serves as simplification of the heat equation in which we get rid of the space variable. For the sake of simplicity we derive *a posteriori* error estimates for homogeneous equation (1.34), but the results can be easily extended to general case of non-zero right-hand side. In order to describe the time discretization corresponding to (1.34), let us introduce a subdivision of the time interval $[0, T]$

$$0 = t_0 < t_1 < \cdots < t_N = T,$$

with time steps $\tau_n = t_{n+1} - t_n$ for $n = 0, \ldots, N-1$ and $\tau = \max\limits_{0 \le n \le N-1} \tau_n$. The discretization considered here corresponds to the Crank-Nicolson method and has the following form for $n = 0, \ldots, N-1$

$$\frac{u^{n+1} - u^n}{\tau_n} + A\frac{u^{n+1} + u^n}{2} = 0, \tag{1.35}$$

where $u^0 = u_0$.

We shall need the following notations for $n = 0, \ldots, N-1$

$$\partial_n u = \frac{u^{n+1} - u^n}{\tau_n}, \quad u^{n+1/2} = \frac{1}{2}(u^{n+1} + u^n), \quad \tau_{n-1/2} = \frac{1}{2}(\tau_n + \tau_{n-1}), \tag{1.36}$$

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

$$\partial_n^2 u = \frac{\partial_n u - \partial_{n-1} u}{\tau_{n-1/2}}. \tag{1.37}$$

We introduce continuous piecewise quadratic approximation in time for $t \in [t_n, t_{n+1}]$ where $n = 1, \ldots, N-1$

$$\tilde{u}_\tau(t) = u^{n+1/2} + (t - t^{n+1/2})\partial_n u + \frac{1}{2}(t - t_n)(t - t_{n+1})\partial_n^2 u. \tag{1.38}$$

Lemma 6. *The following a posteriori error estimate holds between the solution u of problem (1.34) and the time reconstruction $\tilde{u}_\tau$ (1.38), for all t_{n+1}, $n = 1, \dots, N-1$:*

$$\int_{t_1}^{t_{n+1}} A\,|u - \tilde{u}_\tau|^2\,dt + |u(t_{n+1}) - \tilde{u}_\tau(t_{n+1})|^2 \leq |u(t_1) - \tilde{u}_\tau(t_1)|^2 + \sum_{m=1}^{n} \eta^2(t_n), \quad (1.39)$$

where the error estimator $\eta(t_n)$ is defined by

$$\eta^2(t_n) = \tau_n^3 \left(\frac{\tau_{n-1}^2}{48} + \frac{\tau_n^2}{120} \right) A|\partial_n^2 u|^2.$$

Proof. The *a posteriori* analysis relies on an appropriate residual equation for the reconstruction $\tilde{u}_\tau$. Thus, introducing the error between reconstruction $\tilde{u}_\tau$ and solution u to problem (1.34) $\tilde{e} = u - \tilde{u}_\tau$ the residual equation is defined as follows

$$\begin{aligned} \frac{d\tilde{e}}{dt} + A\tilde{e} &= -\frac{d\tilde{u}_\tau(t)}{dt} - A\tilde{u}_\tau(t) \\ &= -(t - t_{n-1/2})\partial_n^2 u - A(t - t_{n-1/2})\partial_n u - A(t - t_{n-1})(t - t_n)\partial_n^2 u. \end{aligned} \quad (1.40)$$

Consider now (1.35) at time steps $n+1$ and n. Subtracting one from another and dividing by $\tau_{n-1/2}$ yields

$$\partial_n^2 u + A\left(\frac{u^{n+1} - u^{n-1}}{\tau_n + \tau_{n-1}} \right) = 0. \quad (1.41)$$

Moreover, note that

$$-\frac{u^{n+1} - u^{n-1}}{\tau_n + \tau_{n-1}} + \partial_n u = \frac{\tau_{n-1}}{2}\partial_n^2 u,$$

so that (1.40) simplifies to

$$\frac{d\tilde{e}}{dt} + A\tilde{e} = -A\frac{\tau_{n-1}}{2}(t - t_{n+1/2})\partial_n^2 u - A\frac{1}{2}(t - t_n)(t - t_{n+1})\partial_n^2 u. \quad (1.42)$$

Thus multiplying (1.42) by $\tilde{e}$, using the fact that $ab \leq \frac{1}{2}a^2 + \frac{1}{2}b^2$ and the Cauchy-Schwarz inequality we obtain

$$\frac{d|\tilde{e}|^2}{dt} + A|\tilde{e}|^2 \leq A\left| \frac{\tau_{n-1}}{2}(t - t_{n+1/2})\partial_n^2 u + \frac{1}{2}(t - t_n)(t - t_{n+1})\partial_n^2 u \right|^2. \quad (1.43)$$

Integrating the last inequality from t_n to t_{n+1} we obtain

$$\int_{t_n}^{t_n+1} A\,|\tilde{e}|^2\,dt + |\tilde{e}(t_{n+1})|^2 \leq |\tilde{e}(t_n)|^2 + \tau_n^3 \left(\frac{\tau_{n-1}^2}{48} + \frac{\tau_n^2}{120} \right) A|\partial_n^2 u|^2. \quad (1.44)$$

Summing up this inequality on n leads to the desired result. □

1.2.2 *A posteriori* error estimates in space and time

Let $u = u(x,t) : \Omega \times [0,T] \to \mathbb{R}$ be the solution to

$$\begin{cases} \dfrac{\partial u}{\partial t} - \Delta u = f, & \text{in } \Omega \times]0,T], \\ u = 0, & \text{on } \partial\Omega \times]0,T], \\ u(\cdot,0) = u_0, & \text{in } \Omega, \end{cases} \tag{1.45}$$

where f, u_0, v_0 are given functions. We suppose henceforth $f \in L^2(0,T;H^{-1}(\Omega))$, $u_0 \in L^2(\Omega)$ and seek a solution $u \in W$ [Eva10] with

$$W = \left\{ w \in L^2\left(0,T;H_0^1(\Omega)\right) \text{and} \frac{\partial w}{\partial t} \in L^2\left(0,T;H^{-1}(\Omega)\right) \right\}, \tag{1.46}$$

such that $u(x,0) = u_0$ and

$$\left\langle \frac{\partial u}{\partial t}, \varphi \right\rangle + (\nabla u, \nabla \varphi) = (f, \varphi), \ \forall \varphi \in H_0^1(\Omega), \tag{1.47}$$

where $\langle \cdot, \cdot \rangle$ denotes the duality pairing between $H^{-1}(\Omega)$ and $H_0^1(\Omega)$ and the parentheses $(\cdot,\cdot)$ stand for the inner product in $L^2(\Omega)$. Let V_h be the usual finite element space of continuous, piecewise linear functions on the mesh $\mathcal{T}_h$ defined in (1.2). We set the initial condition to $u_h^0 = r_h u^0$. For each $n = 1, \ldots, N$ we compute $u_h^n \in V_h$ such that $\forall \varphi_h \in V_h$

$$\int_\Omega \left(\frac{u_h^{n+1} - u_h^n}{\tau_n} \right) \varphi_h dx + \frac{1}{2} \int_\Omega (\nabla u_h^{n+1} + \nabla u_h^n) \cdot \nabla \varphi_h dx = \frac{1}{2} \int_\Omega (f^{n+1} + f^n) \varphi_h dx. \tag{1.48}$$

From here on, f^n is an abbreviation for $f(\cdot, t_n)$. In this section the following notations will be used for the first order discrete derivatives in time for $n = 0, \ldots, N-1$:

$$\partial_n u_h = \frac{u_h^n - u_h^{n-1}}{\tau_{n-1}}, \quad u_h^{n+1/2} = \frac{1}{2}(u_h^{n+1} + u_h^n), \tag{1.49}$$

and for the second order discrete derivatives for $n = 1, \ldots, N-1$:

$$\partial_n^2 u_h = \frac{\partial_{n+1} u_h - \partial_n u_h}{\tau_{n-1/2}}. \tag{1.50}$$

We start from introducing the continuous piecewise linear approximation in time for $t \in [t_n, t_{n+1}]$ where $n = 0, \ldots, N-1$

$$u_{h\tau}(x,t) = u_h^{n+1/2} + (t - t^{n+1/2}) \partial_n u_h. \tag{1.51}$$

We also introduce continuous piecewise quadratic approximation in time for $t \in [t_n, t_{n+1}]$ where $n = 1, \ldots, N-1$

$$\tilde{u}_{h\tau}(x,t) = u_h^{n+1/2} + (t - t^{n+1/2})\partial_n u_h + \frac{1}{2}(t - t_n)(t - t_{n+1})\partial_n^2 u_h. \tag{1.52}$$

Exactly the second order reconstruction allows an *a posteriori* error estimate of optimal second order in time to be obtained. We also set for $t \in [t_n, t_{n+1}]$ where $n = 1, \ldots, N-1$

$$\tilde{f} = f^{n+1/2} + (t - t^{n+1/2})\frac{f^{n+1} - f^{n-1}}{\tau_n + \tau_{n-1}}.$$

The *a posteriori* analysis relies on an appropriate residual equation for the reconstruction $\tilde{u}_{h\tau}$. We set $e = u - u_{h\tau}$ and $\tilde{e} = u - \tilde{u}_{h\tau}$. To derive the *a posteriori* error estimate in time, like in the case of ordinary differential equation, we need the following properties of the Crank-Nicolson scheme and of the reconstruction $\tilde{u}_{h\tau}$.

Proposition 5. *For all $\varphi_h \in V_h$ and for $t \in [t_n, t_{n+1}]$ where $n = 1, \ldots, N-1$ we have*

$$\int_\Omega \frac{\partial \tilde{u}_{h\tau}}{\partial t}\varphi_h dx + \int_\Omega \nabla u_{h\tau} \cdot \nabla \varphi_h dx = \int_\Omega \tilde{f}\varphi_h dx + (t - t^{n+1/2})\frac{\tau_{n-1}}{2}\int_\Omega \nabla \partial_n^2 u_h \cdot \nabla \varphi_h dx. \tag{1.53}$$

We will now present the error indicator based on $\tilde{u}_{h\tau}$.

Isotropic error estimate

We will present isotropic error estimate for the error measured in a slightly stronger norm than before, namely

$$u \mapsto \left(\int_0^T |u(t)|^2_{H^1(\Omega))} + \int_0^T \left\| \frac{\partial u}{\partial t}(t) \right\|^2_{H^{-1}(\Omega))} \right)^{1/2}.$$

The choice of the norm is motivated by the fact that the upper and the lower bounds of optimal order can be proved for the space and time indicators.

Lemma 7 (Upper bound)**.** *Let u be the solution of problem (1.45) and $(u_h^n)_{0 \le n \le N}$ be the discrete solution given by Crank-Nicolson method (1.35). Then the following global upper bound of the error holds*

$$\int_{t^1}^T |\tilde{e}|^2_{H^1(\Omega))} + \int_{t^1}^T \left\| \frac{\partial \tilde{e}}{\partial t} \right\|^2_{H^{-1}(\Omega))} \le ||e(\cdot, t^1)||^2_{L^2(\Omega)} + C \sum_{n=1}^{N-1} \left\{ \sum_{K \in \mathcal{T}_h} (\eta^S_{K,n})^2 + (\eta^T_n)^2 + \int_{t^n}^{t^{n+1}} \left\| f - \tilde{f} \right\|^2_{L^2(\Omega)} dt + \sum_{K \in \mathcal{T}_h} \frac{h_k^2 \tau_n^3}{12} \left\| \partial_n^2 u_h \right\|^2_{L^2(K)} \right\}, \tag{1.54}$$

where the traditional local error indicator in space is defined $\forall K \in \mathcal{T}_h$ by

$$(\eta^S_{K,n})^2 = \int_{t^n}^{t^{n+1}} \left(h_K^2 \left\| f - \partial_n u_h + \Delta u_{h\tau} \right\|^2_{L^2(K)} + h_K \left\| [\nabla u_{h\tau} \cdot \mathbf{n}] \right\|^2_{L^2(\partial K)} \right) dt,$$

and the error indicator in time is defined by

$$(\eta_n^T)^2 = \left\{ \frac{\tau_{n-1}^2 \tau_n^3}{48} + \frac{\tau_n^5}{120} \right\} \left\| \nabla \partial_n^2 u_h \right\|_{L^2(\Omega)}^2 . \tag{1.55}$$

Proof. Follows from Proposition 5 and standard estimates for the Scott-Zhang interpolation I_h on isotropic meshes (1.8). □

Lemma 8 (Lower bound)**.** *Let u be the solution of problem (1.45) and $(u_h^n)_{0\leq n\leq N}$ be the discrete solution given by Crank-Nicolson method (1.35). Then the following local in time lower bounds $\forall n \geq 1$ hold*

$$\frac{1}{3}(\eta_n^T)^2 \leq \int_{t^n}^{t^{n+1}} |\tilde{e}|_{H^1(\Omega)}^2 + \int_{t^n}^{t^{n+1}} \left\| \frac{\partial \tilde{e}}{\partial t} \right\|_{H^{-1}(\Omega)}^2 + \int_{t^n}^{t^{n+1}} \left\| f - \tilde{f} \right\|_{H^{-1}(\Omega)}^2 dt, \tag{1.56}$$

and

$$\sum_{K\in\mathcal{T}_h} (\eta_{K,n}^S)^2 \leq C \left(\int_{t^n}^{t^{n+1}} |\tilde{e}|_{H^1(\Omega)}^2 + \int_{t^n}^{t^{n+1}} \left\| \frac{\partial \tilde{e}}{\partial t} \right\|_{H^{-1}(\Omega)}^2 \right), \tag{1.57}$$

where constant C depends on the aspect ratio of any triangle $K \in \mathcal{T}_h$

Proof. We first prove the lower bound (1.56). The identity in Proposition 5 can be rewritten as

$$\begin{aligned} \frac{1}{2}[(t - t^{n+1/2})\tau_{n-1} - (t-t^n)(t-t^{n+1})] \int_\Omega \nabla \partial_n^2 u_h \cdot \nabla v_h \, dx \\ = \int_\Omega \frac{\partial \tilde{e}}{\partial t} v_h \, dx + \int_\Omega \nabla \tilde{e} \cdot \nabla v_h \, dx - \int_\Omega (f - \tilde{f}) v_h \, dx. \end{aligned}$$

This is valid for any $v_h \in V_h^0$. In particular, taking $v_h = \partial_n^2 u_h$ yields

$$\begin{aligned} \frac{1}{4}[(t - t^{n+1/2})\tau_{n-1} - (t-t^n)(t-t^{n+1})]^2 ||\nabla \partial_n^2 u_h||_{L^2(\Omega)}^2 \\ \leq 3 \left(\left\| \frac{\partial \tilde{e}}{\partial t} \right\|_{H^{-1}(\Omega)}^2 + \|\tilde{e}\|_{H^1(\Omega)}^2 + ||f - \tilde{f}||_{H^{-1}(\Omega)}^2 \right). \end{aligned}$$

Integrating in time from t^n to t^{n+1} we obtain the lower bound (1.56).

Next we prove the lower bound (1.57). Introducing a piecewise linear function $h(x)$ measuring the local meshsize, i.e. for example

$$h(x_i) = \frac{\sum_{K\in\Delta_{x_i}} h_K |K|}{|\Delta_{x_i}|} \text{ at any node } x_i,$$

and the bubble functions b_K on any triangle K we observe

$$\sum_{K\in\mathcal{T}_h} h_K^2 \left\| f_h + \Delta u_{h\tau} - \partial_n u_h \right\|_{L^2(K)}^2 \leqslant C \sum_{K\in\mathcal{T}_h} \int_K (f_h + \Delta u_{h\tau} - \partial_n u_h) \, v_h h^2,$$

with $v_h = (f_h + \Delta u_{h\tau} - \partial_n u_h)\, b_K$ on K. Then, assuming for brevity $f_h = f = -\Delta u + \dfrac{\partial_t u}{\partial t}$, we get by integration by parts

$$\sum_{K\in\mathcal{T}_h} h_K^2 \, \|f_h + \Delta u_{h\tau} - \partial_n u_h\|^2_{L^2(K)} \leqslant C \sum_{K\in\mathcal{T}_h} \left(\int_K \nabla\tilde{e}\cdot\nabla(v_h h^2) + \int_K \frac{\partial\tilde{e}}{\partial t} v_h h^2 \right).$$

Let w be the solution to

$$-\Delta w = \frac{\partial\tilde{e}}{\partial t} \quad \text{on}\,\Omega, \qquad w = 0 \quad \text{on}\,\partial\Omega,$$

so that $|w|_{H^1(\Omega)} = \left|\dfrac{\partial\tilde{e}}{\partial t}\right|_{H^{-1}(\Omega)}$ and

$$\begin{aligned}\sum_{K\in\mathcal{T}_h} \int_K \frac{\partial\tilde{e}}{\partial t} v_h h^2 = \int_\Omega (-\Delta w) v_h h^2 = \int_\Omega \nabla w\cdot\nabla(v_h h^2) \\ \leqslant \left\|\frac{\partial\tilde{e}}{\partial t}\right\|_{H^{-1}(\Omega)} \left(\sum_{K\in\mathcal{T}_h} \|\nabla(v_h h^2)\|^2_{L^2(K)} \right)^{1/2}.\end{aligned}$$

Hence

$$\begin{aligned}\sum_{K\in\mathcal{T}_h} h_K^2 \, \|f_h + \Delta u_{h\tau} - \partial_n u_h\|^2_{L^2(K)} \\ \leqslant C \left(|\tilde{e}|_{H^1(\Omega)} + \left\|\frac{\partial\tilde{e}}{\partial t}\right\|_{H^{-1}(\Omega)} \right) \left(\sum_{K\in\mathcal{T}_h} \|\nabla(v_h h^2)\|^2_{L^2(K)} \right)^{1/2}.\end{aligned}$$

We have on any K

$$\max_x h^2(x) \leqslant C h_K^2 \quad \text{and} \quad \max_x |\nabla h|^2(x) \leqslant C h_K,$$

which implies

$$\begin{aligned}\|\nabla(v_h h^2)\|_{L^2(K)} &\leqslant C(h_K^2 \|\nabla v_h\|_{L^2(K)} + h_K \|v_h\|_{L^2(K)}) \\ &\leqslant C h_K \|v_h\|_{L^2(K)} \leqslant C h_K \|f_h + \Delta u_{h\tau} - \partial_n u_h\|_{L^2(K)}.\end{aligned}$$

Finally,

$$\begin{aligned}\sum_{K\in\mathcal{T}_h} h_K^2 \|f_h + \Delta u_{h\tau} - \partial_t u_{h\tau}\|^2_{L^2(K)} \\ \leqslant C \left(|\tilde{e}|_{H^1(\Omega)} + \left\|\frac{\partial\tilde{e}}{\partial t}\right\|_{H^{-1}(\Omega)} \right) \left(\sum_{K\in\mathcal{T}_h} h_K^2 \|f_h + \Delta u_{h\tau} - \partial_n u_h\|^2_{L^2(K)} \right)^{1/2},\end{aligned}$$

or

$$\sum_{K\in\mathcal{T}_h} h_K^2 \|f_h + \Delta u_{h\tau} - \partial_t u_{h\tau}\|^2_{L^2(K)} \leqslant C \left(|\tilde{e}|^2_{H^1(\Omega)} + \left\|\frac{\partial\tilde{e}}{\partial t}\right\|_{H^{-1}(\Omega)} \right).$$

The part of the estimator with jump residuals is handled similarly. □

Anisotropic error estimate

Lemma 9. *Suppose that the mesh $\mathcal{T}_h$ is such that there exists a constant c independent of the mesh size and aspect ration such that*

$$\lambda_{1,K}^2 \left(\boldsymbol{r}_{1,K}^T G_K(\tilde{e}) \boldsymbol{r}_{1,K}\right) \leq c\lambda_{2,K}^2 \left(\boldsymbol{r}_{2,K}^T G_K(\tilde{e}) \boldsymbol{r}_{2,K}\right) \quad \forall K \in \mathcal{T}_h,$$

where G_K is defined by (1.15). Then there is a constant C independent of the mesh size and aspect ratio such that

$$\int_{t^1}^{T} \|\nabla e\|_{L^2(\Omega)}^2 dt + \|\nabla e(\cdot,T)\|_{L^2(\Omega)}^2 \leq \|\nabla e(\cdot,t^1)\|_{L^2(\Omega)}^2 + C \sum_{n=1}^{N-1} \Bigg\{ \sum_{K\in\mathcal{T}_h} (\eta_{K,n}^{S,A})^2 + (\eta_n^T)^2 + \int_{t^n}^{t^{n+1}} \left\| f - \tilde{f} \right\|_{H^{-1}(\Omega)}^2 dt + \sum_{K\in\mathcal{T}_h} \frac{\lambda_{2,K}^2 \tau_n^3}{12} \left\| \partial_n^2 u_h \right\|_{L^2(K)}^2 \Bigg\}, \quad (1.58)$$

where the local anisotropic space error estimator $\eta_{K,n}^{S,A}$ is defined by

$$(\eta_{K,n}^{S,A})^2 = \int_{t^n}^{t^{n+1}} \left(\|f - \partial_n u_h + \Delta \tilde{u}_{h\tau}\|_{L^2(K)} + \frac{1}{2\lambda_{2,K}^{1/2}} \| \left[\boldsymbol{n} \cdot \nabla \tilde{u}_{h\tau}\right] \|_{L^2(\partial K)} \right) \omega_K(\tilde{e}) dt,$$

where ω_K is defined by (1.14). The time error estimator η_n^T is defined by (1.55).

Proof. Follows from Proposition 9 and the interpolation estimates of Proposition 3. □

Remark 2. *The second term in the error estimate of Lemma 9 corresponds to anisotropic residual-based space error estimator. The last three terms are used to estimate the error due to the time discretization.*

Remark 3. *The estimate of Lemma 9 is not a traditional a posteriori estimate since it involves the gradient of the exact solution u. A way to approximate the gradient of u by a computable quantity was explained in subsection 1.1.3. The resulting error estimator was proved very efficient for the adaptive algorithm for the Crank-Nicolson scheme [LPP09].*

1.3 The wave equation

We consider initial boundary-value problem for the wave equation. Let $u = u(x,t) : \Omega \times [0,T] \to \mathbb{R}$ be the solution to

$$\begin{cases} \dfrac{\partial^2 u}{\partial t^2} - \Delta u = f, & \text{in } \Omega \times]0,T], \\ u = 0, & \text{on } \partial\Omega \times]0,T], \\ u(\cdot,0) = u_0, & \text{in } \Omega, \\ \dfrac{\partial u}{\partial t}(\cdot,0) = v_0, & \text{in } \Omega, \end{cases} \quad (1.59)$$

where f, u_0, v_0 are given functions. Note that if we introduce the auxiliary unknown $v = \frac{\partial u}{\partial t}$ then model (1.59) can be rewritten as the following first-order in time system

$$\begin{cases} \dfrac{\partial u}{\partial t} - v = 0, & \text{in } \Omega \times]0,T], \\ \dfrac{\partial v}{\partial t} - \Delta u = f, & \text{in } \Omega \times]0,T], \\ u = 0, & \text{on } \partial\Omega \times]0,T], \\ u(\cdot,0) = u_0,\ v(\cdot,0) = v_0, & \text{in } \Omega. \end{cases} \tag{1.60}$$

The possibility to rewrite the problem (1.59) as a first-order in time system (1.60) is crucial for deriving the time error estimates in [BS05; Geo+16]. The above problem (1.59) has the following weak formulation [Eva10]: find a function

$$\begin{aligned} u \in L^2\left(0,T;H_0^1(\Omega)\right) \text{ such that } \frac{\partial u}{\partial t} \in L^2\left(0,T;L^2(\Omega)\right), \\ \frac{\partial^2 u}{\partial t^2} \in L^2\left(0,T;H^{-1}(\Omega)\right), \text{ with given } f \in L^2(0,T;L^2(\Omega)), \\ u_0 \in H_0^1(\Omega),\ v_0 \in L^2(\Omega) \end{aligned} \tag{1.61}$$

and

$$\left\langle \frac{\partial^2 u}{\partial t^2}, \varphi \right\rangle + (\nabla u, \nabla \varphi) = (f, \varphi), \ \forall \varphi \in H_0^1(\Omega), \tag{1.62}$$

where $\langle \cdot, \cdot \rangle$ denotes the duality pairing between $H^{-1}(\Omega)$ and $H_0^1(\Omega)$ and the parentheses $(\cdot, \cdot)$ stand for the inner product in $L^2(\Omega)$. Moreover, the following regularity result is proved in Chap. 7, Sect. 2, Theorem 5 of [Eva10]

Theorem 10. *Assume that function u (1.61) is the solution of problem (1.62). Then in fact*

$$u \in C^0\left(0,T;H_0^1(\Omega)\right), \ \frac{\partial u}{\partial t} \in C^0\left(0,T;L^2(\Omega)\right), \ \frac{\partial^2 u}{\partial t^2} \in C^0\left(0,T;H^{-1}(\Omega)\right).$$

Let us now discretize (1.59) or, equivalently, (1.60) in space using the finite element method. We use the standard finite element space $V_h \subset H_0^1(\Omega)$ defined in (1.2). Then, the finite element discretization of (1.62) is to find $u_h \in H^2(0,T;V_h)$ such that

$$\int_\Omega \frac{\partial^2 u_h}{\partial t^2} \varphi_h + \int_\Omega \nabla u_h \cdot \nabla \varphi_h = \int_\Omega f \varphi_h, \quad \forall \varphi_h \in V_h, \tag{1.63}$$

and $u_h(0) = \hat{I}_h u_0$, $\dfrac{\partial u_h}{\partial t}(0) = \hat{I}_h v_0$ where $\hat{I}_h$ is some interpolant.

In the sequel we shall work with several time discretizations that will be specified later: the backward Euler scheme [BS05], the Cosine-type scheme [Geo+16] and the Newmark scheme with the particular choice for the parameters, namely $\beta = 1/4$, $\gamma = 1/2$ [BW76].

1.3.1 Time adaptivity for the wave equation discretized by the backward Euler scheme

The aim of this section is to present *a posteriori* error estimates for the wave equation discretized by the backward Euler scheme in time [BS05]. At first we give the main idea of obtaining error estimates in time on the example of ordinary differential equation of second order in time. At the second stage we present the main result of the work [BS05] for the fully discrete problem discretized in space by the finite element method. For the sake of brevity, we restrict the presentation to the case of non-changeable meshes. We finish the subsection by presenting two new alternative error estimates in space and time. The first error estimate is comparable with Bernardi-Süli error estimator and based on the same technique with some slight modification, meanwhile the second error estimator is of optimal order in space.

A posteriori error estimate for a second order ordinary differential equation

Let us consider first the following ordinary differential equation

$$\begin{cases} \dfrac{d^2u(t)}{dt^2} + Au(t) = f(t), \quad t \in [0;T]\,, \\ u(0) = u_0, \\ \dfrac{du}{dt}(0) = v_0, \end{cases} \tag{1.64}$$

with a constant $A > 0$. Like in [BS05], we restrict ourselves to zero right-hand side function $f = 0$. This problem serves as simplification of the wave equation in which we get rid of the space variable.

Note that if we introduce the auxiliary unknown $v = \dfrac{du}{dt}$ then (1.64) can be rewritten as the first order system. Indeed, using the vector notation $\boldsymbol{U}(t) = \begin{pmatrix} u(t) \\ v(t) \end{pmatrix}$ we obtain

$$\begin{cases} \dfrac{d}{dt}\boldsymbol{U} + \begin{pmatrix} 0 & -1 \\ A & 0 \end{pmatrix} \boldsymbol{U} = 0, \quad t \in]0;T]\,, \\ \boldsymbol{U}(0) = \begin{pmatrix} u_0 \\ v_0 \end{pmatrix}. \end{cases} \tag{1.65}$$

Applying Euler's backward scheme to system (1.65) with $\boldsymbol{U}^n = \begin{pmatrix} u^n \\ v^n \end{pmatrix}$ where $v^n = \dfrac{u^n - u^{n-1}}{\tau_n}$ for $n = 1, \ldots, N$ we obtain the following discrete problem

$$\begin{cases} \dfrac{\boldsymbol{U}^{n+1} - \boldsymbol{U}^n}{\tau_n} + \begin{pmatrix} 0 & -1 \\ A & 0 \end{pmatrix} \boldsymbol{U}^{n+1} = 0, \quad n = 1, \ldots, N-1, \\ \boldsymbol{U}^0 = \begin{pmatrix} u_0 \\ v_0 \end{pmatrix}. \end{cases} \tag{1.66}$$

Note that Euler's backward scheme for system (1.64) consists of finding a family $(u^n)_{0\leq n\leq N}$ such that

$$\begin{cases} \dfrac{u^{n+1}-u^n}{\tau_n} - \dfrac{u^n-u^{n-1}}{\tau_{n-1}} + A\tau_n u^{n+1} = 0, \quad n = 1,\dots,N-1, \\ \dfrac{u^1-u^0}{\tau_0} = A\tau_0 u^1 + v_0, \\ u^0 = u_0. \end{cases} \tag{1.67}$$

It is easy to show that this system coincides with problem (1.66).

We shall measure the error in the following norm

$$e = \max_{0\leq n\leq N} \left(\frac{1}{A}\left|v^n - u'(t_n)\right|^2 + |u^n - u(t_n)|^2 \right)^{1/2}.$$

We start from *a priori* error estimate that follows from stability properties of Euler's backward scheme (1.67).

Lemma 11. *The following a priori estimate holds between the solution u of problem (1.64) and the solution $(u^n)_{0\leq n\leq N}$ of problem (1.67), for all $t_{n+1}, n = 0,\dots,N-1$:*

$$\frac{1}{A}\left|v^{n+1} - u'(t_{n+1})\right|^2 + \left|u^{n+1} - u(t_{n+1})\right|^2 \leq 2|\tau^2| \left(\int_0^{t_{n+1}} \left(\frac{1}{A}\left|v''(t)\right| + \left|u''(t)\right| \right) \right)^2. \tag{1.68}$$

where $|\tau|$ is the maximum of the $\tau_n,\ n = 0,\dots,N-1$.

Proof. We first note that on taking the inner product of the first line of system (1.66) with the vector $\begin{pmatrix} u^{n+1} \\ \frac{1}{A}v^{n+1} \end{pmatrix}$ leads to the stability estimate:

$$\frac{1}{A}\left|v^{n+1}\right|^2 + \left|u^{n+1}\right|^2 \leq 2\left(\frac{1}{A}\left|v^1\right|^2 + \left|u^1\right|^2 \right). \tag{1.69}$$

The *a priori* error estimate relies on (1.69). Indeed, the same technique allows us to obtain easily the following estimate:

$$\begin{aligned} \frac{1}{A}\left|v^{n+1} - u'(t_{n+1})\right|^2 + \left|u^{n+1} - u(t_{n+1})\right|^2 &\leq 2\left(\frac{1}{A}\left|v^1 - u'(t_1)\right|^2 + \left|u^1 - u(t_1)\right|^2 \right) \\ &+ 2\Big(\sum_{m=1}^{n} \Big(\frac{1}{A}\left| \frac{v(t_{n+1}) - v(t_n)}{\tau_n} - v'(t_{n+1}) \right| \\ &+ \left| \frac{u(t_{n+1}) - u(t_n)}{\tau_n} - u'(t_{n+1}) \right| \Big) \Big)^2. \end{aligned} \tag{1.70}$$

Using Taylor's formula we obtain

$$\frac{u(t_{n+1})-u(t_n)}{\tau_n}-u'(t_{n+1})=-\frac{1}{\tau_n}\int_{t_n}^{t_{n+1}}u''(s)(s-t_n)ds. \tag{1.71}$$

A similar argument is used to evaluate the other terms on the right-hand side of the inequality (1.70). This leads to the desired estimate. □

The next step is to derive *a posteriori* bound on the error.

Lemma 12. *The following a posteriori error estimate holds between the solution u of problem (1.64) and the solution $(u^n)_{0\leq n\leq N}$ of problem (1.67), for all $t_{n+1}, n=0,\ldots,N-1$:*

$$\frac{1}{A}\left|v^{n+1}-u'(t_{n+1})\right|+\left|u^{n+1}-u(t_{n+1})\right|\leq\frac{1}{2}\sum_{m=0}^{n}\eta_m + \frac{1}{A}\left|v^0-u'(t_0)\right|+\left|u^0-u(t_0)\right|, \tag{1.72}$$

where for each n, $n=1,\ldots,N-1$, the refinement indicator η_n is defined by

$$\eta_n=\tau_n A^{1/2}\left|u^{n+1}-u^n\right|+\tau_n\left|\frac{u^{n+1}-u^n}{\tau_n}-\frac{u^n-u^{n-1}}{\tau_{n-1}}\right|. \tag{1.73}$$

Proof. At first we introduce continuous piecewise linear reconstruction in time for $t\in[t_n,t_{n+1}]$ for $n=0,\ldots,N-1$ by

$$u_\tau(t)=u^{n+1}+(t-t^{n+1})\frac{u^{n+1}-u^n}{\tau_n}. \tag{1.74}$$

Similarly, we introduce piecewise linear reconstruction in time $v_\tau(t)$ based on v^n, $n=0,\ldots,N$.

The *a posteriori* analysis relies on an appropriate residual equation for the linear reconstruction $\boldsymbol{U}_\tau=\begin{pmatrix}u_\tau\\ v_\tau\end{pmatrix}$. We have

$$\begin{cases}\dfrac{d}{dt}\boldsymbol{U}_\tau+\begin{pmatrix}0&-1\\ A&0\end{pmatrix}\boldsymbol{U}_\tau=\begin{pmatrix}R_u\\ R_v\end{pmatrix}, \quad t\in\,]0;T]\,,\\ \boldsymbol{E}(0)=\mathbf{0}.\end{cases} \tag{1.75}$$

where the residual quantities R_u and R_v for $t\in[t_n;t_{n+1}]$ and $n=0,\ldots,N-1$ are defined by

$$\begin{aligned}R_u(t)&=v_\tau-v^{n+1},\\ R_v(t)&=A(u^{n+1}-u_\tau).\end{aligned}$$

Thus, introducing the error between reconstruction $\boldsymbol{U}_\tau$ and solution $\boldsymbol{U}$ to problem (1.65) :

$$\boldsymbol{E}=\boldsymbol{U}-\boldsymbol{U}_\tau=\begin{pmatrix}E_u\\ E_v\end{pmatrix}=\begin{pmatrix}u-u_\tau\\ v-v_\tau\end{pmatrix}, \tag{1.76}$$

denoting

$$Z(t) = \left(\frac{1}{A} |E_v|^2 + |E_u|^2 \right)^{1/2}$$

and taking the inner product of (1.75) with $\begin{pmatrix} E_u \\ \frac{1}{A} E_v \end{pmatrix}$ we obtain

$$\frac{1}{2} \frac{dZ^2}{dt} = R_u E_u + \frac{1}{A} R_v E_v \leq \left(|R_u|^2 + \frac{1}{A} |R_v|^2 \right)^{1/2} Z,$$

whence

$$\frac{dZ}{dt} \leq \left(|R_u|^2 + \frac{1}{A} |R_v|^2 \right)^{1/2} \leq |R_u| + \frac{1}{A^{1/2}} |R_v| .$$

We integrate the last inequality from t_1 to t_{n+1} and obtain

$$Z(t_{n+1}) \leq Z(t_1) + \int_0^{t_{n+1}} \left(|R_u| + \frac{1}{A^{1/2}} |R_v| \right) ds. \tag{1.77}$$

Next, to bound the first integral above we observe that

$$\begin{aligned} \frac{1}{A^{1/2}} \int_{t_m}^{t_{m+1}} |R_v| ds &= A^{1/2} \int_{t_m}^{t_{m+1}} |u^{m+1} - u_\tau| \\ &= A^{1/2} |u^{m+1} - u^m| \int_{t_m}^{t_{m+1}} \frac{t_{m+1} - s}{\tau_m} ds \\ &= \frac{\tau_m}{2} A^{1/2} |u^{m+1} - u^m|, \quad t \in [t_m, t_{m+1}] , \ m = 1, \ldots, n. \end{aligned}$$

Similar argument is used to evaluate the second integral in (1.77)

$$\begin{aligned} \int_{t_m}^{t_{m+1}} |R_u| ds &= \int_{t_m}^{t_{m+1}} |v^{m+1} - v_\tau| = |v^{m+1} - v^m| \int_{t_m}^{t_{m+1}} \frac{t_{m+1} - s}{\tau_m} ds \\ &= \frac{\tau_m}{2} \left| \frac{u^{m+1} - u^m}{\tau_m} - \frac{u^m - u^{m-1}}{\tau_{m-1}} \right|, \quad t \in [t_m, t_{m+1}] , \ m = 1, \ldots, n. \end{aligned}$$

Combining these bounds leads to the desired result. □

Remark 4. *Comparing the a priori error estimate (1.68) with the a posteriori one (1.72) one sees that the error estimator is essentially the same in both cases. Indeed, the a priori error estimate contains the second and the third derivatives from the solution of problem (1.64). Their discrete counterparts in the a posteriori error estimate (1.72) are the second and the third discrete derivatives respectively.*

Remark 5. *To the best of our knowledge, there is no known upper bound for the refinement error estimator $\sum_{n=0}^{N-1} \eta_n$ from Lemma 12. Indeed, the upper bound from ([BS05]) is derived*

when the error is measured in absolutely different norm, namely

$$\hat{e} = \sum_{n=0}^{N-1} \left(\frac{1}{A} \left| v^n - u'(t_n) \right|^2 + |u^n - u(t_n)|^2 + A^{1/2} \left| \int_{t^n}^{t^{n+1}} (u - u_\tau)(s) ds \right| + \left| \int_{t^n}^{t^{n+1}} (v - v_\tau)(s) ds \right| \right)^{1/2}. \quad (1.78)$$

***A posteriori* error estimates in space and time**

We start from the fully discretized problem of initial boundary-value problem for the wave equation (1.59). The Euler backward method consist in seeking $u_h^n \in V_h$ for $n = 1, \ldots, N$ such that

$$\begin{cases} \left(\dfrac{u_h^{n+1} - u_h^n}{\tau_n} - \dfrac{u_h^n - u_h^{n-1}}{\tau_{n-1}}, \varphi_h \right) + \tau_n \left(\nabla u_h^{n+1}, \nabla \varphi_h \right) = 0, \\ \qquad\qquad \forall \varphi_h \in V_h, \quad n = 1, \ldots, N-1, \\ \left(\dfrac{u_h^1 - u_h^0}{\tau_0}, \varphi_h \right) = \tau_0 \left(\nabla u_h^0, \nabla \varphi_h \right) + (v_h^0, \varphi_h), \quad \forall \varphi_h \in V_h, \end{cases} \quad (1.79)$$

and u_h^0, v_h^0 are some approximations to u_0, v_0.

We now present the *a posteriori* bound for the fully discretized problem (1.79) following [BS05]. The *a posteriori* error estimate admits the decoupling of the error committed in the temporal and spatial discretization.

We suppose that the time discretization is such that the time regularity parameter

$$\sigma_\tau = \max_{1 \le m \le N-1} \frac{\tau_m}{\tau_{m-1}},$$

is bounded. The *a posteriori* error estimate in time is based on the technique presented above.

Local error indicator associated with the spatial discretization defined by

$$\eta_{n,K} = h_K \left\| \frac{u_h^{n+1} - u_h^n}{\tau_n} - \frac{u_h^n - u_h^{n-1}}{\tau_{n-1}} \right\|_{L^2(K)} + \frac{\tau_n}{2} \sum_{E \in E_K} h_E^{1/2} \| \left[\boldsymbol{n} \cdot \nabla u_h^{n+1} \right] \|_{L^2(E)}. \quad (1.80)$$

Note that the space error estimator is similar to the standard residual-based bound for elliptic equation, see Proposition 2.

Lemma 13. *Let u be the solution of problem (1.59) and $(u_h^n)_{0 \le n \le N}$ be the discrete solution given by backward Euler scheme (1.79). Then the following a posteriori error estimate holds*

for all t_{n+1}, $n = 1, \ldots, N-1$:

$$\left\| v_h^{n+1} - \frac{\partial u}{\partial t}(t_{n+1}) \right\|_{H^{-1}(\Omega)} + \left\| u_h^{n+1} - u(t_{n+1}) \right\|_{L^2(\Omega)} \tag{1.81}$$
$$\leq c \left(\sum_{m=1}^{n} \left(\eta_m + (1+\sigma_\tau)(n-m) \left(\sum_{K \in \mathcal{T}_h} \eta_{m,K}^2 \right)^{1/2} \right) + (1+\sigma_\tau) n D_{\tau h} \right),$$

where the term $D_{\tau h}$, given by

$$D_{\tau h} = \tau_0 \left\| \nabla u_0 \right\|_{L^2(\Omega)} + \tau_0^2 \left\| \nabla v_0 \right\|_{L^2(\Omega)} + \left\| u_0 - u_h^0 \right\|_{L^2(\Omega)} + \tau_0 \left\| v_0 - v_h^0 \right\|_{L^2(\Omega)} + \left\| v_0 - v_h^0 \right\|_{H^{-1}(\Omega)},$$

only depends on the data.

Remark 6. *From the estimate (1.81) it is clear that there is a loss of optimality with respect to τ: the loss grows linearly with n, where n is the counter of the time level t_n. Moreover, it's certainly not optimal with respect to space. Indeed, the error*

$$e = \left\| v_h^{n+1} - \frac{\partial u}{\partial t}(t_n) \right\|_{H^{-1}(\Omega)} + \left\| u_h^{n+1} - u(t_{n+1}) \right\|_{L^2(\Omega)},$$

is of order h^2 while the estimator (1.81) gives only order h.

Remark 7. *In Chapter 2 we shall present a posteriori error estimates for the wave equation discretized in time by the Newmark method for the error measured in the alternative norm*

$$u \mapsto \max_{t \in [0,T]} \left(\left\| \frac{\partial u}{\partial t}(t) \right\|_{L^2(\Omega)}^2 + |u(t)|_{H^1(\Omega)}^2 \right)^{1/2}.$$

The motivation of our choice is to avoid in practice the computation of H^{-1}-norm of the error. Nevertheless, the a posteriori error analysis from Chapters 2 can be easily extended for the error measured in the Bernardi-Süli norm

$$u \mapsto \max_{t \in [0,T]} \left(\left\| \frac{\partial u}{\partial t}(t) \right\|_{H^{-1}(\Omega)}^2 + |u(t)|_{L^2(\Omega)}^2 \right)^{1/2}.$$

Alternative *a posteriori* error estimates

We present two alternative space-time error estimates using the approach that was developed for the Newmark time discretization (see Chapters 2, 3).

Theorem 14. *Let u be the solution of problem (1.59) and $(u_h^n)_{0\le n\le N}$ be the discrete solution given by backward Euler scheme (1.79). Then the following a posteriori error estimate holds*

$$\left(\left\|v_h^n - \frac{\partial u}{\partial t}(t_n)\right\|^2_{H^{-1}(\Omega)} + ||u_h^n - u(t_n)||^2_{L^2(\Omega)}\right)^{1/2} \le \left(\|v_h^0 - v_0\|^2_{L^2(\Omega)} + |u_h^0 - u_0|^2_{H^1(\Omega)}\right)^{1/2} + \eta_S(t_N) + \sum_{k=0}^{N-1} \tau_k \eta_T(t_k), \quad (1.82)$$

where the space indicator is defined by

$$\eta_S(t_N) = C_1 \sum_{n=0}^{N} \tau_n \left(\sum_{K\in\mathcal{T}_h} h_K^2 \|\partial_{n+1/2} v_h\|^2_{L^2(K)} + \sum_{E\in\mathcal{E}_h} h_E \left(\|[\boldsymbol{n}\cdot\nabla u_h^n]\|^2_{L^2(E)} + \|[\boldsymbol{n}\cdot\nabla u_h^{n+1}]\|^2_{L^2(E)} \right) \right)^{\frac{1}{2}}, \quad (1.83)$$

and the time error indicator is

$$\eta_T(t_n) = \frac{\tau_n}{2}\left(C_2 |\partial_{n+1/2} u_h|^2_{H^1(\Omega)} + \|\partial_{n+1/2} v_h\|_{L^2(\Omega)} \right)^{1/2}. \quad (1.84)$$

Here $C_1,\ C_2$ are constants depending only on the mesh regularity.

Proof. In the following, we adopt the vector notation $\boldsymbol{U}(t,x) = \begin{pmatrix} u(t,x) \\ v(t,x) \end{pmatrix}$ where $v = \partial u/\partial t$. Then, assuming $f = 0$, we have

$$\left(\frac{\partial \boldsymbol{U}}{\partial t}, \boldsymbol{\Phi}\right) + (\mathcal{A}\boldsymbol{U}, \boldsymbol{\Phi}) = 0, \quad \forall \boldsymbol{\Phi} \in (H_0^1(\Omega))^2, \quad (1.85)$$

where $\mathcal{A} = \begin{pmatrix} 0 & -1 \\ A & 0 \end{pmatrix}$, $A = -\Delta$, i.e. $(Au, v) = (\nabla u, \nabla v)$ and

$$(\boldsymbol{U}, \boldsymbol{\Phi}) = \left(\begin{pmatrix} u \\ v \end{pmatrix}, \begin{pmatrix} \varphi \\ \psi \end{pmatrix} \right) := (u, \varphi) + (v, \psi).$$

Similarly, Euler scheme can be rewritten as

$$\left(\frac{\boldsymbol{U}_h^{n+1} - \boldsymbol{U}_h^n}{\tau_n}, \boldsymbol{\Phi}_h\right) + \left(\mathcal{A}_h \boldsymbol{U}_h^{n+1}, \boldsymbol{\Phi}_h\right) = 0, \quad \forall \boldsymbol{\Phi}_h \in V_h^2, \quad (1.86)$$

where $\boldsymbol{U}_h^n = \begin{pmatrix} u_h^n \\ v_h^n \end{pmatrix}$, $\mathcal{A}_h = \begin{pmatrix} 0 & -1 \\ A_h & 0 \end{pmatrix}$ with $(A_h u_h, v_h) = (\nabla u_h, \nabla v_h)$.

The *a posteriori* analysis relies on an appropriate residual equation for the linear reconstruction $\boldsymbol{U}_{h\tau} = \begin{pmatrix} u_{h\tau} \\ v_{h\tau} \end{pmatrix}$. We define it for $t \in [t_n, t_{n+1}]$, $n = 1, \ldots, N-1$ as

$$\boldsymbol{U}_{h\tau}(t) = \boldsymbol{U}_h^{n+1} + (t - t_{n+1})\partial_{n+1/2}\boldsymbol{U}_h, \tag{1.87}$$

so that we have the following differential equation

$$\left(\frac{\partial \boldsymbol{U}_{h\tau}}{\partial t}, \boldsymbol{\Phi}_h\right) + (\mathcal{A}_h \boldsymbol{U}_{h\tau}, \boldsymbol{\Phi}_h) \\ = \left(\mathcal{A}_h\left(\boldsymbol{U}_{h\tau} - \boldsymbol{U}_h^{n+1}\right), \boldsymbol{\Phi}_h\right) = \left((t - t_{n+1})\mathcal{A}_h\partial_{n+1/2}\boldsymbol{U}_h, \boldsymbol{\Phi}_h\right). \tag{1.88}$$

Introduce the error between reconstruction $\boldsymbol{U}_{h\tau}$ and solution $\boldsymbol{U}$ to problem (1.85) :

$$\boldsymbol{E} = \boldsymbol{U}_{h\tau} - \boldsymbol{U}, \tag{1.89}$$

or, component-wise

$$\boldsymbol{E} = \begin{pmatrix} E_u \\ E_v \end{pmatrix} = \begin{pmatrix} u_{h\tau} - u \\ v_{h\tau} - v \end{pmatrix}.$$

Taking the difference between (1.86) and (1.85) we obtain the residual differential equation for the error valid for $t \in [t_n, t_{n+1}]$, $n = 1, \ldots, N-1$

$$(\partial_t \boldsymbol{E}, \boldsymbol{\Phi}) + (\mathcal{A}\boldsymbol{E}, \boldsymbol{\Phi}) = \left(\frac{\partial \boldsymbol{U}_{h\tau}}{\partial t}, \boldsymbol{\Phi} - \boldsymbol{\Phi}_h\right) + (\mathcal{A}\boldsymbol{U}_{h\tau}, \boldsymbol{\Phi} - \boldsymbol{\Phi}_h) \\ + \left(\frac{\partial \boldsymbol{U}_{h\tau}}{\partial t}, \boldsymbol{\Phi}_h\right) + (\mathcal{A}_h \boldsymbol{U}_{h\tau}, \boldsymbol{\Phi}_h) = \left(\frac{\partial \boldsymbol{U}_{h\tau}}{\partial t}, \boldsymbol{\Phi} - \boldsymbol{\Phi}_h\right) + (\mathcal{A}\boldsymbol{U}_{h\tau}, \boldsymbol{\Phi} - \boldsymbol{\Phi}_h) \\ + \left((t - t_{n+1})\mathcal{A}_h\partial_{n+1/2}\boldsymbol{U}_h, \boldsymbol{\Phi}_h\right), \quad \forall \boldsymbol{\Phi}_h \in V_h^2. \tag{1.90}$$

Now we take $\boldsymbol{\Phi} = \begin{pmatrix} E_u \\ A^{-1}E_v \end{pmatrix}$, $\boldsymbol{\Phi}_h = \begin{pmatrix} P_h E_u \\ I_h A^{-1} E_v \end{pmatrix}$ where P_h is defined in (1.3) and I_h is a Scott-Zhang interpolation operator as before. Noting that $(\mathcal{A}\boldsymbol{E}, \boldsymbol{\Phi}) = 0$ and

$$\left(\frac{\partial u_{h\tau}}{\partial t} - v_{h\tau}, E_u - P_h E_u\right) = 0,$$

we get

$$\left(E_u, \frac{\partial E_u}{\partial t}\right) + \left(\frac{\partial E_v}{\partial t}, A^{-1}E_v\right) = \left(\frac{\partial v_{h\tau}}{\partial t}, (I - I_h)A^{-1}E_v\right) \\ + \left(\nabla u_{h\tau}, \nabla\left((I - I_h)A^{-1}E_v\right)\right) + \left((t - t_{n+1})\nabla\partial_{n+1/2}u_h, \nabla I_h A^{-1}E_v\right) \\ - \left((t - t_{n+1})\partial_{n+1/2}v_h, E_u\right).$$

Integrating in time from 0 to some $t^* \geq 0$ yields

$$\begin{aligned}\frac{1}{2}\left(||E_u||^2_{L^2(\Omega)} + \|E_v\|^2_{H^{-1}(\Omega)}\right)(t^*) &= \frac{1}{2}\left(||E_u||^2_{L^2(\Omega)} + \|E_v\|_{H^{-1}(\Omega)}\right)(0) \\ &+ \int_0^{t^*}\left(\frac{\partial v_{h\tau}}{\partial t}, (I-I_h)A^{-1}E_v\right)dt + \int_0^{t^*}\left(\nabla u_{h\tau}, \nabla\left((I-I_h)A^{-1}E_v\right)\right)dt \\ &+ \sum_{n=0}^{N}\int_{t_n}^{t_{n+1}\wedge t^*}(t-t_{n+1})\left[\left(\nabla\partial_{n+1/2}u_h, \nabla I_h A^{-1}E_v\right) - \left(\partial_{n+1/2}v_h, E_u\right)\right]dt \\ &\coloneqq I + II + III. \quad (1.91)\end{aligned}$$

Let

$$Z(t) = \sqrt{||E_u||^2_{L^2(\Omega)} + \|E_v\|^2_{H^{-1}(\Omega)}},$$

and assume that t^* is the point in time where Z attains its maximum and $t^* \in (t_n, t_{n+1}]$ for some n.

We thus get for the first and second terms in (1.91) by integration by parts

$$\begin{aligned}I + II \leqslant \int_0^{t^*}\sum_{K\in\mathcal{T}_h}\left\|\frac{\partial v_{h\tau}}{\partial t} + \Delta u_{h\tau}\right\|_{L^2(K)}\left\|(I-I_h)A^{-1}E_v\right\|_{L^2(K)}dt \\ + \frac{1}{2}\int_0^{t^*}\sum_{K\in\mathcal{T}_h}\|[\boldsymbol{n}\cdot\nabla u_{h\tau}]\|_{L^2(\partial K)}\left\|(I-I_h)A^{-1}E_v\right\|_{L^2(\partial K)}dt.\end{aligned}$$

By interpolation bounds

$$\left\|(I-I_h)A^{-1}E_v\right\|_{L^2(K)} + {h_K}^{1/2}\left\|(I-I_h)A^{-1}E_v\right\|_{L^2(\partial K)} \leqslant Ch_K\left|A^{-1}E_v\right|_{H^1(\omega_K)},$$

and Cauchy-Schwarz inequality we have

$$\begin{aligned}I + II \leqslant C\int_0^{t^*}\Bigg(\sum_{K\in\mathcal{T}_h}h_K^2\left\|\frac{\partial v_{\tau h}}{\partial t} + \Delta u_{h\tau}\right\|^2_{L^2(K)} \\ + \sum_{E\in\mathcal{E}_h}h_E\|[\boldsymbol{n}\cdot\nabla u_{h\tau}]\|^2_{L^2(E)}\Bigg)^{1/2}\left|A^{-1}E_v\right|_{H^1(\Omega)}dt \\ \leqslant C\sum_{n=0}^{N}\tau_n\Bigg(\sum_{K\in\mathcal{T}_h}h_K^2\|\partial_{n+1/2}v_h\|^2_{L^2(K)} \\ + \sum_{E\in\mathcal{E}_h}h_E\left(\|[\boldsymbol{n}\cdot\nabla u_h^n]\|^2_{L^2(E)} + \|[\boldsymbol{n}\cdot\nabla u_h^{n+1}]\|^2_{L^2(E)}\right)\Bigg)^{1/2}Z(t^*). \quad (1.92)\end{aligned}$$

Indeed, at any time t

$$\left|A^{-1}E_v\right|_{H^1(\Omega)} \leq \|E_v\|_{H^{-1}(\Omega)} \leq Z(t) \leq Z(t^*).$$

We turn now to the third term in (1.91)

$$III \leqslant \sum_{n=0}^{N} \frac{\tau_n^2}{2} \Big[|\partial_{n+1/2} u_h|_{H^1(\Omega)} \max_{t\in[t_n,t_{n+1}]} \left| I_h A^{-1} E_v \right|_{H^1(\Omega)} + \|\partial_{n+1/2} v_h\|_{L^2(\Omega)} \max_{t\in[t_n,t_{n+1}]} \|E_u\|_{L^2(\Omega)} \Big]$$
$$\leq \sum_{n=0}^{N} \frac{\tau_n^2}{2} \left[C|\partial_{n+1/2} u_h|^2_{H^1(\Omega)} + \|\partial_{n+1/2} v_h\|^2_{L^2(\Omega)} \right]^{1/2} Z(t^*). \quad (1.93)$$

We have used here the bounds $\left|I_h A^{-1} E_v\right|_{H^1(\Omega)} \leq C \left|A^{-1} E_v\right|_{H^1(\Omega)} \leq \|E_v\|_{H^{-1}(\Omega)}$. Inserting the estimates for integrals $I + II$ and III into (1.91) we obtain (1.82). □

Remark 8. *The obtained error estimator is sharper then Bernardi-Süli estimator (1.81) in the sense that its space estimator doesn't depend on the counter of the time level n. However it's still not optimal with respect to the error in space.*

Theorem 15. *Let u be the solution of problem (1.59) and $(u_h^n)_{0\leq n\leq N}$ be the discrete solution given by backward Euler scheme (1.79). Then the following a posteriori error estimate holds*

$$\left(\left| v_h^n - \frac{\partial u}{\partial t}(t_n) \right|^2_{H^{-1}(\Omega)} + ||u_h^n - u(t_n)||^2_{L^2(\Omega)} \right)^{1/2}$$
$$\leq (\|v_h^0 - v_0\|^2_{L^2(\Omega)} + |u_h^0 - u_0|^2_{H^1(\Omega)})^{1/2} + \hat{\eta}_S(t_N) + \sum_{k=0}^{N-1} \tau_k \eta_T(t_k), \quad (1.94)$$

where the time error indicator η_T is given by (1.84) and the space indicator is

$$\hat{\eta}_S(t_N) = C_1 \max_{0\leqslant n\leqslant N} \Bigg[\sum_{K\in\mathcal{T}_h} h_K^4 \|\partial_{n+1/2} v_h\|^2_{L^2(K)} + \sum_{E\in\mathcal{E}_h} h_E^3 \left(\|[\boldsymbol{n}\cdot\nabla u_h^n]\|^2_{L^2(E)} + \|[\boldsymbol{n}\cdot\nabla u_h^{n+1}]\|^2_{L^2(E)} \right) \Bigg]^{1/2}$$
$$+ C_2 \sum_{n=0}^{N-1} \tau_n \left[\sum_{E\in\mathcal{E}_h} h_E^3 |[\boldsymbol{n}\cdot\nabla\partial_{n+1/2} u_h]|^2_{L^2(E)} \right]^{1/2}$$
$$+ C_3 \sum_{n=1}^{N} \frac{\tau_n + \tau_{n-1}}{2} \left[\sum_{K\in\mathcal{T}_h} h_K^4 ||\partial_n^2 v_h||^2_{L^2(K)} \right]^{1/2}. \quad (1.95)$$

Here $C_1,\ C_2,\ C_3$ are constants depending only on the mesh regularity.

Proof. We follow the proof of the first theorem until (1.91). We treat term III here as before but propose new bounds for $I + II$. For this, we observe

$$E_v(t) = \partial_t E_u(t) + (t - t_{n+1})\partial_{n+1/2} v_h, \quad (1.96)$$

for $t \in [t_n, t_{n+1}]$. Thus, the sum $I + II$ in (1.91) can be rewritten as

$$\begin{aligned} I + II = \int_0^{t^*} \left(\frac{\partial v_{\tau h}}{\partial t}, (I - I_h)A^{-1}\partial_t E_u \right) dt \\ + \int_0^{t^*} \left(\nabla u_{h\tau}, \nabla \left((I - I_h)A^{-1}\partial_t E_u \right) \right) dt \\ + \sum_{n=0}^{N} \int_{t_n}^{t_{n+1}\wedge t^*} (t - t_{n+1}) \Bigg[\left(\frac{\partial v_{h\tau}}{\partial t}, (I - I_h)A^{-1}\partial_{n+1/2} v_h \right) dt \\ + (\nabla u_{h\tau}, \nabla((I - I_h)A^{-1}\partial_{n+1/2} v_h))dt \Bigg] = I_1 + II_1 + III_1. \quad (1.97) \end{aligned}$$

We now integrate by parts with respect to time in the terms $I_1 + II_1$. Let us do it for the first term:

$$\begin{aligned} I_1 = \sum_{n=0}^{N} \int_{t_n}^{t_{n+1}\wedge t^*} \left(\frac{\partial v_{h\tau}}{\partial t}, \frac{\partial}{\partial t}(E_u - I_h E_u) \right) dt \\ = \left(\frac{\partial v_{h\tau}}{\partial t}, (I - I_h)A^{-1}E_u \right)(t^*) - \left(\frac{\partial v_{\tau h}}{\partial t}, (I - I_h)A^{-1}E_u \right)(0) \\ - \sum_{n=1}^{N} \left(\left[\frac{\partial v_{h\tau}}{\partial t} \right]_{t_n}, (I - I_h)A^{-1}E_u(t_n) \right) \\ - \sum_{n=0}^{N} \int_{t_n}^{t_{n+1}\wedge t^*} \left(\frac{\partial^2 v_{h\tau}}{\partial t^2}, (I - I_h)A^{-1}E_u \right) dt. \quad (1.98) \end{aligned}$$

Here $[\cdot]_{t_n}$ denotes the jump with respect to time, i.e.

$$[w]_{t_n} = \lim_{t \to t_n^+} w(t) - \lim_{t \to t_n^-} w(t).$$

Using the same trick in the other terms we can finally write

$$\begin{aligned} I_1 + II_1 = \left(\frac{\partial v_{h\tau}}{\partial t}, (I - I_h)A^{-1}E_u \right)(t^*) + \left(\nabla u_{h\tau}, \nabla(I - I_h)A^{-1}E_u \right)(t^*) \\ - \left(\frac{\partial v_{h\tau}}{\partial t}, (I - I_h)A^{-1}E_u \right)(0) - \left(\nabla u_{h\tau}, \nabla(I - I_h)A^{-1}E_u \right)(0) \\ - \sum_{n=1}^{N} \left(\tau_n^* \partial_n^2 v_h, (I - I_h)A^{-1}E_u(t_n) \right) \\ - \sum_{n=0}^{N} \int_{t_n}^{t_{n+1}\wedge t^*} \left(\nabla \frac{\partial u_{h\tau}}{\partial t}, \nabla(I - I_h)A^{-1}E_u \right) dt. \end{aligned}$$

We have used here a simple expression for the jump of time of $\partial v_{h\tau}/\partial t$

$$\left[\frac{\partial v_{h\tau}}{\partial t}\right]_{t_n} = \tau_{n-1/2}\partial_n^2 v_h, \tag{1.99}$$

and noted that $u_{h\tau}$ is continuous in time and $\dfrac{\partial^2 v_{h\tau}}{\partial t^2} = 0$ on (t_n, t_{n+1}).

Integration by parts element by element over Ω and interpolation estimates yield

$$\begin{aligned}
&I_1 + II_1 \\
&\leq C\left[\sum_{K\in\mathcal{T}_h} h_K^4 \left|\left|\frac{\partial v_{h\tau}}{\partial t}\right|\right|^2_{L^2(K)} + \sum_{E\in\mathcal{E}_h} h_E^3 |[\boldsymbol{n}\cdot\nabla u_{h\tau}]|^2_{L^2(E)}\right]^{1/2} (t^*)\left|A^{-1}E_u\right|_{H^2(\Omega)}(t^*) \\
&\quad + \left[\sum_{K\in\mathcal{T}_h} h_K^4 \left|\left|\frac{\partial v_{h\tau}}{\partial t}\right|\right|^2_{L^2(K)} + \sum_{E\in\mathcal{E}_h} h_E^3 |[\boldsymbol{n}\cdot\nabla u_{h\tau}]|^2_{L^2(E)}\right]^{1/2} (0)\left|A^{-1}E_u\right|_{H^2(\Omega)}(0) \\
&\qquad + C\sum_{n=1}^{N}\tau_n^*\left[\sum_{K\in\mathcal{T}_h} h_K^4 ||\partial_n^2 v_h||^2_{L^2(K)}\right]^{1/2} \left|A^{-1}E_u\right|_{H^2(\Omega)}(t_n) \\
&\quad + C\sum_{n=0}^{N}\tau_n\left[\sum_{E\in\mathcal{E}_h} h_E^3 |||[\boldsymbol{n}\cdot\nabla\partial_{n+1/2}u_h]|||^2_{L^2(E)}\right]^{1/2} \max_{t\in(t_n,t_{n+1})}\left|A^{-1}E_u\right|_{H^2(\Omega)}(t) \\
&\qquad\qquad \leq C\hat{\eta}_S(t_{N+1})Z(t^*),
\end{aligned}$$

We have used here the bounds

$$\left|A^{-1}E_u\right|_{H^2(\Omega)}(t) \leqslant C||E_u||_{L^2(\Omega)}(t) \leqslant Z(t) \leqslant Z(t^*)$$

for all $t \in [0, t^*]$.

We turn now to the third term in (1.97). Integration by parts element by element gives

$$\begin{aligned}
III_1 &\leqslant \sum_{n=0}^{N}\int_{t_n}^{t_{n+1}\wedge t^*}(t_{n+1}-t)\left[\sum_{K\in\mathcal{T}_h} h_K^4 \left|\left|\frac{\partial v_{h\tau}}{\partial t}\right|\right|^2_{L^2(K)}\right. \\
&\qquad \left. + \sum_{E\in\mathcal{E}_h} h_E^3 |||[\boldsymbol{n}\cdot\nabla u_{h\tau}]|||^2_{L^2(E)}\right]^{1/2} \left|A^{-1}\partial_{n+1/2}v_h\right|_{H^2(\Omega)} dt \\
&\qquad \leq C\hat{\eta}_S(t_{N+1})\sum_{n=0}^{N}\frac{\tau_n^2}{2}\|\partial_{n+1/2}v_h\|_{L^2(\Omega)} \leq C\hat{\eta}_S(t_{N+1})\sum_{n=0}^{N} t_n\eta_T(t_n).
\end{aligned}$$

We have thus obtained

$$Z(t^*)^2 \leqslant C\left(Z(0)+\hat{\eta}_S(t_{N+1})+\sum_{n=0}^{N} t_n\eta_T(t_n)\right)Z(t^*) + C\hat{\eta}_S(t_{N+1})\sum_{n=0}^{N} t_n\eta_T(t_n).$$

This entails

$$Z(t^*) \leqslant C\left(Z(0)+\hat{\eta}_S(t_{N+1})+\sum_{n=0}^{N} t_n\eta_T(t_n)\right).$$

□

Remark 9. *The obtained error estimator has a more complex structure than the original Bernardi-Süli estimator (1.81), but it is of optimal second order with respect to the error in space. Moreover, it doesn't depend on the counter of the time level n.*

1.3.2 *A posteriori* error estimates for the leap-frog method

It is well known that the Euler discretization is diffusive and thus rarely used for the wave equation. In this subsection we show optimal order *a posteriori* error estimates for general cosine-type second order methods [BDS85], [BD89] controlling the time discretization error from [Geo+16]. However this error estimate derived only for the case of constant time step and the possibility of extending this approach to the case of non-constant time step remains yet to be seen.

The *a posteriori* error bound is derived for the error in L^∞-in-time-energy-in-space norm. It is possible to derive it for problem (1.64) with non-zero right-hand side $f(t)$ and positive definite, self-adjoint, linear operator $\mathcal{A}$ instead of a constant A as in [Geo+16] ; nevertheless for the sake of brevity we formulate the error bound for zero right-hand side and constant $A \geq 0$. Following [Geo+16], in two-step cosine method, for $n = 1, \ldots, N-1$, we seek approximations u^{n+1} such that

$$\partial_n^2 u + A\left[q_1 u^{n+1} - 2p_1 u^n + q_1 u^{n-1}\right] = 0, \tag{1.100}$$

and we set $p_1 = q_1 - 1/2$ for second order accuracy. The time step τ is supposed to be constant. The approach is based on the rewriting the scheme as the one-step system on staggered time grids and the using of the appropriate time reconstruction adapted from [AMN06]. Indeed, the cosine methods (1.100) can be reformulated as a system in two staggered grids. We introduce the auxiliary variable

$$v^{n+1/2} = \partial_n u, \quad n = 0, \ldots, N-1 \tag{1.101}$$

and we set $v^{-1/2} = 2v_0 - v^{1/2}$. We define $u^{-1} = u_0 - \tau v^{-1/2}$ and introduce the notation

$$\partial_{n+1/2} v = v^{n+1/2} - v^{n-1/2}, \quad n = 0, \ldots, N-1. \tag{1.102}$$

We shall need the piecewise linear interpolant $\overline{u} : [0,T] \to \mathbb{R}$ of the sequence $(u^{n+1/2})_{-1\leq n\leq N-1}$, the piecewise linear interpolant $\tilde{u} : [-\tau, T] \to \mathbb{R}$ of the sequence $(u^n)_{-1\leq n\leq N}$, the piecewise linear interpolant $\overline{u} : [0, t_{N-1}] \to \mathbb{R}$ of the sequence

$(v^n)_{0\leq n\leq N-1}$ and the piecewise linear interpolant $\tilde{v} : [\tau/2, t_{N-1/2}] \to \mathbb{R}$ of the sequence $(v^{n+1/2})_{-1\leq n\leq N-1}$. In what follows we use the notation

$$\begin{aligned} &u^{n+1/2} = u(t_{n+1/2}),\\ &v^n = v(t_n), \quad n = 0, \ldots, N-1. \end{aligned} \tag{1.103}$$

We also define a piecewise constant midpoint interpolator $\tilde{I}_0$ on $(t_{n-1/2}, t_{n+1/2})$ for $n = 0, \ldots, N-1$ and a piecewise constant midpoint interpolator I_0 on (t_{n-1}, t_n) for $n = 1, \ldots, N-1$. Therefore we are ready to rewrite (1.100) as

$$\begin{aligned} &\frac{d\tilde{u}}{dt} - I_0\overline{v} = R_v,\\ &\frac{d\tilde{v}}{dt} - A\tilde{I}_0\overline{u} = R_u, \end{aligned} \tag{1.104}$$

where

$$\begin{aligned} R_u(t) \mid_{(t_{n-1/2}, t_{n+1/2}]} &= \frac{1-4q_1}{4} A(u^{n+1} - 2u^n + u^{n+1}),\\ R_v(t) \mid_{(t_n, t_{n+1}]} &= -\frac{1}{4}(v^{n+3/2} - 2v^{n+1/2} + v^{n-1/2}). \end{aligned} \tag{1.105}$$

We introduce on each interval $\left(t^{n-1/2}, t^{n+1/2}\right]$, for $n = 0, \ldots, N-1$, the reconstruction $\hat{v}$ of $\tilde{v}$ by

$$\hat{v}(t) := v^{n-1/2} + \int_{t_{n-1/2}}^{t} (-A\overline{u} + R_u)dt, \tag{1.106}$$

and the reconstruction $\hat{u}$ of $\tilde{u}$ by

$$\hat{u}(t) := u^{n-1} + \int_{t_{n-1}}^{t} (\overline{v} + R_v)dt, \tag{1.107}$$

Lemma 16. *Let u be the solution of (1.64). Then the following a posteriori error estimate holds*

$$\begin{aligned} \max_{t\in[0,t^N]} \left(A\left|u-\hat{u}\right|^2 + \left|\frac{du}{dt} - \hat{v}\right|^2 \right) \leq 2 \left(A\left|u-\hat{u}\right|^2 + \left|\frac{du}{dt} - \hat{v}\right|^2 \right)(0)\\ + 4\int_0^{t^N} \left(A\left|R_2\right|^2 + \left|R_1\right|^2 \right) dt, \end{aligned} \tag{1.108}$$

with

$$R_1 = -A(\hat{u} - \overline{u}) - R_u, \tag{1.109}$$

$$R_2 = \hat{v} - \overline{v} - R_v. \tag{1.110}$$

Remark 10. *In Chapter 4 we present a numerical comparison between the time error estimator (1.108) and the new alternative time error estimators proposed in Chapters 2, 3 in the case of the Newmark time discretization with two choices of coefficients: $\beta = 1/4,\ \gamma = 1/2$ and $\beta = 0,\ \gamma = 1/2$.*

1.3.3 *A posteriori* error estimator for the finite element discretization of the wave equation

We now present the space isotropic error estimates in the $L^2(0,T;H^1(\Omega))$ norm for the semi discretized wave equation (1.63) from [Pic10]. The approach is based on the elliptic reconstruction technique introduced in [MN03] for parabolic problems. More precisely, given u_h solution of (1.63) we introduce the elliptic reconstruction $U \in L^2(0,T;H_0^1(\Omega))$ defined by

$$\int_\Omega \frac{\partial^2 u_h}{\partial t^2}\varphi + \int_\Omega \nabla U \cdot \nabla \varphi = \int_\Omega f\varphi \quad \forall \varphi \in H_0^1(\Omega), \tag{1.111}$$

Then, the error $u - u_h$ can be estimated by controlling $u - U$ and $U - u_h$. Next the approach to isotropic finite elements as above is used (see Lemma 1).

Lemma 17. *Let u be the solution of the wave equation (1.62) and u_h be the solution of the semi discretized problem (1.63). There exist some constants C_1, C_2 such that*

$$\int_0^T \int_\Omega |\nabla(u-u_h)|^2 \le C_1 \left(\int_0^T \sum_{K\in\mathcal{T}_h} \eta_{K,1}^2 + \sum_{K\in\mathcal{T}_h} \eta_{K,3}^2 \right) + C_2 \left(\int_0^T \sum_{K\in\mathcal{T}_h} \eta_{K,2}^2 + \sum_{K\in\mathcal{T}_h} \eta_{K,4}^2 \right). \tag{1.112}$$

where $\eta_{K,1}^2$ is defined by

$$\eta_{K,1}^2 = h_K^4 \left\| f - \frac{\partial^2 u_h}{\partial t^2} \right\|_{L^2(K)}^2 + h_K^3 \left\| [\boldsymbol{n} \cdot \nabla u_h] \right\|_{L^2(\partial K)}^2, \tag{1.113}$$

$\eta_{K,2}^2$ is defined by

$$\eta_{K,2}^2 = h_K^4 \left\| \frac{\partial^2 f}{\partial t^2} - \frac{\partial^4 u_h}{\partial t^4} \right\|_{L^2(K)}^2 + h_k^3 \left\| \left[\boldsymbol{n} \cdot \nabla \frac{\partial^2 u_h}{\partial t^2} \right] \right\|_{L^2(\partial K)}^2, \tag{1.114}$$

$\eta_{K,3}^2$ is defined by

$$\eta_{K,3}^2 = h_K^4 \left\| f(0) - \frac{\partial^2 u_h}{\partial t^2}(0) \right\|_{L^2(K)}^2 + h_K^3 \left\| [\boldsymbol{n} \cdot \nabla u_h(0)] \right\|_{L^2(\partial K)}^2 + \frac{h_K^4}{\rho_K^2} |u_h(0)|_{H^2(K)}^2, \tag{1.115}$$

and $\eta_{K,4}^2$ is defined by

$$\eta_{K,4}^2 = h_K^4 \left\| \frac{\partial f}{\partial t}(0) - \frac{\partial^3 u_h}{\partial t^3}(0) \right\|_{L^2(K)}^2 + h_k^3 \left\| \left[\boldsymbol{n} \cdot \nabla \frac{\partial u_h}{\partial t}(0) \right] \right\|_{L^2(\partial K)}^2 + \frac{h_K^4}{\rho_K^2} |v_h(0)|_{H^2(K)}^2. \tag{1.116}$$

We now recall briefly from [Pic10] the main ideas of deriving the *a posteriori* error bound from above. The proof is based on the fact that $u - u_h = u - U + U - u_h$ and thus the error $u - u_h$ cab be estimated by controlling $u - U$ and $U - u_h$. We have

$$\int_0^T \int_\Omega |\nabla(u - u_h)|^2 \leq 2 \int_0^T \int_\Omega |\nabla(u - U)|^2 + 2 \int_0^T \int_\Omega |\nabla(U - u_h)|^2 = I + II. \quad (1.117)$$

For the second term in (1.117) from the definition of the elliptic reconstruction (1.111) and from the scheme (1.63) we have

$$\int_\Omega |\nabla(U - u_h)|^2 = \int_\Omega \left(f - \frac{\partial^2 u_h}{\partial t^2} \right) (U - u_h - \varphi_h) - \int_\Omega \nabla u_h \cdot \nabla(U - u_h - \varphi_h), \quad \forall \varphi_h \in V_h.$$

Thus, using the standard techniques of integration by parts over each triangle K and the interpolation estimates of Proposition 2 it's easy to show that

$$\int_0^T \int_\Omega |\nabla(U - u_h)|^2 \leq C_1 \int_0^T \sum_{K \in \mathcal{T}_h} \eta_{K,1}^2.$$

The first term in (1.117) can be estimated as

$$\int_0^T \int_\Omega |\nabla(u - U)|^2 \leq 2T \int_\Omega \left| \frac{\partial}{\partial t}(u - U)(0) \right|^2 + 2T \int_\Omega |\nabla(u - U)(0)|^2 + T^2 \int_0^T \int_\Omega \left| \frac{\partial^2}{\partial t^2}(U - u_h) \right|^2. \quad (1.118)$$

It follows from the definition of the elliptic reconstruction (1.111) derivated twice with respect to time. Next the technique from [MN03] and isotropic interpolation results are used and thus the constants $C_1,\ C_2$ in the estimate (1.112) depend on the aspect ratio.

Remark 11. *We present also analogous anisotropic error estimate based on the interpolation estimates of Proposition 3.*

Lemma 18. *Let u be the solution of the wave equation (1.62) and u_h be the solution of the semi discretized problem (1.63). There exists a constant C_1 independent of the mesh size and aspect ratio and a constant C_2 such that*

$$\int_0^T \int_\Omega |\nabla(u - u_h)|^2 \leq C_1 \left(\int_0^T \sum_{K \in \mathcal{T}_h} \eta_{K,1}^2 + \sum_{K \in \mathcal{T}_h} \eta_{K,3}^2 \right) + C_2 \left(\int_0^T \sum_{K \in \mathcal{T}_h} \eta_{K,2}^2 + \sum_{K \in \mathcal{T}_h} \eta_{K,4}^2 \right). \quad (1.119)$$

where $\eta^2_{K,1}$ *is defined by*

$$\eta^2_{K,1} = \left(\left\| f - \frac{\partial^2 u_h}{\partial t^2} \right\|_{L^2(K)} + \frac{1}{2\lambda^{1/2}_{2,K}} \| \left[\boldsymbol{n} \cdot \nabla u_h \right] \|_{L^2(\partial K)} \right) \times \omega_K (U - u_h), \quad (1.120)$$

$\eta^2_{K,2}$, $\eta^2_{K,4}$ *are defined by (1.114) and (1.116) respectively. And* $\eta^2_{K,3}$ *is defined by*

$$\begin{aligned} \eta^2_{K,3} = \Bigg(& \left\| f(0) - \frac{\partial^2 u_h}{\partial t^2}(0) \right\|_{L^2(K)} \\ & + \frac{1}{2\lambda^{1/2}_{2,K}} \| \left[\boldsymbol{n} \cdot \nabla u_h(0) \right] \|_{L^2(\partial K)} \Bigg) \times \omega_K ((u - U)(0)) \\ & + \frac{\lambda^4_{1,K}}{\lambda^2_{2,K}} \int_K (\boldsymbol{r}^T_{1,K} H(u_0) \boldsymbol{r}_{1,K})^2 + 2\lambda^2_{1,K} \int_K (\boldsymbol{r}^T_{1,K} H(u_0) \boldsymbol{r}_{2,K})^2 \\ & + \lambda^2_{2,K} \int_K (\boldsymbol{r}^T_{2,K} H(u_0) \boldsymbol{r}_{2,K})^2, \end{aligned} \quad (1.121)$$

Herein above $H(v)$ *is the Hessian matrix*

$$H(v) = \begin{pmatrix} \dfrac{\partial^2 v}{\partial x_1^2} & \dfrac{\partial^2 v}{\partial x_1 \partial x_2} \\ \dfrac{\partial^2 v}{\partial x_1 \partial x_2} & \dfrac{\partial^2 v}{\partial x_2^2} \end{pmatrix}. \quad (1.122)$$

Remark 12. *In practice in [Pic10] only the first term from the right-hand side in (1.112)*

$$\int_0^T \sum_{K \in \mathcal{T}_h} \eta^2_{K,1}$$

is used as error indicator. The terms

$$\int_0^T \sum_{K \in \mathcal{T}_h} \eta^2_{K,2} + \sum_{K \in \mathcal{T}_h} \eta^2_{K,4}$$

are of higher order and the term

$$\sum_{K \in \mathcal{T}_h} \eta^2_{K,3}$$

is neglected due to the incorrect behavior at the numerical tests.

Chapter 2

Time and space *a posteriori* error estimators for the wave equation discretized in time by a second order scheme

---

The motivation of this chapter is to obtain *a posteriori* error estimates of optimal order in time and space for the fully discrete wave equation in energy norm discretized with the Newmark scheme in time (equivalent to a cosine method as presented in [Geo+16]) and with finite elements in space. We adopt the particular choice for the parameters in the Newmark scheme, namely $\beta = 1/4$, $\gamma = 1/2$ [BW76]. This choice of parameters is popular since it provides a conservative method with respect to the energy norm. Another interesting feature of this variant of the method, which is in fact essential for our analysis, is the fact that the method can be reinterpreted as the Crank-Nicolson discretization of the reformulation of the governing equation in the first-order system, as in [Bak76]. We are thus able to use the techniques stemming from *a posteriori* error analysis for the Crank-Nicolson discretization of the heat equation in [LPP09], based on a piecewise quadratic polynomial in time reconstruction of the numerical solution. This leads to optimal *a posteriori* error estimate in time and also allows us to easily recover the estimates in space. The resulting estimates are referred to as the 3-point estimator since our quadratic reconstruction is drawn through the values of the discrete solution at 3 points in time. The reliability of 3-point estimator is proved theoretically for general regular meshes in space and non-uniform meshes in time. It is also illustrated by numerical experiments.

We do not provide a proof of the optimality (efficiency) of our error estimators in space and time. However, we are able to prove that the time estimator is of optimal order at least on sufficiently smooth solutions, quasi-uniform meshes in space and uniform meshes in time. The most interesting finding of this analysis is the crucial importance of the way in which the initial conditions are discretized (elliptic projections): a straightforward discretization, such as the nodal interpolation, may ruin the error estimators while providing quite acceptable numerical solution. Numerical experiments confirm these theoretical findings and demonstrate that our error estimators are of optimal order in space and time, even in situation not accessible to the current theory (non quasi-uniform meshes, not constant time steps). Thus the estimators can be used to construct an adaptive algorithm both in time and in space.

The outline of the chapter is as follows. We present the governing equations, the discretization and *a priori* error estimates in Section 2.1. In Section 2.2, *a posteriori* error estimate is derived and some considerations concerning the optimality of time estimators are given. Numerical results are analyzed in Section 2.4.

Chapter contents

2.1 The Newmark scheme for the wave equation and *a priori* error analysis

We consider initial boundary-value problem for the wave equation (1.59) and its weak formulation (1.62). We discretize (1.62) in space using the finite element method and in time using an appropriate marching scheme. As was already mentioned in Section 1, the finite element discretization of (1.62) is to find $u_h \in H^2(0,t;V_h)$ such that

$$\int_\Omega \frac{\partial^2 u_h}{\partial t^2}\varphi_h + \int_\Omega \nabla u_h \cdot \nabla \varphi_h = \int_\Omega f\varphi_h, \quad \forall \varphi_h \in V_h, \tag{2.1}$$

and $u_h(0) = \hat{I}_h u_0$, $\dfrac{\partial u_h}{\partial t}(0) = \hat{I}_h v_0$ where $\hat{I}_h$ is some interpolant. Let us introduce a subdivision of the time interval $[0,T]$

$$0 = t_0 < t_1 < \cdots < t_N = T,$$

with time steps $\tau_n = t_{n+1} - t_n$ for $n = 0,\dots,N-1$ and $\tau = \max\limits_{0\le n\le N-1} \tau_n$. Following [Bak76], by applying Crank-Nicolson discretization to both equations in (1.60) we get a second order in time scheme. The fully discretized method is as follows: taking $u_h^0, v_h^0 \in V_h$ as some approximations to u_0, v_0 compute $u_h^n, v_h^n \in V_h$ for $n = 0,\dots,N-1$ from the system

$$\frac{u_h^{n+1} - u_h^n}{\tau_n} - \frac{v_h^{n+1} + v_h^n}{2} = 0, \tag{2.2}$$

$$\left(\frac{v_h^{n+1} - v_h^n}{\tau_n}, \varphi_h\right) + \left(\nabla \frac{u_h^{n+1} + u_h^n}{2}, \nabla\varphi_h\right) = \left(\frac{f^{n+1} + f^n}{2}, \varphi_h\right), \quad \forall \varphi_h \in V_h. \tag{2.3}$$

As before, f^n is an abbreviation for $f(\cdot, t_n)$, but not for u_h^n or v_h^n.

Note that we can eliminate v_h^n from (2.2)-(2.3) and rewrite the scheme (2.2)-(2.3) in terms of u_h^n only. This results in the following method: given approximations $u_h^0, v_h^0 \in V_h$ of u_0, v_0 compute $u_h^1 \in V_h$ from

$$\left(\frac{u_h^1 - u_h^0}{\tau_0}, \varphi_h\right) + \left(\nabla\frac{\tau_0(u_h^1 + u_h^0)}{4}, \nabla\varphi_h\right) = \left(v_h^0 + \frac{\tau_0}{4}(f^1 + f^0), \varphi_h\right), \quad \forall \varphi_h \in V_h, \quad (2.4)$$

and then compute $u_h^{n+1} \in V_h$ for $n = 1, \ldots, N-1$ from equation

$$\left(\frac{u_h^{n+1} - u_h^n}{\tau_n} - \frac{u_h^n - u_h^{n-1}}{\tau_{n-1}}, \varphi_h\right) + \left(\nabla\frac{\tau_n(u_h^{n+1} + u_h^n) + \tau_{n-1}(u_h^n + u_h^{n-1})}{4}, \nabla\varphi_h\right) = \left(\frac{\tau_n(f^{n+1} + f^n) + \tau_{n-1}(f^n + f^{n-1})}{4}, \varphi_h\right), \quad \forall \varphi_h \in V_h. \quad (2.5)$$

This equation is derived by multiplying (2.3) by $\tau_n/2$, doing the same at the previous time step, taking the sum of the two results and observing

$$\frac{v_h^{n+1} - v_h^{n-1}}{2} = \frac{v_h^{n+1} - v_h^n}{2} + \frac{v_h^n - v_h^{n-1}}{2} = \frac{u_h^{n+1} - u_h^n}{\tau_n} - \frac{u_h^n - u_h^{n-1}}{\tau_{n-1}},$$

by (2.2).

We have thus recovered the Newmark scheme [New59a; RT83] with coefficients $\beta = \frac{1}{4}, \gamma = \frac{1}{2}$ as applied to the wave equation (1.59). Note that the presentation of this scheme in [New59a] and in the subsequent literature on applications in structural mechanics is a little bit different, but the present form (2.4)-(2.5) can be found, for example, in [RT83]. It is easy to see that for any $u_h^0, v_h^0 \in V_h$, both schemes (2.2)-(2.3) and (2.4)-(2.5) provide the same unique solution $u_h^n, v_h^n \in V_h$ for $n = 1, \ldots, N$. In the case of scheme (2.4)-(2.5), v_h^n can be reconstructed from u_h^n recursively with the formula

$$v_h^{n+1} = 2\frac{u_h^{n+1} - u_h^n}{\tau_n} - v_h^n. \quad (2.6)$$

In this chapter we shall use the following notations

$$u_h^{n+1/2} := \frac{u_h^{n+1} + u_h^n}{2}, \quad \partial_{n+1/2} u_h := \frac{u_h^{n+1} - u_h^n}{\tau_n}, \quad \partial_n u_h := \frac{u_h^{n+1} - u_h^{n-1}}{\tau_n + \tau_{n-1}}, \quad (2.7)$$
$$\partial_n^2 u_h := \frac{1}{\tau_{n-1/2}}\left(\frac{u_h^{n+1} - u_h^n}{\tau_n} - \frac{u_h^n - u_h^{n-1}}{\tau_{n-1}}\right) \text{ with } \tau_{n-1/2} := \frac{\tau_n + \tau_{n-1}}{2}.$$

We apply this notations to all quantities indexed by a superscript, so that, for example, $f^{n+1/2} = (f^{n+1} + f^n)/2$. We also denote $u(x, t_n)$, $v(x, t_n)$ by u^n, v^n so that, for example, $u^{n+1/2} = \left(u^{n+1} + u^n\right)/2 = \left(u(x, t_{n+1}) + u(x, t_n)\right)/2$.

We turn now to *a priori* error analysis for the scheme (2.2)-(2.3). We shall measure the error in the following norm

$$u \mapsto \max_{t\in[0,T]} \left(\left\| \frac{\partial u}{\partial t}(t) \right\|^2_{L^2(\Omega)} + |u(t)|^2_{H^1(\Omega)} \right)^{1/2}. \tag{2.8}$$

Here and in what follows, we use the notations $u(t)$ and $\dfrac{\partial u}{\partial t}(t)$ as a shorthand for, respectively, $u(\cdot,t)$ and $\dfrac{\partial u}{\partial t}(\cdot,t)$. The norms and semi-norms in Sobolev spaces $H^k(\Omega)$ are denoted, respectively, by $\|\cdot\|_{H^k(\Omega)}$ and $|\cdot|_{H^k(\Omega)}$. We call (2.8) the energy norm referring to the underlying physics of the studied phenomenon. Indeed, the first term in (2.8) may be assimilated to the kinetic energy and the second one to the potential energy.

Note that *a priori* error estimates for scheme (2.2)-(2.3) can be found in [Bak76; Dup73; RT83]. We are going to construct *a priori* error estimates following the ideas of [Bak76] but we measure the error in a different norm, namely the energy norm (2.8), and present the estimate in a slightly different manner, foreshadowing the upcoming *a posteriori* estimates.

Theorem 19. *Let u be a smooth solution of the wave equation (1.59) and u_h^n, v_h^n be the discrete solution of the scheme (2.2)-(2.3). If $u_0 \in H^2(\Omega)$, $v_0 \in H^1(\Omega)$ and the approximations to the initial conditions are chosen such that $\|v_h^0 - v_0\|_{L^2(\Omega)} \le Ch|v_0|_{H^1(\Omega)}$ and $|u_h^0 - u_0|_{H^1(\Omega)} \le Ch|u_0|_{H^2(\Omega)}$, then the following a priori error estimate holds*

$$\begin{aligned}
&\max_{0\le n\le N} \left(\left\| v_h^n - \frac{\partial u}{\partial t}(t_n) \right\|^2_{L^2(\Omega)} + |u_h^n - u(t_n)|^2_{H^1(\Omega)} \right)^{1/2} \\
&\qquad \le Ch\left(|v_0|_{H^1(\Omega)} + |u_0|_{H^2(\Omega)}\right) + C\sum_{n=0}^{N-1} \tau_n^2 \left(\int_{t_n}^{t_{n+1}} \left| \frac{\partial^3 u}{\partial t^3} \right|_{H^1(\Omega)} dt \right. \\
&\left. \qquad + \int_{t_n}^{t_{n+1}} \left\| \frac{\partial^4 u}{\partial t^4} \right\|_{L^2(\Omega)} dt \right) + Ch\left(\int_{t_0}^{t_N} \left| \frac{\partial^2 u}{\partial t^2} \right|_{H^1(\Omega)} dt + \sum_{n=0}^{N} \tau_n' \left| \frac{\partial u}{\partial t}(t_n) \right|_{H^2(\Omega)} \right. \\
&\qquad \left. + \left| \frac{\partial u}{\partial t}(t_N) \right|_{H^1(\Omega)} + |u(t_N)|_{H^2(\Omega)} \right),
\end{aligned} \tag{2.9}$$

with a constant $C > 0$ depending only on the regularity of the mesh $\mathcal{T}_h$. We have set here $\tau_n' = \tau_{n-1/2}$ for $1 < n < N-1$ and $\tau_0' = \tau_0$, $\tau_N' = \tau_N$.

Proof. Let us introduce $e_u^n = u_h^n - \Pi_h u^n$ and $e_v^n = v_h^n - I_h v^n$ where Π_h is the H_0^1-orthogonal projection operator defined in (1.3) and I_h is the Scott-Zhang interpolation operator as before.

Observe that for $\varphi_h, \psi_h \in V_h$ the following equations hold

$$\left(\nabla \partial_{n+1/2} e_u, \nabla \varphi_h\right) - \left(\nabla e_v^{n+1/2}, \nabla \varphi_h\right) = -\left(\nabla\left(\partial_{n+1/2} u - I_h v^{n+1/2}\right), \nabla \varphi_h\right), \tag{2.10}$$

$$\left(\partial_{n+1/2} e_v, \psi_h\right) + \left(\nabla e_u^{n+1/2}, \nabla \psi_h\right) = \left(\left(\frac{\partial^2 u}{\partial t^2}\right)^{n+1/2} - I_h\left(\partial_{n+1/2} v\right), \psi_h\right). \tag{2.11}$$

The last equation is a direct consequence of (2.3) together with the governing equation (1.59) evaluated at times t_n and t_{n+1}. In accordance with the conventions above, we have denoted here

$$\left(\frac{\partial^2 u}{\partial t^2}\right)^{n+1/2} := \frac{1}{2}\left(\frac{\partial^2 u}{\partial t^2}(t_n) + \frac{\partial^2 u}{\partial t^2}(t_{n+1})\right).$$

Equation (2.10) is obtained from (2.2) taking the gradient of both sides, multiplying by $\nabla \varphi_h$ and integrating over Ω.

Putting $\varphi_h = e_u^{n+1/2}$ and $\psi_h = e_v^{n+1/2}$ and taking the sum of (2.10)–(2.11) yields

$$\frac{|e_u^{n+1}|^2_{H^1(\Omega)} - |e_u^n|^2_{H^1(\Omega)} + \|e_v^{n+1}\|^2_{L^2(\Omega)} - \|e_v^n\|^2_{L^2(\Omega)}}{2\tau_n} = -\left(\nabla R_1^n, \nabla e_u^{n+1/2}\right) + \left(R_2^n, e_v^{n+1/2}\right), \tag{2.12}$$

with

$$R_1^n = \partial_{n+1/2} u - I_h v^{n+1/2} \text{ and } R_2^n = \left(\frac{\partial^2 u}{\partial t^2}\right)^{n+1/2} - I_h\left(\partial_{n+1/2} v\right).$$

Set

$$E^n = \left(|e_u^n|^2_{H^1(\Omega)} + \|e_v^n\|^2_{L^2(\Omega)}\right)^{1/2},$$

so that equality (2.12) with Cauchy-Schwarz inequality entails

$$\frac{(E^{n+1})^2 - (E^n)^2}{2\tau_n} \le \left(|R_1^n|^2_{H^1(\Omega)} + \|R_2^n\|^2_{L^2(\Omega)}\right)^{1/2} \frac{E^{n+1} + E^n}{2},$$

which implies

$$E^{n+1} - E^n \le \tau_n\left(|R_1^n|_{H^1(\Omega)} + \|R_2^n\|_{L^2(\Omega)}\right).$$

Summing this over n from 0 to $N-1$ gives

$$(|e_u^N|^2_{H^1(\Omega)} + \|e_v^N\|^2_{L^2(\Omega)})^{1/2} \le (|e_u^0|^2_{H^1(\Omega)} + \|e_v^0\|^2_{L^2(\Omega)})^{1/2} + \sum_{n=0}^{N-1} \tau_n(|R_1^n|_{H^1(\Omega)} + \|R_2^n\|_{L^2(\Omega)}). \tag{2.13}$$

We have the following estimates for R_1^n and R_2^n

$$|R_1^n|_{H^1(\Omega)} \leq C\tau_n \int_{t_n}^{t_{n+1}} \left| \frac{\partial^3 u}{\partial t^3} \right|_{H^1(\Omega)} dt + Ch \left(\left| \frac{\partial u}{\partial t}(t^n) \right|_{H^2(\Omega)} + \left| \frac{\partial u}{\partial t}\left(t^{n+1}\right) \right|_{H^2(\Omega)} \right), \tag{2.14}$$

$$\|R_2^n\|_{L^2(\Omega)} \leq C\tau_n \int_{t_n}^{t_{n+1}} \left\| \frac{\partial^4 u}{\partial t^4} \right\|_{L^2(\Omega)} dt + C\frac{h}{\tau_n} \int_{t_n}^{t_{n+1}} \left| \frac{\partial^2 u}{\partial t^2} \right|_{H^1(\Omega)} dt. \tag{2.15}$$

The proof of (2.14)–(2.15) is quite standard, but tedious. For brevity, we provide here only the proof of estimate (2.15): we rewrite the definition of R_2^n recalling that $v = \partial u/\partial t$ and using the Taylor expansion around $t = t_{n+1/2}$ as follows

$$R_2^n = \frac{1}{2}\left(\frac{\partial^2 u}{\partial t^2}(t_{n+1}) + \frac{\partial^2 u}{\partial t^2}(t_n) \right) - \frac{1}{\tau_n}\left(\frac{\partial u}{\partial t}(t_{n+1}) - \frac{\partial u}{\partial t}(t_n) \right)$$
$$+\frac{1}{\tau_n}(I - I_h)\left(\frac{\partial u}{\partial t}(t_{n+1}) - \frac{\partial u}{\partial t}(t_n) \right) = \int_{t_{n+1/2}}^{t_{n+1}} \left(\frac{t_{n+1} - t}{2} - \frac{(t_{n+1} - t)^2}{2\tau_n} \right) \frac{\partial^4 u}{\partial t^4} dt$$
$$- \int_{t_n}^{t_{n+1/2}} \left(\frac{t_n - t}{2} + \frac{(t_n - t)^2}{2\tau_n} \right) \frac{\partial^4 u}{\partial t^4} dt + \frac{1}{\tau_n}(I - I_h) \int_{t_n}^{t_{n+1}} \frac{\partial^2 u}{\partial t^2} dt.$$

Taking the $L^2(\Omega)$ norm on both sides and applying the projection error estimates (1.5) we obtain (2.15).

Substituting (2.14)–(2.15) into (2.13) yields

$$\left(\left|e_u^N\right|_{H^1(\Omega)}^2 + \left\|e_v^N\right\|_{L^2(\Omega)}^2 \right)^{1/2} \leq \left(\left|e_u^0\right|_{H^1(\Omega)}^2 + \left\|e_v^0\right\|_{L^2(\Omega)}^2 \right)^{1/2}$$
$$+ C\sum_{n=0}^{N-1} \tau_n^2 \left(\int_{t_n}^{t_{n+1}} \left| \frac{\partial^3 u}{\partial t^3} \right|_{H^1(\Omega)} dt + \int_{t_n}^{t_{n+1}} \left\| \frac{\partial^4 u}{\partial t^4} \right\|_{L^2(\Omega)} dt \right)$$
$$+ Ch \int_0^{t_N} \left| \frac{\partial^2 u}{\partial t^2} \right|_{H^1(\Omega)} dt + Ch \sum_{n=0}^{N} \tau_n' \left| \frac{\partial u}{\partial t}(t_n) \right|_{H^2(\Omega)}.$$

Applying the triangle inequality and estimate (1.5) in the above inequality we get

$$\left(\left\| v_h^N - \frac{\partial u}{\partial t}(t_N) \right\|_{L^2(\Omega)}^2 + \left|u_h^N - u(t_N)\right|_{H^1(\Omega)}^2 \right)^{1/2} \leq \left(\left|e_u^N\right|_{H^1(\Omega)}^2 + \left\|e_v^N\right\|_{L^2(\Omega)}^2 \right)^{1/2}$$
$$+ \left(\left\| (I - I_h)\frac{\partial u}{\partial t}(t_N) \right\|_{L^2(\Omega)}^2 + \left|(I - \Pi_h)\, u(t_N)\right|_{H^1(\Omega)}^2 \right)^{1/2}. \tag{2.16}$$

which implies (2.9) since we can safely assume that the maximum of the error in (2.9) is attained at the final time t_N (if not, it suffices to redeclare the time where the maximum is attained as t_N). □

Remark 13. *Estimate (2.9) is of order h in space which is due to the the presence of H^1 term*

in the norm in which we measure the error. One sees easily that essentially the proof above gives the estimate of order h^2, multiplied by the norms of the exact solution in more regular spaces, if the target norm is changed to

$$\max_{0\le n\le N}\left\|v_h^n-\frac{\partial u}{\partial t}(t_n)\right\|_{L^2(\Omega)}.$$

One would rely then on the estimate

$$\|v-\Pi_h v\|_{L^2(\Omega)}\le Ch^2|v|_{H^2(\Omega)},$$

for the orthogonal projection error and one would obtain

$$\begin{aligned}\left\|v_h^N-\frac{\partial u}{\partial t}(t_N)\right\|_{L^2(\Omega)}&\le\left\|v_h^0-v_0\right\|^2_{L^2(\Omega)}+Ch^2\left|v_0\right|_{H^2(\Omega)}\\&+\sum_{n=0}^{N-1}\tau_n^2\left(\int_{t_n}^{t_{n+1}}\left|\frac{\partial^3u}{\partial t^3}\right|_{H^1(\Omega)}dt+\int_{t_n}^{t_{n+1}}\left\|\frac{\partial^4u}{\partial t^4}\right\|_{L^2(\Omega)}dt\right)\\&+Ch^2\left(\int_{t_0}^{t_N}\left|\frac{\partial^2u}{\partial t^2}\right|_{H^2(\Omega)}dt+\left|\frac{\partial u}{\partial t}(t_N)\right|_{H^2(\Omega)}\right).\end{aligned}\tag{2.17}$$

2.2 *A posteriori* error estimates for the wave equation in the "energy" norm

Our aim here is to derive *a posteriori* bounds in time and space for the error measured in the norm (2.8) and discuss some considerations about upper bound for *a posteriori* error estimator in time.

2.2.1 The 3-point *a posteriori* error estimator: upper bound for the error

The basic technical tool in deriving time error estimator is the piecewise quadratic (in time) reconstruction of the discrete solution, already used in [LPP09] in a similar context, see Section 1.

Definition 2.2.1. *Let u_h^n be the discrete solution given by the scheme (2.5). Then, the piecewise quadratic reconstruction $\tilde{u}_{h\tau}(t):[0,T]\to V_h$ is constructed as the continuous in time function that is equal on $[t_n,t_{n+1}]$, $n\ge 1$, to the quadratic polynomial in t that coincides with u_h^{n+1} (respectively u_h^n, u_h^{n-1}) at time t_{n+1} (respectively t_n, t_{n-1}). Moreover, $\tilde{u}_{h\tau}(t)$ is defined on $[t_0,t_1]$ as the quadratic polynomial in t that coincides with u_h^2 (respectively u_h^1, u_h^0) at time t_2 (respectively t_1, t_0). Similarly, we introduce piecewise quadratic reconstruction $\tilde{v}_{h\tau}(t):[0,T]\to V_h$ based on v_h^n defined by (2.6) and $\tilde{f}_\tau(t):[0,T]\to L^2(\Omega)$ based on $f(t_n,\cdot)$.*

Our quadratic reconstructions $\tilde{u}_{h\tau}$, $\tilde{v}_{h\tau}$ are thus based on three points in time (normally looking backwards in time, with the exemption of the initial time slab $[t_0,t_1]$). This is why the error estimator derived in the following theorem using Definition 2.2.1 will be referred to as the 3-point estimator.

Theorem 20. *The following a posteriori error estimate holds between the solution u of the wave equation (1.59), or equivalently (1.60), and the discrete solution u_h^n given by (2.4), (2.5) for all $t_n,\ 0 \leq n \leq N$ with v_h^n given by (2.6):*

$$\left(\left\| v_h^n - \frac{\partial u}{\partial t}(t_n) \right\|_{L^2(\Omega)}^2 + |u_h^n - u(t_n)|_{H^1(\Omega)}^2 \right)^{1/2}$$
$$\leq \left(\left\| v_h^0 - v_0 \right\|_{L^2(\Omega)}^2 + \left| u_h^0 - u_0 \right|_{H^1(\Omega)}^2 \right)^{1/2}$$
$$+ \eta_S(t_n) + \sum_{k=1}^{n} \tau_{k-1} \eta_T(t_{k-1}) + \int_0^{t_n} \|f - \tilde{f}_\tau\|_{L^2(\Omega)} dt. \quad (2.18)$$

The error indicator in time for $k = 1, \ldots, N-1$ is

$$\eta_T(t_k) = \left(\frac{1}{12}\tau_k^2 + \frac{1}{8}\tau_{k-1}\tau_k \right) \left(\left| \partial_k^2 v_h \right|_{H^1(\Omega)} + \left\| \partial_k^2 f_h - z_h^k \right\|_{L^2(\Omega)}^2 \right)^{1/2}, \quad (2.19)$$

where z_h^k is such that

$$\left(z_h^k, \varphi_h \right) = (\nabla \partial_k^2 u_h, \nabla \varphi_h), \quad \forall \varphi_h \in V_h, \quad (2.20)$$

and

$$\eta_T(t_0) = \left(\frac{5}{12}\tau_0^2 + \frac{1}{2}\tau_1\tau_0 \right) \left(\left| \partial_1^2 v_h \right|_{H^1(\Omega)} + \left\| \partial_1^2 f_h - z_h^1 \right\|_{L^2(\Omega)}^2 \right)^{1/2}. \quad (2.21)$$

The space indicator is defined by

$$\eta_S(t_n) = C_1 \max_{0 \leqslant t \leqslant t_n} \left[\sum_{K \in \mathcal{T}_h} h_K^2 \left\| \frac{\partial \tilde{v}_{h\tau}}{\partial t} - \Delta \tilde{u}_{h\tau} - f \right\|_{L^2(K)}^2 \right.$$
$$\left. + \sum_{E \in \mathcal{E}_h} h_E \left| [n \cdot \nabla \tilde{u}_{h\tau}] \right|_{L^2(E)}^2 \right]^{1/2}$$
$$+ C_2 \sum_{m=0}^{n-1} \int_{t_m}^{t_{m+1}} \left[\sum_{K \in \mathcal{T}_h} h_K^2 \left\| \frac{\partial^2 \tilde{v}_{h\tau}}{\partial t^2} - \Delta \frac{\partial \tilde{u}_{h\tau}}{\partial t} - \frac{\partial f}{\partial t} \right\|_{L^2(K)}^2 \right.$$
$$\left. + \sum_{E \in \mathcal{E}_h} h_E \left\| \left[n \cdot \nabla \frac{\partial \tilde{u}_{h\tau}}{\partial t} \right] \right\|_{L^2(E)}^2 \right]^{1/2} dt$$
$$+ C_3 \sum_{m=1}^{n-1} \tau_{m-1} \left[\sum_{K \in \mathcal{T}_h} h_K^2 \left\| \partial_m^2 v_h - \partial_{m-1}^2 v_h \right\|_{L^2(K)}^2 \right]^{1/2}. \quad (2.22)$$

here $C_1,\ C_2,\ C_3$ are constants depending only on the mesh regularity, $[\cdot]$ stands for a jump on an edge $E \in \mathcal{E}_h$, and $\tilde{u}_{h\tau}$, $\tilde{v}_{h\tau}$ are given by Definition 2.2.1.

Proof. In the following, we adopt the vector notation

$$U(t,x) = \begin{pmatrix} u(t,x) \\ v(t,x) \end{pmatrix},$$

where $v = \partial u/\partial t$. Note that the first equation in (1.60) implies that

$$\left(\nabla\frac{\partial u}{\partial t}, \nabla\varphi\right) - (\nabla v, \nabla\varphi) = 0, \quad \forall\varphi \in H_0^1(\Omega),$$

by taking its gradient, multiplying it by $\nabla\varphi$ and integrating over Ω. Thus, system (1.60) can be rewritten in the vector notations as

$$b\left(\frac{\partial U}{\partial t}, \Phi\right) + (\mathcal{A}\nabla U, \nabla\Phi) = b(F, \Phi), \quad \forall\Phi \in (H_0^1(\Omega))^2, \tag{2.23}$$

where $\mathcal{A} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, $F = \begin{pmatrix} 0 \\ f \end{pmatrix}$ and

$$b(U, \Phi) = b\left(\begin{pmatrix} u \\ v \end{pmatrix}, \begin{pmatrix} \varphi \\ \psi \end{pmatrix}\right) := (\nabla u, \nabla\varphi) + (v, \psi).$$

Similarly, Newmark scheme (2.2)–(2.3) can be rewritten as

$$b\left(\frac{U_h^{n+1} - U_h^n}{\tau_n}, \Phi_h\right) + \left(\mathcal{A}\nabla\frac{U_h^{n+1} + U_h^n}{2}, \nabla\Phi_h\right) = b\left(F^{n+1/2}, \Phi_h\right), \quad \forall\Phi_h \in V_h^2, \tag{2.24}$$

where $U_h^n = \begin{pmatrix} u_h^n \\ v_h^n \end{pmatrix}$ and $F^{n+1/2} = \begin{pmatrix} 0 \\ f^{n+1/2} \end{pmatrix}$.

The *a posteriori* analysis relies on an appropriate residual equation for the quadratic reconstruction $\tilde{U}_{h\tau} = \begin{pmatrix} \tilde{u}_{h\tau} \\ \tilde{v}_{h\tau} \end{pmatrix}$. From the definition of the quadratic reconstruction $\tilde{U}_{h\tau}$ we have for $t \in [t_n, t_{n+1}]$, $n = 1, \ldots, N-1$

$$\tilde{U}_{h\tau}(t) = U_h^{n+1} + (t - t_{n+1})\partial_{n+1/2}U_h + \frac{1}{2}(t - t_{n+1})(t - t_n)\partial_n^2 U_h, \tag{2.25}$$

so that, after some simplifications,

$$\begin{aligned} b\left(\frac{\partial\tilde{U}_{h\tau}}{\partial t}, \Phi_h\right) + (\mathcal{A}\nabla\tilde{U}_{h\tau}, \nabla\Phi_h) &= b\left((t - t_{n+1/2})\partial_n^2 U_h + F^{n+1/2}, \Phi_h\right) \\ &+ \left((t - t_{n+1/2})\mathcal{A}\nabla\partial_{n+1/2}U_h + \frac{1}{2}(t - t_{n+1})(t - t_n)\mathcal{A}\nabla\partial_n^2 U_h, \nabla\Phi_h\right). \end{aligned} \tag{2.26}$$

Consider now (2.24) at time steps n and $n-1$. Subtracting one from another and dividing by $\tau_{n-1/2}$ yields

$$b\left(\partial_n^2 U_h, \Phi_h\right) + (\mathcal{A}\nabla\partial_n U_h, \nabla\Phi_h) = b\left(\partial_n F, \Phi_h\right),$$

or

$$b\left(\partial_n^2 U_h, \Phi_h\right) + \left(\mathcal{A}\nabla\left(\partial_{n+1/2}U_h - \frac{\tau_{n-1}}{2}\partial_n^2 U_h\right), \nabla\Phi_h\right) = b\left(\partial_n F, \Phi_h\right),$$

so that (2.26) simplifies to

$$b\left(\frac{\partial \tilde{U}_{h\tau}}{\partial t}, \Phi_h\right) + \left(\mathcal{A}\nabla \tilde{U}_{h\tau}, \Phi_h\right) = \left(p_n \mathcal{A}\nabla \partial_n^2 U_h, \nabla \Phi_h\right) + b\left(\left(t - t_{n+1/2}\right)\partial_n F + F^{n+1/2}, \Phi_h\right) = \left(p_n \mathcal{A}\nabla \partial_n^2 U_h, \nabla \Phi_h\right) + b\left(\tilde{F}_\tau - p_n \partial_n^2 F, \Phi_h\right), \quad (2.27)$$

where

$$p_n = \frac{\tau_{n-1}}{2}(t - t_{n+1/2}) + \frac{1}{2}(t - t_{n+1})(t - t_n), \quad (2.28)$$
$$\tilde{F}_\tau(t) = F_h^{n+1} + (t - t_{n+1})\partial_{n+1/2} F + \frac{1}{2}(t - t_{n+1})(t - t_n)\partial_n^2 F.$$

Introduce the error between reconstruction $\tilde{U}_{h\tau}$ and solution U to problem (2.23) :

$$E = \tilde{U}_{h\tau} - U, \quad (2.29)$$

or, component-wise

$$E = \begin{pmatrix} E_u \\ E_v \end{pmatrix} = \begin{pmatrix} \tilde{u}_{h\tau} - u \\ \tilde{v}_{h\tau} - v \end{pmatrix}.$$

Taking the difference between (2.27) and (2.23) we obtain the residual differential equation for the error valid for $t \in [t_n, t_{n+1}]$ with $n = 1, \ldots, N-1$

$$b(\partial_t E, \Phi) + (\mathcal{A}\nabla E, \nabla \Phi) = b\left(\frac{\partial \tilde{U}_{\tau h}}{\partial t} - F, \Phi - \Phi_h\right) + \left(\mathcal{A}\nabla \tilde{U}_{\tau h}, \nabla(\Phi - \Phi_h)\right) + \left(p_n \mathcal{A}\nabla \partial_n^2 U_h, \nabla \Phi_h\right) + b\left(\tilde{F}_\tau - F - p_n \partial_n^2 F, \Phi_h\right), \quad \forall \Phi_h \in V_h^2. \quad (2.30)$$

Now we take $\Phi = E$, $\Phi_h = \begin{pmatrix} \Pi_h E_u \\ I_h E_v \end{pmatrix}$ where Π_h is the H_0^1-orthogonal projection operator defined in (1.4) and I_h is a Scott-Zhang interpolation operator. Noting that $(\mathcal{A}\nabla E, \nabla E) = 0$ and

$$\left(\nabla \frac{\partial \tilde{u}_{h\tau}}{\partial t}, \nabla(E_u - \Pi_h E_u)\right) = \left(\nabla \tilde{v}_{h\tau}, \nabla\left(E_u - \Pi_h E_u\right)\right) = 0.$$

Introducing operator $A_h : V_h \to V_h$ such that

$$(A_h w_h, \varphi_h) = (\nabla w_h, \nabla \varphi_h), \quad \forall \varphi_h \in V_h, \quad (2.31)$$

we get

$$\left(\frac{\partial E_v}{\partial t}, E_v\right) + \left(\nabla E_u, \nabla \frac{\partial E_u}{\partial t}\right) = \left(\frac{\partial \tilde{v}_{\tau h}}{\partial t} - f, E_v - \Pi_h E_v\right) + \left(\nabla \tilde{u}_{\tau h}, \nabla\left(E_v - I_h E_v\right)\right) + \left(p_n\left(A_h \partial_n^2 u_h - \partial_n^2 f_h\right), I_h E_v\right) - \left(p_n \nabla \partial_n^2 v_h, \nabla E_u\right) + \left(\tilde{f}_\tau - f, I_h E_v\right).$$

Note that equation similar to (3.21) also holds for $t \in [t_0, t_1]$

$$b(\partial_t E, \Phi) + (\mathcal{A}\nabla E, \nabla\Phi) = b\left(\frac{\partial \tilde{U}_{\tau h}}{\partial t} - F, \Phi - \Phi_h\right) + \left(\mathcal{A}\nabla\tilde{U}_{\tau h}, \nabla(\Phi - \Phi_h)\right) + \left(p_1\mathcal{A}\nabla\partial_1^2 U_h, \nabla\Phi_h\right) + b\left(\tilde{F}_\tau - F - p_1\partial_1^2 F, \Phi_h\right). \quad (2.32)$$

That follows from the definition of the piecewise quadratic reconstruction $\tilde{u}_{h\tau}(t)$ for $t \in [t_0, t_1]$. Integrating (3.21) and (2.32) in time from 0 to some $t^* \geq t_1$ yields

$$\begin{aligned}
&\frac{1}{2}\left(|E_u|^2_{H^1(\Omega)} + \|E_v\|^2_{L^2(\Omega)}\right)(t^*) = \frac{1}{2}\left(|E_u|^2_{H^1(\Omega)} + \|E_v\|^2_{L^2(\Omega)}\right)(0) \\
&+ \int_0^{t^*}\left(\frac{\partial\tilde{v}_{\tau h}}{\partial t} - f, E_v - I_h E_v\right)dt + \int_0^{t^*}(\nabla\tilde{u}_{\tau h}, \nabla(E_v - I_h E_v))\,dt \\
&+ \int_{t_1}^{t^*}\left[\left(p_n\left(A_h\partial_n^2 u_h - \partial_n^2 f_h\right), I_h E_v\right) - \left(p_n\nabla\partial_n^2 v_h, \nabla E_u\right) + \left(\tilde{f}_\tau - f, I_h E_v\right)\right]dt \\
&+ \int_0^{t_1}\left[\left(p_1\left(A_h\partial_1^2 u_h - \partial_1^2 f_h\right), {}_h\, E_v\right) - \left(p_1\nabla\partial_1^2 v_h, \nabla E_u\right) + \left(\tilde{f}_\tau - f, I_h E_v\right)\right]dt \\
&\qquad := I + II + III + IV. \qquad (2.33)
\end{aligned}$$

Let

$$Z(t) = \sqrt{|E_u|^2_{H^1(\Omega)} + \|E_v\|^2_{L^2(\Omega)}},$$

and assume that t^* is the point in time where Z attains its maximum and $t^* \in (t_n, t_{n+1}]$ for some n. Observe

$$(I - \tilde{I}_h)E_v = (I - \tilde{I}_h)(\tilde{v}_{h\tau} - v) = (I - \tilde{I}_h)\left(\frac{\partial\tilde{u}_{h\tau}}{\partial t} - \frac{\partial u}{\partial t}\right) = \frac{\partial}{\partial t}(I - \tilde{I}_h)E_u,$$

since $(I - \tilde{I}_h)\varphi_h = 0$ for any $\varphi_h \in V_h$. We thus get for the first and second terms in (3.24)

$$I + II = \int_0^{t^*}\left(\frac{\partial\tilde{v}_{\tau h}}{\partial t} - f, \frac{\partial}{\partial t}(E_u - I_h E_u)\right)dt + \int_0^{t^*}\left(\nabla\tilde{u}_{\tau h}, \frac{\partial}{\partial t}\nabla(E_u -_h E_u)\right)dt.$$

We now integrate by parts with respect to time in the two integrals above. Let us do it for the first term:

$$\begin{aligned}\int_0^{t^*} \left(\frac{\partial \tilde{v}_{\tau h}}{\partial t} - f, \frac{\partial}{\partial t}(E_u - I_h E_u)\right) dt \\ = \sum_{m=0}^{n} \int_{t_m}^{\min(t_{m+1},t^*)} \left(\frac{\partial \tilde{v}_{\tau h}}{\partial t} - f, \frac{\partial}{\partial t}(E_u - I_h E_u)\right) dt \\ = \left(\frac{\partial \tilde{v}_{\tau h}}{\partial t} - f, E_u - \tilde{I}_h E_u\right)(t^*) - \sum_{m=1}^{n} \left(\left[\frac{\partial \tilde{v}_{\tau h}}{\partial t}\right]_{t_m}, (E_u - \tilde{I}_h E_u)(t_n)\right) \\ - \sum_{m=0}^{n} \int_{t_m}^{\min(t_{m+1},t^*)} \left(\frac{\partial^2 \tilde{v}_{\tau h}}{\partial t^2} - \frac{\partial f}{\partial t}, E_u - \tilde{I}_h E_u\right) dt.\end{aligned}$$

Here $[\cdot]_{t_n}$ denotes the jump with respect to time, i.e.

$$[w]_{t_n} = \lim_{t \to t_n^+} w(t) - \lim_{t \to t_n^-} w(t).$$

Using the same trick in the other term we can finally write

$$\begin{aligned}I + II = \left(\frac{\partial \tilde{v}_{\tau h}}{\partial t} - f, E_u - \tilde{I}_h E_u\right)(t^*) + (\nabla \tilde{u}_{\tau h}, \nabla(E_u - I_h E_u))(t^*) \\ - \sum_{m=1}^{n} \left(\left[\frac{\partial \tilde{v}_{\tau h}}{\partial t}\right]_{t_m}, (E_u - \tilde{I}_h E_u)(t_n)\right) \\ - \sum_{m=0}^{n} \int_{t_m}^{\min(t_{m+1},t^*)} \left(\frac{\partial^2 \tilde{v}_{\tau h}}{\partial t^2} - \frac{\partial f}{\partial t}, E_u - \tilde{I}_h E_u\right) dt \\ - \sum_{m=0}^{n} \int_{t_m}^{min(t_{m+1},t^*)} \left(\nabla \frac{\partial \tilde{u}_{\tau h}}{\partial t}, \nabla(E_u - I_h E_u)\right) dt. \qquad (2.34)\end{aligned}$$

We have used here a simple expression for the jump of time of $\partial \tilde{v}_{h\tau}/\partial t$

$$\left[\frac{\partial \tilde{v}_{h\tau}}{\partial t}\right]_{t_n} = \tau_{n-1} 2(\partial_n^2 v_h - \partial_{n-1}^2 v_h), \qquad (2.35)$$

and noted that $\tilde{u}_{h\tau}$ is continuous in time.

Integration by parts element by element over Ω and interpolation estimates (1.8) yield

$$
\begin{aligned}
I + II \leq C_1 \Bigg[\sum_{K \in \mathcal{T}_h} h_K^2 \left\| \frac{\partial \tilde{v}_{h\tau}}{\partial t} - \Delta \tilde{u}_{h\tau} - f \right\|_{L^2(K)}^2 \\
+ \sum_{E \in \mathcal{E}_h} h_E \left\| [n \cdot \nabla \tilde{u}_{h\tau}] \right\|_{L^2(E)}^2 \Bigg]^{1/2} (t^*) |E_u|_{H^1(\Omega)}(t^*) \\
+ C_1 \Bigg[\sum_{K \in \mathcal{T}_h} h_K^2 \left\| \frac{\partial \tilde{v}_{h\tau}}{\partial t} - \Delta \tilde{u}_{h\tau} - f \right\|_{L^2(K)}^2 \\
+ \sum_{E \in \mathcal{E}_h} h_E \left\| [n \cdot \nabla \tilde{u}_{h\tau}] \right\|_{L^2(E)}^2 \Bigg]^{1/2} (0) |E_u|_{H^1(\Omega)}(0) \\
+ C_2 \sum_{m=1}^{n} \frac{\tau_{m-1}}{2} \left[\sum_{K \in \mathcal{T}_h} h_K^2 \left\| \partial_m^2 v_h - \partial_{m-1}^2 v_h \right\|_{L^2(K)}^2 \right]^{1/2} |E_u|_{H^1(\Omega)}(t_m) \\
+ C_3 \sum_{m=0}^{n} \int_{t_m}^{min(t_{m+1}, t^*)} \Bigg[\sum_{K \in \mathcal{T}_h} h_K^2 \left\| \frac{\partial^2 \tilde{v}_{h\tau}}{\partial t^2} - \Delta \frac{\partial \tilde{u}_{\tau h}}{\partial t} - \frac{\partial f}{\partial t} \right\|_{L^2(K)}^2 \\
+ \sum_{E \in \mathcal{E}_h} h_E \left\| \left[n \cdot \nabla \frac{\partial \tilde{u}_{\tau h}}{\partial t} \right] \right\|_{L^2(E)}^2 \Bigg]^{1/2} (t) |E_u|_{H^1(\Omega)}(t) dt.
\end{aligned}
$$

We turn now to the third term in (3.24)

$$
\begin{aligned}
III &= \int_{t_1}^{t^*} \{ (p_n (A_h \partial_n^2 u_h - \partial_n^2 f_h), I_h E_v) - \left(p_n \nabla \partial_n^2 v_h, \nabla E_u \right) + (\tilde{f}_\tau - f, I_h E_v) \} dt \\
&\leq C \sum_{m=1}^{n} \Bigg[\left(\int_{t_m}^{t_{m+1}} |p_m| dt \right) \left(\left\| \partial_m^2 f_h - A_h \partial_m^2 u_h \right\|_{L^2(\Omega)} + \left| \partial_m^2 v_h \right|_{H^1(\Omega)} \right) \\
&\qquad + \int_{t_m}^{t_{m+1}} \left\| f - \tilde{f}_\tau \right\|_{L^2(\Omega)} dt \Bigg] Z(t^*),
\end{aligned}
$$

with

$$
\int_{t_m}^{t_{m+1}} |p_m| dt \leq \frac{1}{12} \tau_m^3 + \frac{1}{8} \tau_{m-1} \tau_m^2.
$$

We have used here the bounds $|E_u|_{H^1(\Omega)}(t) \leqslant Z(t) \leqslant Z(t^*)$ and $\|E_v\|_{L^2(\Omega)} \leqslant Z(t) \leqslant Z(t^*)$ for all $t \in [0, t^*]$. Similar reasoning for the fourth term in (3.24) give us

$$
\begin{aligned}
IV &= \int_{t_0}^{t_1} \{ (p_1 (A_h \partial_1^2 u_h - \partial_1^2 f_h), I_h E_v) - \left(p_1 \nabla \partial_1^2 v_h, \nabla E_u \right) + (\tilde{f}_\tau - f, I_h E_v) \} dt \\
&\leq C \Bigg[\left(\int_{t_0}^{t_1} |p_1| dt \right) \left(\left\| \partial_1^2 f_h - A_h \partial_1^2 u_h \right\|_{L^2(\Omega)} + \left| \partial_1^2 v_h \right|_{H^1(\Omega)} \right) \\
&\qquad + \int_{t_0}^{t_1} \left\| f - \tilde{f}_\tau \right\|_{L^2(\Omega)} dt \Bigg] Z(t^*),
\end{aligned}
$$

where

$$\int_{t_0}^{t_1} |p_1| dt \leq \frac{5}{12}\tau_0^3 + \frac{1}{2}\tau_1\tau_0^2.$$

Applying the same bounds for $|E_u|_{H^1(\Omega)}(t)$ and $\|E_v\|_{L^2(\Omega)} \leqslant Z(t)$ to the estimates for integrals $I + II$, inserting them into (3.24) and noting that $A_h\partial_k^2 u_h = z_h^k$ we obtain (2.18). □

Remark 14. *Comparing the a priori estimate (2.9) with the a posteriori one (2.18) one sees that the time error indicator is essentially the same in both cases. Indeed, the term*

$$\int_{t_n}^{t_{n+1}} \left\| \frac{\partial^4 u}{\partial t^4} \right\|_{L^2(\Omega)} dt$$

can be rewritten as

$$\int_{t_n}^{t_{n+1}} \left\| \frac{\partial^2 f}{\partial t^2} + \Delta \frac{\partial^2 u}{\partial t^2} \right\|_{L^2(\Omega)} dt$$

and it's discrete counterpart is in 2.19 and 2.21. Note also that the last term in (2.18) is negligible, at least if f the sufficiently smooth in time, since $\|f - \tilde{f}_\tau\|_{L^2(\Omega)} = O(\tau_n^3)$ for $t \in (t_n, t_{n+1})$.

Moreover, in view of a posteriori estimate some of the terms are of the higher order τh^2, so that neglecting the higher order terms, a posteriori space error estimator can be reduced to the two first lines in (2.22), i.e.

$$\eta_S^{(1)}(t_k) = C_1 \max_{0\leqslant t\leqslant t_k} \left[\sum_{K\in\mathcal{T}_h} h_K^2 \left\| \frac{\partial \tilde{v}_{h\tau}}{\partial t} - \Delta \tilde{u}_{h\tau} - f \right\|_{L^2(K)}^2 + \sum_{E\in\mathcal{E}_h} h_E \|[n\cdot\nabla \tilde{u}_{h\tau}]\|_{L^2(E)}^2 \right]^{1/2}(t), \tag{2.36}$$

$$\eta_S^{(2)}(t_k) = C_2 \sum_{m=0}^{k} \int_{t_m}^{t_{m+1}} \left[\sum_{K\in\mathcal{T}_h} h_K^2 \left\| \frac{\partial^2 \tilde{v}_{h\tau}}{\partial t^2} - \Delta \frac{\partial \tilde{u}_{h\tau}}{\partial t} - \frac{\partial f}{\partial t} \right\|_{L^2(K)}^2 + \sum_{E\in\mathcal{E}_h} h_E \left\| \left[n\cdot\nabla \frac{\partial \tilde{u}_{h\tau}}{\partial t} \right] \right\|_{L^2(E)}^2 \right]^{1/2}(t)dt. \tag{2.37}$$

2.2.2 Optimal rate of the 3-point error estimator

We do not have a lower bound for our error estimators in space and time. Note that such a bound is not available even in a simpler setting of Euler discretization in time, cf. [BS05]. We are going to prove a partial result in the direction of optimality, namely that the indicator of error in time provides the estimate of order τ^2 at least on sufficiently smooth solutions. For this, we should examine if the quantities $\partial_n^2 f_h - A_h\partial_n^2 u_h$ and $\partial_n^2 v_h$ remain bounded in L^2 and H^1 norms respectively. This will be achieved in Lemma 23 assuming that the initial conditions are discretized in a specific way, via the H_0^1-orthogonal projection.

We restrict ourselves to the constant time steps $\tau_n = \tau$ and introduce the notations

$$\partial_n^0 u_h = u_h^{n+1}, \qquad \partial_n^{j+1} u_h = \frac{\partial_n^j u_h - \partial_{n-1}^j u}{\tau}, \qquad j = 0, 1, \ldots, \quad n \geq j-1,$$
$$\bar{\partial}_n^0 u_h = \frac{u_h^{n+1} + u_h^n}{2}, \qquad \bar{\partial}_n^{j+1} u_h = \frac{\bar{\partial}_n^j u_h - \bar{\partial}_{n-1}^j u_h}{\tau}, \qquad j = 0, 1, \ldots, \quad n \geq j.$$

The Crank-Nicolson scheme for first-order system (2.2), (2.3) for $n \geq 0$ is written with these notations as

$$\partial_n^1 u_h - \bar{\partial}_n^0 v_h = 0, \tag{2.38}$$
$$\left(\partial_n^1 v_h, \varphi_h\right) + \left(\nabla \bar{\partial}_n^0 u_h, \nabla \varphi_h\right) = \left(\bar{\partial}_n^0 f_h, \varphi_h\right), \quad \forall \varphi_h \in V_h \tag{2.39}$$

where f_h^n, $n \geq 0$, are the L^2-orthogonal projection of $f(t_n, \cdot)$ on V_h. The following lemma provides a higher regularity result on the discrete level, i.e. the boundedness of terms $\partial_n^2 f_h - A_h \partial_n^2 u_h$ and $\partial_n^2 v_h$ for any $j \in \mathbb{N}^0$.

Lemma 21. *Let u_h^n and v_h^n be the solution to (2.2), (2.3) for $n \geq 0$. One has then for all $j \in \mathbb{N}^0$, $N \in \mathbb{N}$, $N \geq j$*

$$\left(\left\| \partial_N^j f_h - A_h \partial_N^j u_h \right\|_{L^2(\Omega)}^2 + \left| \partial_N^j v_h \right|_{H^1(\Omega)}^2 \right)^{1/2} \leq \left(\left\| \partial_j^j f_h - A_h \partial_j^j u_h \right\|_{L^2(\Omega)}^2 + \left| \partial_j^j v_h \right|_{H^1(\Omega)}^2 \right)^{1/2} + \tau \sum_{n=j+1}^{N} \left\| \partial_n^{j+1} f \right\|_{L^2(\Omega)}. \tag{2.40}$$

Proof. Starting from (2.38), (2.39), taking the differences between steps n and $n-1$ and then making an induction on $j = 0, 1, \ldots$ one arrives at

$$\partial_n^{j+1} u_h = \bar{\partial}_n^j v_h, \tag{2.41}$$
$$\partial_n^{j+1} v_h = \bar{\partial}_n^j f_h - A_h \bar{\partial}_n^j u_h. \tag{2.42}$$

One can also prove that $\forall w_h^n \in V_h$ and $j = 0, 1, \ldots$

$$\bar{\partial}_n^j w_h = \frac{\partial_n^j w_h + \partial_{n-1}^j w_h}{2}. \tag{2.43}$$

Indeed, this is obvious for $j = 0$ and then it follows for any j by induction.

Taking the inner product of (2.42) with $\tau A_h \partial_n^{j+1} u_h - \tau \partial_n^{j+1} f_h$, using (2.43) and definition of ∂_n^{j+1} we obtain

$$\left(\partial_n^{j+1} v_h, \tau A_h \partial_n^{j+1} u_h - \tau \partial_n^{j+1} f_h\right) = \left(\bar{\partial}_n^j f_h - A_h \bar{\partial}_n^j u_h, \tau A_h \partial_n^{j+1} u_h - \tau \partial_n^{j+1} f_h\right) = -\frac{\left\| \partial_n^j f_h - A_h \partial_n^j u_h \right\|_{L^2(\Omega)}^2}{2} + \frac{\left\| \partial_{n-1}^j f_h - A_h \partial_{n-1}^j u_h \right\|_{L^2(\Omega)}^2}{2}.$$

Now we apply (2.43) and (2.41) to the left-hand side above

$$\begin{aligned}\left(\partial_n^{j+1} v_h, \tau A_h \partial_n^{j+1} u_h - \tau \partial_n^{j+1} f_h\right) & \\ &= \left(\partial_n^j v_h - \partial_{n-1}^j v_h, A_h \partial_n^{j+1} u_h\right) - \left(\partial_n^{j+1} v_h, \tau \partial_n^{j+1} f_h\right) \\ &= \frac{\left|\partial_n^j v_h\right|^2_{H^1(\Omega)} - \left|\partial_{n-1}^j v_h\right|^2_{H^1(\Omega)}}{2} - \left(\partial_n^{j+1} v_h, \tau \partial_n^{j+1} f_h\right).\end{aligned}$$

Thus

$$\begin{aligned}\frac{\left|\partial_n^j v_h\right|^2_{H^1(\Omega)} - \left|\partial_{n-1}^j v_h\right|^2_{H^1(\Omega)}}{2} - \left(\partial_n^{j+1} v_h, \tau \partial_n^{j+1} f_h\right) &= -\frac{\left\|\partial_n^j f_h - A_h \partial_n^j u_h\right\|^2_{L^2(\Omega)}}{2} \\ &\quad + \frac{\left\|\partial_{n-1}^j f_h - A_h \partial_{n-1}^j u_h\right\|^2_{L^2(\Omega)}}{2}.\end{aligned}$$

We recall by (2.42)

$$\begin{aligned}\tau \partial_n^{j+1} v_h &= \tau \left(\bar\partial_n^j f_h - A_h \bar\partial_n^j u_h\right) \\ &= \frac{\tau}{2}\left(\partial_n^j f_h + \partial_{n-1}^j f_h - A_h \partial_{n-1}^j u_h - A_h \partial_{n-1}^j u_h\right),\end{aligned}$$

and hence

$$\begin{aligned}&\left\|\partial_n^j f_h - A_h \partial_n^j u_h\right\|^2_{L^2(\Omega)} + \left|\partial_n^j v_h\right|^2_{H^1(\Omega)} - \left\|\partial_{n-1}^j f_h - A_h \partial_{n-1}^j u_h\right\|^2_{L^2(\Omega)} - \left|\partial_{n-1}^j v_h\right|^2_{H^1(\Omega)} \\ &\qquad \le \tau \left\|\partial_n^{j+1} f_h\right\|_{L^2(\Omega)} \left(\left\|\partial_n^j f_h - A_h \partial_n^j u_h\right\|_{L^2(\Omega)} + \left\|\partial_{n-1}^j f_h - A_h \partial_{n-1}^j u_h\right\|_{L^2(\Omega)}\right).\end{aligned}$$

Denoting

$$Z_n = \left(\left\|\partial_n^j f_h - A_h \partial_n^j u_h\right\|^2_{L^2(\Omega)} + \left|\partial_n^j v_h\right|^2_{H^1(\Omega)}\right)^{1/2},$$

the last inequality can be rewritten as

$$\begin{aligned}Z_n^2 - Z_{n-1}^2 \le \tau \left\|\partial_n^{j+1} f_h\right\|_{L^2(\Omega)} &\left(\left\|\partial_n^j f_h - A_h \partial_n^j u_h\right\|_{L^2(\Omega)}\right. \\ &\left. + \left\|\partial_{n-1}^j f_h - A_h \partial_{n-1}^j u_h\right\|_{L^2(\Omega)}\right) \le \tau \left\|\partial_n^{j+1} f_h\right\|_{L^2(\Omega)} (Z_n + Z_{n-1}),\end{aligned}$$

so that

$$Z_n - Z_{n-1} \le \tau \left\|\partial_n^{j+1} f_h\right\|_{L^2(\Omega)}.$$

Summing this over n we get (2.40). □

In order to take into account the initial conditions, we shall need the following auxiliary result about stability properties of operator A_h defined by (2.31) and the L^2-orthogonal projection $P_h : L^2(\Omega) \to V_h$ defined by (1.3).

Lemma 22. *Assuming the mesh $\mathcal{T}_h$ to be quasi-uniform, there exist $C > 0$ depending only on the regularity of $\mathcal{T}_h$ such that*

$$\forall v \in H_0^1(\Omega) \quad : \quad |P_h v|_{H^1(\Omega)} \leq C|v|_{H^1(\Omega)}, \tag{2.44}$$

$$\forall v \in H^2(\Omega) \cap H_0^1(\Omega) \quad : \quad \|A_h P_h v\|_{L^2(\Omega)} \leq C|v|_{H^2(\Omega)}. \tag{2.45}$$

Proof. Let $v \in H_0^1(\Omega)$. Using a Scott-Zhang interpolant $I_h : H_0^1(\Omega) \to V_h$, an inverse inequality and the stability properties of I_h, we observe

$$|P_h v|_{H^1(\Omega)} \leq |P_h v - I_h v|_{H^1(\Omega)} + |I_h v|_{H^1(\Omega)} \leq \frac{C}{h}\|P_h v - I_h v\|_{L^2(\Omega)} + |v|_{H^1(\Omega)}.$$

Then, from approximation properties of the Scott-Zhang operator (2) we have

$$\|P_h v - v\|_{L^2(\Omega)} \leq \|I_h v - v\|_{L^2(\Omega)} \leq Ch|v|_{H^1(\Omega)} \leq Ch|v|_{H^1(\Omega)},$$

which entails (2.44). We assume now $v \in H^2(\Omega) \cap H_0^1(\Omega)$ and use a similar idea to prove (2.45)

$$(A_h P_h v, \varphi_h) = (\nabla (P_h - I_h) v, \nabla \varphi_h) + (\nabla I_h v, \nabla \varphi_h). \tag{2.46}$$

We can bound the first term in the right-hand side of (2.46) using the inverse inequality and the approximation properties of I_h

$$(\nabla (P_h - I_h) v, \nabla \varphi_h) \leq \frac{C}{h^2}\|P_h v - I_h v\|_{L^2(\Omega)}\|\varphi_h\|_{L^2(\Omega)} \leq C|v|_{H^2(\Omega)}\|\varphi_h\|_{L^2(\Omega)}.$$

To deal with the second term in the right-hand side of (2.46), we integrate by parts over all the triangles of the mesh and recall that $\Delta\varphi_h = 0$ on any triangle, so that

$$(\nabla I_h v, \nabla \varphi_h) = \sum_{E \in \mathcal{E}_h} \int_E \left[\frac{\partial I_h v}{\partial n}\right] \varphi_h \leq \sum_{E \in \mathcal{E}_h} \left\| \left[\frac{\partial I_h v}{\partial n}\right] \right\|_{L^2(E)} \|\varphi_h\|_{L^2(E)}.$$

Using the standard estimate

$$\|\varphi_h\|_{L^2(E)} \leq \frac{C}{\sqrt{h}}\|\varphi_h\|_{L^2(\omega_E)},$$

the property of Scott-Zhang interpolant combining with scaling

$$\left\| \left[\frac{\partial I_h v}{\partial n}\right] \right\|_{L^2(E)} \leq C\sqrt{h}|v|_{H^2(\omega_E)},$$

for all $E \in \mathcal{E}_h$ and substituting it to (2.46) leads to

$$(A_h P_h v, \varphi_h) \leq C|v|_{H^2(\Omega)}\|\varphi_h\|_{L^2(\Omega)}.$$

Taking here $\varphi_h = A_h P_h v$, we obtain desired result (2.45). □

Remark 15. *Our proof of Lemma 22 uses inverse inequalities and is thus restricted to the quasi-uniform meshes $\mathcal{T}_h$. The first estimate (2.44) is actually established in ([BPS02]) under much milder hypotheses on the mesh compatible with usual mesh refinement techniques.*

We conjecture that the second estimate (2.45) also holds under similar assumptions. Some numerical examples in this direction are given at the end of Subsection 2.4.2.

We are now able to complete the estimate of Lemma 21 in the case $j = 2$ which is pertinent to our *a posteriori* analysis.

Lemma 23. *Let u_h^n be the solution to (2.4), (2.5) on a quasi-uniform mesh with*

$$u_h^0 = \Pi_h u^0,\ v_h^0 = \Pi_h v^0, \tag{2.47}$$

where Π_h is the H_0^1-orthogonal projection on V_h. One has for all $N \geq 1$

$$\begin{aligned}\left(\left\|\partial_N^2 f_h - A_h \partial_N^2 u_h\right\|_{L^2(\Omega)}^2 + \left|\partial_N^2 v_h\right|_{H^1(\Omega)}^2\right)^{1/2} \\ \leq C\left(\left|\frac{\partial^3 u}{\partial t^3}(0)\right|_{H^1(\Omega)} + \left|\frac{\partial^2 u}{\partial t^2}(0)\right|_{H^2(\Omega)} + \max_{t\in[0,2\tau]}\left\|\frac{\partial^2 f}{\partial t^2}(t)\right\|_{L^2(\Omega)}\right) \\ + \int_0^{t_N}\left\|\frac{\partial^3 f}{\partial t^3}\right\|_{L^2(\Omega)} dt,\end{aligned} \tag{2.48}$$

with a constant $C > 0$ independent of h, τ, N.

Proof. Set

$$Z = 2\left(I + \frac{\tau^2}{4}A_h\right)^{-1}\left(I - \frac{\tau^2}{4}A_h\right).$$

Then scheme (2.5) for $n \geq 1$ can be rewritten as

$$u_h^{n+1} = Zu_h^n - u_h^{n-1} + \tau^2\left(I + \frac{\tau^2}{4}A_h\right)^{-1}\bar{f}_h^n.$$

Moreover, the initial step (2.4) can be written as

$$\frac{u_h^1 - u_h^0 - \tau v_h^0}{\tau^2} + A_h\frac{u_h^1 + u_h^0}{4} = \bar{f}_h^0 := \frac{f_h^1 + f_h^0}{4}.$$

This gives the following expressions for u_h^1, u_h^2:

$$\begin{aligned}u_h^1 &= \tau^2\left(I + \frac{\tau^2}{4}A_h\right)^{-1}\left(\bar{f}_h^0 + \frac{1}{\tau}v_h^0\right) + \frac{1}{2}Zu_h^0, \\ u_h^2 &= \tau^2\left(I + \frac{\tau^2}{4}A_h\right)^{-1}\left(Z\left(\bar{f}_h^0 + \frac{1}{\tau}v_h^0\right) + \bar{f}_h^1\right) + \left(\frac{1}{2}Z^2 - I\right)u_h^0.\end{aligned}$$

Thus,

$$\begin{aligned}\partial_1^2 f_h - A_h\partial_1^2 u_h = \partial_1^2 f_h - \frac{A_h^2 Z}{2\left(I + \frac{\tau^2}{4}A_h\right)}u_h^0 \\ - A_h\left(I + \frac{\tau^2}{4}A_h\right)^{-1}\left((Z - 2I)\left(\bar{f}_h^0 + \frac{1}{\tau}v_h^0\right) + \bar{f}_h^1\right),\end{aligned}$$

and

$$\partial_1^2 v_h = -A_h \frac{u_h^2 - u_h^0}{2\tau} + \frac{f_h^2 - f_h^0}{2\tau} = -\frac{A_h}{2\tau}\left(\frac{1}{2}Z^2 - 2I\right)u_h^0 - \frac{A_h}{2\tau}\tau^2\left(I + \frac{\tau^2}{4}A_h\right)^{-1}\left(Z\left(\bar{f}_h^0 + \frac{1}{\tau}v_h^0\right) + \bar{f}_h^1\right) + \frac{f_h^2 - f_h^0}{2\tau}.$$

After some tedious calculations, this can be rewritten as

$$\partial_1^2 f_h - A_h\partial_1^2 u_h = -\frac{1}{2}\frac{Z}{\left(I + \frac{\tau^2}{4}A_h\right)^2}\left(A_h^2 u_h^0 - A_h f_h^0\right) + \frac{\tau A_h}{\left(I + \frac{\tau^2}{4}A_h\right)^2}\left(A_h v_h^0 - \partial_0^1 f_h\right) + \left(I + \frac{\tau^2}{4}A_h\right)^{-1}\partial_1^2 f_h, \quad (2.49)$$

and

$$\partial_1^2 v_h = -\frac{\tau}{\left(I + \frac{\tau^2}{4}A_h\right)^2}\left(A_h^2 u_h^0 - A_h f_h^0\right) + \frac{Z}{2\left(I + \frac{\tau^2}{4}A_h\right)}\left(A_h v_h^0 - \partial_0^1 f_h\right) - \frac{\tau}{2\left(I + \frac{\tau^2}{4}A_h\right)}\partial_1^2 f_h. \quad (2.50)$$

Since A_h is a symmetric positive definite operator, we have

$$\|R(\tau^2 A_h)v_h\|_{L^2(\Omega)} \leq C\|v_h\|_{L^2(\Omega)},$$

for any $v_h \in V_h$ and any rational function R with the degree of nominator less that or equal to that of the denominator and a constant C depending only on R.

Similarly, using the fact

$$|v_h|_{H^1(\Omega)} = (A_h v_h, v_h)^{\frac{1}{2}} = \left\|A_h^{1/2} v_h\right\|_{L^2(\Omega)},$$

for any $v_h \in V_h$, one can observe

$$\|\tau A_h R(\tau^2 A_h)v_h\|_{L^2(\Omega)} \leq C\|A_h^{1/2} v_h\|_{L^2(\Omega)} = C|v_h|_{H^1(\Omega)}.$$

for any rational function R with the degree of nominator less than that of the denominator and a constant C depending only on R.

Applying these estimates to (2.50) yields

$$\|\partial_1^2 f_h - A_h\partial_1^2 u_h\|_{L^2(\Omega)} \leq C\left(\|A_h^2 u_h^0 - A_h f_h^0\|_{L^2(\Omega)} + \left|A_h v_h^0 - \frac{\partial f_h}{\partial t}(0)\right|_{H^1(\Omega)} + \left\|\frac{\tau A_h}{\left(I + \frac{\tau^2}{4}A_h\right)^2}\left(\frac{\partial f_h}{\partial t}(0) - \partial_0^1 f_h\right)\right\|_{L^2(\Omega)} + \|\partial_1^2 f_h\|_{L^2(\Omega)}\right).$$

Since

$$\partial_0^1 f_h = \frac{\partial f_h}{\partial t}(0) + \frac{1}{\tau}\int_0^\tau (\tau - s)\frac{\partial^2 f}{\partial t^2}(s)ds,$$

we have

$$\left\| \frac{\tau A_h}{\left(I + \frac{\tau^2}{4}A_h\right)^2}\left(\frac{\partial f_h}{\partial t}(0) - \partial_0^1 f_h\right)\right\|_{L^2(\Omega)} \leq \max_{t\in[0,\tau]}\left\|\frac{\tau^2 A_h}{\left(I + \frac{\tau^2}{4}A_h\right)^2}\frac{\partial^2 f_h}{\partial t^2}(t)\right\|_{L^2(\Omega)}$$
$$\leq C\max_{t\in[0,\tau]}\left\|\frac{\partial^2 f_h}{\partial t^2}(t)\right\|_{L^2(\Omega)}.$$

Noting finally that $\|\partial_1^2 f_h\|_{L^2(\Omega)}$ can be bounded by the maximum of $\left\|\frac{\partial^2 f}{\partial t^2}(t)\right\|_{L^2(\Omega)}$ over time interval $[0, 2\tau]$, we arrive at

$$\|\partial_1^2 f_h - A_h\partial_1^2 u_h\|_{L^2(\Omega)} \leq C\left(\left\|A_h^2 u_h^0 - A_h f_h^0\right\|_{L^2(\Omega)} + \left|A_h v_h^0 - \frac{\partial f_h}{\partial t}(0)\right|_{H^1(\Omega)}\right.$$
$$\left. + \max_{t\in[0,2\tau]}\left\|\frac{\partial^2 f}{\partial t^2}(t)\right\|_{L^2(\Omega)}\right).$$

By a similar reasoning we can also bound $\left|\partial_1^2 v_h\right|_{H^1(\Omega)}$ by the same quantity as in the right-hand side of the equation above. For this, we take the H^1 norm on both sides of (2.50) and observe for the first term on the right hand side

$$\left|\frac{\tau}{\left(I + \frac{\tau^2}{4}A_h\right)^2}\left(A_h^2 u_h^0 - A_h f_h^0\right)\right|_{H^1(\Omega)} = \left\|\frac{\tau A_h^{1/2}}{\left(I + \frac{\tau^2}{4}A_h\right)^2}\left(A_h^2 u_h^0 - A_h f_h^0\right)\right\|_{L^2(\Omega)}$$
$$\leq C\left\|A_h^2 u_h^0 - A_h f_h^0\right\|_{L^2(\Omega)}.$$

The other terms can be treated similarly so that, skipping some details, we obtain

$$\left(\left\|\partial_1^2 f_h - A_h\partial_1^2 u_h\right\|_{L^2(\Omega)}^2 + \left|\partial_1^2 v_h\right|_{H^1(\Omega)}^2\right)^{1/2} \leq C\left(\left\|A_h^2 u_h^0 - A_h f_h^0\right\|_{L^2(\Omega)}\right.$$
$$\left. + \left|A_h v_h^0 - \frac{\partial f_h}{\partial t}(0)\right|_{H^1(\Omega)} + \max_{t\in[0,2\tau]}\left\|\frac{\partial^2 f}{\partial t^2}(t)\right\|_{L^2(\Omega)}\right). \quad (2.51)$$

We can now invoke the estimate of Lemma 21 with $j = 2$ and combine it with (2.51). This gives

$$\left(\left\|\partial_N^2 f_h - A_h \partial_N^2 u_h\right\|^2_{L^2(\Omega)} + \left|\partial_N^2 v_h\right|^2_{H^1(\Omega)}\right)^{1/2} \leq \sum_{n=3}^{N} \tau \left\|\partial_n^3 f\right\|_{L^2(\Omega)} + C\left(\left\|A_h^2 u_h^0 - A_h f_h^0\right\|_{L^2(\Omega)} + \left|A_h v_h^0 - \frac{\partial f_h}{\partial t}(0)\right|_{H^1(\Omega)} + \max_{t\in[0,\tau]} \left\|\frac{\partial^2 f}{\partial t^2}(t)\right\|_{L^2(\Omega)}\right). \quad (2.52)$$

The first term in right-hand side in (2.52) can be easily bounded by $\int_0^{t_N} \left\|\frac{\partial^3 f}{\partial t^3}\right\|_{L^2(\Omega)} dt$. The remaining terms in the middle line of (2.52) are bounded using Lemma 22 and the relation $A_h \Pi_h = -P_h \Delta$ as follows

$$\left\|A_h^2 u_h^0 - A_h f_h^0\right\|_{L^2(\Omega)} = \left\|A_h P_h(-\Delta u^0 - f^0)\right\|_{L^2(\Omega)} = \left\|A_h P_h \frac{\partial^2 u}{\partial t^2}(0)\right\|_{L^2(\Omega)} \leq C \left|\frac{\partial^2 u}{\partial t^2}(0)\right|_{H^2(\Omega)},$$

and

$$\left|A_h v_h^0 - \frac{\partial f_h}{\partial t}(0)\right|_{H^1(\Omega)} = \left|P_h\left(-\Delta v^0 - \frac{\partial f}{\partial t}(0)\right)\right|_{H^1(\Omega)} \leq \left|P_h \frac{\partial^3 u}{\partial t^3}(0)\right|_{H^1(\Omega)} \leq C \left|\frac{\partial^3 u}{\partial t^3}(0)\right|_{H^1(\Omega)}.$$

This gives (2.48). □

Remark 16. *Note that in Lemma 23 the approximation of initial conditions and right-hand-side function is crucial for boundedness of higher order discrete derivatives and consequently to optimality of our time and space error estimators. We illustrate this fact with some numerical examples in Subsection 2.4.2.*

Corollary 24. *Let u be the solution of wave equation (1.59) and*

$$\frac{\partial^3 u}{\partial t^3}(0) \in H^1(\Omega),\ \frac{\partial^2 u}{\partial t^2}(0) \in H^2(\Omega),$$
$$\frac{\partial^2 f}{\partial t^2}(t) \in L^\infty(0,T;L^2(\Omega)),\ \frac{\partial^3 f}{\partial t^3}(t) \in L^2(0,T;L^2(\Omega)).$$

Suppose that mesh $\mathcal{T}_h$ is quasi-uniform and the mesh in time is uniform ($t_k = k\tau$). Then, the 3-point time error estimator $\eta_T(t_k)$ defined by (2.19), (2.21) is of order τ^2, i.e.

$$\eta_T(t_k) \leq C\tau^2. \quad (2.53)$$

with a positive constant C depending only on u, f, and the mesh regularity.

Proof. Follows immediately from Lemma 23. □

2.3 Anisotropic estimate

The estimate of Theorem 20 can be rewritten for anisotropic case using usual technique from [FP01], [FP03].

Theorem 25. *There exists constant c independent of the mesh size and aspect ratio such that the following a posteriori error estimate holds between the solution u of the wave equation (1.59), or equivalently (1.60), and the discrete solution u_h^n given by (2.4)–(2.5) for all t_n, $0 \leq n \leq N$ with v_h^n given by (2.6):*

$$\begin{aligned}\left\| v_h^n - \frac{\partial u}{\partial t}(t_n) \right\|^2_{L^2(\Omega)} &+ |u_h^n - u(t_n)|^2_{H^1(\Omega)} \\ &\leq \left\| v_h^0 - v_0 \right\|^2_{L^2(\Omega)} + \left| u_h^0 - u_0 \right|^2_{H^1(\Omega)} \\ &+ \eta_{AN}^2(t_n) + \left(\sum_{k=1}^{n} \tau_{k-1} \eta_T(t_{k-1}) \right)^2 + \left(\int_0^{t_n} \|f - \tilde{f}_\tau\|_{L^2(\Omega)} dt \right)^2 . \quad (2.54)\end{aligned}$$

The error indicator in time $\eta_T(t_k)$ is defined by (2.21) for $k = 0$ and by (2.19) for $k = 1, \ldots, N-1$. The anisotropic space error estimator is defined by

$$\begin{aligned}\eta_{AN}^2(t_n) = c \Bigg(& \max_{0 \leqslant t \leqslant t_n} \Bigg[\sum_{K \in \mathcal{T}_h} \Bigg(\left\| \frac{\partial \tilde{v}_{h\tau}}{\partial t} - \Delta \tilde{u}_{h\tau} - f \right\|_{L^2(K)} \\ & \qquad + \frac{1}{2\lambda_{2,K}^{1/2}} \left\| [n \cdot \nabla \tilde{u}_{h\tau}] \right\|_{L^2(\partial K)} \Bigg) \times \omega_K(E_u) \Bigg] \\ & + \sum_{m=0}^{n-1} \int_{t_m}^{t_{m+1}} \Bigg[\sum_{K \in \mathcal{T}_h} \Bigg(\left\| \frac{\partial^2 \tilde{v}_{h\tau}}{\partial t^2} - \Delta \frac{\partial \tilde{u}_{h\tau}}{\partial t} - \frac{\partial f}{\partial t} \right\|_{L^2(K)} \\ & \qquad + \frac{1}{2\lambda_{2,K}^{1/2}} \left\| \left[n \cdot \nabla \frac{\partial \tilde{u}_{h\tau}}{\partial t} \right] \right\|_{L^2(\partial K)} \Bigg) \times \omega_K(E_u) \Bigg] dt \\ & + \sum_{m=1}^{n-1} \tau_{m-1} \sum_{K \in \mathcal{T}_h} \left(\left\| \partial_m^2 v_h - \partial_{m-1}^2 v_h \right\|_{L^2(K)} \times \omega_K \left(E_u(t_m) \right) \right) \Bigg) . \quad (2.55)\end{aligned}$$

Proof. We start by reproducing the proof of Theorem 20 up to equation (3.24).

The first two terms in (3.24) are responsible for the space estimator and can be dealt with the anisotropic estimates from [FP01], [FP03]. Indeed, using Proposition

3 and Cauchy-Schwarz inequality we obtain

$$
\begin{aligned}
I+II = &\left(\frac{\partial \tilde{v}_{\tau h}}{\partial t} - f, E_u - \tilde{I}_h E_u\right)(t^*) + \left(\nabla \tilde{u}_{\tau h}, \nabla(E_u - I_h E_u)\right)(t^*) \\
&- \sum_{m=1}^{n} \left(\left[\frac{\partial \tilde{v}_{\tau h}}{\partial t}\right]_{t_m}, (E_u - \tilde{I}_h E_u)(t_n)\right) \\
&- \sum_{m=0}^{n} \int_{t_m}^{\min(t_{m+1}, t^*)} \left(\frac{\partial^2 \tilde{v}_{\tau h}}{\partial t^2} - \frac{\partial f}{\partial t}, E_u - \tilde{I}_h E_u\right) dt \\
&- \sum_{m=0}^{n} \int_{t_m}^{min(t_{m+1}, t^*)} \left(\nabla \frac{\partial \tilde{u}_{\tau h}}{\partial t}, \nabla(E_u - I_h E_u)\right) dt \leq \eta_{AN}^2(t_n).
\end{aligned}
$$

The third and the fourth term in (2.34) are responsible for the time estimator. We have

$$
\begin{aligned}
III = &\int_{t_1}^{t^*} \Big[\big(p_n(A_h \partial_n^2 u_h - \partial_n^2 f_h), I_h E_v\big) - \big(p_n \nabla \partial_n^2 v_h, \nabla E_u\big) + (\tilde{f}_\tau - f, I_h E_v)\Big] dt \\
&\leq C \sum_{m=1}^{n} \Bigg[\left(\int_{t_m}^{t_{m+1}} |p_m| dt \right) \left(\left\| \partial_m^2 f_h - A_h \partial_m^2 u_h \right\|_{L^2(\Omega)} + \left| \partial_m^2 v_h \right|_{H^1(\Omega)} \right) \\
&\quad + \int_{t_m}^{t_{m+1}} \left\| f - \tilde{f}_\tau \right\|_{L^2(\Omega)} dt \Bigg] Z(t^*) \leq \frac{C}{2p} \Bigg(\left(\sum_{k=1}^{n} \tau_{k-1} \eta_T(t_{k-1}) \right)^2 \\
&\quad + \left(\int_0^{t_n} \|f - \tilde{f}_\tau\|_{L^2(\Omega)} dt \right)^2 \Bigg) + p Z^2(t^*).
\end{aligned}
$$

with p_m is given by (2.28) and thus

$$
\int_{t_m}^{t_{m+1}} |p_m| dt \leq \frac{1}{12} \tau_m^3 + \frac{1}{8} \tau_{m-1} \tau_m^2.
$$

We have used here the stability properties of the Scott-Zhang interpolation operators, see Proposition 4, and the bounds $|E_u|_{H^1(\Omega)}(t) \leqslant Z(t) \leqslant Z(t^*)$ and $\|E_v\|_{L^2(\Omega)} \leqslant Z(t) \leqslant Z(t^*)$ for all $t \in [0, t^*]$. Using similar reasoning for the fourth term in (3.24) and choosing appropriate p we obtain (2.54). □

The estimate of Theorem 25 is not a traditional *a posteriori* error estimate since it involves $\omega_K(E_u)$ and hence the gradient of the exact solution. We have already explained in Chapter 1 a way to approximate the gradient of u by a computable quantity. Note, however, that a slight change in the proof of the preceding theorem could produce an alternative estimate so that E_u no longer appears in the right hand side although it is hidden in the error distribution hypothesis 2.56:

Theorem 26. *Let the mesh $\mathcal{T}_h$ be such that there exists a constant c independent of the mesh size and aspect ratio such that*

$$
\lambda_{1,K}^2 \left(\boldsymbol{r}_{1,K}^T G_K(E_u) \boldsymbol{r}_{1,K} \right) \leq c \lambda_{2,K}^2 \left(\boldsymbol{r}_{2,K}^T G_K(E_u) \boldsymbol{r}_{2,K} \right). \tag{2.56}
$$

Here $\lambda_{i,K}$, $\boldsymbol{r}_{i,K}$ and G_K are defined in Subsection 1.0.2. Then there exists constant C independent of the mesh size and aspect ratio such that the a posteriori error estimate (2.54) holds

between the solution u of the wave equation (1.59), or equivalently (1.60), and the discrete solution u_h^n given by (2.4)–(2.5) for all t_n, $0 \leq n \leq N$ with v_h^n given by (2.6). The error indicator in time $\eta_T(t_k)$ is defined by (2.21) for $k = 0$ and by (2.19) for $k = 1, \ldots, N-1$. The anisotropic space error estimator is defined by

$$\begin{aligned}\eta_{AN}^2(t_n) = C\Bigg(&\max_{0\leqslant t\leqslant t_n}\Bigg[\sum_{K\in\mathcal{T}_h}\Bigg(\lambda_{2,K}^2\left\|\frac{\partial\tilde{v}_{h\tau}}{\partial t}-\Delta\tilde{u}_{h\tau}-f\right\|_{L^2(K)}^2\\ &+\lambda_{2,K}\left\|[n\cdot\nabla\tilde{u}_{h\tau}]\right\|_{L^2(\partial K)}^2\Bigg)\Bigg]\\ &+\sum_{m=0}^{n-1}\int_{t_m}^{t_{m+1}}\Bigg[\sum_{K\in\mathcal{T}_h}\Bigg(\lambda_{2,K}^2\left\|\frac{\partial^2\tilde{v}_{h\tau}}{\partial t^2}-\Delta\frac{\partial\tilde{u}_{h\tau}}{\partial t}-\frac{\partial f}{\partial t}\right\|_{L^2(K)}^2\\ &+\lambda_{2,K}\left\|\left[n\cdot\nabla\frac{\partial\tilde{u}_{h\tau}}{\partial t}\right]\right\|_{L^2(\partial K)}^2\Bigg)\Bigg]dt\\ &+\sum_{m=1}^{n-1}\tau_{m-1}\sum_{K\in\mathcal{T}_h}\lambda_{2,K}^2\left\|\partial_m^2 v_h-\partial_{m-1}^2 v_h\right\|_{L^2(K)}^2\Bigg).\end{aligned} \tag{2.57}$$

Proof. Considering the inequality (2.56) implies that

$$\omega_K(E_u) \leq \tilde{C}\lambda_{2,K}\left\|\nabla E_u\right\|_{L^2(K)} \quad \forall K \in \mathcal{T}_h, \tag{2.58}$$

where $\tilde{C}$ is independent of the mesh size and aspect ratio. Using (2.58) in (2.55) leads to the final result. □

2.4 Numerical results

2.4.1 The 3-point error estimator on structured mesh

We now report numerical results for initial boundary-value problem for wave equation with uniform time steps when using 3-point time error estimator (2.19), (2.21). We compute space estimators (2.36) and (2.37) in practice as follows:

$$\begin{aligned}\eta_S^{(1)}(t_N) &= \max_{1\leq n\leq N-1}\left[\sum_{K\in\mathcal{T}_h}h_K^2\left\|\partial_n v_h-f_h^n\right\|_{L^2(K)}^2+\sum_{E\in\mathcal{E}_h}h_E\|[n\cdot\nabla u_h^n]\|_{L^2(E)}^2\right]^{1/2},\\ \eta_S^{(2)}(t_N) &= \sum_{n=1}^{N-1}\tau_n\Bigg[\sum_{K\in\mathcal{T}_h}h_K^2\left\|\partial_n^2 v_h-\partial_n f_h\right\|_{L^2(K)}^2\\ &\qquad+\sum_{E\in\mathcal{E}_h}h_E\left\|[n\cdot\nabla\partial_n u_h]\right\|_{L^2(E)}^2\Bigg]^{1/2}.\end{aligned} \tag{2.59}$$

The quality of our error estimators in space and time is determined by following effectivity index:

$$ei = \frac{\eta_T+\eta_S}{e}.$$

h	τ	ei	η_T	η_S	$\eta_S^{(1)}$	$\eta_S^{(2)}$	N_0	e
1/160	$\sqrt{h}$	13.74	0.114	0.37	0.12	0.24	97.79	0.035
1/320	$\sqrt{h}$	13.58	0.054	0.18	0.061	0.12	97.59	0.017
1/640	$\sqrt{h}$	13.42	0.026	0.092	0.031	0.062	97.5	0.0088
1/160	h	16.98	0.00062	0.37	0.12	0.24	97.79	0.021
1/320	h	16.97	0.00015	0.18	0.062	0.12	97.59	0.011
1/640	h	16.97	3.82e-05	0.092	0.031	0.062	97.5	0.005

TABLE 2.1: Results for case (a). The quantity N_0 is defined in (2.63) and provided here for future reference.

h	τ	ei	η_T	η_S	$\eta_S^{(1)}$	$\eta_S^{(2)}$	e
1/320	1/20	13.05	2.03	12.15	6.13	6.02	1.09
1/320	1/40	12.11	0.92	12.27	6.15	6.11	1.09
1/320	1/80	11.62	0.37	12.29	6.16	6.13	1.09
1/640	1/20	12.14	0.51	6.09	3.07	3.02	0.54
1/640	1/40	11.68	0.23	6.13	3.08	3.05	0.54
1/640	1/80	11.64	0.096	6.15	3.08	3.07	0.54

TABLE 2.2: Results for case (b).

The true error is

$$e = \max_{0 \leqslant n \leqslant N} \left(\left\| v_h^n - \frac{\partial u}{\partial t}(t_n) \right\|_{L^2(\Omega)}^2 + |u_h^n - u(t_n)|_{H^1(\Omega)}^2 \right)^{1/2}. \tag{2.60}$$

Consider the problem (1.59) with $\Omega = (0,1) \times (0,1)$, $T = 1$ and the exact solution u given by

$$\begin{aligned} &\text{case (a)} \quad u(x,y,t) = \cos(\pi t)\sin(\pi x)\sin(\pi y), \\ &\text{case (b)} \quad u(x,y,t) = \cos(0.5\pi t)\sin(10\pi x)\sin(10\pi y), \\ &\text{case (c)} \quad u(x,y,t) = \cos(15\pi t)\sin(\pi x)\sin(\pi y). \end{aligned}$$

We interpolate initial conditions and right-hand-side function with nodal interpolation. Structured meshes in space (see Fig. 2.1) are used in all the experiments of this section. Numerical results are reported in Tables 2.1–2.3. Note that these cases and the meshes in space in time in the following numerical experiments are chosen so that the error in case (a) should be due to both time and space discretization, that in case (b) comes mainly from the space discretization, and that in case (c) mainly from the time discretization.

Referring to Table 2.1, we observe from first three rows that setting $h = \tau^2$ the error is divided by 2 each time h is divided by 2, consistent with $e \sim O(\tau^2 + h)$. The space error estimator and the time error estimator behave similarly and thus provide a good representation of the true error. The effectivity index tends to a constant value. In rows 4-6, we choose $h = \tau$ in order to insure that the discretization in time

h	τ	ei	η_T	η_S	$\eta_S^{(1)}$	$\eta_S^{(2)}$	e
$1/160$	$1/80$	73.98	55.92	4.17	0.75	3.41	0.81
$1/320$	$1/80$	71.42	55.92	2.08	0.38	1.71	0.81
$1/640$	$1/80$	70.13	55.93	1.04	0.19	0.85	0.81
$1/160$	$1/160$	87.44	14.15	3.78	0.15	3.63	0.21
$1/320$	$1/160$	78.22	14.15	1.89	0.076	1.82	0.21
$1/640$	$1/160$	73.61	14.15	0.95	0.038	0.91	0.21

TABLE 2.3: Results for case (c).

gives an error of higher order than that in space, i.e. $O(h^2)$ vs. $O(h)$, respectively. Our estimators capture well this behavior of the two parts of the error.

In Table 2.2, in order to illustrate the sharpness of the space estimator, we take case (b) where the error is mainly due to the space discretization. We can see from this table that the space error estimator η_S behaves as the true error. Indeed, for a given space step, η_S does not depend on the time step τ, and for constant τ, η_S is divided by two when the space step h is divided by two.

Finally, we consider case (c), Table 2.3. We observe that the time error estimator η_T behaves as the true error, when the error is mainly due to the time discretization.

We therefore conclude that our time and space error estimators are sharp in the regime of constant time steps and structured space meshes. They separate well the two sources of the error and can be thus used for the mesh adaptation in space and time.

Remark 17. *As said already, the space estimator η_S behaves as $O(h)$ in the numerical experiments reported in Tables 2.1-2.2. The situation is slightly different in Table 2.3. Indeed, the first part of space error estimator $\eta_S^{(1)}$ behaves here as $O(\tau^2 h)$. This can be explained by the fact that, as seen from the definitions (2.36), (2.37), both $\eta_S^{(1)}$ and $\eta_S^{(2)}$ are also influenced by discretization in time. In general, in the leading order in h and τ, one can conjecture $\eta_S^{(1,2)} = Ah + B h\tau^2$ with case dependent A and B. The second term $Bh\tau^2$ is asymptotically negligible but it can become visible in some situations where the solution is highly oscillating in time and the mesh in time is not sufficiently refined, as indeed observed with $\eta_S^{(1)}$ in Table 2.3. Fortunately, its value is small compared to the time error estimator and thus we can hope that this effect is not essential for mesh refinement.*

2.4.2 The 3-point error estimator on unstructured mesh

We turn now to the numerical experiments on unstructured Delaunay meshes, cf. Fig. 2.1 (right). These experiments will reveal the dependence of the error estimators on approximation of initial conditions and of the right-hand side f. Indeed, as noted in Subsection 2.2.2, these approximations should be chosen carefully to ensure the optimality of our error estimators.

We consider the test case from the previous subsection with the exact solution u given by case (a). We test two different ways to approximate the initial conditions and the right-hand side: nodal interpolation

$$u_h^0 = I_h u^0,\ v_h^0 = I_h v^0,\ f_h^n = I_h f^n,\ 0 \le n \le N, \tag{2.61}$$

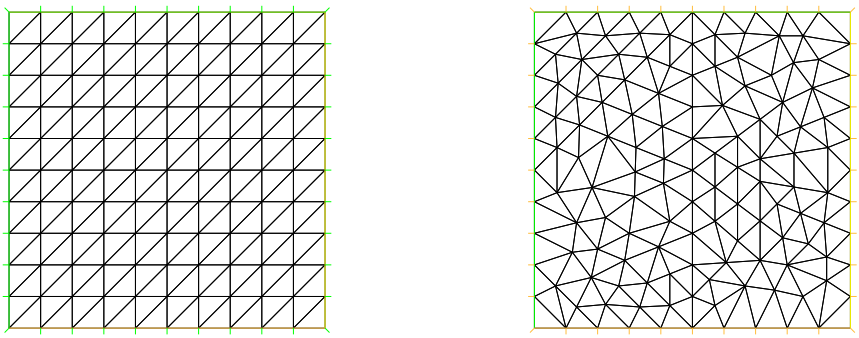


FIGURE 2.1: Structured (on the left) and unstructured (on the right) a 10×10 meshes of the unit square.

h	τ	ei	$\hat{ei}$	η_T	$\hat{\eta}_T$	η_S	$\eta_S^{(1)}$	$\eta_S^{(2)}$	N_0	e
1/160	$\sqrt{h}$	75	29	2.1	.61	.33	.094	.23	934718	.033
1/320	$\sqrt{h}$	120.74	43.83	1.76	.53	.17	.047	.13	3.31e+06	.016
1/640	$\sqrt{h}$	244.56	85.93	1.89	.59	.11	.023	.082	1.44e+07	.0082
1/160	h	196.92	156.38	1.61	.93	1.73	.096	1.63	934718	.017
1/320	h	353.63	281.47	1.43	.83	1.49	.047	1.45	3.31e+06	.088
1/640	h	751.43	598.9	1.54	.9	1.59	.023	1.56	1.44e+07	.0042

TABLE 2.4: Results for case (a), constant time steps, unstructured Delaunay meshes, nodal interpolation of the initial conditions and f as in (2.61).

and orthogonal projections as in Lemma 23

$$u_h^0 = \Pi_h u^0,\ v_h^0 = \Pi_h v^0,\ f_h^n = P_h f^n,\ 0 \le n \le N. \tag{2.62}$$

The results are reported in Tables 2.4 and 2.5. The meshes, the time steps and other details of the numerical algorithm, are exactly the same in these two tables. We observe that the errors are very similar as well and conclude therefore that the accuracy of the method does not depend on the manner in which the initial conditions and f are approximated, either (2.61) or (2.62).

On the contrary, the behavior of error estimators is quite different in the two

h	τ	ei	$\hat{ei}$	η_T	$\hat{\eta}_T$	η_S	$\eta_S^{(1)}$	$\eta_S^{(2)}$	N_0	e
1/160	$\sqrt{h}$	12.29	11.44	.115	.087	.28	.094	.19	98.48	.032
1/320	$\sqrt{h}$	12.13	11.56	.054	.045	.14	.047	.094	98.18	.016
1/640	$\sqrt{h}$	12.	11.62	.027	.023	.071	.024	.047	98.27	.0081
1/160	h	17.4	17.4	.00062	.00061	.29	.095	.19	98.48	.017
1/320	h	17.25	17.25	.00015	.00015	.14	.047	.094	98.18	.082
1/640	h	17.28	17.28	3.83e-05	3.81e-05	.071	.023	.047	98.27	.0041

TABLE 2.5: Results for case (a), constant time steps, unstructured Delaunay meshes, orthogonal projection of the initial conditions and f as in (2.62).

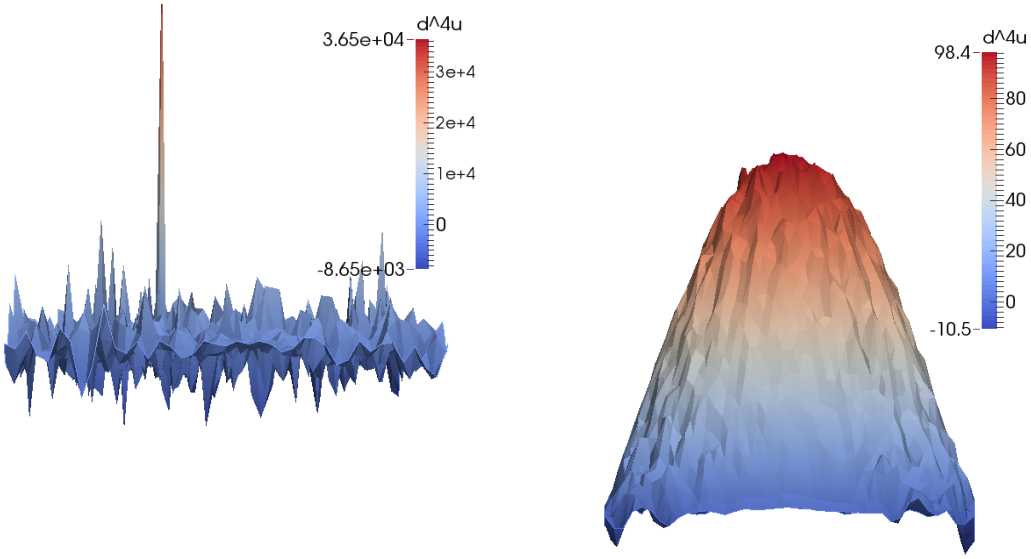


FIGURE 2.2: $\partial_4^4 u_h$ for different discretization of the initial conditions: on the left (see Table 2.4) we take u_h^0 as the nodal interpolation of u_0 while on the right (see Table 2.5) $u_h^0 = \Pi_h u_0$, $h = 0.125$, $\tau = 0.025$.

cases. From Table 2.4 (nodal interpolation), we see that both time error estimators η_T, $\hat{\eta}_T$ blow up with mesh refinement, while the second part of the space estimator $\eta_S^{(2)}$ behaves (non optimally) like $O(\tau + h)$. Only the first part of the space estimator $\eta_S^{(1)}$ behaves as the true error. Such a strange behavior of our estimators indicates the unboundedness of higher order discrete derivatives in time. Indeed, the estimators η_T, $\hat{\eta}_T$ and $\eta_S^{(2)}$ contain high order discrete derivatives $\partial_n^2 f_h - A_h \partial_n^2 u_h$, $\partial_n^4 u_h$ and $\partial_n^2 v_h$ respectively. These error estimators can be of the optimal order only if all these derivatives are uniformly bounded. We recall that this property was examined in Lemma 23 and its proof hinges on the boundedness of

$$N_0 = \left\| A_h^2 u_h^0 - A_h f_h^0 \right\|_{L^2(\Omega)} . \tag{2.63}$$

However, as reported in Table 2.4, N_0 also blows up under the nodal interpolation of initial conditions and of the right-hand side. This is not surprising given that the boundedness of N_0 in Lemma 23 is a consequence of Lemma 22 and thus it is not guaranteed if one replaces projections (2.62) by nodal interpolation (2.61). On the other hand, the results in Table 2.5 corresponding to interpolation by projection (2.62) confirm the order $O(\tau^2 + h)$ for our error estimators, consistently with the theory developed in Lemmas 23 and 22.

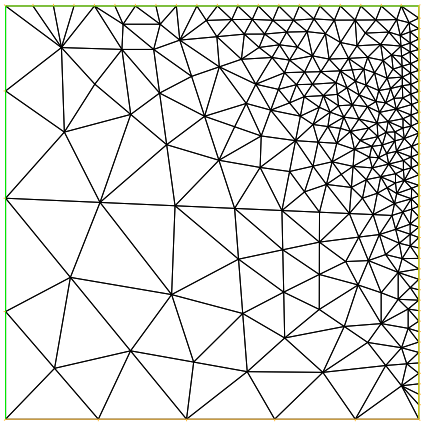
The huge difference between the two data approximations can be also seen by looking directly at $\partial_4^4 u_h$. We report this quantity in Fig. 2.2 for the case (a) on a mesh with $h = 0.0125$ and time step $\tau = 0.025$ at $t = t_4 = 0.1$. On the left picture (nodal interpolation) we see that $\partial_4^4 u_h$ contains a lot of severe spurious oscillations, while the right picture (projection of initial conditions) contains a reasonable and quite smooth

Mesh	τ	ei	$\hat{ei}$	M_1	M_2	N_0	e
case (1)	1/10	17.15	16.98	10.39	2.25	102.59	.37
case (2)	1/20	17.15	17.05	9.99	2.22	98.62	.099
case (3)	1/40	17.15	17.12	9.97	2.22	98.45	.025

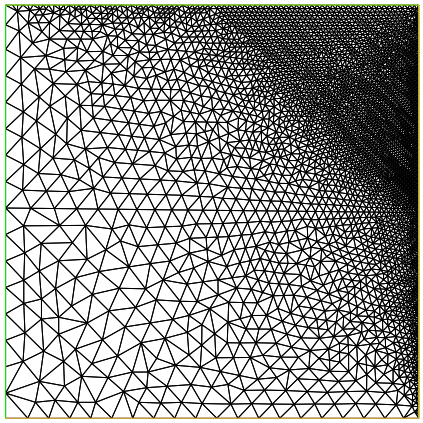
TABLE 2.6: Results for case (a), constant time step, unstructured Delaunay mesh, orthogonal projection of the initial conditions and f as in (2.62), $M_1 = \|A_h P_h u^0\|_{L^2(\Omega)}$, $M_2 = \|P_h u^0\|_{H^1(\Omega)}$.

approximation of $\dfrac{\partial^4 u}{\partial t^4}$. This is another manifestation of the critical importance of the choice of an approximation of initial conditions and of the right-hand side for our error estimators. We note that such a phenomenon was not observed for the heat equation [LPP09]. We also recall from Table 2.1 that space and time error estimators provide a good representation of the true error on a structured mesh even under the nodal interpolation. Note that the quantity defined by (2.63) remains also bounded on the structured mesh.

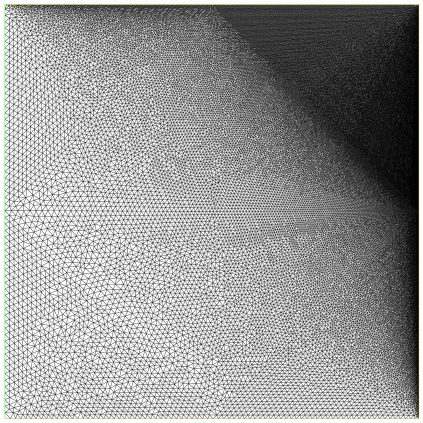
We recall that the theory of Subsection 2.2.2, in particular Lemma 22, are established under the quasi-uniform mesh assumption. We conclude this article by a numerical test on non quasi-uniform meshes in order to asses the stability of operators A_h and P_h. We apply our numerical method to (1.59) with the exact solution u from case (a) on meshes from Fig. 2.3. The results are given in Table 2.6. We see that space and time error estimators provide a good representation of the true error, like in examples from Tables 2.1 and 2.5 with quasi-uniform meshes. Moreover, we observe stability for terms $\|A_h P_h u^0\|_{L^2(\Omega)}$, $\|P_h u^0\|_{H^1(\Omega)}$, and consequently N_0. This indicates that our error indicators may be useful for time and space adaptivity on rather general meshes.



case (1), 486 triangles



case (2), 7535 triangles



case (3), 121299 triangles

FIGURE 2.3: non quasi-uniform meshes (see Table 2.6).

Chapter 3

An easily computable error estimator in space and time for the wave equation

In this chapter, we are interested in easily computable error estimator in space and time for the wave equation. The 3-point time error estimator proposed in previous chapter contains the Laplacian of the discrete solution which should be computed via auxiliary finite element problems at each time step. This requires thus a non-negligible extra work in comparison with computing the discrete solution itself. In the present chapter, we propose an alternative time error estimator that avoids these additional computations.

In deriving our *a posteriori* estimates, we follow first the approach of previous chapter. First of all, we use the fact that the Newmark method can be reinterpreted as the Crank-Nicolson discretization of the reformulation of the governing equation as the first-order system, as in [Bak76]. We then use the techniques stemming from *a posteriori* error analysis for the Crank-Nicolson discretization of the heat equation in [LPP09], based on a piecewise quadratic polynomial in time reconstruction of the numerical solution. Finally, in a departure from Theorem 20, we replace the second derivatives in space (Laplacian of the discrete solution) in the error estimate with the forth derivatives in time by reusing the governing equation. This leads to the new *a posteriori* error estimate in time and also allows us to easily recover the error estimates in space that turn out to be the same as those of Theorem 20. The resulting estimate is referred to as the 5-point estimator since it contains the fourth order finite differences in time and thus involves the discrete solution at 5 points in time at each time step.

Like in the case of the 3-point estimator, we are able to prove that the new 5-point estimator is reliable on general regular meshes in space and non-uniform meshes in time (with constants depending on the regularity of meshes in both space and time). Moreover, the 5-point estimator is proved to be of optimal order at least on sufficiently smooth solutions, quasi-uniform meshes in space and uniform meshes in time, again reproducing the results known for the 3-point estimator. Numerical experiments demonstrate that the 3-point and the 5-point error estimators produce very similar results in the majority of test cases. Both turn out to be of optimal order in space and time, even in situations not accessible to the current theory (non quasi-uniform meshes, not constant time steps). It should be therefore possible to use the

new estimator for mesh adaptation in space and time. In fact, the best strategy in practice may be to combine both estimators to take benefit from the strengths of each of them: the relative cheapness of the 5-point one, and the better numerical behavior of the 3-point estimator under abrupt changes of the mesh.

The outline of the chapter is as follows. In Section 3.1, the 5-point *a posteriori* error estimator for the fully discrete wave problem is derived. Numerical experiments for the 3-point and the 5-point error estimators on several test cases are presented in Section 3.2.

Chapter contents

3.1 The 5-point *a posteriori* error estimator

As it is already mentioned, the time error estimator from Theorem (20) contains a finite element approximation to the Laplacian of u_h^k, i.e. z_h^k given by (2.20). This is unfortunate because z_h^k should be computed by solving an additional finite element problem that implies additional computational effort. Keeping in mind that the term $\partial_n^2 f_h - z_h^n$ in (2.19) is a discretization of $\partial^2 f/\partial t^2 + \Delta u = \partial^4 u/\partial t^4$ at the time t_n our goal now is to avoid the second derivatives in space in the error estimates and replace them with the forth derivatives in time.

We introduce a "fourth order finite difference in time" ∂_n^4 defined by

$$\partial_n^4 w_h \;=\; \frac{8}{\tau_n+\tau_{n-1}+\tau_{n-2}+\tau_{n-3}}\left(\frac{\partial_n^2 w_h-\partial_{n-1}^2 w_h}{\tau_n+\tau_{n-2}}-\frac{\partial_{n-1}^2 w_h-\partial_{n-2}^2 w_h}{\tau_{n-1}+\tau_{n-3}}\right), \quad (3.1)$$

on any sequence $\{w_h^n\}_{n=0,1,\ldots} \in V_h$. This can be rewritten as a composition of two second order finite difference operators

$$\partial_n^4 w_h = \hat{\partial}_n^2 \partial^2 w_h, \quad (3.2)$$

where $\partial^2 w_h$ is the standard finite difference applied to w_h, and $\hat{\partial}_n^2$ is a modified second order finite difference defined by

$$\hat{\partial}_n^2 w_h = \frac{2}{(\hat{t}_n-\hat{t}_{n-2})}\left(\frac{w_h^n-w_h^{n-1}}{\hat{t}_n-\hat{t}_{n-1}}-\frac{w_h^{n-1}-w_h^{n-2}}{\hat{t}_{n-1}-\hat{t}_{n-2}}\right), \quad (3.3)$$

$$\hat{t}_n = \frac{t_{n+1}+t_{n-1}}{2},$$

on any sequence $\{w_h^n\}_{n=0,1,\ldots} \in V_h$. Note that a lower subscript "n" is lacking from $\partial^2 w_h$ in (3.2) consistent with the fact that $\hat{\partial}_n^2$ is applied there to the sequence

$\{\partial_n^2 w_h\}_{n=0,1,\dots}$ rather than to a single instance of $\partial_n^2 w_h$. In full detail, (3.2) should be interpreted as $\partial_n^4 w_h = \hat{\partial}_n^2 w_h$ with $w_h^n = \partial_n^2 w_h$.

Remark 18. *In the case of constant time steps $\tau_n = \tau$, (3.1) is reduced to*

$$\partial_n^4 w_h = \frac{w_h^{n+1} - 4w_h^n + 6w_h^{n-1} - 4w_h^{n-2} + w_h^{n-3}}{\tau^4}.$$

It is thus indeed a standard finite difference approximation to the fourth derivative. In particular, it is exact on polynomials (in time) of degree up to 4. However, a standard fourth order finite difference in the general case of non constant time steps would be given by the divided differences

$$\begin{aligned}\tilde{\partial}_n^4 w_h &= 4![w_h^{n-3}, \dots, w_h^{n+1}] \\ &= \frac{12}{\tau_n + \tau_{n-1} + \tau_{n-2} + \tau_{n-3}} \left(\frac{\partial_n^2 w_h - \partial_{n-1}^2 w_h}{\tau_n + \tau_{n-1} + \tau_{n-2}} - \frac{\partial_{n-1}^2 w_h - \partial_{n-2}^2 w_h}{\tau_{n-1} + \tau_{n-2} + \tau_{n-3}} \right).\end{aligned}$$

Clearly, the formulas $\partial_n^4 w_h$ and $\tilde{\partial}_n^4 w_h$, although similar, do not coincide in general, and consequently $\partial_n^4 w_h$ is not necessarily consistent with the fourth derivative in time of w_h. Definition (3.1) may seem thus artificial and counter-intuitive. We shall see however that it arises naturally in the analysis of Newmark scheme, cf. forthcoming Lemma 27. Indeed, in order to "differentiate" in time the averaged quantities $\bar{w}_h^n$ defined by (3.4) and present the scheme (2.5), cf. also (3.12), one needs to employ the modified second order finite difference $\hat{\partial}_n^2$, which shall be composed further with ∂_n^2 to give rise $\partial^4{}_n$.

For any sequence $\{w_h^n\}_{n=0,1,\dots} \in V_h$, we denote

$$\bar{w}_h^n = \frac{\tau_n(w_h^{n+1} + w_h^n) + \tau_{n-1}(w_h^n + w_h^{n-1})}{4\tau_{n-1/2}}. \tag{3.4}$$

Consistently with the conventions above, $\bar{w}_h$ will stand for any sequence $\{\bar{w}_h^n\}_{n=0,1,\dots}$. The following technical lemma establish a connection between second order discrete derivatives $\hat{\partial}_n^2$ and ∂_n^2.

Lemma 27. *For all integer $n = 3, \dots N-1$ there exist coefficients α_k, $k = n-2, n-1, n$ such that for all $\{w_h^n\}$*

$$\hat{\partial}_n^2 \bar{w}_h = \sum_{k=n-2}^{n} \alpha_k \partial_k^2 w_h. \tag{3.5}$$

Moreover

$$|\alpha_k| \le c, \text{ for } k = n-2, n-1, n, \text{ and } \sum_{k=n-2}^{n} \alpha_k \ge C,$$

where c and C are positive constants depending only on the mesh regularity in time, i.e. on $\max_{k\ge 0}\left(\frac{\tau_{k+1}}{\tau_k} + \frac{\tau_k}{\tau_{k+1}}\right)$.

Proof. We first note that relation (3.5) does not contain any derivatives in space and thus it should hold at any point $x \in \Omega$. Consequently, it is sufficient to prove this lemma assuming that w_h^n, $\partial_n^2 w_h$, etc. are real numbers, i.e. replacing V_h by $\mathbb{R}$. This is the assumption adopted in this proof. We shall thus drop the sub-indexes h everywhere. Furthermore, it will be convenient to reinterpret w^n in (3.2), (3.3) and (3.4) as

the values of a real valued function $w(t)$ at $t = t_n$. We shall also use the notations like $\bar{w}^n$, $\partial_n^2 w$, and so on, where w is a continuous function on $\mathbb{R}$, always assuming $w^n = w(t_n)$.

Observe that $\hat{\partial}_n^2 \bar{w}$ is a linear combination of 5 numbers $\{w^{n-3}, \dots, w^{n+1}\}$. Thus, it is enough to check equality (3.5) on any 5 continuous functions $\phi_{(k)}(t)$, $k = n-3, \dots, n+1$, such that the vector of values of $\phi_{(k)}$ at times t_l, $l = n-3, \dots, n+1$, form a basis of $\mathbb{R}^5$. For fixed n, let us choose these functions as

$$\phi_{(k)}(t) = \begin{cases} \dfrac{t - t_{k-1}}{\tau_{k-1}}, & \text{if } t < t_k, \\ \dfrac{t_{k+1} - t}{\tau_k}, & \text{if } t \geq t_k, \end{cases} \qquad k = n-3, \dots, n+1. \tag{3.6}$$

First we notice that for every linear function $u(t)$ on $[t_{n-3}, t_{n+1}]$ we have

$$\hat{\partial}_n^2 \bar{u} = \partial_n^2 u = 0.$$

Thus, we get immediately

$$\hat{\partial}_n^2 \bar{\phi}_{(n-3)} = \partial_n^2 \phi_{(n-3)} = 0,$$

and

$$\hat{\partial}_n^2 \bar{\phi}_{(n+1)} = \partial_n^2 \phi_{(n+1)} = 0,$$

so that (3.5) is fulfilled on functions $\phi_{(n-3)}$, $\phi_{(n+1)}$ with any coefficients α_k, $k = n-2, n-1, n$. Now we want to provide coefficients α_k, $k = n-2, n-1, n$ for which (3.5) is fulfilled on functions $\phi_{(n-2)}$, $\phi_{(n-1)}$ and $\phi_{(n)}$. For brevity, we demonstrate the idea only for function $\phi_{(n)}(t)$. Function $\phi_{(n)}(t)$ is linear on $[t_{n-3}, t_n]$ and thus

$$\partial_{n-2}^2 \phi_{(n)} = 0, \ \partial_{n-1}^2 \phi_{(n)} = 0.$$

From direct computations it is easy to show that

$$\partial_n^2 \phi_{(n)} \sim \frac{1}{\tau_n^2}, \ \bar{\phi}_{(n)} \sim 1, \ \hat{\partial}_n^2 \bar{\phi}_{(n)} \sim \frac{1}{\tau_n^2}$$

where $\sim$ hides some factors that can be bounded by constants depending only on the mesh regularity. Thus we are able to establish expression for coefficient

$$\alpha_n = \frac{\hat{\partial}_n^2 \bar{\phi}_{(n)}}{\partial_n^2 \phi_{(n)}} \leq C.$$

Similar reasoning for function $\phi_{(n-1)}$ and $\phi_{(n-2)}$ shows that

$$\alpha_{n-1} = \frac{\hat{\partial}_n^2 \bar{\phi}_{(n-1)}}{\partial_{n-1}^2 \phi_{(n-1)}} \leq C \text{ and } \alpha_{n-2} = \frac{\hat{\partial}_n^2 \bar{\phi}_{(n-2)}}{\partial_{n-2}^2 \phi_{(n-2)}} \leq C.$$

The next step is to show boundedness from below of $\sum_{k=n-2}^{n} \alpha_k$. We will show it by applying equality (3.5) to the second order polynomial function $s(t) = \dfrac{t^2}{2}$. Using

a Taylor expansion of $s(t)$ around $\hat{t}_n$ in the definition of $\bar{s}^n$ gives

$$\bar{s}^n = \frac{\tau_n(\hat{t}_n^2 + \hat{t}_n\tau_{n-1} + \frac{1}{4}(\tau_n^2 + \tau_{n-1}^2)) + \tau_{n-1}(\hat{t}_n^2 - \hat{t}_n\tau_n + \frac{1}{4}(\tau_n^2 + \tau_{n-1}^2))}{2(\tau_n + \tau_{n-1})}$$
$$= \frac{\hat{t}_n^2}{2} + \frac{1}{8}\left(\tau_n^2 + \tau_{n-1}^2\right).$$

Substituting this into the definition of $\hat{\partial}_n^2\bar{s}$ we obtain

$$\hat{\partial}_n^2\bar{s} = \frac{\frac{\bar{s}^n - \bar{s}^{n-1}}{\hat{t}_n - \hat{t}_{n-1}} - \frac{\bar{s}^{n-1} - \bar{s}^{n-2}}{\hat{t}_{n-1} - \hat{t}_{n-2}}}{\left(\hat{t}_n - \hat{t}_{n-2}\right)/2} = 1 + \frac{1}{8}\left(\frac{2\frac{\tau_n^2 - \tau_{n-2}^2}{\tau_n + \tau_{n-2}} - 2\frac{\tau_{n-1}^2 - \tau_{n-3}^2}{\tau_{n-1} + \tau_{n-3}}}{\frac{1}{4}(\tau_n + \tau_{n-1} + \tau_{n-2} + \tau_{n-3})}\right)$$
$$= 1 + \frac{\tau_n - \tau_{n-1} - \tau_{n-2} + \tau_{n-3}}{\tau_n + \tau_{n-1} + \tau_{n-2} + \tau_{n-3}}.$$

Using (3.5) and the fact that $\partial_k^2 s = 1$ for $k = n-2, n-1, n$ we note that

$$1 + \frac{\tau_n - \tau_{n-1} - \tau_{n-2} + \tau_{n-3}}{\tau_n + \tau_{n-1} + \tau_{n-2} + \tau_{n-3}} = \sum_{k=n-2}^{n} \alpha_k.$$

This implies $\displaystyle\sum_{k=n-2}^{n} \alpha_k \geq C.$ □

Lemma 28. *Let $w_h^n, s_h^n \in V_h$ be such that*

$$\frac{w_h^{n+1} - w_h^n}{\tau_n} - \frac{s_h^n + s_h^{n+1}}{2} = 0, \ \forall n \geq 0. \tag{3.7}$$

For all $n \geq 3$ there exist coefficients β_k, $k = n-2, n-1, n$ such that

$$\sum_{k=n-2}^{n} \alpha_k \partial_k^2 w_h = \left(\sum_{k=n-2}^{n} \alpha_k\right) \partial_n^2 w_h - \tau_n \sum_{k=n-2}^{n} \beta_k \partial_k^2 s_h, \tag{3.8}$$

where coefficients α_k, $k = n-2, n-1, n$ are introduced in Lemma 27. Moreover

$$|\beta_k| \leq C, \ k = n-2, n-1, n,$$

where C is a positive constant depending only on the mesh regularity in time, i.e. on $\max_{k\geq 0}\left(\frac{\tau_{k+1}}{\tau_k} + \frac{\tau_k}{\tau_{k+1}}\right)$.

Proof. Like in the proof of Lemma 27, we assume $V_h = \mathbb{R}$, drop the sub-indexes h and interpret w^n, s^n as the values of continuous real valued functions $w(t)$, $s(t)$ at $t = t_n$. Using (3.7) and notations (2.7) implies $\partial_k^2 w = \partial_k s$. Now, we are able to rewrite (3.8) in terms of s^n only

$$\sum_{k=n-2}^{n} \alpha_k \partial_k s = \left(\sum_{k=n-2}^{n} \alpha_k\right) \partial_n s - \tau_n \sum_{k=n-2}^{n} \beta_k \partial_k^2 s. \tag{3.9}$$

As in the proof of Lemma 27 we take into account the fact that equation (3.9) should hold for every 5 numbers $\{s^{n-3}, \dots, s^{n+1}\}$ and therefore it's enough to check equality (3.9) on 5 linearly independent piecewise linear functions $\phi_{(k)}$ introduced by (3.6). Using the reasoning as in Lemma 27 leads to desired result (3.8). □

We can now prove an *a posteriori* error estimate involving $\partial_n^4 u_h$. Since the latter is computed through 5 points in time $\{t_{n-3}, \dots, t_{n+1}\}$, we shall refer to this approach as the 5-point estimator. For the same reason, this estimator is only applicable from the time t_4. The error at first 3 time steps should be thus measured differently, for example using the 3-point estimator from Theorem 20. The deriving of 5-point time error estimator is based on main ideas from the proof of Theorem 20 using Lemma 27.

Theorem 29. *The following a posteriori error estimate holds between the solution u of the wave equation (1.59) and the discrete solution u_h^n given by (2.4)–(2.5) for all t_n, $4 \le n \le N$ with v_h^n given by (2.6):*

$$\left(\left\| v_h^n - \frac{\partial u}{\partial t}(t_n) \right\|_{L^2(\Omega)}^2 + |u_h^n - u(t_n)|_{H^1(\Omega)}^2 \right)^{1/2}$$
$$\le \left(\left\| v_h^3 - \frac{\partial u}{\partial t}(t_3) \right\|_{L^2(\Omega)}^2 + \left| u_h^3 - u(t_3) \right|_{H^1(\Omega)}^2 \right)^{1/2}$$
$$+ \eta_S(t_N) + C \sum_{k=3}^{N-1} \tau_k \hat{\eta}_T(t_k) + C \sum_{k=3}^{N-1} \tau_k \hat{\eta}_T^{h.o.t.}(t_k) + \int_{t_3}^{t_n} \|f - \tilde{f}_\tau\|_{L^2()\Omega} dt, \quad (3.10)$$

where the space indicator is defined by (2.22) and the time error indicator is

$$\hat{\eta}_T(t_k) = \left(\frac{1}{12}\tau_k^2 + \frac{1}{8}\tau_{k-1}\tau_k \right) \left(\left| \partial_k^2 v_h \right|_{H^1(\Omega)} + \left\| \partial_k^4 u_h \right\|_{L^2(\Omega)} \right), \quad (3.11)$$

with additional higher order terms

$$\hat{\eta}_T^{h.o.t.}(t_k) = \tau_k^3 \left\| \partial_k^2 \dot{f}_h - A_h \partial_k^2 v_h \right\|_{L^2(\Omega)},$$

where $\dot{f}_h^n$ satisfy

$$\frac{f_h^{n+1} - f_h^n}{\tau_n} = \frac{\dot{f}_h^n + \dot{f}_h^{n+1}}{2}.$$

The constant $C > 0$ depends only on the mesh regularity in time, i.e. on $\max_{k \ge 0} \left(\frac{\tau_{k+1}}{\tau_k} + \frac{\tau_k}{\tau_{k+1}} \right)$.

Proof. We note first of all that it is sufficient to prove the Theorem for the final time, i.e. $n = N$ because the statement for the general case $n \le N$ will follow by resetting the final time N to n. We can rewrite scheme (2.5) as

$$\partial_n^2 u_h + A_h \bar{u}_h^n = \bar{f}_h^n, \quad (3.12)$$

for $n = 0, \dots, N-1$ where $f_h^n = P_h f(t_n, \cdot)$ and operator A_h defined in (3.22). Taking a linear combination of instances of (3.12) at steps $n, n-1, n-2$ with appropriate

coefficients gives

$$\partial_n^4 u_h + A_h \hat{\partial}_n^2 \bar{u}_h = \hat{\partial}_n^2 \bar{f}_h. \tag{3.13}$$

Using the definition of operator $\hat{\partial}_n^2$ and re-introducing v_h^n by (2.2) leads to

$$\hat{\partial}_n^2 \bar{u}_h = \sum_{k=n-2}^{n} \alpha_k \partial_k^2 u_h = \left(\sum_{k=n-2}^{n} \alpha_k \right) \partial_n^2 u_h - \tau_n \sum_{k=n-2}^{n} \beta_k \partial_k^2 v_h.$$

with coefficients α_k, β_k introduced in Lemma 27 and Lemma 28. Moreover, by Lemma 27 $\gamma = \left(\sum_{k=n-2}^{n} \alpha_k\right)^{-1}$ is positive and bounded so that

$$\partial_n^2 u_h = \gamma \hat{\partial}_n^2 \bar{u}_h + \tau_n \sum_{k=n-2}^{n} \gamma_k \partial_k^2 v_h,$$

with $\gamma_k = \gamma \beta_k$ that are all uniformly bounded on regular meshes in time. Similarly,

$$\partial_n^2 f_h = \hat{\partial}_n^2 \bar{f}_h + \tau_n \sum_{k=n-2}^{n} \gamma_k \partial_k^2 \dot{f}_h.$$

Thus,

$$\begin{aligned} \partial_n^2 f_h - A_h \partial_n^2 u_h &= \hat{\partial}_n^2 \bar{f}_h - A_h \hat{\partial}_n^2 \bar{u}_h + \tau_n \sum_{k=n-2}^{n} \gamma_k \left(\partial_k^2 \dot{f}_h - A_h \partial_k^2 v_h \right) \\ &= \partial_n^4 u_h + \tau_n \sum_{k=n-2}^{n} \gamma_k \left(\partial_k^2 \dot{f}_h - A_h \partial_k^2 v_h \right). \end{aligned} \tag{3.14}$$

We can now reproduce the proof of Theorem 20. In the following, we adopt the vector notation

$$U(t,x) = \begin{pmatrix} u(t,x) \\ v(t,x) \end{pmatrix},$$

where $v = \partial u / \partial t$. Note that the first equation in (1.60) implies that

$$\left(\nabla \frac{\partial u}{\partial t}, \nabla \varphi \right) - (\nabla v, \nabla \varphi) = 0, \quad \forall \varphi \in H_0^1(\Omega),$$

by taking its gradient, multiplying it by $\nabla \varphi$ and integrating over Ω. Thus, system (1.60) can be rewritten in the vector notations as

$$b\left(\frac{\partial U}{\partial t}, \Phi \right) + (\mathcal{A} \nabla U, \nabla \Phi) = b(F, \Phi), \quad \forall \Phi \in (H_0^1(\Omega))^2, \tag{3.15}$$

where $\mathcal{A} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, $F = \begin{pmatrix} 0 \\ f \end{pmatrix}$ and

$$b(U, \Phi) = b\left(\begin{pmatrix} u \\ v \end{pmatrix}, \begin{pmatrix} \varphi \\ \psi \end{pmatrix} \right) := (\nabla u, \nabla \varphi) + (v, \psi).$$

Similarly, Newmark scheme (2.2)–(2.3) can be rewritten as

$$b\left(\frac{U_h^{n+1}-U_h^n}{\tau_n},\Phi_h\right)+\left(\mathcal{A}\nabla\frac{U_h^{n+1}+U_h^n}{2},\nabla\Phi_h\right) = b\left(F^{n+1/2},\Phi_h\right),\quad \forall\Phi_h\in V_h^2, \quad (3.16)$$

where $U_h^n=\begin{pmatrix}u_h^n\\ v_h^n\end{pmatrix}$ and $F^{n+1/2}=\begin{pmatrix}0\\ f^{n+1/2}\end{pmatrix}$.

The *a posteriori* analysis relies on an appropriate residual equation for the quadratic reconstruction

$$\tilde{U}_{h\tau}=\begin{pmatrix}\tilde{u}_{h\tau}\\ \tilde{v}_{h\tau}\end{pmatrix}.$$

We have thus for $t\in[t_n,t_{n+1}]$, $n=1,\dots,N-1$

$$\tilde{U}_{h\tau}(t)=U_h^{n+1}+(t-t_{n+1})\partial_{n+1/2}U_h+\frac{1}{2}(t-t_{n+1})(t-t_n)\partial_n^2U_h. \quad (3.17)$$

so that, after some simplifications,

$$b\left(\frac{\partial\tilde{U}_{h\tau}}{\partial t},\Phi_h\right)+(\mathcal{A}\nabla\tilde{U}_{h\tau},\nabla\Phi_h)=b\left((t-t_{n+1/2})\partial_n^2U_h+F^{n+1/2},\Phi_h\right) + \left((t-t_{n+1/2})\mathcal{A}\nabla\partial_{n+1/2}U_h+\frac{1}{2}(t-t_{n+1})(t-t_n)\mathcal{A}\nabla\partial_n^2U_h,\nabla\Phi_h\right). \quad (3.18)$$

Consider now (3.16) at the time steps n and $n-1$. Subtracting one from another and dividing by $\tau_{n-1/2}$ yields

$$b\left(\partial_n^2U_h,\Phi_h\right)+(\mathcal{A}\nabla\partial_nU_h,\nabla\Phi_h)=b\left(\partial_nF,\Phi_h\right),$$

or

$$b\left(\partial_n^2U_h,\Phi_h\right)+\left(\mathcal{A}\nabla\left(\partial_{n+1/2}U_h-\frac{\tau_{n-1}}{2}\partial_n^2U_h\right),\nabla\Phi_h\right)=b\left(\partial_nF,\Phi_h\right),$$

so that (3.18) simplifies to

$$b\left(\frac{\partial\tilde{U}_{h\tau}}{\partial t},\Phi_h\right)+\left(\mathcal{A}\nabla\tilde{U}_{h\tau},\Phi_h\right) = \left(p_n\mathcal{A}\nabla\partial_n^2U_h,\nabla\Phi_h\right)+b\left(\left(t-t_{n+1/2}\right)\partial_nF+F^{n+1/2},\Phi_h\right) = \left(p_n\mathcal{A}\nabla\partial_n^2U_h,\nabla\Phi_h\right)+b\left(\tilde{F}_\tau-p_n\partial_n^2F,\Phi_h\right), \quad (3.19)$$

where

$$p_n=\frac{\tau_{n-1}}{2}(t-t_{n+1/2})+\frac{1}{2}(t-t_{n+1})(t-t_n),$$
$$\tilde{F}_\tau(t)=F_h^{n+1}+(t-t_{n+1})\partial_{n+1/2}F+\frac{1}{2}(t-t_{n+1})(t-t_n)\partial_n^2F.$$

Introduce the error between reconstruction $\tilde{U}_{h\tau}$ and solution U to problem (3.15) :

$$E = \tilde{U}_{h\tau} - U, \tag{3.20}$$

or, component-wise

$$E = \begin{pmatrix} E_u \\ E_v \end{pmatrix} = \begin{pmatrix} \tilde{u}_{h\tau} - u \\ \tilde{v}_{h\tau} - v \end{pmatrix}.$$

Taking the difference between (3.19) and (3.15) we obtain the residual differential equation for the error valid for $t \in [t_n, t_{n+1}]$, $n = 1, \ldots, N-1$

$$\begin{aligned} b(\partial_t E, \Phi) + (\mathcal{A}\nabla E, \nabla\Phi) = b\left(\frac{\partial \tilde{U}_{\tau h}}{\partial t} - F, \Phi - \Phi_h\right) \\ + \left(\mathcal{A}\nabla\tilde{U}_{\tau h}, \nabla(\Phi - \Phi_h)\right) + \left(p_n \mathcal{A}\nabla\partial_n^2 U_h, \nabla\Phi_h\right) \\ + b\left(\tilde{F}_\tau - F - p_n\partial_n^2 F, \Phi_h\right), \quad \forall \Phi_h \in V_h^2. \end{aligned} \tag{3.21}$$

Now we take $\Phi = E$, $\Phi_h = \begin{pmatrix} \Pi_h E_u \\ \tilde{I}_h E_v \end{pmatrix}$, note that $(\mathcal{A}\nabla E, \nabla E) = 0$ and

$$\left(\nabla\frac{\partial\tilde{u}_{h\tau}}{\partial t}, \nabla(E_u - \Pi_h E_u)\right) = (\nabla\tilde{v}_{h\tau}, \nabla\left(E_u - \Pi_h E_u\right)) = 0.$$

Introducing operator $A_h : V_h \to V_h$ such that

$$(A_h w_h, \varphi_h) = (\nabla w_h, \nabla\varphi_h), \quad \forall \varphi_h \in V_h, \tag{3.22}$$

we get

$$\begin{aligned} \left(\frac{\partial E_v}{\partial t}, E_v\right) + \left(\nabla E_u, \nabla\frac{\partial E_u}{\partial t}\right) = \left(\frac{\partial\tilde{v}_{\tau h}}{\partial t} - f, E_v - \Pi_h E_v\right) \\ + \left(\nabla\tilde{u}_{\tau h}, \nabla\left(E_v - \tilde{I}_h E_v\right)\right) + \left(p_n\left(A_h\partial_n^2 u_h - \partial_n^2 f_h\right), \tilde{I}_h E_v\right) \\ - \left(p_n\nabla\partial_n^2 v_h, \nabla E_u\right) + \left(\tilde{f}_\tau - f, \tilde{I}_h E_v\right). \end{aligned}$$

Note that equation similar to (3.21) also holds for $t \in [t_0, t_1]$

$$\begin{aligned} b(\partial_t E, \Phi) + (\mathcal{A}\nabla E, \nabla\Phi) = b\left(\frac{\partial\tilde{U}_{\tau h}}{\partial t} - F, \Phi - \Phi_h\right) + \left(\mathcal{A}\nabla\tilde{U}_{\tau h}, \nabla(\Phi - \Phi_h)\right) \\ + \left(p_1\mathcal{A}\nabla\partial_1^2 U_h, \nabla\Phi_h\right) + b\left(\tilde{F}_\tau - F - p_1\partial_1^2 F, \Phi_h\right). \end{aligned} \tag{3.23}$$

That follows from the definition of the piecewise quadratic reconstruction $\tilde{u}_{h\tau}(t)$ for $t \in [t_0, t_1]$. Integrating (3.21) and (3.23) in time from 0 to some $t^* \geq t_3$ yields

$$\begin{aligned}\frac{1}{2}\left(|E_u|^2_{H^1(\Omega)} + \|E_v\|^2_{L^2(\Omega)}\right)(t^*) &= \frac{1}{2}\left(|E_u|^2_{H^1(\Omega)} + \|E_v\|^2_{L^2(\Omega)}\right)(0) \\ &+ \underbrace{\int_0^{t^*}\left(\frac{\partial \tilde{v}_{\tau h}}{\partial t} - f, E_v - \tilde{I}_h E_v\right) dt}_{I} + \underbrace{\int_0^{t^*}\left(\nabla \tilde{u}_{\tau h}, \nabla(E_v - \tilde{I}_h E_v)\right) dt}_{II} \\ &+ \underbrace{\int_0^{t_1}\left[\left(p_1\left(A_h\partial_1^2 u_h - \partial_1^2 f_h\right), \tilde{I}_h E_v\right) - \left(p_1 \nabla \partial_1^2 v_h, \nabla E_u\right) + \left(\tilde{f}_\tau - f, \tilde{I}_h E_v\right)\right] dt}_{III}.\end{aligned} \tag{3.24}$$

We set

$$Z(t) = \sqrt{|E_u|^2_{H^1(\Omega)} + \|E_v\|^2_{L^2(\Omega)}},$$

and assume that $t^* \in [t_3, t_N]$ is the point in time where Z attains its maximum and $t^* \in (t_n, t_{n+1}]$ for some n. For the first and second terms in (3.24) we have

$$\begin{aligned} I + II \leq\ & C_1\left[\sum_{K\in\mathcal{T}_h} h_K^2 \left\|\frac{\partial \tilde{v}_{h\tau}}{\partial t} - \Delta \tilde{u}_{h\tau} - f\right\|^2_{L^2(K)}\right. \\ &\left. + \sum_{E\in\mathcal{E}_h} h_E \left\|[n\cdot\nabla \tilde{u}_{h\tau}]\right\|^2_{L^2(E)}\right]^{1/2}(t^*)|E_u|_{H^1(\Omega)}(t^*) \\ &+ C_1\left[\sum_{K\in\mathcal{T}_h} h_K^2 \left\|\frac{\partial \tilde{v}_{h\tau}}{\partial t} - \Delta \tilde{u}_{h\tau} - f\right\|^2_{L^2(K)}\right. \\ &\left. + \sum_{E\in\mathcal{E}_h} h_E \left\|[n\cdot\nabla \tilde{u}_{h\tau}]\right\|^2_{L^2(E)}\right]^{1/2}(0)|E_u|_{H^1(\Omega)}(0) \\ &+ C_2 \sum_{m=1}^{n} \frac{\tau_{m-1}}{2}\left[\sum_{K\in\mathcal{T}_h} h_K^2 \left\|\partial_m^2 v_h - \partial_{m-1}^2 v_h\right\|^2_{L^2(K)}\right]^{1/2} |E_u|_{H^1(\Omega)}(t_m) \\ &+ C_3 \sum_{m=0}^{n} \int_{t_m}^{min(t_{m+1}, t^*)}\left[\sum_{K\in\mathcal{T}_h} h_K^2 \left\|\frac{\partial^2 \tilde{v}_{h\tau}}{\partial t^2} - \Delta\frac{\partial \tilde{u}_{\tau h}}{\partial t} - \frac{\partial f}{\partial t}\right\|^2_{L^2(K)}\right. \\ &\left. + \sum_{E\in\mathcal{E}_h} h_E \left\|\left[n\cdot\nabla\frac{\partial \tilde{u}_{\tau h}}{\partial t}\right]\right\|^2_{L^2(E)}\right]^{1/2}(t)|E_u|_{H^1(\Omega)}(t)dt. \end{aligned}$$

Indeed, it follows from integration by parts with respect to time, see the proof of Theorem 3.2. The third term in (3.24) is responsible for the time estimator. It can be written as

$$III = \sum_{n=3}^{N-1} \int_{t_m}^{\min(t_{m+1},t^*)} \left[\left(p_m \left(\partial_m^4 u_h + \tau_m \sum_{k=m-2}^{m} \gamma_k \left(\partial_k^2 \dot{f}_h - A_h \partial_k^2 v_h \right) \right), \tilde{I}_h E_v \right) - (p_m \nabla \partial_m^2 v_h, \nabla E_u) + (\tilde{f}_\tau - f, \tilde{I}_h E_v) \right] dt. \quad (3.25)$$

Recalling that $Z(t^*)$ is the maximum of $Z(t)$ and using the estimate $\|\tilde{I}_h E_v\|_{L^2(\Omega)} \leq C\|E_v\|_{L^2(\Omega)}$ we continue as

$$III \leq Z(t^*) \int_{t_m}^{min(t_{m+1},t^*)} \sum_{m=3}^{n-1} |p_m| dt \left(C\|\partial_m^4 u_h\|_{L^2(\Omega)} + |\nabla \partial_m^2 v_h|_{H^1(\Omega)} + C\tau_m \sum_{k=m-2}^{m} \gamma_k \|\partial_k^2 \dot{f}_h - A_h \partial_k^2 v_h\|_{L^2(\Omega)} \right) + Z(t^*) \int_{t_3}^{t_n} \|f - \tilde{f}_\tau\|_{L^2(\Omega)}.$$

Noting

$$\int_{t_m}^{t_{m+1}} |p_m| dt \leq \frac{1}{12}\tau_m^3 + \frac{1}{8}\tau_{m-1}\tau_m^2,$$

we can finally bound III as

$$III \leq \left(C \sum_{k=3}^{N-1} \tau_k \hat{\eta}_T(t_k) + C \sum_{k=3}^{N-1} \tau_k \hat{\eta}_T^{h.o.t.}(t_k) + \int_{t_3}^{t_n} \|f - \tilde{f}_\tau\|_{L^2()\Omega} dt \right) Z(t^*)$$

Summing together the estimates on the terms I, II, III, and recalling $Z(t^*) \geq Z(t_N)$ yields (29) at the final time t_N. □

Remark 19. *The terms $\hat{\eta}_T^{h.o.t.}(t_k)$ in (29) are of higher order than $\hat{\eta}_T(t_k)$. We propose therefore to ignore $\hat{\eta}_T^{h.o.t.}(t_k)$ in practice together with the integral of $f - \tilde{f}_\tau$, and to use $\hat{\eta}_T(t_k)$ as the indicator of error due to the discretization in time. The following Theorem shows that that the latter is indeed of optimal order τ^2, at least for sufficiently smooth solutions, on quasi-uniform meshes in space and uniform meshes in time.*

Theorem 30. *Let u be the solution of wave equation (1.59) and*

$$\frac{\partial^3 u}{\partial t^3}(0) \in H^1(\Omega)\,, \frac{\partial^2 u}{\partial t^2}(0) \in H^2(\Omega),$$

$$\frac{\partial^2 f}{\partial t^2}(t) \in L^\infty(0,T;L^2(\Omega))\,, \frac{\partial^3 f}{\partial t^3}(t) \in L^2(0,T;L^2(\Omega)).$$

Suppose that mesh $\mathcal{T}_h$ is quasi-uniform and the mesh in time is uniform ($t_k = k\tau$). Then, the 5-point time error estimator $\hat{\eta}_T(t_k)$ defined by (29) is of order τ^2, i.e.

$$\hat{\eta}_T(t_k) \leq C\tau^2.$$

with a positive constant C depending only on u, f, and the mesh regularity.

Proof. The result follows from Theorem 2.2.2 by using (3.13) and Lemma 27

$$\left\|\partial_n^4 u_h\right\|_{L^2(\Omega)} = \left\|\hat{\partial}_n^2 \bar{f}_h - A_h \hat{\partial}_n^2 \bar{u}_h\right\|_{L^2(\Omega)} = \left\|\sum_{k=n-2}^{n} \alpha_k \left(\partial_k^2 f_h - A_h \partial_k^2 u_h\right)\right\|_{L^2(\Omega)}.$$

□

Remark 20. *Note, that as in the case of the 3-point error estimator, the approximation of initial conditions and right-hand-side function is crucial for optimality of our time and space error estimators.*

3.2 Numerical results

3.2.1 The 5-point error estimator for a second order ordinary differential equation

We first present numerical comparison of the 3-point error estimator and the 5-point error estimator at the case of ordinary differential equation of second order (1.64). The Newmark scheme reduces in this case to

$$\begin{aligned}
\frac{u^{n+1}-u^n}{\tau_n} - \frac{u^n - u^{n-1}}{\tau_{n-1}} + A\frac{\tau_n(u^{n+1}+u^n)+\tau_{n-1}(u^n+u^{n-1})}{4} &= \\
= \frac{\tau_n(f^{n+1}+f^n)+\tau_{n-1}(f^n+f^{n-1})}{4},\ 1 \le n \le N-1,& \\
\frac{u^1-u^0}{\tau_0} = v_0 - \frac{\tau_0}{4}A(u^1+u^0) + \frac{\tau_0}{4}(f^1+f^0),& \\
u^0 = u_0.&
\end{aligned}$$

and the error becomes

$$e = \max_{0\le n\le N}\left(\left|v^n - u'(t_n)\right|^2 + A\left|u^n - u(t_n)\right|^2\right)^{1/2}.$$

The 3-point *a posteriori* error estimate simplifies to this form:

$$\begin{aligned}
e \le \sum_{k=0}^{n-1}\tau_k\eta_T(t_k) = \tau_0\left(\frac{5}{12}\tau_0^2 + \frac{1}{2}\tau_0\tau_1\right)\sqrt{A(\partial_1^2 v)^2 + (\partial_1^2 f - A\partial_1^2 u)^2} \\
+ \sum_{k=1}^{n-1}\tau_k\left(\frac{1}{12}\tau_k^2 + \frac{1}{8}\tau_{k-1}\tau_k\right)\sqrt{A(\partial_k^2 v)^2 + (\partial_k^2 f - A\partial_k^2 u)^2},
\end{aligned}$$

$\forall n:\ 0 \le n \le N$. Similarly the 5-point *a posteriori* error estimate simplifies to this form:

$$e \le \sum_{k=3}^{n-1}\hat{\eta}_T(t_k) = \sum_{k=3}^{n-1}\tau_k\left(\frac{1}{12}\tau_k^2 + \frac{1}{8}\tau_{k-1}\tau_k\right)\sqrt{A(\partial_k^2 v)^2 + (\partial_k^4 u)^2},$$

$\forall n:\ 4 \le n \le N$.

A	N	η_T	$\hat{\eta}_T$	e	ei_T	$\hat{ei}_T$
100	100	.21	.203	.085	2.47	2.39
100	1000	.0021	.0021	8.34e-04	2.5	2.49
100	10000	2.08e-05	2.08e-05	8.35e-06	2.5	2.5
1000	100	20.51	19.47	8.35	2.46	2.33
1000	1000	.209	.208	.084	2.5	2.49
1000	10000	.0021	.0021	8.33e-04	2.5	2.5
10000	100	1.68e+03	1.4e+03	200	8.38	6.98
10000	1000	20.8	20.7	8.34	2.5	2.49
10000	10000	.208	.208	.083	2.5	2.5

TABLE 3.1: Effective indices for constant time steps and $f = 0$.

We define the following effectivity indices in order to measure the quality of our estimators η_T and $\hat{\eta}_T$

$$ei_T = \frac{\eta_T}{e}, \quad \hat{ei}_T = \frac{\hat{\eta}_T}{e}.$$

We present in Table 3.1 the results for equation (1.64) setting $f = 0$, the exact solution $u = cos(\sqrt{A}t)$, final time $T = 1$, and using constant time steps $\tau = T/N$. We observe that the 3-point and the 5-point estimators are divided by about 100 when the time step τ is divided by 10. The true error e also behaves as $O(\tau^2)$ and hence both time error estimators behave as the true error.

In order to check behavior of time error estimators for variable time step (see Table 3.2) we take the previous example with time step $\forall n : \ 0 \leq n \leq N$

$$\tau_n = \begin{cases} 0.1\tau_*, & \text{if } mod(n, 2) = 0, \\ \tau_*, & \text{if } mod(n, 2) = 1, \end{cases} \tag{3.26}$$

where τ_* is a given fixed value. As in the case of constant time step we have the equivalence between the true error and both estimated errors. We have plotted on Fig 3.1 evolution in time of the values $\sum_{k=0}^{n-1} \eta_T(t_k)$ and $\sum_{k=3}^{n-1} \hat{\eta}_T(t_k)$ compared to e.

Table 3.3 contains the results for even more non-uniform time step $\forall n : \ 0 \leq n \leq N$

$$\tau_n = \begin{cases} 0.01\tau_*, & \text{if } mod(n, 2) = 0, \\ \tau_*, & \text{if } mod(n, 2) = 1, \end{cases} \tag{3.27}$$

on otherwise the same test case. Note that in the case when $A = 100$ and $N = 19800$ the 5-points error estimator $\hat{\eta}_T$ blows up, while the 3-point estimator behaves as the true error. This effect is consistent with Theorem 20. Indeed, the constants in the bounds of this Theorem may depend on the meshes regularity in time.

Our conclusion is thus that for toy model classic and alternative *a posteriori* error estimators are sharp on both constant and variable time grids.

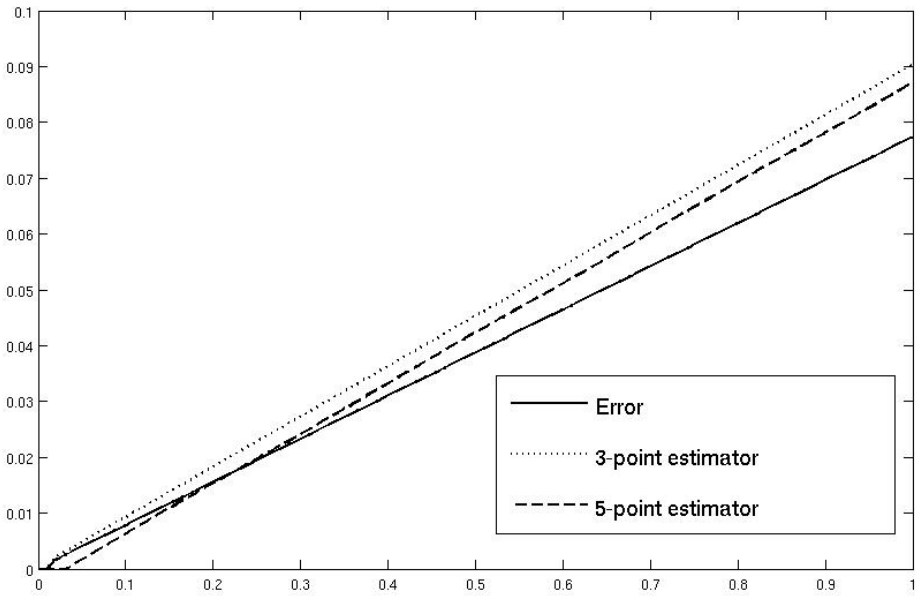


FIGURE 3.1: Evolution in time of the 3-point and the 5-point time estimators for variable time step (3.26), $A = 100$, $N = 180$, $T = 1$.

A	N	η_T	$\hat{\eta}_T$	e	ei_T	$\hat{ei}_T$
100	180	.09	.087	.077	1.17	1.13
100	1816	8.85e-04	8.82 e-04	7.59e-04	1.17	1.16
100	18180	8.83e-06	8.83e-06	7.6e-06	1.16	1.16
1000	180	8.91	8.52	7.6	1.17	1.13
1000	1816	.089	.088	.076	1.17	1.16
1000	18180	8.84e-04	8.83e-04	7.59e-04	1.16	1.16
10000	180	802.84	725.1	200	4.01	3.63
10000	1816	8.84	8.8	7.58	1.17	1.16
10000	18180	.088	.088	.076	1.16	1.16

TABLE 3.2: Effective indices for variable time step (3.26) and $f = 0$.

3.2.2 The 5-point error estimator for the wave equation on unstructured mesh

We now report numerical results for initial boundary-value problem for the wave equation with non-uniform time steps when using the 3-point time error estimator (2.19, 2.21) and the 5-point time error estimator (29).

We compute two parts of the space estimator (2.22) in practice by (2.59). The quality of our error estimators in space and time is determined by following effectivity indices:

$$ei = \frac{\eta_T + \eta_S}{e}, \quad \hat{ei} = \frac{\hat{\eta}_T + \eta_S}{e},$$

A	N	η_T	$\hat{\eta}_T$	e	ei_T	$\hat{ei}_T$
100	196	.086	.083	.084	1.02	0.98
100	1978	8.39e-04	8.36 e-04	8.26e-04	1.02	1.01
100	19800	8.38e-06	1.82e-05	8.1e-06	1.03	2.24
1000	196	8.47	8.1	8.26	1.02	0.98
1000	1978	.083	.084	.0827	1.02	1.01
1000	19800	8.37e-04	8.37e-04	8.26e-04	1.01	1.01
10000	196	764.2	691.7	200	3.82	3.46
10000	1978	8.39	8.35	8.25	1.02	1.01
10000	19800	.084	.084	.083	1.01	1.01

TABLE 3.3: Effective indices for variable time step (3.27) and $f = 0$.

where the first index contains the 3-point time error estimator and space estimator, while the last measures the grade of the 5-point time error estimator and space estimator. The true error is defined in (2.60).

Consider the problem (1.59) with $\Omega = (0,1) \times (0,1),\ T = 1$ and the exact solution u given by

$$u(x, y, t) = e^{-100r^2(x,y,t)}, \tag{3.28}$$

where

$$r^2(x, y, t) = (x - 0.3 - 0.4t^2)^2 + (y - 0.3 - 0.4t^2)^2. \tag{3.29}$$

Thus, u is a Gaussian function, whose center moves from point $(0.3, 0.3)$ at $t = 0$ to point $(0.7, 0.7)$ at $t = 1$. The transport velocity $0.8t(1,1)^T$ is peaking at $t = 1$. We choose non-uniform time step as

$$\tau_n = \frac{\tau_0}{\sqrt{t_n}},$$

for $n = 1, \ldots, N-1$. We interpolate initial conditions with H_0^1-orthogonal projection

$$u_h^0 = \Pi_h u^0,\ v_h^0 = \Pi_h v^0,$$

and right-hand-side function with L^2-orthogonal projection as it noted in previous chapter

$$f_h^n = P_h f^n,\ 0 \le n \le N.$$

Unstructured Delaunay meshes in space are used in all the experiments. Numerical results are reported in Table 3.4. Note that this case is chosen so that the non-uniform time step is required, see Fig.3.2.

Referring to Table 3.4, we observe that when setting initial time step as $\tau^2 \sim O(h)$ the error is divided by 2 each time h is divided by 2, consistent with $e \sim O(\tau^2 + h)$. The space error estimator and the two time error estimators behave similarly and thus provide a good representation of the true error. Both effectivity indices tend to a constant value.

We therefore conclude that our space and time error estimators are sharp in the regime of non-uniform time steps and Delaunay space meshes. They separate well

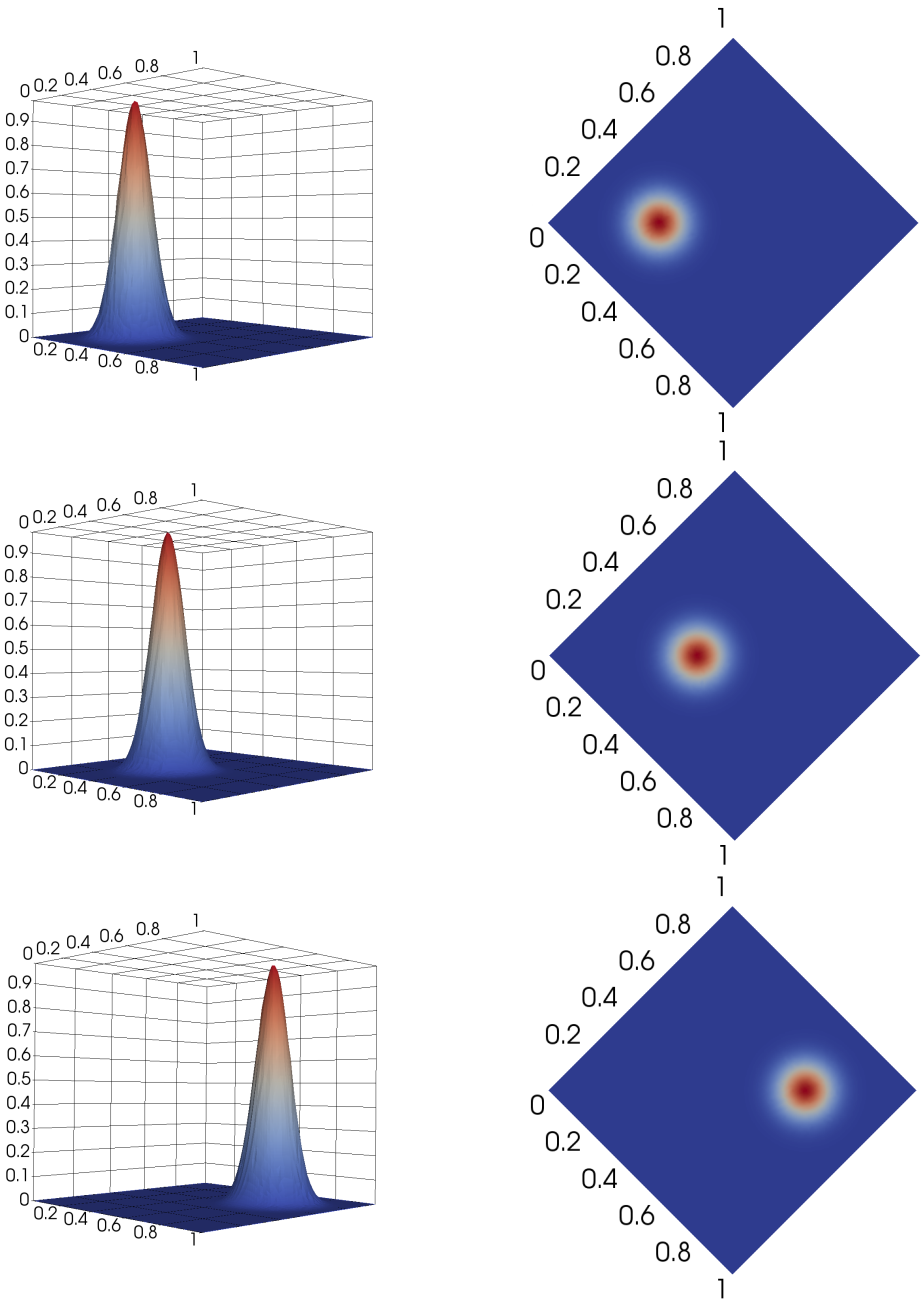


FIGURE 3.2: Solution (3.28). From up to down: time $t = 0,\ 0.5,\ 1$.

h	τ_0	ei	$\hat{ei}$	η_T	$\hat{\eta}_T$	η_S	τ_F	N_{ts}	e
.05	.01	4.85	4.83	.096	.088	2.55	.0063	105	.58
.025	.0071	5.39	5.38	.054	.051	1.39	.0045	149	.27
.0125	.005	5.94	5.93	.028	.026	.72	.0032	210	.13
.00625	.0035	5.94	5.94	.014	.013	.36	.0022	297	.065
.003125	.0025	5.94	5.94	.0067	.0065	.18	.0016	421	.032

TABLE 3.4: Results for case (3.28), non-uniform time step.

the two sources of the error and can be thus used for the mesh adaptation in space and time. In particularly, the 3-point and the 5-point time estimators become more and more close to each other when h and τ tend to 0.

Chapter 4

Numerical study and comparisons for a general second order Newmark scheme

---

The goal of this section is to extend results our *a posteriori* error analysis to general second order Newmark schemes. For the sake of brevity we work only with the case of second order ordinary differential equation, not taking into account the space discretization. We derive a *a posteriori* error estimator for a general Newmark scheme of second order ($\gamma = 1/2$, β arbitrary), generalizing Theorem 32. This is the main result of this section. Numerical experiments confirm that the convergence rate of the time error estimator is similar to that of the true error, taking explicit ($\beta = 0$) and implicit conservative ($\beta = 1/4$) second order Newmark schemes as examples.

Another time error estimator for a general Newmark scheme has been already proposed in [Geo+16]. This estimator is based on a reformulation of the scheme as a discretization of the first order system on staggered grids, while our estimator uses rather the analogy with Crank-Nicolson discretization. The estimator from [Geo+16] is thus limited to uniform meshes in time, while our 3-point time error estimator is derived for non constant time steps and can thus be used for mesh adaptation in time. We present numerical comparison between our time error estimator and that of [Geo+16] in the case of constant time step. We conclude that both two time estimators are sharp and reliable.

The outline of the chapter is as follows. We present the general 3-point time error estimator and its derivation in the first section. In the second section we give an explicit formula to compute the staggered grids time error estimator from [Geo+16]. Numerical experiments for several test cases are given in the last section.

Chapter contents

4.1 The 3-point error estimator

We turn back to the second order ordinary differential equation, cf. (1.64)

$$\begin{cases} \dfrac{d^2u(t)}{dt^2} + Au(t) = f(t), \quad t \in [0;T]\,, \\ u(0) = u_0, \\ u'(0) = v_0, \end{cases} \tag{4.1}$$

with a constant $A > 0$.

The general Newmark scheme [New59b] for it can be written as follows: given u_0, v_0 compute u^{n+1}, v^{n+1} for $n = 0, \ldots, N-1$ from

$$u^{n+1} = u^n + \tau_n v^n + \tau_n^2 \frac{1-2\beta}{2} w^n + \tau_n^2 \beta w^{n+1}, \tag{4.2}$$

$$v^{n+1} = v^n + (1-\gamma)\tau_n w^n + \gamma \tau_n w^{n+1}, \tag{4.3}$$

where u^n is an approximation for $u(t_n)$, v^n is an approximation for $u'(t_n)$, $\tau_n = t_{n+1} - t_n$ is the time step, w^n is a shorthand for $f^n - Au^n$ with $f^n = f(t_n)$, β and γ are some coefficients such that $0 \le \gamma \le 1$ and $0 \le \beta \le \dfrac{1}{2}$.

In order to obtain the second order method, we adopt the particular choice for the parameter $\gamma = \dfrac{1}{2}$ in (4.3). Thus we have

$$u^{n+1} = u^n + \tau_n v^n + \tau_n^2 \frac{1-2\beta}{2} w^n + \tau_n^2 \beta w^{n+1}, \tag{4.4}$$

$$v^{n+1} = v^n + \tau_n \frac{w^n + w^{n+1}}{2}, \tag{4.5}$$

for $n = 0, \ldots, N-1$. Hereafter we call the scheme (4.4), (4.5) as general second order Newmark scheme. Setting β to various values gives a wide range of schemes. The popular choices are $\beta = 0$ that gives an explicit scheme and $\beta = \dfrac{1}{4}$ that gives an implicit conservative scheme with respect to the energy norm.

Further we show that a general second order Newmark scheme (4.4)–(4.5) for equation 4.1 can be rewritten with two different ways: as a cosine-type second order method or as a perturbed Crank-Nicolson scheme.

First reformulation - cosine scheme

Dividing (4.4) by τ_n, doing the same at the previous time step and subtracting one from another, yields to

$$\frac{u^{n+1}-u^n}{\tau_n} - \frac{u^n - u^{n-1}}{\tau_{n-1}} = v^n - v^{n-1} + \tau_n \frac{1-2\beta}{2} w^n + \tau_n \beta w^{n+1} - \tau_{n-1} \frac{1-2\beta}{2} w^{n-1} - \tau_{n-1} \beta w^n.$$

Using (4.5) at the previous time step to eliminate v^n and recalling that $w^n = f^n - Au^n$, we get

$$\frac{u^{n+1}-u^n}{\tau_n} - \frac{u^n - u^{n-1}}{\tau_{n-1}} + A\left(\tau_{n-1}\beta u^{n-1} + (\tau_{n-1}+\tau_n)\frac{1-2\beta}{2}u^n + \tau_n\beta u^{n+1}\right) = \tau_{n-1}\beta f^{n-1} + (\tau_{n-1}+\tau_n)\frac{1-2\beta}{2}f^n + \tau_n\beta f^{n+1}, \quad (4.6)$$

for $n = 0, \dots, N-1$. This is the cosine method, cf. [Geo+16] with $q_1 = \beta$. When $\beta = 0$ the above method is called the leap-frog method.

Second reformulation - Crank-Nicolson scheme

We are now going to rearrange the equations (4.4)–(4.5) so that they look like a discretization of the first order system for $u(t)$ and $v(t) = u'(t)$. The resulting scheme turns out to be a perturbation of Crank-Nicolson scheme, which will be useful in the subsequent analysis.

We start from rewriting (4.5) as

$$v^n = \frac{v^{n+1}+v^n}{2} - \tau_n\frac{w^n + w^{n+1}}{4},$$

and substitute this into (4.4) to get

$$\frac{u^{n+1}-u^n}{\tau_n} = \frac{v^{n+1}+v^n}{2} + \tau_n\left(\beta - \frac{1}{4}\right)(w^{n+1} - w^n).$$

Thus, the original scheme can be also written as

$$\frac{u^{n+1}-u^n}{\tau_n} - \frac{v^{n+1}+v^n}{2} = \delta^n, \quad (4.7)$$

$$\frac{v^{n+1}-v^n}{\tau_n} + A\frac{u^{n+1}+u^n}{2} = \frac{f^{n+1}+f^n}{2}, \quad (4.8)$$

for $n = 0, \dots, N-1$ with

$$\delta^n = \tau_n\left(\beta - \frac{1}{4}\right)(f^{n+1} - Au^{n+1} - f^n + Au^n).$$

Remark 21. *The fact that for wave equation (4.1) the Newmark method (4.4), (4.5) can be rewritten as the Crank-Nicolson scheme for the first-order system (4.7), (4.8) plays the key role in a posteriori error analysis for the second order methods for the wave equation. Indeed, we already saw that the derivation of the 3-point and the 5-point error estimators in sections 2, 3 is based on the scheme (4.7), (4.8).*

4.1.1 *A priori* error estimate

We turn now to *a priori* error analysis for the scheme (4.4)-(4.5). In this chapter we use the following notations

$$u^{n+\frac{1}{2}} = \frac{u^{n+1}+u^n}{2}, \quad \partial_n u = \frac{u^{n+1}-u^{n-1}}{\tau_n+\tau_{n-1}},$$
$$\partial_n^2 u = \left(\frac{\tau_{n-1}+\tau_n}{2}\right)^{-1}\left(\frac{u^{n+1}-u^n}{\tau_n} - \frac{u^n-u^{n-1}}{\tau_{n-1}}\right),$$
$$\partial_{n+\frac{1}{2}} u = \frac{u^{n+1}-u^n}{\tau_n}, \quad \tau_{n-\frac{1}{2}} = \frac{\tau_n+\tau_{n-1}}{2}, \quad f^{n+\frac{1}{2}} = \frac{f^{n+1}+f^n}{2}.$$

which we shall apply to any quantities numbered with a superscript n.

As we already mentioned the schemes above with $\gamma = \frac{1}{2}$ are of 2nd order in time for both u^n and v^n. This is not particularly evident if one looks at the reformulation (4.6), but it is clear from the original form (4.4)–(4.5) or (4.7)–(4.8). The precise statement is given in the following theorem. We shall need there a CFL-type condition on the time step

$$\tau_n^2\left(\frac{1}{4}-\beta\right) < \frac{1}{A}. \tag{4.9}$$

This restriction is of course automatically satisfied for $\beta \geq \frac{1}{4}$.

Theorem 31. *Let u be a smooth solution of the ordinary differential equation (4.1) and u^n, v^n be the discrete solution of the scheme (2.2)-(2.3). Assume that (4.9) holds for all the time steps τ_n. Then, the following a priori error estimate holds*

$$\max_{0\leq n\leq N}\left(\left|v^n-u'(t_n)\right|^2+|u^n-u(t_n)|^2\right)^{1/2} \leq C\sum_{n=0}^{N-1}\tau_n^2\left(\int_{t_n}^{t_{n+1}}\sqrt{A}\left|u'''(t)\right|dt+\int_{t_n}^{t_{n+1}}\left|u''''(t)\right|dt\right) \tag{4.10}$$

with a constant C which does not depend on the mesh or on the final time.

Proof. We follow the proof of *a priori* error estimate from Section 2. Let us introduce $e_u^n = u^n - u(t_n)$ and $e_v^n = v^n - v(t_n)$. We also introduce the notation

$$B = 1 + A\tau_n^2\left(\beta - \frac{1}{4}\right),$$

and note $B > 0$ thanks to (4.9).

Observe that the following equations hold

$$B\partial_{n+1/2}e_u - e_v^{n+1/2} = \tau_n^2\left(\beta - \frac{1}{4}\right)\frac{f(t_{n+1}) - f(t_n)}{\tau_n} - B\frac{u(t_{n+1}) - u(t_n)}{\tau_n} + \frac{v(t_{n+1}) + v(t_n)}{2}, \tag{4.11}$$

$$\partial_{n+1/2}e_v + Ae_u^{n+1/2} = \frac{u''(t_{n+1}) + u''(t_n)}{2} - \frac{v(t_{n+1}) - v(t_n)}{\tau_n}. \tag{4.12}$$

The first equation above is a direct consequence of (4.7). The last equation is a direct consequence of (4.8) together with the governing equation (4.1) evaluated at times t_n and t_{n+1}.

Multiplying (4.11) by $Ae_u^{n+1/2}$ and (4.12) by $e_v^{n+1/2}$ and taking the sum of (4.11) and (4.12) yields

$$\frac{AB|e_u^{n+1}|^2 - AB|e_u^n|^2 + |e_v^{n+1}|^2 - |e_v^n|^2}{2\tau_n} = AR_1^n e_u^{n+1/2} + R_2^n e_v^{n+1/2}, \tag{4.13}$$

with

$$R_1^n = \left(\tau_n^2\left(\beta - \frac{1}{4}\right)\frac{u''(t_{n+1}) - u''(t_n)}{\tau_n} - \frac{u(t_{n+1}) - u(t_n)}{\tau_n} + \frac{v(t_{n+1}) + v(t_n)}{2}\right),$$

$$R_2^n = \left(\frac{u''(t_{n+1}) + u''(t_n)}{2} - \frac{v(t_{n+1}) - v(t_n)}{\tau_n}\right).$$

Set

$$E^n = \left(AB\,|e_u^n|^2 + |e_v^n|^2\right)^{1/2},$$

so that equality (4.13) with Cauchy-Schwarz inequality entails

$$\frac{(E^{n+1})^2 - (E^n)^2}{2\tau_n} \le \left(A|R_1^n|^2 + |R_2^n|^2\right)^{1/2}\frac{E^{n+1} + E^n}{2},$$

which implies

$$E^{n+1} - E^n \le \tau_n\left(\sqrt{A}|R_1^n| + |R_2^n|\right).$$

Summing this over n from 0 to $N-1$ gives

$$(AB|e_u^N|^2 + |e_v^N|^2)^{1/2} \le \sum_{n=0}^{N-1}\tau_n(\sqrt{A}|R_1^n| + |R_2^n|).$$

or

$$(A|e_u^N|^2 + |e_v^N|^2)^{1/2} \le c\sum_{n=0}^{N-1}\tau_n(\sqrt{A}|R_1^n| + |R_2^n|) \tag{4.14}$$

with $c = \max(1, 1/B)$.

We have the following estimates for R_1^n and R_2^n

$$|R_1^n| \leq C\tau_n \int_{t_n}^{t_{n+1}} \left|u'''(t)\right| dt, \tag{4.15}$$

$$|R_2^n| \leq C\tau_n \int_{t_n}^{t_{n+1}} \left|u''''(t)\right| dt, \tag{4.16}$$

which can be proved by Taylor expansion similarly to analogous estimates for (2.14)–(2.15) in *a priori* estimates from Chapter 2. Substituting (4.15)–(4.16) into (4.14) yields (4.10). □

4.1.2 *A posteriori* error estimate

Let us adopt the vector notations

$$U^n = \begin{pmatrix} u^n \\ v^n \end{pmatrix}, F^n = \begin{pmatrix} \delta^n \\ \frac{f^{n+1}+f^n}{2} \end{pmatrix},$$

so that the scheme (4.7)–(4.8) can be written as

$$\partial_{n+\frac{1}{2}} U + \mathcal{A} U^{n+\frac{1}{2}} = F^n, \tag{4.17}$$

Let us introduce the piecewise quadratic reconstruction $\tilde{U}_\tau$ of U^n such that $\tilde{U}_\tau(t)$ is a continuous function that is equal on $[t_n, t_{n+1}]$, $n \geq 1$ to the quadratic polynomial that coincides with U^{n+1} (resp. U^n, U^{n-1}) at time t_{n+1} (resp. t_n, t_{n-1}). We have thus for $t \in [t_n, t_{n+1}]$, $n \geq 1$

$$\tilde{U}_\tau(t) = U^{n+\frac{1}{2}} + (t - t_{n+\frac{1}{2}})\partial_{n+\frac{1}{2}} U + \frac{1}{2}(t - t_{n+1})(t - t_n)\partial_n^2 U. \tag{4.18}$$

Moreover, $\tilde{U}_\tau(t)$ is defined on $[t_0, t_1]$ as the quadratic polynomial that coincides with U^2 (resp. U^1, U^0) at time t_2 (resp. t_1, t_0). Similarly we introduce piecewise quadratic reconstruction $\tilde{f}_\tau(t)$ based on f^n.

We are now ready to derive time *a posteriori* error estimate for general Newmark scheme of second order.

Theorem 32. *The following a posteriori error estimate holds between the solution u of problem (4.1) and the solution u^n provided by the scheme (4.6) or, equivalently, by (4.7)–(4.8)*

$$\left(\left|v^n - u'(t_n)\right|^2 + A|u^n - u(t_n)|^2\right)^{\frac{1}{2}} \leq \sum_{k=0}^{n-1} \tau_k \eta_T(t_k), \tag{4.19}$$

where the error indicator for $k = 1, \ldots, N-1$ is

$$\eta_T(t_k) = \left(\frac{1}{12}\tau_k + \frac{1}{8}\tau_{k-1}\right) \tau_k \sqrt{A(\partial_k^2 v)^2 + (\partial_k^2 f - A\partial_k^2 u)^2} + \int_{t_k}^{t_{k+1}} \left|f - \tilde{f}_\tau\right| + \int_{t_k}^{t_{k+1}} \sqrt{A}\,|R_{\delta^k}|\,, \tag{4.20}$$

and

$$\eta_T(t_0) = \left(\frac{5}{12}\tau_0 + \frac{1}{2}\tau_1\right)\tau_0\sqrt{A(\partial_1^2 v)^2 + (\partial_1^2 f - A\partial_1^2 u)^2} + \int_{t_0}^{t_1}\left|f - \tilde{f}_\tau\right| + \int_{t_0}^{t_1}\sqrt{A}\,|R_{\delta^1}|\,, \quad (4.21)$$

where

$$R_{\delta^n} = \delta^n + (t - t_{n+\frac{1}{2}})\frac{\delta^n - \delta^{n-1}}{\tau_{n-\frac{1}{2}}}. \quad (4.22)$$

Proof. We use the scheme (4.17) to derive the differential equation for the reconstruction $\tilde{U}_\tau$ for $t \in [t_n, t_{n+1}]$, $n \geq 1$

$$\frac{d\tilde{U}_\tau}{dt} + \mathcal{A}\tilde{U}_\tau = (t - t_{n+\frac{1}{2}})\left[\partial_n^2 U + \mathcal{A}\partial_{n+\frac{1}{2}}U\right] + \frac{1}{2}(t - t_{n+1})(t - t_n)\mathcal{A}\partial_n^2 U + F^n.$$

Now, reusing the scheme on two time slabs, we get

$$\partial_n^2 U + \mathcal{A}\partial_n U = \frac{F^n - F^{n-1}}{\tau_{n-\frac{1}{2}}},$$

or

$$\partial_n^2 U + \mathcal{A}\left(\partial_{n+\frac{1}{2}}U - \frac{\tau_{n-1}}{2}\partial_n^2 U\right) = \frac{F^n - F^{n-1}}{\tau_{n-\frac{1}{2}}},$$

so that the differential equation can be rewritten as

$$\frac{d\tilde{U}_\tau}{dt} + \mathcal{A}\tilde{U}_\tau = \left(\frac{\tau_{n-1}}{2}(t - t_{n+\frac{1}{2}}) + \frac{1}{2}(t - t_{n+1})(t - t_n)\right)\mathcal{A}\partial_n^2 U + (t - t_{n+\frac{1}{2}})\frac{F^n - F^{n-1}}{\tau_{n-\frac{1}{2}}} + F^n.$$

Let us introduce the vectorial error

$$E(t) = \tilde{U}_\tau(t) - U(t) = \begin{pmatrix} E_u(t) \\ E_v(t) \end{pmatrix}, \quad (4.23)$$

where

$$U(t) = \begin{pmatrix} u(t) \\ u'(t) \end{pmatrix}. \quad (4.24)$$

Then, for $t \in (t_n, t_{n+1}]$

$$\frac{dE}{dt} + \mathcal{A}E = p_n(t)\mathcal{A}\partial_n^2 U + F^n + (t - t_{n+\frac{1}{2}})\frac{F^n - F^{n-1}}{\tau_{n-\frac{1}{2}}} - F(t), \quad (4.25)$$

where

$$p_n(t) = \frac{\tau_{n-1}}{2}(t - t_{n+\frac{1}{2}}) + \frac{1}{2}(t - t_{n+1})(t - t_n).$$

We take the scalar product of previous equality with $\begin{pmatrix} AE_u \\ E_v \end{pmatrix}$ to obtain

$$\frac{1}{2}\frac{d(A(E_u)^2+(E_v)^2)}{dt} = p_n(t)\begin{pmatrix} -\partial_n^2 v \\ A\partial_n^2 u - \partial_n^2 f \end{pmatrix}\cdot\begin{pmatrix} AE_u \\ E_v \end{pmatrix} + \begin{pmatrix} \delta^n + (t-t_{n+\frac{1}{2}})\frac{\delta^n-\delta^{n-1}}{\tau_{n-\frac{1}{2}}} \\ \tilde{f}_\tau - f \end{pmatrix}\cdot\begin{pmatrix} AE_u \\ E_v \end{pmatrix}.$$

Denoting

$$Z(t) = \sqrt{A(E_u(t))^2+(E_v(t))^2},$$

we get by Cauchy-Schwartz inequality for $t \in [t_n, t_{n+1}]$, $n \geq 1$

$$Z\frac{dZ}{dt} \leq \Big[A\left(-p_n(t)\partial_n^2 v + \delta^n + (t-t_{n+\frac{1}{2}})\frac{\delta^n-\delta^{n-1}}{\tau_{n-\frac{1}{2}}}\right)^2 + (p_n(t)(A\partial_n^2 u - \partial_n^2 f) + \tilde{f}_\tau - f)^2\Big]^{\frac{1}{2}} Z.$$

From the definition of the piecewise quadratic reconstruction $\tilde{U}_\tau(t)$ for $t \in [t_0, t_1]$ follows that a similar inequality holds for $t \in [t_0, t_1]$

$$Z\frac{dZ}{dt} \leq \Big[A\left(-p_1(t)\partial_1^2 v + \delta^1 + (t-t_{\frac{3}{2}})\frac{\delta^1-\delta^0}{\tau_{\frac{1}{2}}}\right)^2 + (p_1(t)(A\partial_1^2 u - \partial_1^2 f) + \tilde{f}_\tau - f)^2\Big]^{\frac{1}{2}} Z.$$

Simplify two previous inequalities by Z on both sides and integrate to get (4.19). □

Remark 22. *The potentially dangerous feature in the estimate (4.19) is in the terms with δ^k. Indeed, they contain $A^{\frac{3}{2}}u^n$ which is $\sqrt{A}$ times more than other terms. Note that this terms are not present in the case $\beta = \frac{1}{4}$. In Section 4.3 we numerically compare the cases $\beta = \frac{1}{4}$ and $\beta = 0$ in terms of the accuracy of the estimator and the accuracy of the scheme itself if A is big.*

4.2 The staggered grids error estimator

In this subsection we give the explicit formula to compute the staggered grids *a posteriori* error estimator from (1.108)

$$\eta_T^{SG} = 2\left(A|u-\tilde{u}|^2 + |u'-\tilde{v}|^2\right)^{1/2}(0) + 4\int_0^{t^N}\left(A|R_2|^2 + |R_1|^2\right)^{1/2} dt. \quad (4.26)$$

Taking into account the fact that the reconstruction $\hat{u}(t)$ is defined on $(t_{n-1}, t_n]$ for $n = 1, \ldots, N-1$ while $\overline{u}(t)$ and R_u are defined on $(t_{n-1/2}, t_{n+1/2}]$ for $n = 0, \ldots, N-1$ (respectively, the reconstruction $\hat{v}(t)$ is defined on $(t_{n-1/2}, t_{n+1/2}]$ for $n = 0, \ldots, N-1$ while $\overline{v}(t)$ and R_v are defined on $(t_{n-1}, t_n]$ for $n = 1, \ldots, N-1$) we compute *a*

posteriori error estimator as follows:

$$\eta_T^{SG}\,|_{(t_{n-1},t_n]} = \int_{t_{n-1}}^{t_{n-1/2}} \Big(A\,|R_2|^2 + |R_1|^2\Big)^{1/2}\,dt + \int_{t_{n-1/2}}^{t_n} \Big(A\,|R_2|^2 + |R_1|^2\Big)^{1/2}\,dt, \quad (4.27)$$

for $n = 1, \ldots, N-1$. We compute (1.109) using linear reconstructions $\tilde{u}$ and $\tilde{v}$:

$$\begin{aligned} R_1 &= -A(\hat{u} - \tilde{u}) - A(\tilde{u} - \overline{u}) - R_u, \\ R_2 &= \hat{v} - \tilde{v} + \tilde{v} - \overline{v} - R_v, \end{aligned} \quad (4.28)$$

where

$$\begin{aligned} \overline{u}\,|_{(t_{n-1/2},t_{n+1/2}]}\,(t) &= u^{n-1/2} + (t - t_{n-1/2})\frac{u^{n+1/2} - u^{n-1/2}}{\tau}, \\ \overline{v}\,|_{(t_{n-1},t_n]}\,(t) &= v^{n-1} + (t - t_{n-1})\frac{v^n - v^{n-1}}{\tau}, \\ \tilde{u}\,|_{(t_{n-1},t_n]}\,(t) &= u^{n-1} + (t - t_{n-1})\frac{u^n - u^{n-1}}{\tau}, \\ \tilde{v}\,|_{(t_{n-1/2},t_{n+1/2}]}\,(t) &= v^{n-1/2} + (t - t_{n-1/2})\frac{v^{n+1/2} - v^{n-1/2}}{\tau}. \end{aligned} \quad (4.29)$$

Thus from (4.29), (1.106) and (1.107) we have:

$$\hat{v}\,|_{(t_{n-1/2},t_{n+1/2}]}\,(t) = v^{n-1/2} - A\Big((t - t_{n-1/2})u^{n-1/2} + (t - t_{n-1/2})^2\frac{u^{n+1/2} - u^{n-1/2}}{2\tau}\Big) + R_u(t - t_{n-1/2}), \quad (4.30)$$

$$\hat{u}\,|_{(t_{n-1},t_n]}\,(t) = u^{n-1} + (t - t_{n-1})v^{n-1} + (t - t_{n-1})^2\frac{v^n - v^{n-1}}{2\tau} + R_v(t - t_{n-1}). \quad (4.31)$$

Combining (4.31), (4.30), (4.29) and (4.28) we obtain:

$$R_1\,|_{(t_{n-1},t_{n-1/2}]}\,(t) = -A\Big((t - t_{n-1})^2\frac{v^n - v^{n-1}}{2\tau} + R_v(t - t_{n-1}) - \frac{u^n - 2u^{n-1} + u^{n-2}}{4}\Big) - R_u,$$

$$R_1\,|_{(t_{n-1/2},t_n]}\,(t) = -A\Big((t - t_{n-1})^2\frac{u^{n+1} - u^n - u^{n-1} + u^{n-2}}{4\tau^2} + R_v(t - t_{n-1}) + \frac{u^{n+1} - 2u^n + u^{n-1}}{4}\Big) - R_u,$$

$$R_2\,|_{(t_{n-1},t_{n-1/2}]}\,(t) = -A(t-t_{n-3/2})^2\frac{u^{n-1/2}-u^{n-3/2}}{2\tau} + (t-t_{n-3/2})\frac{u^n-2u^{n-1}+u^{n-2}}{4} + A(t-t_{n-3/2})\frac{u^{n-1}-u^{n-2}}{2} - (t-t_{n-1})\frac{v^{n+1/2}-2v^{n-1/2}+v^{n-3/2}}{2\tau} - R_v,$$

$$R_2\,|_{(t_{n-1/2},t_n]}\,(t) = -A(t-t_{n-1/2})^2\frac{u^{n+1/2}-u^{n-1/2}}{2\tau} + A(t-t_{n-1/2})\frac{u^n-u^{n-1}}{2} + (t-t_{n-1/2})\frac{u^{n+1}-2u^n+u^{n-1}}{4} - \frac{v^{n+1/2}-2v^{n-1/2}+v^{n-3/2}}{2} + (t-t_{n-1})\frac{v^{n+1/2}-2v^{n-1/2}+v^{n-3/2}}{2\tau} - R_v.$$

Remark 23. *Note that $A^{\frac{3}{2}}u^n$ are also presented in the staggered grids estimator (4.26), even if $\beta = \frac{1}{4}$. Indeed, we see that the reconstruction $\hat{v}(t)$ contains $A\overline{u}$. Then, $\hat{v}(t)$ is used in the error estimate inside the term $\left\|\sqrt{A}\mathcal{R}_2\right\|$ with $\mathcal{R}_2 = \hat{v} - \overline{v} - R_v$ so that the error estimate has a contribution of the type $A^{\frac{3}{2}}\overline{u}$.*

The 3-point time error estimator is different from (1.108). An advantage of our approach is that it is readily applied to the case of non constant τ_n, while the estimator from [Geo+16] is constructed only for constant time steps. Furthermore, we remind that the 3-point estimator does not involve the terms like $A^{\frac{3}{2}}u^n$ in the important special case $\beta = \frac{1}{4}$ which may lead to sharper estimates in the case of the full discretization of the wave equation in space and time.

A numerical comparison between the staggered grids error estimator and the 3-point error estimator is provided in the next section.

4.3 Numerical study of the 3-point error estimator

The objective of this section is to study the performance of the 3-point *a posteriori* error estimator (4.19) and to compare it with the staggered grids error estimator (4.26). We introduce the notation:

$$e = \max_{0\le n\le N}\left(\left|v^n - u'(t_n)\right|^2 + A\left|u^n - u(t_n)\right|^2\right)^{1/2}.$$

We restrict ourselves with constant time step τ and right-hand side function $f = 0$ (4.1). In this case from Theorem 32 the 3-point *a posteriori* error estimate has the

following form:

$$e \leq \eta_T = \tau\Bigg(\frac{11}{12}\tau^2\sqrt{A(\partial_1^2 v)^2 + (\partial_1^2 f - A\partial_1^2 u)^2} + \frac{1}{\tau}\int_{t_0}^{t_1}\sqrt{A}\left|\delta^1 + (t - t_{\frac{3}{2}})\frac{\delta^1 - \delta^0}{\tau}\right| dt\Bigg) + \sum_{k=1}^{N-1}\tau\Bigg(\frac{5}{24}\tau^2\sqrt{A(\partial_k^2 v)^2 + (\partial_k^2 f - A\partial_k^2 u)^2} + \frac{1}{\tau}\int_{t_k}^{t_{k+1}}\sqrt{A}\left|\delta^k + (t - t_{k+\frac{1}{2}})\frac{\delta^k - \delta^{k-1}}{\tau}\right| dt\Bigg), \quad (4.32)$$

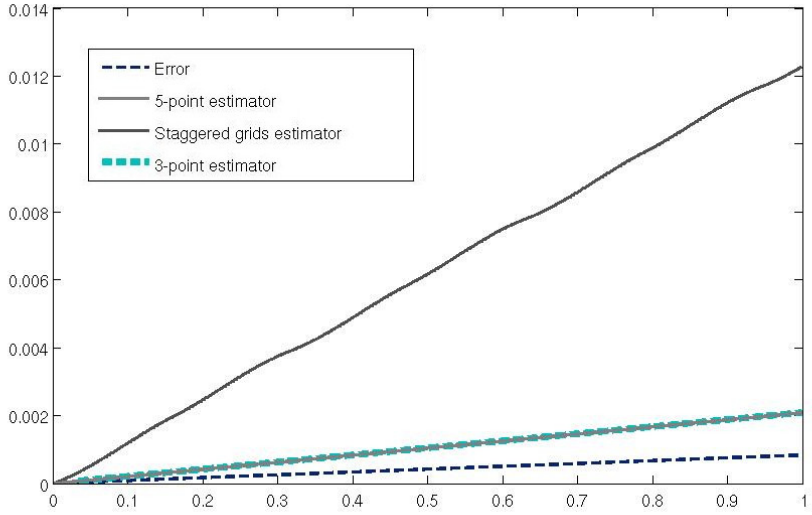


FIGURE 4.1: Evolution in time of the 3-point, the 5-point and the staggered grids time error estimators in comparison with true error, $\beta = 1/4$, $A = 100$, $N = 1000$, $T = 1$.

In order to measure the quality of the 3-point error estimator η_T and the staggered grids error estimator η_T^{SG}, we define the following effectivity indeces

$$ei_T = \frac{\eta_T}{e}, \quad ei_T^{SG} = \frac{\eta_T^{SG}}{e}.$$

We study the test case when the exact solution is given by

$$u = cos(\sqrt{A}t),$$

we take the final time $T = 1$. In Table 4.1 we present the convergence results for implicit Newmark scheme when $\beta = \frac{1}{4}$. We observe that the 3-point and the staggered grids estimators are divided by about 100 when the time step τ is divided by 10. The true error e also behaves as $O(\tau^2)$ and hence both time error estimators behave as the true error.

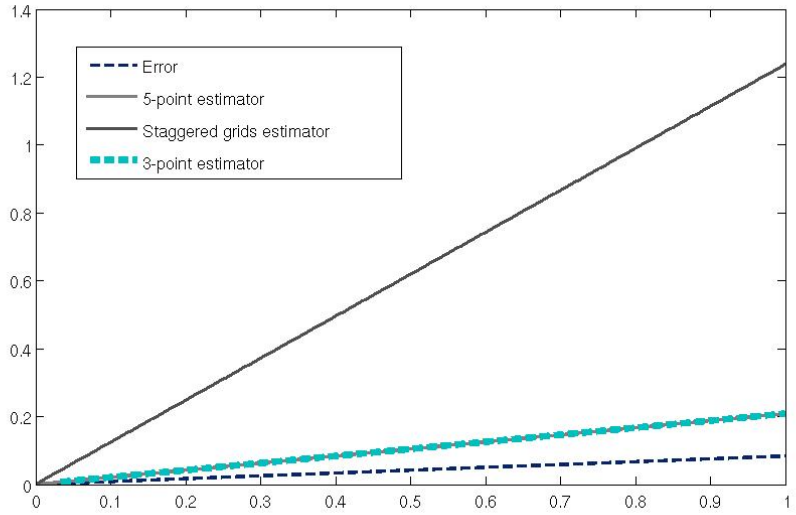


FIGURE 4.2: Evolution in time of the 3-point, the 5-point and the staggered grids time error estimators in comparison with true error, $\beta = 1/4$, $A = 10000$, $N = 10000$, $T = 1$.

We have plotted on Fig. 4.1 and on Fig. 4.2 the evolution in time of the values η_T and η_T^{SG} compared to the actual error e.

A	N	η_T	$\hat{\eta}_T$	η_T^{SG}	e	ei_T	$\hat{ei}_T$	ei_T^{SG}
100	100	.21	.203	1.22	.085	2.47	2.39	14.33
100	1000	.0021	.0021	.012	8.34e-04	2.5	2.49	14.7
100	10000	2.08e-05	2.08e-05	1.22e-04	8.35e-06	2.5	2.5	14.7
1000	100	20.51	19.47	119.15	8.35	2.46	2.33	14.27
1000	1000	.209	.208	1.23	.084	2.5	2.49	14.77
1000	10000	.0021	.0021	0.0123	8.33e-04	2.5	2.5	14.81
10000	100	1.68e+03	1.4e+03	9.25e+04	200	8.38	6.98	46.25
10000	1000	20.8	20.7	123.28	8.34	2.5	2.49	14.79
10000	10000	.208	.208	1.23	.083	2.5	2.5	14.84

TABLE 4.1: Convergence results and effective indices, $f = 0$, $\beta = 1/4$.

We now investigate the sharpness of the 3-point time error estimator by performing numerical experiments for explicit Newmark scheme of second order $\beta = 0$. From Table 4.2 we note the equivalence between the true error and both estimated errors. We have plotted on Fig. 4.3 and on Fig. 4.4 the evolution in time of the values η_T and η_T^{SG} compared to the actual error e.

Our conclusion is thus that for toy model the 3-point and the staggered grids *a posteriori* error estimators are sharp on constant time grids. The appreciable difference between the estimators is that the 3-point time error estimator is derived for non constant time step and thus it can be used for mesh adaptation in time.

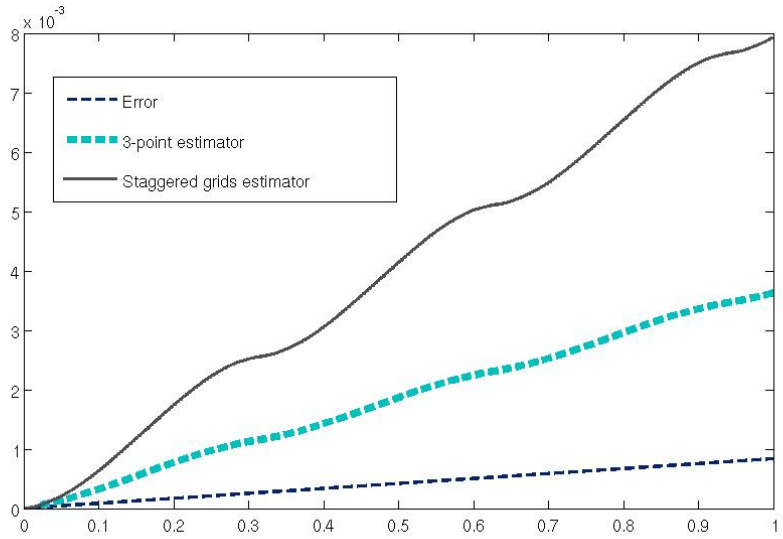


FIGURE 4.3: Evolution in time of the 3-point and the staggered grids time error estimators in comparison with true error, $\beta = 0$, $A = 100$, $N = 1000$, $T = 1$.

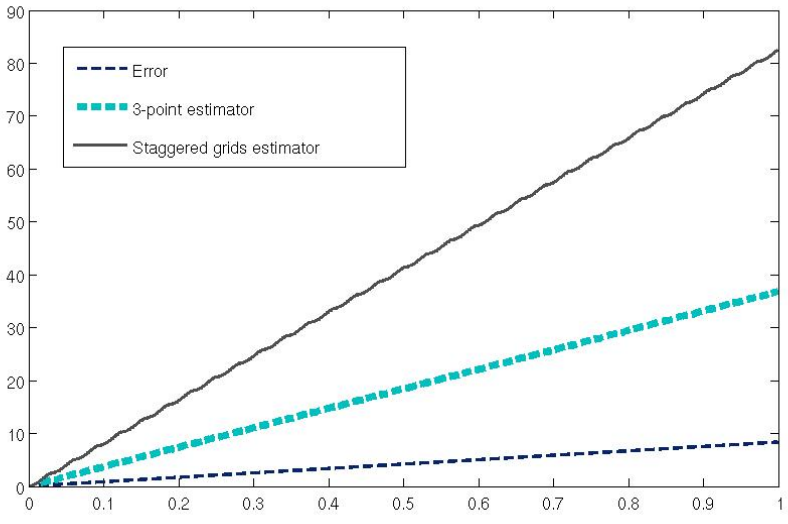


FIGURE 4.4: Evolution in time of the 3-point and the staggered grids time error estimators in comparison with true error, $\beta = 0$, $A = 10000$, $N = 1000$, $T = 1$.

A	N	η_T	η_T^{SG}	e	ei_T	ei_T^{SG}
100	100	.36	.805	.086	4.25	9.31
100	1000	.0036	.0079	8.36 e-04	4.34	9.48
100	10000	3.62e-05	7.92e-05	5.35 e-06	4.34	9.48
1000	100	37.22	85.18	8.59	4.33	9.91
1000	1000	.37	.81	.084	4.38	9.69
1000	10000	.0037	.0081	8.33 e-04	4.4	9.7
10000	100	3.58 e+03	7.59 e+03	208.72	17.19	36.39
10000	1000	36.85	82.55	8.35	4.41	9.88
10000	10000	.37	0.81	.0834	4.41	9.77

TABLE 4.2: Convergence results and effective indices, $f = 0$ and $\beta = 0$.

Chapter 5

An adaptive algorithm in space and in time

The automatic construction of adaptive meshes requires the *a posteriori* error estimates which are now classical in the case of elliptic and parabolic problems discretized in space by the finite elements. *a posteriori* error estimates of optimal order in time and space for the fully discrete wave equation in energy norm discretized with the Newmark scheme in time and with finite elements in space are presented in previous chapters. Using *a posteriori* bounds from there we have implemented an efficient adaptive algorithm.

Time error estimator contains high order discrete derivatives and thus depends from their boundedness. Numerical experiments show that an adaptation of numerical solution from mesh to mesh in general case leads to an interpolation error which cannot be neglected. Indeed, the Newmark scheme in itself produces optimal results, but higher order discrete derivatives in the time *a posteriori* error estimates can blow up with interpolation from mesh to mesh. Thus the question of choosing the interpolation from mesh to mesh is crucially important. Such behavior of time error estimator is not surprising. We have already demonstrated in Chapter 2 the sensibility of our time estimator even to approximation of the initial conditions and the right-hand-side function in the case of non-adapted mesh. This was not the case for elliptic and parabolic problems [Pic03], [LPP09], but the fact that the interpolation error between meshes cannot be neglected for the wave equation was already mentioned in [Pic10]. In this work we present the interpolation from mesh to mesh based on the orthogonal projection which allows to neglect the influence of interpolation on the time error estimator.

The outline of the chapter is as follows. In Section 5.1, we give the description of an adaptive algorithm in space and time. We also present two different ways of choosing new mesh size according to the value of *a posteriori* error estimate in space. Numerical experiments for several test cases are reported in Section 5.2. They confirm the efficiency of our adaptive algorithm. We also discuss the question of choosing an interpolation between meshes.

Chapter contents

5.1 Time and space adaptivity for the wave equation descretized by the Newmark scheme

Although the *a posteriori* error analysis in previous chapters is restricted to the single mesh $\mathcal{T}_h$, we now present an adaptive space-time algorithm which involves several meshes. From now on, for the sake of simplicity we will work only with the first part of the space error estimator :

$$\max_{0\leqslant t_k\leqslant t_n}\left[\sum_{K\in\mathcal{T}_h} h_K^2\left\|\frac{\partial\tilde{v}_{h\tau}}{\partial t}-\Delta\tilde{u}_{h\tau}-f\right\|^2_{L^2(K)}+\sum_{E\in\mathcal{E}_h} h_E\left|[n\cdot\nabla\tilde{u}_{h\tau}]\right|^2_{L^2(E)}\right]^{1/2} = \max_{0\leqslant t_k\leqslant t_n}\eta_S^1(t), \quad (5.1)$$

for $n=0,\ldots,N$. Thus the *a posteriori* error estimate (20) for $n=0,\ldots,N$ takes on the following from:

$$\max_{0\leqslant n\leqslant N}\left(\left\|v_h^n-\frac{\partial u}{\partial t}(t_n)\right\|^2_{L^2(\Omega)}+|u_h^n-u(t_n)|^2_{H^1(\Omega)}\right)^{1/2} \leq C_1\max_{0\leqslant t_k\leqslant t_n}\eta_S^1(t)+\sum_{k=1}^{n}\tau_{k-1}\eta_T(t_{k-1}), \quad (5.2)$$

where the error estimator in time η_T is defined by (2.19) and (2.21). Let set C_1 equal to 1. The goal of our adaptive algorithm is to build meshes in space and in time such that the relative estimated error at time $t_n\in[0,\ldots,t_N]$ is close to a preset tolerance TOL

$$(1-\alpha)TOL\leq\frac{\max\limits_{0\leqslant t_k\leqslant t_n}\eta_S^1(t)+\sum\limits_{k=1}^{n-1}\tau_{k-1}\eta_T(t_{k-1})}{\max\limits_{0\leqslant k\leqslant n}\left(\left\|v_h^k\right\|^2_{L^2(\Omega)}+\left|u_h^k\right|^2_{H^1(\Omega)}\right)^{1/2}}\leq(1+\alpha)TOL, \quad (5.3)$$

where $0<\alpha<1$ is a parameter affecting the number of remeshings, further we set $\alpha=0.25$.

5.1.1 Time mesh

In order to satisfy (5.3) we should construct a mesh such that the error indicator in time $\forall n:\ 1 \leq n \leq N$ satisfies

$$0.5(1-\alpha)TOL \leq \frac{\sum_{k=1}^{n} \tau_{k-1}\eta_T(t_{k-1})}{\max\limits_{0\leqslant k\leqslant n}\left(\left\|v_h^k\right\|^2_{L^2(\Omega)} + \left|u_h^k\right|^2_{H^1(\Omega)}\right)^{1/2}} \leq 0.5(1+\alpha)TOL. \tag{5.4}$$

A sufficient condition to obtain this at final time step t_N is to check that for every time step $n,\ 1 \leq n \leq N$ we have

$$\sum_{k=1}^{n} \tau_{k-1}\eta_T(t_{k-1}) \sim \frac{t_n}{T} 0.5TOL \max_{0\leqslant k\leqslant n}\left(\left\|v_h^k\right\|^2_{L^2(\Omega)} + \left|u_h^k\right|^2_{H^1(\Omega)}\right)^{1/2}. \tag{5.5}$$

Thus, at every time $t_n,\ 1 \leq n \leq N$ we want the time estimator be such that

$$\frac{0.5}{T}(1-\alpha)TOL \leq \frac{\eta_T(t_{n-1})}{\max\limits_{0\leqslant k\leqslant n}\left(\left\|v_h^k\right\|^2_{L^2(\Omega)} + \left|u_h^k\right|^2_{H^1(\Omega)}\right)^{1/2}} \leq \frac{0.5}{T}(1+\alpha). \tag{5.6}$$

If (5.6) is not satisfied we compute new time step τ_n^{new} based on the time error estimator $\eta_T(t_{n-1})$ and old time step τ_n. We know that $\eta_T(t_{n-1}) \sim (\tau_n)^2$ and we want a new time step τ_n^{new} such that the new time error indicator $\eta_T^{new}(t_{n-1}) \sim (\tau_n^{new})^2$ satisfies

$$\eta_T^{new}(t_{n-1}) = \frac{0.5}{T} \max_{0\leqslant k\leqslant n}\left(\left\|v_h^k\right\|^2_{L^2(\Omega)} + \left|u_h^k\right|^2_{H^1(\Omega)}\right)^{1/2}.$$

Thus the explicit formula for time step τ_n is

$$\tau_n^{new} = \min\left(\tau_n\left(\frac{0.5TOL \max\limits_{0\leqslant k\leqslant n}\left(\left\|v_h^k\right\|^2_{L^2(\Omega)} + \left|u_h^k\right|^2_{H^1(\Omega)}\right)^{1/2}}{T\eta_T(t_{n-1})}\right)^{1/2}, T-t_n\right). \tag{5.7}$$

5.1.2 Space mesh

At every time step n we want triangulation $\mathcal{T}_h^n$ such that

$$0.5(1-\alpha)TOL \leq \frac{\max\limits_{0\leqslant t_k\leqslant t_n} \eta_S^1(t)}{\max\limits_{0\leqslant k\leqslant n}\left(\left\|v_h^k\right\|^2_{L^2(\Omega)} + \left|u_h^k\right|^2_{H^1(\Omega)}\right)^{1/2}} \leq 0.5(1+\alpha)TOL. \tag{5.8}$$

Let say for now that for each time t_n, $1 \leq n \leq N$ we want space estimator to be:

$$0.5(1-\alpha)TOL \leq \frac{\eta_S^1(t_n)}{\max\limits_{0\leqslant k\leqslant n}\left(\left\|v_h^k\right\|^2_{L^2(\Omega)} + \left|u_h^k\right|^2_{H^1(\Omega)}\right)^{1/2}} \leq 0.5(1+\alpha)TOL. \tag{5.9}$$

If (5.9) is not satisfied we construct another mesh based on the space error indicator $\eta_S^1(t)$. The question is how to build new mesh size using our space error estimator.

The approach that we shall use is to build mesh size distribution using an analogy between *a priori* and *a posteriori* error estimates. Now we talk only about space adaptivity and suppose that at the current time t_n, $1 \leq n \leq N$ we already have a time step such that (5.6) is satisfied. Thus we are interested only in the space *a posteriori* error estimator.

$$\left(|u-u_{h\tau}|^2_{1,\Omega}+|v-v_{h\tau}|^2_{2,\Omega}\right)(t) \sim \left[\sum_{K\in T_h} h_K^2 \left\|\frac{\partial \tilde{v}_{h\tau}}{\partial t}-\Delta\tilde{u}_{h\tau}-f\right\|^2_{L^2(K)}+\sum_{E\in E_h} h_E\|[n\cdot\nabla\tilde{u}_{h\tau}]\|^2_{L^2(E)}\right](t).$$

Let us introduce the mesh size distribution $h(x)$ such that $h(x)$ at a point x inside a triangle $K\in T_h$ is approximately equal to h_K and the function $\eta(x)$ such that

$$\int_\Omega h^2(x)\eta^2(x)dx = \sum_{K\in T_h} h_K^2 \left\|\frac{\partial \tilde{v}_{h\tau}}{\partial t}-\Delta\tilde{u}_{h\tau}-f\right\|^2_{L^2(K)}+\sum_{E\in E_h} h_E\|[n\cdot\nabla\tilde{u}_{h\tau}]\|^2_{L^2(E)}.$$

Of course, the choice of $\eta(x)$ is not unique. But it is reasonable to require that the equality above hold locally in some sense. For example, this can be achieved by a partition of unity: let $\Phi_i(x)$ be the $P1$ finite element basis function (hat function) associated to the node number i (denoted x_i) and require for all nodes x_i

$$\int_\Omega h^2(x)\eta^2(x)\Phi_i(x)dx = \sum_{K\in T_h} h_K^2\int_K\left(\frac{\partial \tilde{v}_{h\tau}}{\partial t}-\Delta\tilde{u}_{h\tau}-f\right)^2(x)\Phi_i(x)dx + \sum_{E\in E_h} h_E\int_E[n\cdot\nabla\tilde{u}_{h\tau}]^2(x)\Phi_i(x)dx.$$

Approximating $\eta(x)$ with $\eta(x_i)$ over the support of $\Phi_i(x)$ yields

$$\eta(x_i)=\left[\frac{1}{\int_\Omega h^2(x)\Phi_i(x)dx}\left(\sum_{K\in T_h} h_K^2\int_K\left(\frac{\partial \tilde{v}_{h\tau}}{\partial t}-\Delta\tilde{u}_{h\tau}-f\right)^2(x)\Phi_i(x)dx + \sum_{E\in E_h} h_E\int_E[n\cdot\nabla\tilde{u}_{h\tau}]^2(x)\Phi_i(x)dx\right)\right]^{1/2}. \quad (5.10)$$

Now, we can approximate $\eta(x)$ everywhere as the $P1$ FE function taking the values at the nodes given by the formula above. The number of degrees of freedom, i.e. the number of internal vertexes in $\mathcal{T}_h^n$, is given approximately in 2D case by

$$N_{DOF}\sim\int_\Omega\frac{dx}{h^2(x)},$$

since a triangle of size $h(x)$ occupies the area of order $h^2(x)$. On each time step we

want to construct an optimal mesh with the minimal possible N_{DOF} to achieve a given error tolerance, i.e.

$$|u-u_h|^2_{1,\Omega}+|v-v_h|^2_{2,\Omega}=T\tilde{O}L^2=0.25TOL^2\int_\Omega(|\nabla u_{h\tau}|^2+|v_{h\tau}|^2).$$

This is essentially a minimization problem for the mesh size distribution $h(x)$:

$$\begin{array}{c}\min\limits_{h\,\in\,L^2(\Omega)}\\ \int_\Omega h^2(x)\eta^2(x)dx=T\tilde{O}L^2\end{array}\qquad\int_\Omega\frac{dx}{h^2(x)}.$$

The minimum is achieved on a stationary point of the Lagrangian

$$L(h,\lambda)=\int_\Omega\frac{dx}{h^2(x)}+\lambda\left(\int_\Omega h^2(x)\eta^2(x)dx-T\tilde{O}L^2\right),$$

with $h\in L^2(\Omega)$ and $\lambda\in\mathbb{R}$. Taking the variations yields

$$-\int_\Omega\frac{2\delta h(x)dx}{h^3(x)}+\lambda\int_\Omega 2h(x)\delta h(x)\eta^2(x)dx=0,$$

so that the optimal mesh size distribution is

$$h_{opt}(x)=\frac{\alpha}{\sqrt{\eta(x)}},$$

with some $\alpha\in\mathbb{R}$. Finally, recalling the constraint, we get

$$h_{opt}(x)=\frac{T\tilde{O}L}{\sqrt{\int_\Omega\eta(x)dx}}\frac{1}{\sqrt{\eta(x)}}. \tag{5.11}$$

In practice we compute the new mesh size as

$$h^{new}(x)=\min\left(3h^{old}(x),\max\left(\frac{1}{3}h^{old}(x),h_{opt}(x)\right)\right), \tag{5.12}$$

in order to avoid too abruptly changing meshes. Here $h^{old}(x)$ is a current mesh size. Our Adaptive algorithm is summarized in Table 5.1.

Remark 24. *The alternative approach to choosing the mesh size is based on the technique explained in [Pic03]. Using (5.10) we rewrite (5.9) for every P_i node of the mesh:*

$$0.5(1-\alpha)\frac{\sqrt{3}}{N_V}TOL\le\frac{\eta(x_i)}{\max\limits_{0\le k\le n}\left(\left\|v_h^k\right\|^2_{L^2(\Omega)}+\left|u_h^k\right|^2_{H^1(\Omega)}\right)^{1/2}}$$
$$\le 0.5(1+\alpha)\frac{\sqrt{3}}{N_V}TOL, \tag{5.13}$$

where N_V is the number of mesh vertices. Then if the upper bound of (5.9) is not satisfied

Set $\mathcal{T}_h^0,\ u_h^0,\ n=0,\ t=0$	Initialization
Set $\mathcal{T}_h^1 = \mathcal{T}_h^0$	First time step
Compute $u_h^1,\ n=1,\ t=\tau_0$	
Set $n=2,\ \tau_1=\tau_0$	
Do while $t < T$	Time loop
$t := t + \tau_n$	Increment the current time step
Calculate u_h^n on mesh $\mathcal{T}_h^{n-1}$	
Calculate $\eta_S^1(t)$	Space error estimator
If (5.9) is not satisfied	Mesh adaptation
Calculate $\eta(x_i)$ at all nodes x_i of $\mathcal{T}_h^{n-1}$ by (5.10)	
Calculate $h^{new}(x_i)$ at all nodes x_i of $\mathcal{T}_h^{n-1}$ by (5.12)	New mesh size at the mesh vertices
Give h^{new} to the mesh generator and obtain $\mathcal{T}_h^n$	Build new mesh
$t := t - \tau_n$	
Else If (5.6) is not satisfied	Time adaptation
Compute new τ_n^{new} from (5.7)	
$t := t - \tau_n$	
$\tau_n := \tau_n^{new}$	
Else	Go to the next step
$\mathcal{T}_h^n := \mathcal{T}_h^{n-1}$	Same mesh
$\tau_{n+1} = \tau_n$	Same time step
$n := n + 1$	
End If	
End If	
End Do	

TABLE 5.1: Adaptive algorithm.

the we set h corresponding to the node to $\frac{2}{3}h$. Analogously if the lower bound of (5.9) is not satisfied the we set h corresponding to the node to $\frac{3}{2}h$. We give an example of using the approach in the following section.

5.2 Numerical study of the adaptive algorithm

In this section we discuss the question of interpolation between meshes and present numerical study of our adaptive space-time algorithm described in Fig. 5.1 for several test cases.

We take the approximation of initial conditions and of the right-side function as orthogonal projections as in Lemma 23:

$$u_h^0 = \Pi_h u^0,\ v_h^0 = \Pi_h v^0,\ f_h^n = P_h f^n,\ 0 \leq n \leq N.$$

By analogy with initial conditions, after each remeshing in space we use time estimator given by (2.21). In other words after we generate a new mesh we have to wait 2 time steps in order to understand if current time step is sufficient. If it's not the case then we compute new time step and go back for 2 time steps. We do time adaptation only once per each 2 time steps.

Like in the case of heat equation [LPP09], we do the space mesh adaptation only once per each 2 time steps. We denote by $\mathcal{T}_h^n$ and $\mathcal{T}_h^{n+1}$ two meshes at time t_n and t_{n+1} respectively and by V_h^n, V_h^{n+1} the associated finite elements spaces. We shall use interpolation operators

$$\mathcal{I}_h^{n,u} : V_h^n \to V_h^{n+1}, \quad \mathcal{I}_h^{n,u^*} : V_h^n \to V_h^{n+1}, \quad \mathcal{I}_h^{n,v} : V_h^n \to V_h^{n+1},$$

which can be different one from another. Several variants of interpolation will be considered and these operators above will be specified for each variant. In general, if a new mesh has to be build going from t_n to t_{n+1}, we compute $u_h^{n+1} \in V_h^{n+1}$ and $v_h^{n+1} \in V_h^{n+1}$ such that

$$\begin{aligned}
&\left(\frac{u_h^{n+1} - \mathcal{I}_h^{n,u} u_h^n}{\tau_n}, \varphi_h^{n+1}\right) - \left(\frac{v_h^{n+1} + \mathcal{I}_h^{n,v} v_h^n}{2}, \varphi_h^{n+1}\right) = 0, \qquad (5.14)\\
&\left(\frac{v_h^{n+1} - \mathcal{I}_h^{n,v} v_h^n}{\tau_n}, \varphi_h^{n+1}\right) + \left(\nabla\frac{u_h^{n+1} + \mathcal{I}_h^{n,u^*} u_h^n}{2}, \nabla\varphi_h^{n+1}\right) = \left(\frac{f^{n+1} + f^n}{2}, \varphi_h^{n+1}\right).
\end{aligned}$$

for all $\varphi_h^{n+1} \in V_h^{n+1}$.

We consider 4 different interpolation strategies:

1. Motivating ourselves with Lemma 23, we introduce the following interpolation Π_h^n such that for any $v_h^n \in V_h^n$, $\Pi_h^n v_h^n$ belongs to V_h^{n+1} and satisfies :

$$\begin{aligned}
(w_h^n, \varphi_h^n) &= (\nabla v_h^n, \nabla \varphi_h^n),\ \forall \varphi_h^n \in V_h^n\\
\left(\nabla\left(\Pi_h^n v_h^n\right), \nabla\varphi_h^{n+1}\right) &= (r_h^n(w_h^n), \varphi_h^{n+1}),\ \forall \varphi_h^{n+1} \in V_h^{n+1},
\end{aligned}$$

   where r_h^n is the Lagrange interpolant operator on $\mathcal{T}_h^{n+1}$. We choose in (5.14)

$$\mathcal{I}_h^{n,u} = \mathcal{I}_h^{n,u*} = \mathcal{I}_h^{n,v} = \Pi_h^n.$$

2. We introduce the following interpolation $\mathcal{P}_h^n$ such that for any $v_h^n \in V_h^n$, $\mathcal{P}_h^n v_h^n$ belongs to V_h^{n+1} and satisfies :

$$\left(\mathcal{P}_h^n v_h^n, \varphi_h^{n+1}\right) = (r_h^n(w_h^n), \varphi_h^{n+1}),\ \forall \varphi_h^{n+1} \in V_h^{n+1}.$$

We choose in (5.14)

$$\mathcal{I}_h^{n,u} = \mathcal{I}_h^{n,u*} = \Pi_h^n,\ \mathcal{I}_h^{n,v} = \mathcal{P}_h^n.$$

3. Like in [Pic10], we choose in (5.14)

$$\mathcal{I}_h^{n,u} = \mathcal{I}_h^{n,v} = \mathcal{I}_h^{n,u^*} = r_h^n,$$

4. Like in [LPP09], we choose in (5.14)

$$\mathcal{I}_h^{n,u} = \mathcal{I}_h^{n,v} = r_h^n,$$

$$\left(\nabla \mathcal{I}_h^{n,u^*} u_h^n, \nabla \varphi_h^{n+1}\right) = \left(r_h^n(\tilde{P}_h(u_h^n)), \nabla \varphi_h^{n+1}\right),\ \forall \varphi_h^{n+1} \in V_h^{n+1}.$$

Here $\tilde{P}_h$ is Zienkiewicz-Zhu recovery projection defined in Subsection 1.1.3.

Remark 25. *The motivation of strategy 2 is the following. Using the exact H_0^1 orthogonal projection as $\mathcal{I}_h^{n,u}$, $\mathcal{I}_h^{n,u^*}$ and the exact L^2-orthogonal projection as $\mathcal{I}_h^{n,v}$ in (5.14) and reproducing proof of the a priori error estimate from Chapter 2 we are able to prove the a priori error estimate (2.9). But those operators are difficult to implement in practice since they involve the integrals from functions defined on $\mathcal{T}_h^n$ and $\mathcal{T}_h^{n+1}$. However this variant is close to strategy 2 above.*

The first and the second interpolation strategies in practice behave equivalently and for the sake of brevity further we present the numerical results only for the first interpolation strategy. The third and the fourth strategies leads to non optimal behavior of the time error estimator, namely η_T is not a second order in time anymore, due to the unboundedness of the higher order derivatives. This fact is illustrated numerically bellow. Thus throughout this section we use the first interpolation strategy unless otherwise is explicitly mentioned.

The true error is

$$e = \max_{0 \leqslant n \leqslant N} \left(\left\| v_h^n - \frac{\partial u}{\partial t}(t_n) \right\|_{L^2(\Omega)}^2 + |u_h^n - u(t_n)|_{H^1(\Omega)}^2 \right)^{1/2}.$$

Denote as

$$\eta_S = \eta_S^{(1)}(t_N),$$

$$\eta_T = \sum_{k=0}^{N-1} \tau_k \eta_T(t_k).$$

The quality of our error estimators in space and in time is determined by following effectivity indices:

$$ei^S = \frac{\eta_S}{e},\ \ ei_T = \frac{\eta_T}{e}.$$

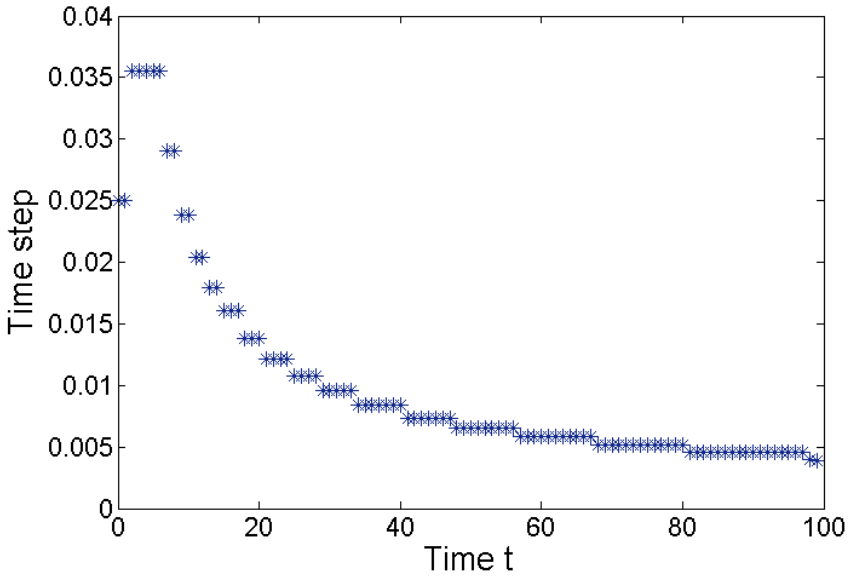


FIGURE 5.1: Changing in time of length of time step for the test case (5.16). $TOL = 0.125$, the mesh size is chosen like in Remark 24.

Consider the problem (1.59) with $\Omega = (0,1) \times (0,1)$, $T = 1$ and the exact solution u given by

$$u(x, y, t) = \cos(\pi t) \sin(\pi x) \sin(\pi y), \tag{5.15}$$

The study of the time and space error estimators when using uniform time steps and mesh size can be found in Chapter 2. We now use the adaptive algorithm described in Table 5.1. Results in Table 5.2 are reported when using several values of TOL where τ_0 - initial time step, NV_{final} - number of vertices in final mesh, NV_{max} - maximal number of vertices, N_{remesh} - number of space remeshings and N_{retime} - number of time remeshings. We observe that that the error divided by 2 each time the tolerance TOL is divided by 2 and both the time error indicator ei_T and the space error indicator ei^S seem to be a good representation for the true error. The number of nodes NV_{final} and $NV_{max}$is approximatively multiplied by 4 when TOL is divided by 2.

Tol	ei^S	ei_T	NV_{final}	NV_{max}	N_{remesh}	N_{retime}	τ_0	e
0.5	3.99	1.49	479	643	3	1	0.05	0.23
0.25	3.85	1.47	1127	2437	3	2	0.05	0.114
0.125	4.17	1.4	4735	9838	4	2	0.05	0.063
0.0625	3.98	1.43	20151	36955	5	2	0.05	0.037

TABLE 5.2: SPACE adaptation and TIME adaptation, elliptic projection of the initial conditions, new interpolation, solution is given by (5.15), $h_x^{start} = h_y^{start} = 0.05$, mesh construction (5.12).

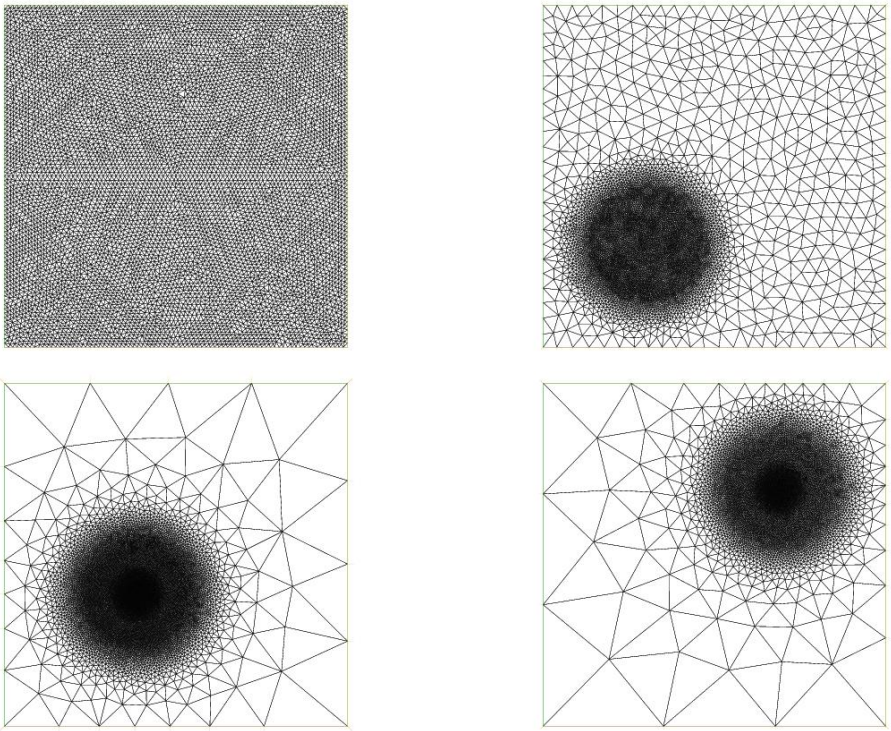


FIGURE 5.2: Example (5.16). Adapted meshes obtained with $TOL = 0.125$, the mesh size is chosen like in Remark 24. From left to right: time $t = 0,\ 0.025,\ 0.499,\ 1$

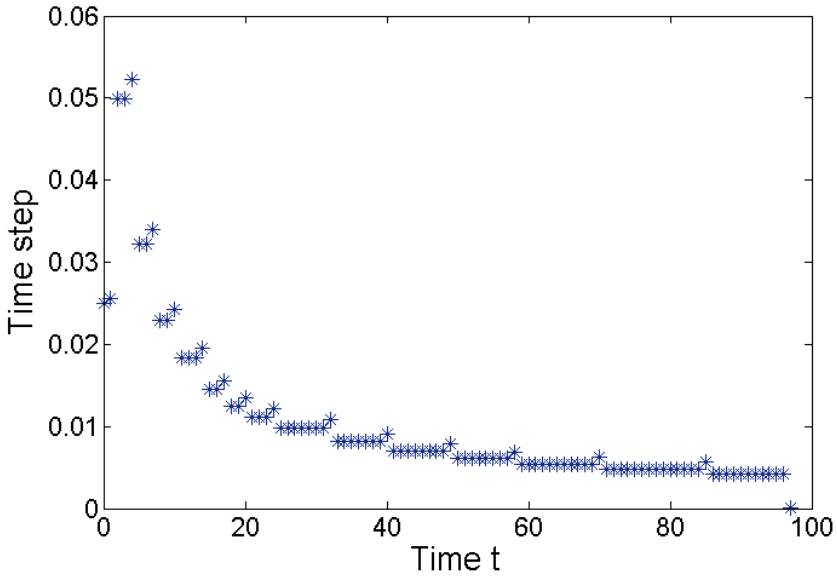


FIGURE 5.3: Changing in time of length of time step for the test case (5.16). $TOL = 0.125$, mesh construction (5.12).

We now consider the case when $\Omega = (0,1) \times (0,1),\ T = 1$ and the exact solution u given by

$$u = e^{-100*((x-0.3-0.4t^2)^2+(y-0.3-0.4t^2)^2)}, \tag{5.16}$$

thus u is a Gaussian function, whose center moves from point (0.3, 0.3) at $t = 0$ to point (0.7,0.7) at $t = 1$. The transport velocity $0.8t$ is peaking at $t = 1$. We consider two ways of choosing the mesh size: from the Remark 24 and from the formula (5.12). Results are reported in Tables 5.3 and 5.4 correspondingly. We investigate the number of vertices, number of remeshings in space and in time for several values of tolerance TOL. In both cases we observe that the time error indicator ei_T and the space error indicator ei^S seem to be a good representation for the true error. The difference between two approaches to choosing the mesh size is only in number of space remeshings and in number of nodes: the approach from the Remark 24 requires more space remeshings but generate the meshes with less nodes then the approach from formula (5.12). We have reported in Figure 5.2 the meshes obtained with the approach from the Remark 24 when the tolerance $TOL = 0.125$. On Figures 5.1 and 5.3 we have also plotted the time step evolution with respect to the number of time steps for both approaches to choosing the mesh size. We observe that in both cases the time step fits the transport velocity.

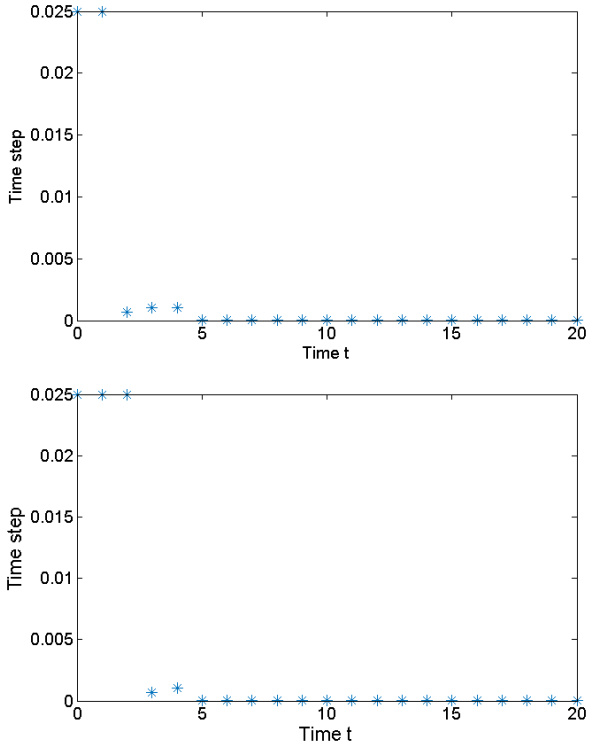


FIGURE 5.4: Changing in time of length of time step for the test case (5.16). $TOL = 0.125$, mesh construction (5.12). Top: intepolation strategy 3 from page 106 is used. Bottom: intepolation strategy 4 from page 106 is used.

We want to highlight the non optimal behavior of time error estimator η_T when the interpolation strategies 3 and 4 from page 106 are used. On Figure 5.4 we have plotted the evolution of the time step during first 20 time steps of adaptive algorithm from Table 5.1. We observe that in both cases the time step evolution is completely different from to that presented on Figures 5.1 and 5.3. The time step $n = 20$ in both

Tol	ei^S	ei_T	NV_{final}	NV_{max}	N_{remesh}	N_{retime}	τ_0	e
0.5	5.63	3.19	664	1008	84	12	0.025	0.17
0.25	5.68	3.0	2318	3844	99	12	0.025	0.086
0.125	5.66	3.0	8919	15628	135	19	0.025	0.044
0.0625	5.65	2.97	34994	60062	189	20	0.025	0.022
0.03125	5.68	3.1	138268	237109	273	21	0.025	0.011

TABLE 5.3: SPACE adaptation and TIME adaptation, elliptic projection of the initial conditions, new interpolation, solution is given by (5.16), $h_x^{start} = h_y^{start} = 0.05$, mesh construction like in Remark 24.

Tol	ei^S	ei_T	NV_{final}	NV_{max}	N_{remesh}	N_{retime}	τ_0	e
0.5	2.17	1.69	911	1156	18	8	0.025	0.61
0.25	5.66	3.08	2815	4606	22	11	0.025	0.12
0.125	5.73	3.08	10731	18490	31	14	0.025	0.059
0.0625	5.71	3.02	44015	74663	34	16	0.025	0.03
0.03125	5.72	3.04	166872	295784	37	18	0.025	0.014

TABLE 5.4: SPACE adaptation and TIME adaptation, elliptic projection of the initial conditions, new interpolation, solution is given by (5.16), $h_x^{start} = h_y^{start} = 0.05$, mesh construction (5.12).

cases is around 2e-06. This significant difference shows that η_T tends to dramatically over-predict the true error. We suspect that this behavior due to the interpolation error after each remeshing that leads to the unboundedness of the higher order derivatives in time. We conclude that the order of convergence for the interpolation strategies 3 and 4 from page 106 is not recovered.

The next solution is given by

$$u = e^{-100*r^2(x,y,t))}, \tag{5.17}$$

where

$$r^2(x,y,t) = ((x - 0.3 - 0.4\beta(t))^2 + (y - 0.3 - 0.4\beta(t))^2),$$

and

$$\beta(t) = 0.5 + 0.5 \tanh\left(\frac{t - 0.5}{0.2}\right).$$

with homogeneous boundary conditions apply on the whole boundary of Ω. As before we take $\Omega =]0,1[^2$and $T = 1$. Results are reported in Table 5.5. We observe that the error at final time is approximatively divided by 2 when TOL is divided by 2. We plot in Figure 5.5 the time step evolution with respect to the number of time steps. We observe that the time step mostly decreases till $n = 75$ and increases until the final time step. Thus the time error indicator ei_T seems to be a good representation for the true error. The total number of vertices at final time is multiplied by 4 as the tolerance is divided by two.

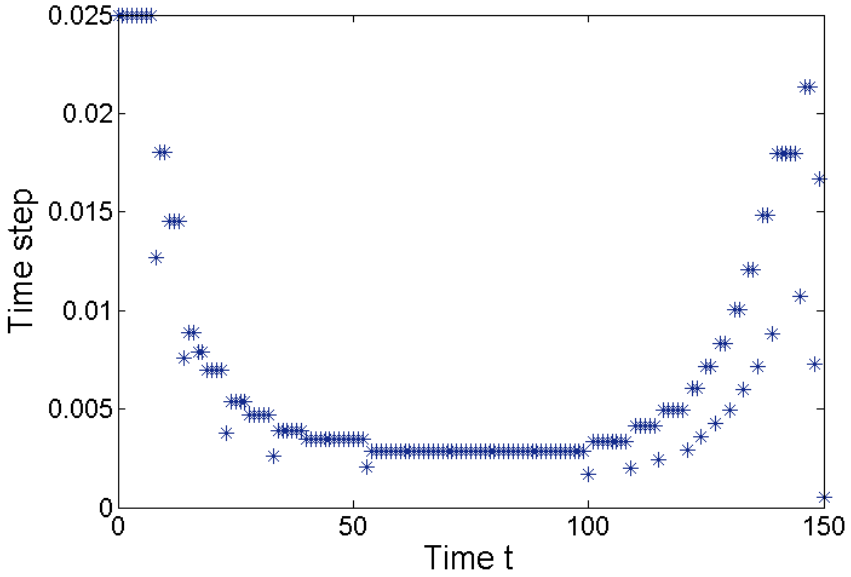


FIGURE 5.5: Changing in time the length of time step for the test case (5.17). $TOL = 0.125$, mesh construction (5.12).

Tol	η^S	η_T	NV_{final}	NV_{max}	N_{remesh}	N_{retime}	τ_0	e
0.5	1.32	0.34	645	1198	28	12	0.025	0.24
0.25	0.65	0.19	2266	4610	32	20	0.025	0.125
0.125	0.33	0.1	9181	18592	35	22	0.025	0.064
0.0625	0.17	0.05	36399	75180	38	22	0.025	0.036
0.03125	0.09	0.03	12453	24673	41	24	0.025	0.018

TABLE 5.5: SPACE adaptation and TIME adaptation, elliptic projection of the initial conditions, new interpolation, solution is given by (5.17), $h_x^{start} = h_y^{start} = 0.05$, mesh construction (5.12).

Chapter 6

CEMRACS project: Parallel in time algorithms for nonlinear iterative methods

This chapter has no connection with the rest of the thesis. It presents the results and accompanying discussion of the project PANLIM realized during CEMRACS 2016 in collaboration with Mustafa Gaja and François-Xavier Roux.

The simulation of complex nonlinear structures via the finite element method is generally based on quasi-static incremental loading procedure leading to successive nonlinear problems solved by Newton-like methods [Ogd97; ZT00] which themselves entail solving multiple ill-conditioned large linear systems. For complex structures, the total number of steps may be very large and, since the procedure is recursive, the computational cost is very large. Reducing the time for such simulations is of great interest for numerical engineering in industry. A recently developed parallel in time method for time dependent problems is the Parareal algorithm [Lio01; MT05] that allows the parallelization of the computation in the temporal domain making use of HPC facilities.

The objective of this work is to present a proposal for the application of the Parareal algorithm for quasi-static nonlinear processes in order to perform these quasi-static steps in parallel by analogy with the Parareal method for time-dependent systems of ordinary or partial differential equations. In this work we describe how this proposal has been realized on a common example of a beam undergoing nonlinear elastic deformation with additional boundary nonlinearity. Numerical results demonstrate the possibility of using the Parareal algorithm on elementary quasi-static problems, although more conclusive numerical tests are needed to provide a substantial evidence for the efficiency of the approach.

The outline of the chapter is as follows. We present the model problem, its spacial discretization and linearization in Section 6.1. Basic idea and precise description of the Parareal algorithm for the case of quasi-static nonlinear problems are laid out in Section 6.2. Numerical experiments for several test cases are reported in Section 6.3. We discuss some reoccurring themes and questions that faced us during this work in Section 6.4. Finally, we draw some concluding remarks in Section 6.5.

Chapter contents

6.1 Nonlinear structures and incremental loading

We treat the boundary value problem for the equations of nonlinear elasticity [Ant05; Sto68] as a minimization problem. For a domain $\Omega = [0,1] \times [0,1] \subset \mathbb{R}^2$ we want to find the displacement field $(u_1, u_2) : \Omega \to \mathbb{R}^2$ that minimizes the functional of total potential energy $J(u_1, u_2)$:

$$\min_{(u_1,u_2)\in V} J(u_1, u_2), \tag{6.1}$$

where V is a suitable function space that satisfies some Dirichlet boundary conditions:

$$(u_1, u_2)(x_1, x_2) = (\bar{u}_1, \bar{u}_2), \;\; (x_1, x_2) \in \Gamma_v, \tag{6.2}$$

where $\bar{u}_1$ and $\bar{u}_2$ are some given constants, Γ_v denotes the part of the boundary on which the displacements are given. The total potential energy is given by the following formula:

$$J(u_1, u_2) = \int_\Omega f(F(u_1, u_2)) - \int_{\Gamma_P} P_\alpha u_2, \tag{6.3}$$

where $F(u_1, u_2) = A(E[u_1, u_2], E[u_1, u_2])$, $A(X, Y)$ is a bilinear symmetric positive form with respect to matrices X and Y, f is a given C^2 function corresponding to hyperelastic constitutive law, P_α is some external force which we apply on the boundary surface $\Gamma_P \subset \partial\Omega$ (for the sake of simplicity we suppose that the force act in the x_2-direction) and

$$E[u_1, u_2] = (E_{ij})_{i=1,2,j=1,2}\,, \; E_{ij} = \frac{1}{2}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} + \frac{\partial u_k}{\partial x_i}\frac{\partial u_k}{\partial x_j}\right), \tag{6.4}$$

is the Green-Saint-Venant strain tensor.

The finite element method is the standard modeling approach to simulate and analyze the behavior of solids [Bra07]. Let us approximate the space V by the finite element space V_h. We thus introduce a regular mesh $\mathcal{T}_h$ on Ω, a triangle K of the mesh $\mathcal{T}_h$ and a standard $P1$ finite element space on it $V_h \subset V$:

$$V_h = \{v_h \in C(\bar{\Omega}) : v_h|_K \in \mathbb{P}_1 \; \forall K \in \mathcal{T}_h\},$$

Thus we obtain the finite element problem of finding the field (u_h, v_h) that satisfy (6.2) and

$$\min_{(u_h,v_h)\in(V_h,V_h)} J(u_h, v_h). \tag{6.5}$$

The solution of the large strain nonlinear problem will require an iterative process and Newton's method forms the basis of practical schemes [ZT00]. We start with an initial guess (u_0, v_0) which is reasonably close to the true correct solution. Let us suppose that we already know the solution (u_{n-1}, v_{n-1}) from the step $n-1$ then we attempt to correct this guess to bring it closer to the proper solution by setting

$$u_n = u_{n-1} + du,\ v_n = v_{n-1} + dv,$$

where the correction term (du, dv) is given by

$$D^2J(u_{n-1}, v_{n-1})((w, s), (du, dv)) \\ = DJ(u_{n-1}, v_{n-1})(w, s)\ \ \forall (w, s) \in (V_h, V_h), \quad (6.6)$$

where $DJ(u, v)$ and $D^2J(u, v)$ are the Jacobian and the Hessian of $J(u, v)$ respectively. Then we check if the magnitude of the correction (du, dv) is small enough, and if not, we go to the next iteration $n+1$ and compute the new correction term $(\tilde{du}, \tilde{dv})$.

6.1.1 Incremental loading

There is no guarantee that Newton's method will converge. It will converge quadratically to the exact solution only if the initial guess is sufficiently close to the correct solution. A common way to avoid this problem is to apply the load in a series of increments instead of all at once [ZT00; Ogd97]. For example, we would like to solve our problem for some certain boundary force F, but numerical test shows us that Newton's method does not converge. We can divide the force into 2 loads: at the beginning we solve our problem with force $F/2$ and after we solve the problem once more but with force F using the solution at the end of the preceding increment as the initial guess. The quality of this guess could be improved by reducing the increment once more. In general, small load increments are essential for a better accuracy of the solution.

In Table 6.1 we show the case of a cantilever rubber beam loaded with a uniformly distributed force $F = 200\ Pa$ and a contact surface. In the case of applying the force all at once, we can't achieve the convergence of Newton's method. The situation is persistent when using 2 and 3 loading steps, whereas we achieve convergence when we start using 4 loading steps and so on. Moreover, this example demonstrates that the number of incremental loads can affect the number of iterations required by Newton's method to achieve the convergence. Indeed, when the force is divided into 7 loading steps we only need 3 Newton iterations to achieve the same tolerance like in the case of 4 loading steps with 4 Newton iterations. This example shows that the incremental loading procedure is required in the case when the initial guess for Newton's method is too far from the exact solution. Note that with incremental loads we obtain essentially a quasi-static nonlinear problem instead of a static one; from now on we have a so-called "pseudo-time" direction which is in fact an additional dimension where it is possible to apply the idea of Parareal method. But before discussing this topic we will introduce a boundary nonlinearity to our problem.

Number of loads	N_I	N_{IC}
1,2,3	no convergence	...
4	4	11
5	4	9
6	4	8
7	3	29
8	3	11

TABLE 6.1: Convergence of Newton's method depending on the number of loading steps. Cantilever rubber beam loaded at the upper bound with a uniformly distributed vertical force $F = 200\ Pa$ with a contact surface, mesh 60×20 points, N_I - average number of Newton iterations, N_{IC} - number of Newton iterations at the loading step when the contact appears for the first time.

6.1.2 Boundary nonlinearity

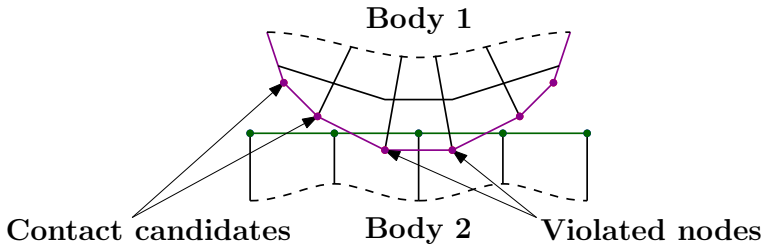
The incremental loading procedure allows us to insert a different form of nonlinearity in our model: a contact problem [Lac13]. Let us consider the cantilever beam:

$$u_1(0, x_2) = 0, u_2(0, x_2) = 0, \quad x_1 \in [0, 1], x_2 \in [0, 1],$$

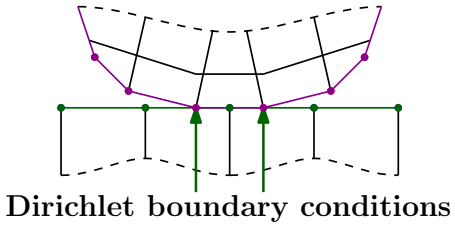
under a vertical load P_α and in one moment the beam touches a solid surface $X_2 = X_{sf}$ with border Γ_S where X_{sf} and (X_1, X_2) are the Eulerian coordinates. We model the appearance of the contact surface with the following algorithm: at every incremental loading step we look for the nodes of our mesh at the border of the beam and check if the vertical displacements of those nodes are equal to or greater than the vertical coordinate of the surface X_{sf} (See Figure 6.1 a), in what follows we call them violated nodes [Kim14]. On the next incremental loading step for every violated node we set the sliding contact boundary condition (a vertex of the finite element mesh can slide over the surface, but cannot move away from it), in other words we take the vertical displacement equal to the displacement corresponding to the contact surface with Dirichlet boundary conditions (See Figure 6.1 b):

$$u_2(x_1, 0) = X_{sf}, \quad \forall x_1 \in \Gamma_S.$$

The number of incremental loadings is very important for detecting the contact surface. From Table 6.1 we see that for the cantilever beam configuration loaded with an uniformly distributed force we achieve the maximum number of the Newton iterations in the load when the contact appears, because we are converging in presence of such strong nonlinear effect like a contact. Thus the accuracy of the solution depends on the nonlinearity involved in the model and the number of taken load increments.



a) Load increment n,



b) Load increment $n+1$.

FIGURE 6.1: Contact surface modulation: we detecte the violated nodes at displacements corresponding to load increment n, then we impose Dirichlet boundary conditions for the violated nodes before the start of load increment $n+1$.

6.2 Parareal method

6.2.1 Parareal method for ODEs

In order to understand the mechanism of the parallelization in time we will briefly explain the basic idea of the parareal method with a simple first order ordinary differential equation [Lio01]:

$$u'(t) = -au(t) \text{ on } [0,T],\ u(0) = u_0, \tag{6.7}$$

where a is a constant. Let us introduce the temporal mesh on the time interval $[0,T]$:

$$0 = T^0 < T^1 < \cdots < T^N = T,$$

with time steps $\Delta T = T/N$. The Parareal algorithm utilizes a coarse solver G to quickly step through time by computing relatively cheap approximate solutions for all time intervals of interest, and then simultaneously refines all of these approximate solutions using an accurate fine solver F. Let $G_{\Delta T}(u_{n-1}, T^{n-1})$ for all $n = 1, \ldots, N$ be a rough approximation of $u(T^n)$ with an initial condition $u(T^{n-1}) = u_{n-1}$, $F_{\Delta T}(u_{n-1}, T^{n-1})$ is a more accurate approximation of the solution $u(T^n)$ with an initial condition $u(T^{n-1}) = u_{n-1}$. Then starting with a coarse approximation U_n^0 at the time points $T^1, T^2, \ldots T^N$, Parareal performs for $k = 0, 1, \ldots$ the correction iteration:

$$U_{n+1}^{k+1} = F_{\Delta T}(U_n^k, T^n) + G_{\Delta T}(U_n^{k+1}, T^n) - G_{\Delta T}(U_n^k, T^n). \tag{6.8}$$

The application of the fine solver to each time interval $[T^n, T^{n+1}]$ is independent of the other time intervals, and thus parallelizable. From (6.8) we see that the refined

solutions are then fed back to the coarse solver, and the iterative cycle continues until all time intervals are converged. The dependencies between the time intervals are carried through the coarse solver which involves stepping sequentially through time. This, of course, represents a sequential process, but it leads to a more rapid solution, assuming that the coarse solver is much faster than the fine solver, and hence, computationally negligible.

In the simple numerical example for equation (6.7) we set $T = 10$ and perform 5 coarse steps each parareal iteration, thus the coarse solution is obtained by the implicit Euler scheme with step $\Delta T = 2$. After that, we introduce a finer mesh with a smaller time step $\Delta t = 0.02$ on each interval $[T^n, T^{n-1}]$, thus we have 100 fine time steps on each interval $[T^n, T^{n-1}]$. We use the same implicit Euler scheme as a fine solver.

We present the error in log scale between the parareal solution and the analytical solution of equation (6.7) for each of the 5 parareal iterations in Figure 6.2. In order to show the convergence of the parareal method at each parareal iteration we also present the error between the solution of Euler's method with 500 time steps and the analytical solution of equation (6.7).

The solution of the first coarse time step is processed by the fine solver using the exact initial state $u(0)$. For every further parareal iteration, the converged accurate solution is computed with initial conditions from the previous iteration, thus the second iteration will give us convergence at least for the first two steps, the third iteration for the first three steps, and so on. Thus the algorithm is guaranteed to converge in at most $K = N$ iterations (see Figure 6.2), though in practice it is possible to achieve much faster convergence with judicious choice of the coarse propagator. The computational complexity of the parareal algorithm is given by the following formula:

$$KNC_{\mathcal{C}} + (K-1)N\frac{\Delta T}{\Delta t}C_{\mathcal{F}}, \tag{6.9}$$

where $C_{\mathcal{C}}$ is the cost to perform the coarse propagator over one time step, $C_{\mathcal{F}}$ is the complexity of one time step for the fine propagator, K is the number of iterations required to achieve the required convergence of the parareal algorithm and N is the number of coarse steps (and the number of processors as well).

Looking at the Parareal algorithm from a different perspective, it is reasonable to presume that the idea of parallelizing the time domain for time-dependent problems can be used for the case of quasi-static problems in order to reduce the computation time.

6.2.2 Parareal method for nonlinear structures

In what follows we shall propose a "time" parallel iterative method for solving the minimization problem (6.5). At first, we introduce NM load increment steps and assume that we compute the coarse solution at points $0, M, 2M, \ldots, NM$ and the fine solution at points $0, 1, 2, \ldots, NM$. Thus we introduce a pseudo-time parameter $\tilde{t} \in [0, \tilde{t}_1, \tilde{t}_2, \ldots, \tilde{t}_{NM}]$ which is used to describe our algorithm via an analogy with a time-dependent problem. For the sake of simplicity we assume that all increments are the same and denote points for the coarse solution as $[0, \tilde{T}_1, \tilde{T}_2, \ldots, \tilde{T}_N]$, in other words $\tilde{t}_M = \tilde{T}_1$, $\tilde{t}_{2M} = \tilde{T}_2$, $\ldots$, $\tilde{t}_{MN} = \tilde{T}_{MN}$. At every point $[0, \tilde{t}_1, \tilde{t}_2, \ldots, \tilde{t}_{NM}]$ we introduce the exact solution of problem (6.5) $(u(\tilde{t}_n), v(\tilde{t}_n))$ for $n \in [0, 1, \ldots, NM]$. We

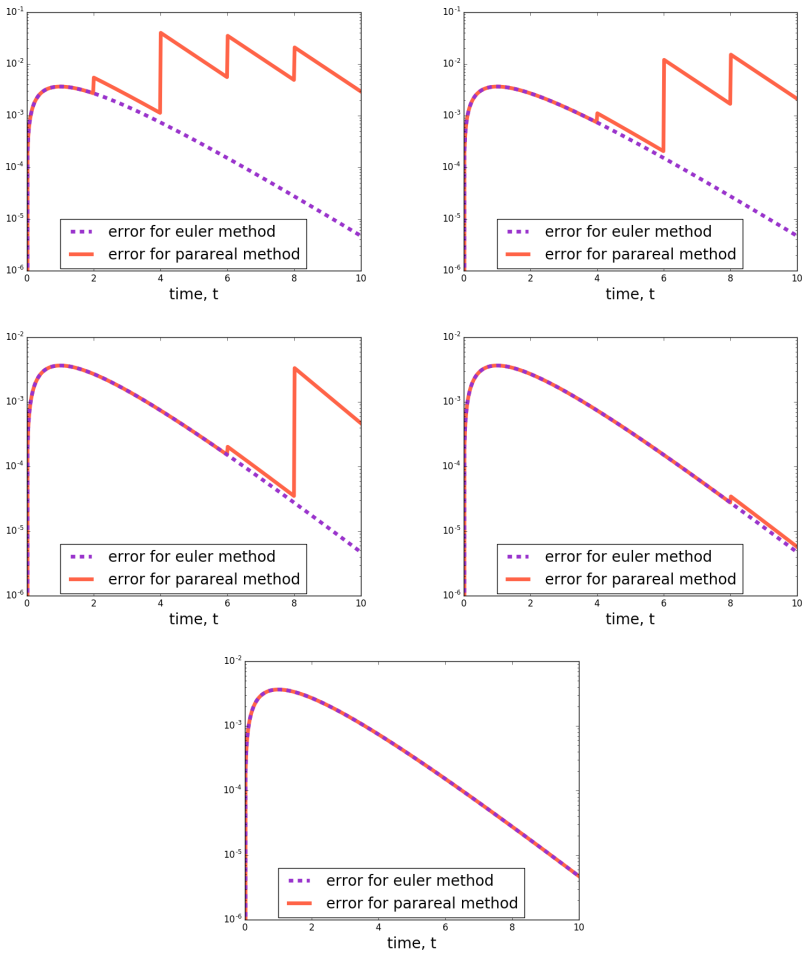


FIGURE 6.2: The error between analytical solution of equation (6.7) and parareal solution (6.8) during 5 iterations of parareal method.

want to find approximate solution

$$(U_n^k, V_n^k) = (U^k(\tilde{t}_n), V^k(\tilde{t}_n)),$$

where index k is the number of parareal iterations. Like in the case of a partial differential equation, let $G(U_{nM}^k, V_{nM}^k)$ for all $n = 0, \ldots, N-1$ be a rough approximation of $(u(\tilde{T}_{n+1}), v(\tilde{T}_{n+1}))$ which we compute starting from initial guess (U_{nM}^k, V_{nM}^k) by using the coarse propagator that is described below. Let $F(U_{nM}^k, V_{nM}^k)$ be a more accurate approximation of the solution $(u(\tilde{T}_{n+1}), v(\tilde{T}_{n+1}))$ which we compute starting from displacements (U_{nM}^k, V_{nM}^k) by using the fine propagator that is described below.

We start from the initial guess $(U_0^0, V_0^0) = (u_0, v_0)$ for Newton's method at the first increment step and obtain an initial coarse solution via

$$\left(U_{nM}^0, V_{nM}^0\right) = G\left(U_{(n-1)M}^0, V_{(n-1)M}^0\right),$$

for $n \in [1, \ldots N]$. That is we obtained Parareal solution of iteration $k = 0$. Hereafter on every Parareal iteration $k \geq 1$ we determine our numerical solution by the formula:

$$(U_{nM}^k, V_{nM}^k) = G\left(U_{(n-1)M}^k, V_{(n-1)M}^k\right) + F\left(U_{(n-1)M}^{k-1}, V_{(n-1)M}^{k-1}\right) - G\left(U_{(n-1)M}^{k-1}, V_{(n-1)M}^{k-1}\right). \quad (6.10)$$

The coarse propagator

Once the solution is known at iteration $k-1$, the coarse problem at iteration $k+1$ is Newton's method for N incremental loads. For all coarse increment steps $n = 1, \ldots, N$ starting with the initial guess for Newton's method as displacements from the previous parareal iteration

$$G^0(U_{nM}^k, V_{nM}^k) = (U_{nM}^{k-1}, V_{nM}^{k-1}),$$

we compute the coarse solution at the Newton iteration j as:

$$G^j\left(U_{nM}^k, V_{nM}^k\right) = G^{j-1}\left(U_{nM}^k, V_{nM}^k\right) + (du_n^j, dv_n^j), \quad (6.11)$$

where the correction term (du_n^j, dv_n^j) by analogy with (6.6) follows from

$$D^2 J\left(G^{j-1}\left(U_{nM}^k, V_{nM}^k\right)\right)((w,s),(du_n^j, dv_n^j)) = DJ\left(G^{j-1}\left(U_{nM}^k, V_{nM}^k\right)\right)(w,s) \quad \forall (w,s) \in (V_h, V_h). \quad (6.12)$$

Let's assume that at the Newton iteration number j^* we achieved the required accuracy for approximate solution, then the coarse propagator is

$$G(U_{nM}^k, V_{nM}^k) = G^{j^*}(U_{nM}^k, V_{nM}^k).$$

The fine propagator

We use a procedure similar to (6.11) and (6.12) to obtain the fine solution $F(U^k_{nM}, V^k_{nM})$ on parareal iteration k. Inside every "coarse" pseudo-time interval $[\tilde{T}_n, \tilde{T}_{n+1}]$ we introduce a local fine solver on fine mesh $\tilde{F}(U^k_{nM+m}, V^k_{nM+m})$ to emphasize the fact that this solver propagates the solution on fine increment steps. We compute the solution with the same Newton's method for M incremental loads using the displacements from the previous Parareal iteration as an initial guess for the first incremental load

$$\tilde{F}^0\left(U^k_{nM}, V^k_{nM}\right) = \tilde{F}\left(U^{k-1}_{nM}, V^{k-1}_{nM}\right).$$

For all the next fine increment steps $m = 1, \ldots, M-1$ we use displacements from the previous increment step as an initial guess for Newton's method, in other words, we start by setting the initial guess as

$$\tilde{F}^0\left(U^k_{nM+m}, V^k_{nM+m}\right) = \tilde{F}\left(U^k_{nM+m-1}, V^k_{nM+m-1}\right).$$

If we already know the fine solution $\tilde{F}^{j-1}\left(U^k_{nM+m}, V^k_{nM+m}\right)$ of the Newton iteration $j-1$ then the solution of iteration j is:

$$\tilde{F}^j\left(U^k_{nM+m}, V^k_{nM+m}\right) = \tilde{F}^{j-1}\left(U^k_{nM+m}, V^k_{nM+m}\right) + (du^j_m, dv^j_m), \tag{6.13}$$

where the correction term (du^j_m, dv^j_m) like in the case of the coarse solver (6.12) is defined by

$$\begin{aligned} &D^2 J\left(\tilde{F}^{j-1}\left(U^k_{nM+m}, V^k_{nM+m}\right)\right)((w,s),(du^j_m, dv^j_m)) \\ &\qquad = DJ\left(\tilde{F}^{j-1}\left(U^k_{nM+m}, V^k_{nM+m}\right)\right)(w,s) \quad \forall (w,s) \in (V_h, V_h). \end{aligned} \tag{6.14}$$

Let's imagine that at the Newton iteration number j^* we achieve the required tolerance then we determine the fine propagator as

$$\tilde{F}\left(U^k_{nM+m}, V^k_{nM+m}\right) = \tilde{F}^{j^*}\left(U^k_{nM+m}, V^k_{nM+m}\right),$$

for all fine increment steps $m = 1, \ldots, M-1$. Thus we determine the fine propagator at coarse time points as

$$F\left(U^k_{nM}, V^k_{nM}\right) = \tilde{F}\left(U^k_{(n+1)M-1}, V^k_{(n+1)M-1}\right).$$

Our algorithm is summarized in Table 6.2.

6.3 Numerical study of the Parareal algorithm for a nonlinear beam

We now turn to numerical examples for our algorithm. What we report here are preliminary results that have to be extended to more complex cases. We run our experiments for four different cases with different boundary conditions: a beam anchored at one or two ends, with and without a contact; see Figure 1.3. However, here

Set N	Number of coarse steps
Set M	Number of fine steps inside every coarse step
Set K	Number of parareal iterations
Set $(U_0^0, V_0^0) = (u_0, v_0)$	
For $n = 1$ to N	
Compute $(U_n^0, V_n^0) = G(U_{n-1}^0, V_{n-1}^0)$	Initial coarse propagation
End for	
For $k = 1$ to K do	Iterative loop
For $n = 0$ to $N - 1$ do	Parallelizable part
For $m = 0$ to $M - 1$ do	
Compute $\tilde{F}(U_{nM+m}^{k-1}, V_{nM+m}^{k-1})$	Local fine propagator
End for	
Compute $F(U_{nM}^{k-1}, V_{nM}^{k-1})$	Fine propagation
End for	
For $n = 0$ to $N - 1$ do	Sequential part
Compute $G(U_n^{k-1}, V_n^{k-1})$	Coarse propagation
Compute (U_n^k, V_n^k) from (6.10)	Correction
End for	
End for	

TABLE 6.2: Parareal algorithm for nonlinear quasi-static structure.

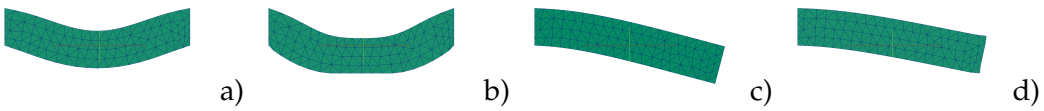


FIGURE 6.3: Beam configuration: a) beam fixed at two ends without a contact, b) beam fixed at two ends with a contact, c) beam fixed at one end without a contact, d) beam fixed at one end with a contact.

we report numerical results only for the case of a one-end fixed beam with a contact, see Figure 1.3 d, since it is an interesting configuration in the sense of the degree of nonlinearity involved. All the results reported hereafter correspond to a vertically uniform distributed external load applied to upper bound of a rubber beam. In the numerical experiments, FreeFem++ is used for the finite element formulation and the linear algebra implementation, the fine solver is parallelized using MPI.

The first experiment considered here is the Parareal algorithm parallelized on 4 cores, 6 coarse steps and 6 fine steps are used, see Table 6.3. In order to measure the quality of the solution we are looking for L^∞ and L^2 norms of displacements. The first line of Table 6.3 corresponds to displacements obtained by applying 36 loading steps for Newton's method. We observe that we have a convergence from the first Parareal iteration, which is reasonable since our numerical example is performed with a sufficiently fine mesh. Indeed, the coarse solver converges and it means that 6 loading steps are enough to reach an accurate solution under the final load and since our process is quasi-static we do not care about the solution at every "time" point. We would like to underline that our aim here is to demonstrate the possibility of using Parareal algorithm on elementary quasi-static problems in order to inspire ourselves for further more realistic and strongly nonlinear models.

The problem formulation for the second experiment is the same as that of the previous experiment except that here we have used a different space mesh size for the coarse and fine propagation. Due to the sequential nature of the coarse propagator we would like it to be as fast and cheap as possible. Since the coarse propagator will be corrected with the fine propagator we can use rough space mesh coarse steps. Results with different coarse and fine mesh sizes are summarized in Table 6.4. First line of Table 6.4 corresponds to displacements obtained with loading steps for direct Newton's method. Note that as in the previous example we have a convergence from the first Parareal iteration.

6.4 Discussion

The results presented in Section 6.3 support the proposal for the feasibility of applying the parareal method for nonlinear structural analysis problems. That said, there is a crucial point that needs to be clarified; the pseudo time stepping is only there in order to ensure the convergence of Newton's method. It is thus more like a stability problem: once we have a converged result then it is the good one. Thus a legitimate question is why bother with several parareal iterations in the first place? For instance, one can think of taking as an initial guess the coarse solution and then apply a fine solver in order to achieve an accurate solution at the last loading step. This line of thought is interesting, but one would be losing the guaranteed theoretical convergence proof laid out in theory [Lio01], as we are unaware of a theory

Parareal iteration	$\|\mathbf{u}\|_{L^\infty}$	$\|u\|_{L^2}$	$\|v\|_{L^2}$	$Abs\,Error$
direct method	0.027394	0.0480169	0.161934	
1	0.0278855	0.0512589	0.168256	4.9e-4
2	0.0280801	0.0521499	0.170242	6.8e-4
3	0.0265591	0.0443302	0.15321	8.3e-4
4	0.0265591	0.0443302	0.15321	8.3e-4
5	0.0265599	0.0443482	0.152963	8.4e-4

TABLE 6.3: Nonlinear beam with one fixed end and a contact. Convergence for fixed mesh 60×20 , $F = 120\ Pa$, 6 coarse steps, 6 fine steps. The absolute error is calculated between the parareal solution and the direct method in L^∞ norm .

Coarse mesh	Fine mesh	Parareal iteration	Number of loads	$\|\mathbf{u}\|_{L^\infty}$	Abs Error
60×20	60×20	direct method	20	0.0278	
60×20	90×30	1	12	0.0272	6e-4
60×20	90×30	2	12	0.0272	6e-4
60×20	60×20	1	12	0.0269	9e-4
60×20	60×20	2	12	0.0271	7e-4

TABLE 6.4: Nonlinear beam with one fixed end and a contact, different mesh sizes for coarse and fine propagation, $F = 100\ Pa$, 12 coarse steps, 12 fine steps. The absolute error is calculated between the parareal solution and the direct method in L^∞ norm .

supporting the aforementioned line of thought.

We don't present results comparing the wall clock time for the regular sequential loading versus the parareal implementation because for the purposes of this work, we had to write files as an output for the coarse and fine solvers and in turn, read them as an initial guess, calculating and propagating the differences and so forth. This process of reading and writing files is naturally computationally expensive, in comparison to the direct sequential case within FreeFem++ as the data points are passed directly from one loading step to another. This comes as a result of not finding a way to go around passing the data points directly without writing them in FreeFem++ for the parareal case between the fine and coarse solver, and we chose to focus on tackling the question of feasibility of the idea. Further implementations would naturally require optimization to avoid this computational bottle-neck. Finally, an interesting recurring question is about the required number of parareal iterations. We think that this point would depend on the model and boundary conditions at hand. Not only, but also on the requirement posed by the application and what is the required level of convergence.

6.5 Concluding remarks

We presented some preliminary numerical results as a proposal to show the feasibility of applying the Parareal algorithm for nonlinear structural analysis problems. The subject of the forthcoming work is to perform more extensive numerical tests to provide a more comprehensive evidence for the feasibility study and then to extend this algorithm to industrial 3D problems and more realistic models using the software package NUMEA provided by Hutchinson.

Bibliography

[Adj02] Slimane Adjerid. "A posteriori finite element error estimation for second-order hyperbolic problems". In: *Computer methods in applied mechanics and engineering* 191.41 (2002), pp. 4699–4719.

[Ain+89] M Ainsworth et al. "Analysis of the Zienkiewicz–Zhu a-posteriori error estimator in the finite element method". In: *International Journal for numerical methods in engineering* 28.9 (1989), pp. 2161–2174.

[AMN06] Georgios Akrivis, Charalambos Makridakis, and Ricardo Nochetto. "A posteriori error estimates for the Crank–Nicolson method for parabolic equations". In: *Mathematics of computation* 75.254 (2006), pp. 511–531.

[Ant05] Stuart S Antman. "Problems in Nonlinear Elasticity". In: *Nonlinear Problems of Elasticity* (2005), pp. 513–584.

[Bak76] Garth A. Baker. "Error estimates for finite element methods for second order hyperbolic equations". In: *SIAM J. Numer. Anal.* 13.4 (1976), pp. 564–576. ISSN: 0036-1429.

[BD89] Laurence A Bales and Vassilios A Dougalis. "Cosine methods for nonlinear second-order hyperbolic equations". In: *Mathematics of computation* 52.186 (1989), pp. 299–319.

[BDS85] Laurence A Bales, Vassilios A Dougalis, and Steven M Serbin. "Cosine methods for second-order hyperbolic equations with time-dependent coefficients". In: *mathematics of computation* 45.171 (1985), pp. 65–89.

[BGR10] Wolfgang Bangerth, Michael Geiger, and Rolf Rannacher. "Adaptive Galerkin finite element methods for the wave equation". In: *Computational Methods in Applied Mathematics Comput. Methods Appl. Math.* 10.1 (2010), pp. 3–48.

[BPS02] James H. Bramble, Joseph E. Pasciak, and Olaf Steinbach. "On the stability of the L^2 projection in $H^1(\Omega)$". In: *Math. Comp.* 71.237 (2002), pp. 147–156. ISSN: 0025-5718. DOI: 10.1090/S0025-5718-01-01314-X. URL: http://dx.doi.org/10.1090/S0025-5718-01-01314-X.

[BR01] Wolfgang Bangerth and Rolf Rannacher. "Adaptive finite element techniques for the acoustic wave equation". In: *Journal of Computational Acoustics* 9.02 (2001), pp. 575–591.

[BR99] Wolfgang Bangerth and Rolf Rannacher. "Finite element approximation of the acoustic wave equation: Error control and mesh adaptation". In: *East West Journal of Numerical Mathematics* 7.4 (1999), pp. 263–282.

[Bra07] Dietrich Braess. *Finite elements: Theory, fast solvers, and applications in solid mechanics*. Cambridge University Press, 2007.

[BS05] Christine Bernardi and Endre Süli. "Time and space adaptivity for the second-order wave equation". In: *Math. Models Methods Appl. Sci.* 15.2 (2005), pp. 199–225. ISSN: 0218-2025. DOI: 10.1142/S0218202505000339. URL: http://dx.doi.org/10.1142/S0218202505000339.

[BW76] Klaus-Jürgen Bathe and Edward L Wilson. *Numerical methods in finite element analysis*. Vol. 197. Prentice-Hall Englewood Cliffs, NJ, 1976.

[Dup73] Todd Dupont. "L^2-estimates for Galerkin methods for second order hyperbolic equations". In: *SIAM journal on numerical analysis* 10.5 (1973), pp. 880–889.

[EG04] A. Ern and J.-L. Guermond. *Theory and practice of finite elements*. Springer, 2004, p. 524. ISBN: 0-387-20574-8.

[EJ91] Kenneth Eriksson and Claes Johnson. "Adaptive finite element methods for parabolic problems. I. A linear model problem". In: *SIAM J. Numer. Anal.* 28.1 (1991), pp. 43–77.

[Eva10] Lawrence C Evans. *Partial differential equations*. American Mathematical Society, 2010.

[FP01] Luca Formaggia and Simona Perotto. "New anisotropic a priori error estimates". In: *Numerische Mathematik* 89.4 (2001), pp. 641–667.

[FP03] Luca Formaggia and Simona Perotto. "Anisotropic error estimates for elliptic problems". In: *Numerische Mathematik* 94.1 (2003), pp. 67–92.

[Geo+16] Emmanuil H. Georgoulis et al. "A Posteriori Error Estimates for Leap-Frog and Cosine Methods for Second Order Evolution Problems". In: *SIAM J. Numer. Anal.* 54.1 (2016), pp. 120–136. ISSN: 0036-1429. DOI: 10.1137/140996318. URL: http://dx.doi.org/10.1137/140996318.

[GLM13] Emmanuil H. Georgoulis, Omar Lakkis, and Charalambos Makridakis. "*A posteriori* $L^\infty(L^2)$-error bounds for finite element approximations to the wave equation". In: *IMA J. Numer. Anal.* 33.4 (2013), pp. 1245–1264. ISSN: 0272-4979. DOI: 10.1093/imanum/drs057. URL: http://dx.doi.org/10.1093/imanum/drs057.

[GLP17a] Olga Gorynina, Alexei Lozinski, and Marco Picasso. "An easily computable error estimator in space and time for the wave equation". In: *Submitted* (2017), available on arXiv:1710.08410.

[GLP17b] Olga Gorynina, Alexei Lozinski, and Marco Picasso. "Time and space adaptivity of the wave equation discretized in time by a second order scheme". In: *Submitted* (2017), available on arXiv:1707.00057.

[Kim14] Nam-Ho Kim. *Introduction to nonlinear finite element analysis*. Springer Science & Business Media, 2014.

[Lac13] Walter Lacarbonara. *Nonlinear structural mechanics: theory, dynamical phenomena and modeling*. Springer Science & Business Media, 2013.

[Lio01] J Lions. "Résolution d'EDP par un schéma en temps «pararéel» A"parareal"in time discretization of PDE's". In: *Academie des Sciences Paris Comptes Rendus Serie Sciences Mathematiques* 332 (2001), pp. 661–668.

[LMP14] Omar Lakkis, Charalambos Makridakis, and Tristan Pryer. "A comparison of duality and energy a posteriori estimates for $L_\infty(0,T;L_2(\Omega)$ in parabolic problems". In: *Mathematics of Computation* (Dec. 2014).

[LPP09] Alexei Lozinski, Marco Picasso, and Virabouth Prachittham. "An anisotropic error estimator for the Crank-Nicolson method: application to a parabolic problem". In: *SIAM J. Sci. Comput.* 31.4 (2009), pp. 2757–2783. ISSN: 1064-8275. DOI: 10.1137/080715135. URL: http://dx.doi.org/10.1137/080715135.

[MN03] Charalambos Makridakis and Ricardo H Nochetto. "Elliptic reconstruction and a posteriori error estimates for parabolic problems". In: *SIAM journal on numerical analysis* 41.4 (2003), pp. 1585–1594.

[MT05] Yvon Maday and Gabriel Turinici. "The parareal in time iterative solver: a further direction to parallel implementation". In: *Lecture Notes in Computational Science and Engineering* 40 (2005), pp. 441–448.

[New59a] Nathan M Newmark. "A method of computation for structural dynamics". In: *Journal of the engineering mechanics division* 85.3 (1959), pp. 67–94.

[New59b] Nathan M Newmark. "A method of computation for structural dynamics". In: *Journal of the engineering mechanics division* 85.3 (1959), pp. 67–94.

[Ogd97] Raymond W Ogden. *Non-linear elastic deformations*. Courier Corporation, 1997.

[Pic03] Marco Picasso. "An anisotropic error indicator based on Zienkiewicz–Zhu error estimator: Application to elliptic and parabolic problems". In: *SIAM Journal on Scientific Computing* 24.4 (2003), pp. 1328–1355.

[Pic10] Marco Picasso. "Numerical study of an anisotropic error estimator in the $L^2(H^1)$ norm for the finite element discretization of the wave equation". In: *SIAM J. Sci. Comput.* 32.4 (2010), pp. 2213–2234. ISSN: 1064-8275. DOI: 10.1137/090778249. URL: http://dx.doi.org/10.1137/090778249.

[Pic98] Marco Picasso. "Adaptive finite elements for a linear parabolic problem". In: *Computer Methods in Applied Mechanics and Engineering* 167.3-4 (1998), pp. 223–237.

[RT83] P.-A. Raviart and J.-M. Thomas. *Introduction à l'analyse numérique des équations aux dérivées partielles*. Collection Mathématiques Appliquées pour la Maitrise. [Collection of Applied Mathematics for the Master's Degree]. Masson, Paris, 1983, p. 224. ISBN: 2-225-75670-8.

[Sto68] James Johnston Stoker. "Nonlinear elasticity". In: (1968).

[SZ90] L. Ridgway Scott and Shangyou Zhang. "Finite element interpolation of nonsmooth functions satisfying boundary conditions". In: *Math. Comp.* 54.190 (1990), pp. 483–493. ISSN: 0025-5718.

[Ver03] R Verfürth. "A posteriori error estimates for finite element discretizations of the heat equation". In: *Calcolo* 40.3 (2003), pp. 195–212.

[Ver94] Rüdiger Verfürth. "A posteriori error estimation and adaptive mesh-refinement techniques". In: *Journal of Computational and Applied Mathematics* 50.1-3 (1994), pp. 67–83.

[Ver96] Rüdiger Verfürth. *A review of a posteriori error estimation and adaptive mesh-refinement techniques*. John Wiley & Sons Inc, 1996.

[ZT00] Olgierd Cecil Zienkiewicz and Robert Leroy Taylor. *The finite element method: solid mechanics*. Vol. 2. Butterworth-heinemann, 2000.

[ZZ87] Olgierd C Zienkiewicz and Jian Z Zhu. "A simple error estimator and adaptive procedure for practical engineerng analysis". In: *International journal for numerical methods in engineering* 24.2 (1987), pp. 337–357.

[ZZ92] Olgierd Cecil Zienkiewicz and Jian Zhong Zhu. "The superconvergent patch recovery and a posteriori error estimates. Part 1: The recovery technique". In: *International Journal for Numerical Methods in Engineering* 33.7 (1992), pp. 1331–1364.