COMPUTATIONAL METHODS IN APPLIED MATHEMATICS, Vol. 10 (2010), No. 1, pp. 3–48

ADAPTIVE GALERKIN FINITE ELEMENT METHODS FOR THE WAVE EQUATION

W. BANGERTH[1] , M. GEIGER[2], AND R. RANNACHER[2]

Abstract — This paper gives an overview of adaptive discretization methods for linear second-order hyperbolic problems such as the acoustic or the elastic wave equation. The emphasis is on Galerkin-type methods for spatial as well as temporal discretization, which also include variants of the Crank-Nicolson and the Newmark finite difference schemes. The adaptive choice of space and time meshes follows the principle of "goal-oriented" adaptivity which is based on a posteriori error estimation employing the solutions of auxiliary dual problems.

2000 Mathematics Subject Classification: 35L05, 65M50, 65M60, 74S05.

Keywords: wave equation, Galerkin method, finite elements, adaptivity, a posteriory error estimation, Newmark scheme, Crank-Nicolson scheme.

1. Introduction

For the numerical solution of linear second-order hyperbolic partial differential equations, e. g. the scalar acoustic wave equation or the elastic wave equation governed by the Lamé-Navier equations of linear elasticity theory, a broad variety of methods is available in order to fully discretize the given equations and subsequently solve the discrete systems. Usually the discretization can be split up into two main components, namely the discretization of spatial variables and the one with respect to time. In this paper, we will primarily, but not exclusively, discuss the latter while considering a rather conventional, though adaptive, spatial discretization. Among the most attractive methods for time discretization are the so-called "continuous Galerkin" (cf. Bales & Lasiecka [1] and French & Peterson [13]) and the "discontinuous Galerkin" (cf. Johnson [24], Grote & al. [17]) schemes. For lowest order, these methods can be identified with certain well-known difference schemes, e. g. the classical trapezoidal Newmark scheme (see Wood [39, 40] and Hughes [19]), the backward Euler scheme and the Crank-Nicolson scheme.

The main topic of this paper are methods for "goal-oriented" a posteriori error estimation and mesh-size adaptation such as the "Dual Weighted Residual (DWR)" approach described in Becker & Rannacher [7] and Bangerth & Rannacher [6]. This method is based on "weighted" a posteriori error estimates for arbitrary error quantities such as point values or line integrals and employs the solutions (generalized Green functions or influence functions) of auxiliary "dual" problems. It depends fundamentally on the *Galerkin character* of the discretization and guides the optimal adjustment of spatial and temporal mesh sizes

[1] *Department of Mathematics, Texas A&M University*, College Station, TX 77843-3368, USA.
[2] *Institute of Applied Mathematics, University of Heidelberg*, 69120 Heidelberg, Germany. E-mail: rannacher@iwr.uni-heidelberg.de

according to the prescribed "goal" of the computation. However, by the same approach, a posteriori error estimates can also be obtained with respect to global norms, e. g. the L^2- or energy norm. In view of the aforementioned close relation between low-order Galerkin methods and finite difference schemes, the DWR approach developed for Galerkin methods is equally applicable to these finite difference schemes.

The DWR method has already been successfully applied in the construction of local mesh-size adaptation for various types of nonstationary partial differential equations (see, e. g., Becker & Rannacher [7], Bangerth & Rannacher [6] and Schmich & Vexler [33]). The *a priori* as well as a posteriori error analysis of Galerkin methods for the wave equation using space-time duality arguments has been initiated by Johnson [24]; see also Bernardi & Süli [8]. In this context the DWR method has been applied to the wave equation in Bangerth [2] and Bangerth & Rannacher [5]. Other strategies for mesh adaptation using heuristic smoothness-based error indicators, such as the "ZZ indicator", have been proposed in Li & Wiberg [37] and Schemann & Bornemann [32]. There have been several attempts to construct a posteriori error estimators for the Newmark scheme during the past decades (see Romero & Lacoma [31] or Schweizerhof & al. [34]), but so far none of these methods fits into the DWR framework and uses its advantages.

The material of this survey paper is mainly based on the the articles of Bangerth & Rannacher [3, 5] and the Diploma theses of Bangerth [2] and Geiger [14]. Its outline is as follows. Section 2 presents the abstract setting of the continuous wave problem and its variational formulations. In Section 3, we discuss the spatial and temporal discretization methods. This includes the equivalence of the Newmark trapezoidal rule to the Crank-Nicolson scheme and in turn the equivalence of the latter to a certain "continuous Galerkin", the so-called "cG(1)" method. Then, Section 4 outlines the DWR method for "goal-oriented" a posteriori error estimation and mesh adaptation and its use for temporal and spatial adaptivity in the Galerkin discretization of the linear wave equation. Finally in Section 5 the results of some numerical tests are presented in order to illustrate the potential of the approach to adaptivity discussed in this paper.

2. The wave equation and its discretization

Throughout this paper, we will consider the second-order hyperbolic PDE system

$$\begin{aligned} \rho(x)\partial_t^2 u(x,t) + \mathcal{A}u(x,t) &= f(x,t) \quad \text{for} \quad (x,t) \in \Omega \times I, \\ u(x,t) &= 0 \quad \text{for} \quad (x,t) \in \partial\Omega_D \times I, \\ \partial_n^{\mathcal{A}} u(x,t) &= 0 \quad \text{for} \quad (x,t) \in \partial\Omega_N \times I, \\ u(x,0) &= u_0^0(x) \quad \text{for} \quad x \in \Omega, \\ \partial_t u(x,0) &= u_0^1(x) \quad \text{for} \quad x \in \Omega, \end{aligned} \tag{2.1}$$

with a positive density function ρ. Here, $I = (0,T]$ denotes a finite time interval, Ω is a bounded convex domain in $\mathbb{R}^n$ $(n \in \{1,2,3\})$ which is for simplicity assumed to be a polygon for $n = 2$ or a polyhedron for $n = 3$. Furthermore, $\partial\Omega_D$ and $\partial\Omega_N$ are disjoint, time-independent parts of the boundary of Ω where we prescribe (homogeneous) Dirichlet and Neumann boundary conditions, respectively. The operator $\partial_n^{\mathcal{A}}$ is the usual directional "normal" derivative associated to the operator $\mathcal{A}$. We assume that $\partial\Omega_D$ has positive measure. The case of inhomogeneous Dirichlet conditions u^D on $\partial\Omega_D$ can be treated the

same way by interpreting u^D as the trace of a sufficiently differentiable function $\tilde{u}$ and solving the differential equation for $v := u - \tilde{u}$ instead of u.

The operator $\mathcal{A}$ is assumed to be a second-order, elliptic spatial differential operator with sufficiently regular coefficients. Representative examples are the (scalar) diffusion operator

$$\mathcal{A}v := -\mathrm{div}(a\nabla v), \tag{2.2}$$

occurring in the acoustic wave equation, and the (vectorial) Lamé-Navier operator

$$\mathcal{A}v := -\mu\Delta v + (\lambda+\mu)\nabla\mathrm{div}\, v, \tag{2.3}$$

governing the elastic wave equation. In the presence of uniform "weak" or "strong" damping the wave equation takes the form

$$\rho(x)\partial_t^2 u(x,t) + \gamma_w\partial_t u(x,t) + \gamma_s\partial_t\mathcal{A}u(x,t) + \mathcal{A}u(x,t) = f(x,t), \tag{2.4}$$

with certain constants $\gamma_w, \gamma_s \geqslant 0$. In most parts of this paper, we will consider the undamped case, i.e., set $\gamma_w = \gamma_s = 0$. The possible extension of results obtained for this situation to the case with damping will be covered by remarks.

Let d be the number of components of the solution function u. For the initial values, we assume $u_0^0 \in H_0^1(\partial\Omega_D;\Omega)^d$ and $u_0^1 \in L^2(\Omega)^d$, where $H_0^1(\partial\Omega_D;\Omega)$ is the space of all H^1-functions vanishing on $\partial\Omega_D$ with dual space denoted by $H^{-1}(\partial\Omega_D;\Omega)^d$. The source function f is assumed to be in $L^2(I;H^{-1}(\partial\Omega_D;\Omega)^d)$. The (scalar or vectorial) L^2 inner product and the corresponding norm are denoted by (u,v) and $\|u\|$, respectively, and the usual 1st-order Sobolev norm by $\|\cdot\|_1$, where no distinction will be made in the notation between the case of scalar- and vector-valued functions.

From now on, we will use the abbreviations $H := L^2(\Omega)^d$ and $V := H_0^1(\partial\Omega_D;\Omega)^d$ with dual space V^*. With this notation, we define the space-time function spaces

$$\mathcal{H} := L^2(I;H), \quad \mathcal{V} := \big\{v \in L^2(I;V) \;\big|\; \partial_t v \in \mathcal{H}\big\}.$$

For simplicity the spatial operator $\mathcal{A}$ is assumed to satisfy a strong coercivity estimate of the form

$$(\mathcal{A}v, v) \geqslant \beta\|v\|_1^2, \quad u \in V, \tag{2.5}$$

with some constant $\beta > 0$. This condition is satisfied for the acoustic and the elastic wave equation due to the Poincaré and the Korn inequality, respectively. Within this framework, it is well known that there exists a unique so-called "weak" (or "variational") solution $u \in \mathcal{V} \cap C(\bar{I};V)$ of the wave equation (2.1) with first-order time derivative $\partial_t u \in \mathcal{H} \cap C(\bar{I};H)$ and second-order time derivative $\partial_t^2 v \in L^2(I,V^*)$; see Lions & Magenes [26], Lions [25], or Wloka [38]. Hence, for the given data the natural solution space $\hat{\mathcal{V}}$ for problem (2.1) is defined by

$$\hat{\mathcal{V}} := \big\{v \in \mathcal{V} \;\big|\; v \in C(\bar{I};V),\ \partial_t v \in C(\bar{I};H),\ \partial_t^2 v \in L^2(I;V^*)\big\}.$$

The second-order evolution equation (2.1) may be equivalently written in the form of a first-order (in time) system for the unknowns $u^0 := u$ and $u^1 := \partial_t u$:

$$\begin{aligned}
\rho(x)\partial_t u^0(x,t) - \rho(x)u^1(x,t) &= 0 \quad \text{for} \quad (x,t) \in \Omega\times I,\\
\rho(x)\partial_t u^1(x,t) + \mathcal{A}u^0(x,t) &= f(x,t) \quad \text{for} \quad (x,t) \in \Omega\times I,\\
u^0(x,t) &= 0 \quad \text{for} \quad (x,t) \in \partial\Omega_D\times I,\\
\partial_n^{\mathcal{A}} u^0(x,t) &= 0 \quad \text{for} \quad (x,t) \in \partial\Omega_N\times I,\\
u^0(x,0) &= u_0^0(x) \quad \text{for} \quad x \in \Omega,\\
u^1(x,0) &= u_0^1(x) \quad \text{for} \quad x \in \Omega.
\end{aligned} \tag{2.6}$$

According to the above remarks this system has a unique (weak) solution in the natural solution space $\hat{\mathcal{V}}^0 \times \hat{\mathcal{V}}^1$, where $\hat{\mathcal{V}}^0 := \hat{\mathcal{V}}$ and

$$\hat{\mathcal{V}}^1 := \big\{ w \in \mathcal{H} \;\big|\; w \in C(\bar{I}; H),\ \partial_t w \in L^2(I; V^*) \big\}.$$

The "mixed" formulation (2.6) is the starting point for Galerkin time discretization as described below, while the "primal" formulation (2.1) is mainly used for finite difference schemes.

An important feature of the wave equation is the conservation of the total energy

$$E(t) := \frac{1}{2}\big\{ \|\partial_t u(t)\|_M^2 + \|u(t)\|_E^2 \big\},$$

where $\|\cdot\|_M$, $\|\cdot\|_E$ are the natural "mass norm" and "energy norm" defined by

$$\|v\|_M^2 := (\rho v, v), \quad v \in H, \qquad \|v\|_E^2 := (\mathcal{A} v, v), \quad v \in V.$$

Indeed, for any weak solution u of (2.1) with no forcing and damping terms, we obtain by multiplying by $\partial_t u$ and observing the boundary conditions that

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{dt}}\big(\|\partial_t u\|_M^2 + \|u\|_E^2 \big) = 0.$$

Integrating this with respect to time yields

$$E(t) = \frac{1}{2}\big\{ \|\partial_t u(t)\|_M^2 + \|u(t)\|_E^2 \big\} = \frac{1}{2}\big\{ \|u_0^1\|_M^2 + \|u_0^0\|_E^2 \big\} = E(0). \tag{2.7}$$

In the presence of damping for strong solutions an "energy inequality" of the following form holds true:

$$E(t) + \int_0^t \big\{ \gamma_w \|\partial_s u(s)\|^2 + \gamma_s \|\partial_s u(s)\|_E^2 \big\}\, \mathrm{d}s \leqslant E(0), \quad t \geqslant 0. \tag{2.8}$$

We will investigate in Section 3 to what extent the various discretization schemes preserve the conservation property (2.7) of the continuous wave equation.

3. Discretization of the wave equation

We begin with an overview of discretization methods for the wave equation. Starting from the continuous model (2.1) or the equivalent system (2.6) there are essentially three different ways to discretization indicated in Fig. 3.1

- In the "Method of Lines" at first the spatial variable is discretized, e.g. by a finite element method, and then the resulting (large) system of ODEs is discretized in time. This approach has the advantage of simple data structures and matrix assembly, and that standard methods from ODE numerics may be used for time discretization. The obvious disadvantage is that the spatial mesh is fixed and therefore adaptation to time-varying features of the solution is prohibited. This is a critical limitation in using "goal-oriented" adaptivity, particularly in the case of time-dependent local "goal quantities".

- In the "Rothe Method" at first the time variable is discretized, e.g. by a finite difference scheme, and then the resulting elliptic PDEs are discretized independently in space. This approach has the advantage of accommodating dynamic spatial mesh adaptation but the disadvantage of rather complex data structures and expensive matrix assembly. In this context the design of higher-order (even second-order) space-time discretization requires some care. In order to reduce the substantial work caused by the mesh transfer from one time level to the next, one may employ hierarchically structured spatial meshes.

- A third option would be to use fully unstructured space-time meshes in a corresponding space-time Galerkin finite element discretization. This would allow to optimally adapt the mesh, for instance, to moving fronts in the space-time domain. However, the practical realization of such a discretization is highly complex, particularly in 3D, and very cost-intensive due to the complicated transfer of data between unstructured spatial meshes. Therefore, this approach is only rarely used in practice (see, however, Dumbser & al. [12] and Castro & al. [10]).

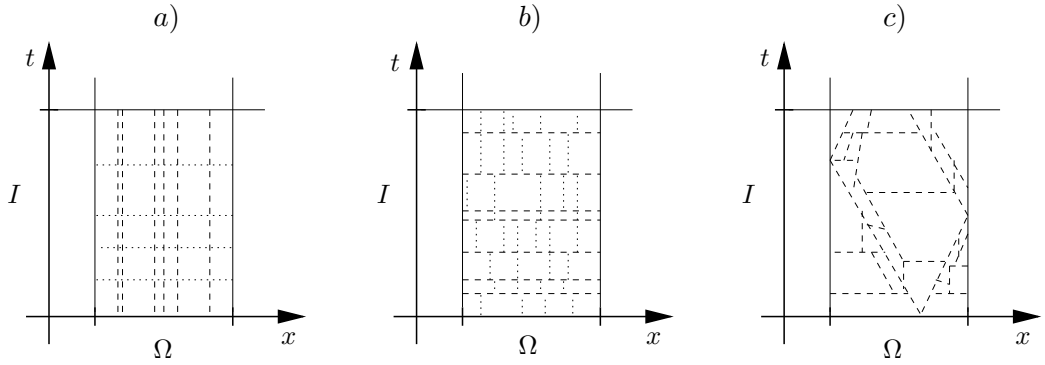


Fig. 3.1. Space-time meshes: structured in space (method of lines) (a); structured in time (Rothe method) (b); unstructured in space and time (c)

Remark 3.1. If the spatial mesh can be kept fixed in time, for space-time Galerkin methods such as the cG(1)/cG(1) or cG(1)/dG(0) method described below, the "Method of Lines" and the "Rothe Method" yield equivalent discretizations. In the discussion below, the Galerkin methods will be considered within the Rothe Method framework, while naturally finite difference methods such as the Newmark scheme are used within the context of the Method of Lines.

3.1. Spatial discretization by the Galerkin finite element method. We start from the "primal" variational equation

$$m(\partial_t^2 u, \varphi) + a(u, \varphi) = (f, \varphi) \quad \forall \varphi \in V, \tag{3.1}$$

which is satisfied on I by the solution $u \in \hat{\mathcal{V}}$ of problem (2.1). Here,

$$m(v, \varphi) := (\rho v, \varphi), \quad a(u, \varphi) := (\mathcal{A}u, \varphi)$$

are symmetric and positive definite bilinear forms, which by continuity are defined on the entire "energy space" V. We now replace the space $V = H_0^1(\partial\Omega_D; \Omega)^d$ by a standard finite-dimensional finite element subspace $V_h = S_h^s(\partial\Omega_D; \Omega)$ of polynomial degree s (i.e. of order $s+1$) where

$$S_h^s(\partial\Omega_D; \Omega) := \big\{ v_h \in C(\bar{\Omega})^d \,\big|\, v_h|_K \in P(K)^d \;\; \forall K \in \mathbb{T}_h, \; v_h|_{\partial\Omega_D} = 0 \big\}.$$

Here, $\mathbb{T}_h = \{K\}$ is a decomposition of $\bar{\Omega}$ into non-overlapping triangles or quadrilaterals in two and tetrahedra or hexahedra in three dimensions satisfying the usual admissibility conditions (cf. Ciarlet [11] or Brenner & Scott [9]). Further, it is assumed that the decompositions $\mathbb{T}_h$ match the given decomposition $\partial\Omega = \overline{\partial\Omega_D} \cup \overline{\partial\Omega_N}$ of the boundary. The local mesh size is $h_K := \mathrm{diam}(K)$ and $h := \max_{K\in\mathbb{T}_h} h_K$. Further, $P(K)$ are certain polynomial spaces containing the full s-degree polynomial spaces $P_s(K)$ or $Q_s(K)$, respectively. The simplest cases, for $s = 1$, are $P(K) = P_1(K)$ ("linear" elements) or $P(K) = Q_1(K)$ ("bilinear" elements in 2D); the latter are exclusively used in the numerical examples below. In this case the cellwise shape functions are constructed by bilinear transformations from a reference cell $\hat{K}$ (unit square) to the "physical" cells $K \in \mathbb{T}_h$.

In the test calculations presented below, we have used either sequences of uniformly refined rectangular meshes or sequences of locally refined meshes involving "hanging nodes" at the edges of neighboring cells (see Fig. 3.2). At these "irregular" nodes the nodal values are eliminated by linear interpolation between the neighboring values at "regular" nodes. For technical purposes, which will be explained below, it may be preferable to use meshes composed of 2×2 patches of cells (see Fig. 3.2).

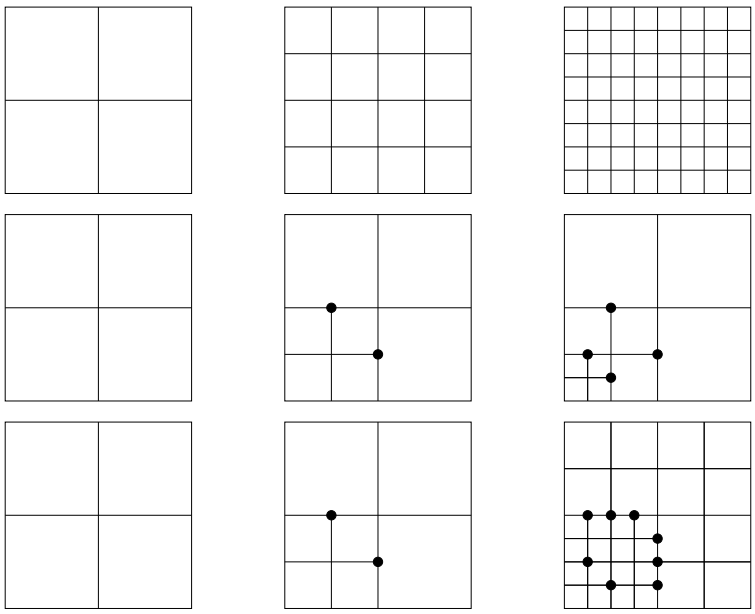


Fig. 3.2. Refined quadrilateral meshes: uniformly (upper row); locally (middle row); locally blockwise (lower row)

After having chosen an appropriate local basis $\{\varphi_k\}_{k=1}^N$ of V_h (usually the so-called "nodal basis"), the spatially semi-discrete formulation of problem (3.1) seeks an approximate solution in the form

$$u_h(x,t) = \sum_{k=1}^{N} y_k(t)\varphi_k(x)$$

where $y_k(t)$ are the (time-dependent) spatial "nodal values" of the finite element function u_h. These are determined by the semi-discrete equation

$$m(\partial_t^2 u_h, \varphi) + a(u_h, \varphi) = (f, \varphi) \quad \forall \varphi \in V_h, \tag{3.2}$$

or equivalently by the linear system of ODEs

$$M\ddot{y}(t) + Ky(t) = F(t), \tag{3.3}$$

for the nodal value vector $y(t) = (y_k(t))_{k=1}^N$. Here,

$$M := \big(m(\varphi_k, \varphi_i)\big)_{i,k=1}^N, \quad K := \big(a(\varphi_k, \varphi_i)\big)_{i,k=1}^N$$

are the so-called "mass matrix" and the "stiffness matrix", respectively, and $F := \big((f, \varphi_i)\big)_{i=1}^N$ is the "force vector". The initial conditions are given in the form

$$u_h(x,0) = \sum_{k=1}^N y_{0k}^0 \varphi_k(x), \quad \partial_t u_h(x,0) = \sum_{k=1}^N y_{0k}^1 \varphi_k(x),$$

where $y_0^0 := (y_{0k}^0)_{k=1}^N$ and $y_0^1 := (y_{0k}^1)_{k=1}^N$ are the nodal value vectors of the L^2 projections $P_h u_0^0$ and $P_h u_0^1$ in V_h of the initial data u_0^0 and u_0^1, respectively, or simply of the corresponding "nodal interpolations" $I_h u_0^0$ and $I_h u_0^1$ in V_h. Clearly, with these initial values the linear system (3.3) possesses a unique solution.

An alternative starting point for semidiscrete formulations begins with the "mixed" variational system

$$\begin{aligned} m(\partial_t u^0, \varphi^1) - m(u^1, \varphi^1) &= 0 \quad \forall \varphi^1 \in H, \\ m(\partial_t u^1, \varphi^0) + a(u^0, \varphi^0) &= (f, \varphi^1) \quad \forall \varphi^0 \in V, \end{aligned} \tag{3.4}$$

which is automatically satisfied on I by the pair $\{u^0, u^1\}$ where $u^0 := u$ and $u^1 := \partial_t u$ and u is the solution of (2.1). With the above nodal basis the corresponding spatially semi-discrete approximation is determined in the form

$$u_h^0(x,t) = \sum_{k=1}^N y_k^0(t) \varphi_k(x), \quad u_h^1(x,t) = \sum_{k=1}^N y_k^1(t) \varphi_k(x),$$

by the semi-discrete system

$$\begin{aligned} m(\partial_t u_h^0, \varphi^1) - m(u_h^1, \varphi^1) &= 0 \quad \forall \varphi^1 \in V_h, \\ m(\partial_t u_h^1, \varphi^0) + a(u_h^0, \varphi^0) &= (f, \varphi^1) \quad \forall \varphi^0 \in V_h. \end{aligned} \tag{3.5}$$

This is equivalent to the system of ODEs

$$M\dot{y}^0(t) - My^1(t) = 0, \quad M\dot{y}^1(t) + Ky^0(t) = F(t). \tag{3.6}$$

Remark 3.2. We note that due to the regularity of the mass matrix M the system (3.6) is equivalent to (3.3) and therefore, for any set of initial data, possesses a unique solution as well. This property depends on the fact that the same finite element ansatz has been chosen for the variables u_h^0 and u_h^1. This choice also implies that both variables strongly vanish on $\partial\Omega_D$ although the original mixed formulation (2.6) did not imply any boundary values for u^1.

3.2. Time discretization by finite difference schemes. We will consider time discretization by some of the most popular finite difference schemes, namely the "one-step-θ schemes", including as special cases the "backward Euler scheme" and the "Crank-Nicolson scheme", and then the class of "Newmark schemes". Here, for notational simplicity, we restrict ourselves to the "Method of Lines" since below these schemes will turn out to be

closely related to Galerkin time-stepping schemes which are more natural within the "Rothe method". At first, for a sequence of discrete time levels

$$0 = t_0 < t_1 < \cdots < t_m < \cdots < t_M = T,$$

we define the time-step lengths $k_m := t_m - t_{m-1}$ and set $k := \max_{m=1,\ldots,M} k_m$.

3.2.1. The one-step-θ schemes. The first-order system (3.6) is taken as starting point for the construction of the so-called "one-step-θ schemes". The approximations at the different time levels t_m are denoted by y_m^0 and y_m^1, respectively. Then, for any parameter value $\theta \in [0,1]$ the classical "one-step-θ scheme" reads as follows:

$$\begin{aligned} M(y_m^0 - y_{m-1}^0) - k_m\big(\theta M y_m^1 + (1-\theta) M y_{m-1}^1\big) &= 0, \\ M(y_m^1 - y_{m-1}^1) + k_m\big(\theta K y_m^0 + (1-\theta) K y_{m-1}^0\big) &= k_m\big(\theta F_m + (1-\theta) F_{m-1}\big). \end{aligned} \tag{3.7}$$

For $\theta = 0$ this scheme corresponds to the "explicit Euler scheme", for $\theta = 1$ to the "implicit Euler scheme",

$$M(y_m^0 - y_{m-1}^0) - k_m M y_m^1 = 0, \quad M(y_m^1 - y_{m-1}^1) + k_m K y_m^0 = k_m F_m. \tag{3.8}$$

and for $\theta = \frac{1}{2}$ to the "Crank-Nicolson scheme",

$$\begin{aligned} M(y_m^0 - y_{m-1}^0) - \tfrac{1}{2} k_m\big(M y_m^1 + M y_{m-1}^1\big) &= 0, \\ M(y_m^1 - y_{m-1}^1) + \tfrac{1}{2} k_m\big(K y_m^0 + K y_{m-1}^0\big) &= \tfrac{1}{2} k_m\big(F_m + F_{m-1}\big). \end{aligned} \tag{3.9}$$

The following properties are well-known from the literature (see Großmann & Roos [16]):

- *Stability:* The one-step-θ scheme is unconditionally stable in the L^2 norm (i.e. without any condition on the time-step sizes k_m) if and only if $\theta \in [\frac{1}{2}, 1]$.
- *Convergence:* The one-step-θ scheme is at least of order one in the time step size k; order two is achieved only for the choice $\theta = \frac{1}{2}$.
- *Energy conservation:* The one-step-θ scheme is energy conserving only for the choice $\theta = \frac{1}{2}$. For $\theta > \frac{1}{2}$ energy loss occurs, while for $\theta < \frac{1}{2}$ the scheme becomes unstable in the L^2 norm.

3.2.2. The Newmark schemes. The system (3.3) of second-order ODEs is taken as starting point for the construction of the Newmark schemes. This scheme attempts to approximate $y(t_m)$, $\dot{y}(t_m)$, $\ddot{y}(t_m)$ by a set of independent variables y_m^0, y_m^1, y_m^2, respectively. Using Taylor expansion up to order three and following the steps described in Wilson [36], one arrives at the fully discrete Newmark system,

$$\begin{aligned} M y_m^2 + K y_m^0 &= F_m, \\ y_m^0 &= y_{m-1}^0 + k_m y_{m-1}^1 + \tfrac{1}{2} k_m^2\big((1-2\beta) y_{m-1}^2 + 2\beta y_m^2\big), \\ y_m^1 &= y_{m-1}^1 + k_m\big((1-\gamma) y_{m-1}^2 + \gamma y_m^2\big), \end{aligned} \tag{3.10}$$

where β and γ are weighting parameters ("Newmark parameters").

All properties of the Newmark scheme (such as order of convergence, stability, discrete energy conservation, etc.) depend on the parameters β and γ. This is well known and can be found in more detail, e. g., in Wood [39, 40] and Hughes [19]. From the literature, we recall the following facts:

- *Stability:* The Newmark scheme is unconditionally stable in the L^2 norm (i.e. without any condition on the time-step sizes k_m) if and only if $2\beta \geqslant \gamma \geqslant \frac{1}{2}$.
- *Convergence:* The Newmark scheme is at least of order one; the order two is achieved only for the choice $\gamma = \frac{1}{2}$.
- *Energy conservation:* The Newmark scheme is energy conserving only for the choice $\gamma = \frac{1}{2}$. For $\gamma > \frac{1}{2}$ energy loss occurs, while for $\gamma < \frac{1}{2}$ the scheme becomes unstable in the L^2 norm.

In view of these properties the choice $2\beta = \gamma = \frac{1}{2}$ appears as particularly attractive in the Newmark scheme and leads to the equations

$$\begin{aligned} My_m^2 + Ky_m^0 &= F_m, \\ y_m^0 &= y_{m-1}^0 + k_m y_{m-1}^1 + \tfrac{1}{4}k_m^2\big(y_{m-1}^2 + y_m^2\big), \\ y_m^1 &= y_{m-1}^1 + \tfrac{1}{2}k_m\big(y_{m-1}^2 + y_m^2\big). \end{aligned} \tag{3.11}$$

This configuration is known as the "average acceleration method" (c.f. Hughes [19]) or the "Newmark trapezoidal rule". Below, we will demonstrate that this scheme is closely related to the Crank-Nicolson scheme described above. The restrictions imposed on the parameters β and γ are sharp, as can easily be verified through simple test problems. For a computational comparison of these methods, we refer to Goudreau & Taylor [15].

Remark 3.3. In order to apply the DWR method, it is crucial to start with a Galerkin formulation of the underlying method. Schweizerhof & al. [34] have derived an interpretation of the Newmark scheme as a Petrov-Galerkin method, for a special choice of the Newmark parameters. However, this Petrov-Galerkin approach is not satisfactory in the DWR framework, because the space of test functions resulting from the construction is merely the linear hull of one single quadratic polynomial. Hence, this interpretation yields no possibility to exploit Galerkin orthogonality, which is an essential feature of the DWR method.

Numerical tests. In order to illustrate the importance of the above restrictions on the parameters β and γ, we consider a simple test problem, namely the one-dimensional wave equation

$$\partial_t^2 u(x,t) - \partial_x^2 u(x,t) = 2t\cos\big(\tfrac{1}{2}\pi x\big) - \tfrac{1}{12}\pi^2\cos\big(\tfrac{1}{2}\pi x\big) \tag{3.12}$$

on the space-time region $(-1,1)\times(0,1]$, with initial conditions chosen so that the exact solution is $u(x,t) = \frac{1}{3}t^3\cos\big(\frac{1}{2}\pi x\big)$. For three different choices of the parameters, we show the behavior of the terms $\|\partial_x(u(T) - y_M^0)\|$ and $\|\partial_t u(T) - y_M^1\|$ at the final time $t_M = T = 1$ as functions of the number of time steps. As predicted, for $\beta = \frac{3}{10}, \gamma = \frac{7}{10}$, we observe instability (Fig. 3.3), for $\beta = 1, \gamma = \frac{9}{10}$ first-order convergence (Fig. 3.4) and for $\beta = \frac{1}{4}, \gamma = \frac{1}{2}$ second-order convergence (Fig. 3.5).

Finally, we consider the homogeneous version of the wave equation (3.12) on the space-time region $(0,1)\times(0,5]$ with the exact solution $u(x,t) = \sin(\pi x)\sin(\pi t)$. For this model the exact total energy is $E(t) = \frac{1}{2}\pi^2 \approx 4.9348$. The behavior in time of the approximate energy is shown in Fig. 3.6. Again as predicted, for $\gamma = \frac{1}{2}$, we observe perfect energy conservation while for $\gamma < \frac{1}{2}$ energy decay occurs and for $\beta < \frac{1}{2}\gamma$ the scheme is rendered unstable.

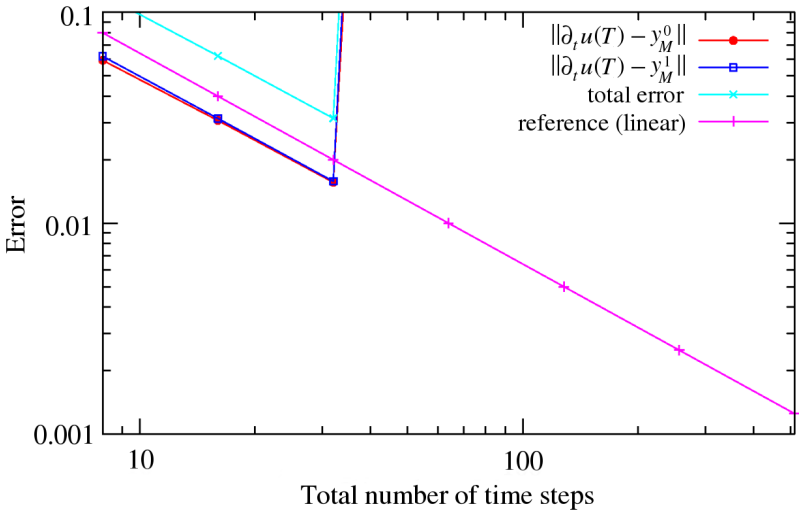


Fig. 3.3. Newmark scheme ($\beta = \frac{3}{10}, \gamma = \frac{7}{10}$): instability

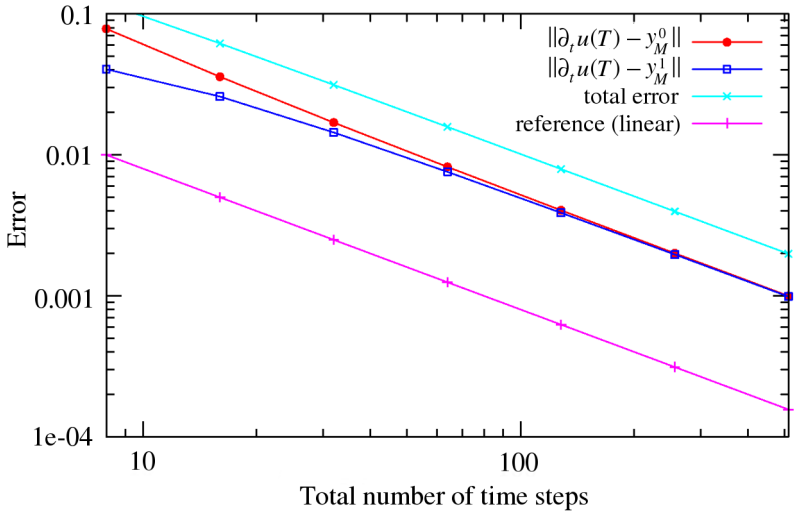


Fig. 3.4. Newmark scheme ($\beta = 1, \gamma = \frac{9}{10}$): first-order convergence

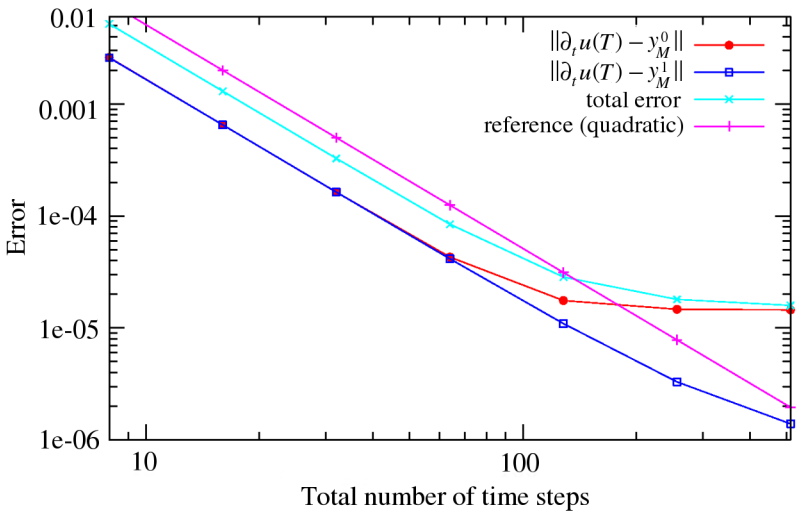
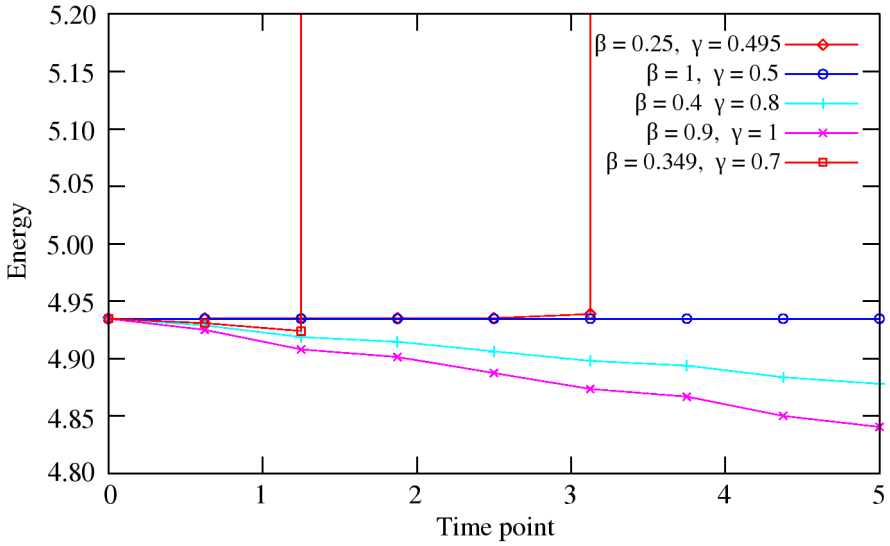


Fig. 3.5. Newmark trapezoidal scheme ($\beta = \frac{1}{4}, \gamma = \frac{1}{2}$): second-order convergence. For very small time steps, the spatial error on the order of 10^{-5} dominates



F i g. 3.6. Newmark scheme for different values of the parameters γ and β: energy conservation and stability properties over 6 400 time steps

3.2.3. Relation between the Newmark and the Crank-Nicolson scheme. Our next goal is to formally establish the algebraic equivalence of the Crank-Nicolson and the trapezoidal Newmark scheme (3.11) obtained from the general Newmark scheme (3.10) by the particular choice $\gamma = 2\beta = \frac{1}{2}$. In view of a remark given at the end of this section, let us show the derivation for the slightly more general case $\gamma = 2\beta$. We will later specialize this for $\gamma = \frac{1}{2}$. Following this, we will give a brief outline of the opposite derivation, i.e. how to obtain the Newmark scheme from the Crank-Nicolson method.

For our considerations, let us start with the general Newmark system (3.10) and fixing $\gamma = 2\beta$. Multiplying the second and third equations by the mass matrix M then yields

$$My_m^2 = F_m - Ky_m^0,$$

$$My_m^0 = My_{m-1}^0 + k_m My_{m-1}^1 + \tfrac{1}{2}k_m^2\big\{(1-\gamma)My_{m-1}^2 + \gamma My_m^2\big\},$$
$$My_m^1 = My_{m-1}^1 + k_m\big((1-\gamma)My_{m-1}^2 + \gamma My_m^2\big).$$

We can use the third equation to replace the term in braces in the second equation. We then use the first equation evaluated at time levels t_m and t_{m-1} to eliminate the remaining occurrences of y_m^2 and y_{m-1}^2. This results in

$$My_m^0 = My_{m-1}^0 + k_m My_{m-1}^1 + \tfrac{1}{2}k_m\big(My_m^1 - My_{m-1}^1\big),$$

$$My_m^1 = My_{m-1}^1 + k_m\big((1-\gamma)(F_{m-1} - Ky_{m-1}^0) + \gamma(F_m - Ky_m^0)\big).$$

From this, we obtain by simple transformations the system

$$My_m^0 = My_{m-1}^0 + \tfrac{1}{2}k_m(My_{m-1}^1 + My_m^1),$$

$$My_m^1 = My_{m-1}^1 + k_m\big(((1-\gamma)F_{m-1} + \gamma F_m) - ((1-\gamma)Ky_{m-1}^0 + \gamma Ky_m^0)\big). \tag{3.13}$$

Here, the first equation already has a trapezoidal rule structure, whereas the second one is a kind of weighted trapezoidal rule or, more precisely, a convex combination of F and Ky

between t_{m-1} and t_m. If we now take $\gamma = \frac{1}{2}$ (which implies $\beta = \frac{1}{4}$ and therefore leads to the "trapezoidal" Newmark scheme), we finally come to the desired result

$$My_m^0 - \tfrac{1}{2}k_m My_m^1 = My_{m-1}^0 + \tfrac{1}{2}k_m My_{m-1}^1,$$

$$My_m^1 + \tfrac{1}{2}k_m Ky_m^0 = My_{m-1}^1 - \tfrac{1}{2}k_m Ky_{m-1}^0 + \tfrac{1}{2}k_m(F_{m-1} + F_m). \tag{3.14}$$

Obviously, this system is equivalent to the Crank-Nicolson scheme (3.9).

Now, we sketch how to obtain the Newmark trapezoidal rule (3.11) from the Crank-Nicolson equations (3.9). As we have seen before, the Crank-Nicolson scheme can also be directly deduced from the two-component system (3.6). Therefore it is essential that we discretize both components using the same basis $\{\varphi_k(x)\}_{k=1}^N$ of the spatial finite element space $V_h \subset V$. By simple algebraic manipulations of equations (3.9), we obtain the following system:

$$My_m^0 - My_{m-1}^0 - k_m My_{m-1}^1 = \tfrac{1}{2}k_m(My_m^1 - My_{m-1}^1),$$

$$My_m^1 - My_{m-1}^1 = \tfrac{1}{2}k_m(F_{m-1} + F_m - Ky_{m-1}^0 - Ky_m^0).$$

Now the only problem is the absence of a variable y_m^2. We can introduce such a variable by setting

$$My_m^2 := F_m - Ky_m^0, \tag{3.15}$$

which simply comes from the spatially semi-discrete system (3.3), and replacing occurrences of $F_m - Ky_m$ in the Crank-Nicolson scheme by My_m^2. Here again it is crucial to use the same basis $\{\varphi_k\}_{k=1}^N$ of V_h as above, because otherwise the mass matrices in (3.15) and in the Crank-Nicolson scheme (3.14) would not be the same. Once we have taken into account equation (3.15), we immediately obtain the trapezoidal Newmark rule (3.11).

Remark 3.4. It can be shown without any difficulty that the equivalence between the Crank-Nicolson scheme and the Newmark trapezoidal rule still holds true if we add damping to the underlying PDE system. In particular, the cases of weak damping (corresponding to a damping term $\partial_t u(x,t)$) and strong damping (realized by adding $\partial_t \mathcal{A}u(x,t)$) are possible, where in the latter case the assumptions about the spaces for the data and the solution have to be properly adjusted.

Remark 3.5. Other choices of the parameters γ and β also lead to variants of the Newmark scheme that can be re-interpreted as widely used finite difference schemes. For example, for $\gamma = 2\beta = 1$ the equations for y^0 and y^1 take the structure of the Crank-Nicolson and the implicit Euler scheme, respectively. We could arrive at this scheme by allowing two different values θ^0, θ^1 in the two equations of system (3.7) and choosing $\theta^0 = \frac{1}{2}, \theta^1 = 1$ (in general, every Newmark scheme with $\gamma = 2\beta$ can be written as a one-step-θ scheme with $\theta^0 = \frac{1}{2}, \theta^1 = \gamma$, as is obvious by comparing (3.13) with (3.7)). As we will see below, the two schemes mentioned are in fact related to continuous and discontinuous Galerkin schemes. Consequently, the rich error estimation theory available for (Petrov-) Galerkin methods can also applied to this version of the Newmark scheme.

3.3. Time discretization by the Galerkin finite element method. After discussing two of the traditional finite difference schemes for time discretization of the spatially semi-discretized wave equation, let us now turn our attention to Galerkin methods for time

discretization. These time discretizations are usually based on a variational form of the mixed formulation (2.6) of the wave equation,

$$\begin{aligned} \rho\partial_t u^0 - \rho u^1 &= 0 \quad \text{in} \quad \Omega\times I,\\ \rho\partial_t u^1 + \mathcal{A}u^0 &= f \quad \text{in} \quad \Omega\times I, \end{aligned} \tag{3.16}$$

for the pair $\{u^0,u^1\}=\{u,\partial_t u\}$ with the boundary and initial conditions remaining as in (2.6). In order to set up corresponding variational formulations for the pairs $\bar u:=\{u^0,u^1\}\in \mathcal{V}\times\mathcal{H}$ we use (non-overlapping) decompositions

$$I=\bigcup_{m=1}^{M} I_m$$

of the time interval $I=(0,T]$ into half-open subintervals $I_m:=(t_{m-1},t_m]$ of length $k_m:=t_m-t_{m-1}$. We set $k:=\max_{m=1,\dots,M}k_m$ and introduce the space-time scalar products

$$((\varphi,\psi)):=\sum_{m=1}^{M}\int_{I_m}(\varphi,\psi)\,\mathrm{d}t,\quad m((\varphi,\psi)):=\sum_{m=1}^{M}\int_{I_m}m(\varphi,\psi)\,\mathrm{d}t,\quad a((\varphi,\psi)):=\sum_{m=1}^{M}\int_{I_m}a(\varphi,\psi)\,\mathrm{d}t.$$

Then, the system (2.6) is equivalent to the following variational problem: Find a pair $\{u^0,u^1\}\in\hat{\mathcal{V}}^0\times\hat{\mathcal{V}}^1$ satisfying

$$\begin{aligned} m((\partial_t u^0,\varphi^1)) - m((u^1,\varphi^1)) + m(u^0(0),\varphi^1(0)) &= m(u_0^0,\varphi^1(0)),\\ m((\partial_t u^1,\varphi^0)) + a((u^0,\varphi^0)) + m(u^1(0),\varphi^0(0)) &= m(u_0^1,\varphi^0(0)) + ((f,\varphi^0)), \end{aligned} \tag{3.17}$$

for all $\{\varphi^1,\varphi^0\}\in\mathcal{W}\times\mathcal{W}$, where the test space is taken as

$$\mathcal{W}:=\big\{v\in L^2(I;V)\;\big|\;v|_{I_m}\in C(\bar I_m;V),\ m=1,\dots,M\big\}.$$

Here, the notation $v|_{I_m}\in C(\bar I_m;V)$ means that $v|_{I_m}$ possesses a continuous continuation to the closure $\bar I_m$ of I_m. In this formulation the initial conditions are imposed in the weak sense. In the following, we will first discretize in time and only then in space, i.e. we will follow the Rothe approach to fully discretized systems.

3.3.1. The "continuous-in-time" Galerkin ($cG(r)$) schemes. Let $P_r(I_n;V)$ denote the space of all polynomial functions of maximum degree r on I_n with values in V. For the time discretization of system (3.17), we introduce the following two finite dimensional subspaces of $L^2(I;V)$, for $r\in\mathbb{N}$:

$$S_k^{r,\mathrm{c}}(I;V):=\big\{p\in C(\bar I;V)^d\;\big|\;p|_{I_m}\in P_r(I_m;V)^d,\ m=1,\dots,M\big\},$$

which will be the space of continuous *trial functions* for the time-discrete variational formulation, and

$$S_k^{r-1,\mathrm{d}}(I;V):=\big\{p\in L^2(I;V)^d\;\big|\;p|_{I_m}\in P_{r-1}(I_m;V)^d,\ m=1,\dots,M\big\},$$

which will be the space of discontinuous *test functions*. Here, the superscripts "c" and "d" refer to the continuity or discontinuity of trial and test functions at time instants t_m, respectively. In the description of the time-discretization schemes, we will use the abbreviated notation $\mathcal{V}_k:=S_k^{r,\mathrm{c}}(I;V)$ and $\mathcal{W}_k:=S_k^{r-1,\mathrm{d}}(I;V)$. Clearly, there holds

$$\mathcal{V}_k\subset\mathcal{V},\quad \mathcal{W}_k\subset\mathcal{W}.$$

Remark 3.6. In general, we will assume the *test functions* to be globally discontinuous in time, and on each subinterval of one polynomial order lower than the (globally continuous) trial functions, because one degree of freedom per subinterval of the trial functions is already fixed by the global continuity condition. By the different polynomial order, we are thus led to a quadratic system of equations, and by the lacking continuity of the test functions the system can be decoupled at each full time step. Hence the resulting scheme may be reinterpreted as a time-stepping method.

With the spaces defined above the time-discrete variational problem seeks a pair $\{u_k^0, u_k^1\} \in \mathcal{V}_k \times \mathcal{V}_k$ satisfying

$$m((\partial_t u_k^0, \varphi^1)) - m((u_k^1, \varphi^1)) + m(u_k^0(0), \varphi^1(0)) = m(u_0^0, \varphi^1(0)),$$

$$m((\partial_t u_k^1, \varphi^0)) + a((u_k^0, \varphi^0)) + m(u_k^1(0), \varphi^0(0)) = m(u_0^1, \varphi^0(0)) + ((f, \varphi^0)), \tag{3.18}$$

for all test pairs $\{\varphi^1, \varphi^0\} \in \mathcal{W}_k \times \mathcal{W}_k$. By adding up the two equations (3.18), we obtain a compact expression for the semi-discrete equations: Find $\hat{u}_k = \{u_k^0, u_k^1\} \in \mathcal{V}_k \times \mathcal{V}_k$ satisfying

$$A(\hat{u}_k, \varphi) = F(\varphi) \quad \forall \varphi = \{\varphi^1, \varphi^0\} \in \mathcal{W}_k \times \mathcal{W}_k, \tag{3.19}$$

with the bilinear form and force term defined, respectively, as follows:

$$A(\hat{u}_k, \varphi) := m((\partial_t u_k^0, \varphi^1)) - m((u_k^1, \varphi^1)) + m(u_{k,0}^0, \varphi_0^1) + m((\partial_t u_k^1, \varphi^0)) + a((u_k^0, \varphi^0)) + m(u_{k,0}^1, \varphi_0^0),$$

$$F(\varphi) := m(u_0^0, \varphi_0^1) + m(u_0^1, \varphi_0^0) + ((f, \varphi^0)).$$

This time discretization may be viewed as a Petrov-Galerkin method with test space $\mathcal{W}_k \times \mathcal{W}_k$ different from the trial space $\mathcal{V}_k \times \mathcal{V}_k$. We note that the discretization (3.19) is strongly consistent with the continuous problem in mixed formulation (2.6), i.e., the exact solution $\hat{u} := \{u^0, u^1\} = \{u, \partial_t u\} \in \hat{\mathcal{V}}^0 \times \hat{\mathcal{V}}^1$ automatically satisfies

$$A(\hat{u}, \varphi) = F(\varphi), \quad \varphi \in \mathcal{W}_k \times \mathcal{W}_k, \tag{3.20}$$

and therefore the following "Galerkin orthogonality" relation holds:

$$A(\hat{u} - \hat{u}_k, \varphi) = 0, \quad \varphi \in \mathcal{W}_k \times \mathcal{W}_k. \tag{3.21}$$

To discretize equation (3.19) in space, we replace the continuous space V in the definition of the above trial and test spaces by the finite element space V_h defined above. Here, for notational simplicity, we restrict ourselves to the "Method of Lines" approach, i. e., the finite element space V_h is kept the same over the whole time interval I. Then, using the fully discrete spaces

$$\mathcal{V}_{hk} := S_k^{r,\mathrm{c}}(I; V_h), \quad \mathcal{W}_{hk} := S_k^{r-1,\mathrm{d}}(I; V_h)$$

the resulting fully discrete problem seeks a pair $U = \{U^0, U^1\} \in \mathcal{V}_{hk} \times \mathcal{V}_{hk}$ satisfying

$$m((\partial_t U^0, \varphi^1)) - m((U^1, \varphi^1)) + m(U_0^0, \varphi_0^1) = m(u_0^0, \varphi_0^1),$$

$$m((\partial_t U^1, \varphi^0)) + a((U^0, \varphi^0)) + m(U_0^1, \varphi_0^0) = m(u_0^1, \varphi_0^0) + ((f, \varphi^0)), \tag{3.22}$$

for all test pairs $\varphi = \{\varphi^1, \varphi^0\} \in \mathcal{W}_{hk} \times \mathcal{W}_{hk}$. As above these equations can be written in compact form: Find $U \in \mathcal{V}_{hk} \times \mathcal{V}_{hk}$ satisfying

$$A(U, \varphi) = F(\varphi) \quad \forall \varphi \in \mathcal{W}_{hk} \times \mathcal{W}_{hk}. \tag{3.23}$$

It is easy to show that the conservation property carries over to the solution U of the spatially discretized equations provided that the meshes do not change between time levels.

Fully discrete problems written in the form (3.22) or (3.23) possess unique solutions. We do not give this argument for general polynomial degree $r \in \mathbb{N}$ but refer to the relevant literature (Johnson [24], Bales & Lasiecka [1] and French & Peterson [13]). For the lowest-order case $r = 1$ this follows from the equivalence of this particular Galerkin method to the well-known Crank-Nicolson scheme, which will be established below.

In shorthand notation, the full discretization of the wave equation represented by (3.23) is denoted as the cG(s)/cG(r) method where "cG" stands for "continuous Galerkin" (i.e. continuous trial functions) and $s, r \in \mathbb{N}$ refer to the local polynomial degrees of the trial functions in space and time, respectively. Below, we will focus on the simplest version of this method, namely the cG(1)/cG(1) method, which uses continuous piecewise linear/n-linear trial functions in space as well as in time. This discretization is of total second order and admits *a priori* error estimates of the form

$$\sup_{t\in I} \left\{ \|(u^0 - U^0)(t)\| + \|(u^1 - U^1)(t)\| \right\} = \mathcal{O}(h^2 + k^2), \tag{3.24}$$

provided that the continuous solution of problem (2.1) is sufficiently smooth (see French & Peterson [13]).

Remark 3.7. Below, it will be shown that if the forcing term is zero or constant in time, the cG(1)/cG(1) method is algebraically equivalent to the Crank-Nicolson scheme. We have already seen that the latter in turn is equivalent to the trapezoidal Newmark scheme provided that the same finite element basis is used in the spatial discretization. Hence, any result known for one of these discretization methods immediately carries over to the other two schemes. Since here the order of the spatial finite element ansatz does not explicitly occur, the above arguments also hold for general cG(s)/cG(1) methods with $s \in \mathbb{N}$.

Remark 3.8. The Rothe method underlying an adaptive space-time discretization first discretizes in time and leaves the task of spatial discretization as a second step. This opens up the possibility of using different finite element meshes for different time steps, for example to resolve a wave front that moves through the domain. In this context, the use of tensor-product space-time meshes is advisable in order to facilitate the separate local adaptation of spatial and temporal mesh sizes. However, this is not without practical difficulties: In the cG(r) time discretization the *trial functions* have to be continuous in time. For tensor-product space-time meshes this limits the flexibility in adapting the spatial meshes (general "remeshing" or cell shifting is prohibited) and requires the introduction of spurious "hanging nodes" in the spatial meshes $\mathbb{T}_h^m$ (see Fig. 3.2). In these "irregular" nodal points the unknowns are eliminated by interpolating values at neighboring "regular" nodal points. Accordingly, in the time slabs $\bar{\Omega} \times I_m$ the trial functions are defined on spatial meshes which are combinations of the meshes $\mathbb{T}_h^{m-1}$ and $\mathbb{T}_h^m$ at the two end points t_{m-1} and t_m. In contrast to that, the "test functions", which are allowed to be discontinuous in time, are defined in the time slabs $\bar{\Omega} \times I_m$ on the spatial meshes corresponding to $\mathbb{T}_h^m$, i.e. the meshes at the right end points t_m. Furthermore, in this discretization temporal trial functions have support in both time intervals I_m, I_{m+1} adjacent to time instants t_m. On each time interval I_m, the trial functions defined at time instant t_{m-1} therefore overlap with test functions defined at time instant t_m, which implies that we also have to form space-time integrals of trial functions from $P^r(I_m, V_h^{m-1})$ defined on a mesh $\mathbb{T}_h^{m-1}$ times test functions from

$P^{r-1}(I_m, V_h^m)$ defined on a mesh $\mathbb{T}_h^m$. This integration could be done on a subdivision of Ω that consists of the "union" of the two triangulations,

$$\mathbb{T}_h^{m-1,m} := \left\{ \omega = K \cap K' \;\middle|\; K \in \mathbb{T}_h^{m-1}, K' \in \mathbb{T}_h^m, K \cap K' \neq \varnothing \right\}.$$

In two dimensions this set is the subdivision of $\bar{\Omega}$ by the union of mesh lines of $\mathbb{T}_h^{m-1}$ and $\mathbb{T}_h^m$. On the other hand, due to the irregular structure of the elements of $\mathbb{T}_h^{m-1,m}$, computations are only feasible with reasonable effort if the grids $\mathbb{T}_h^{m-1}$ and $\mathbb{T}_h^m$ are related in some way. The only choice of meshes that allows to evaluate such integrals efficiently involves hierarchically refined grids where $\mathbb{T}_h^{m-1,m}$ is the set of most refined cells from the two grids. Details on the practical implementation of such approaches can be found in Bangerth [2], Bangerth & Rannacher [3,5] and Schmich & Vexler [33]; see also Süli & Wilkins [35] and Bernardi & Süli [8].

3.3.2. Relation between the $cG(1)/cG(1)$ and Crank-Nicolson schemes. Due to the discontinuity in time of the test functions, the global space-time problem (3.22) can be written in the form of a time-stepping scheme. Let us now consider the case $r = 1$, i.e. trial functions are linear in time. Starting from the L^2 projections into V_h of the initial values u_0^0, u_0^1, the local problem on the subinterval $I_m = (t_{m-1}, t_m]$ reads

$$\int_{I_m} \left\{ m(\partial_t U^0, \varphi^1) - m(U^1, \varphi^1) \right\} \mathrm{d}t = 0 \quad \forall \varphi^1 \in P_0(I_m; V_h)^d,$$
$$\int_{I_m} \left\{ m(\partial_t U^1, \varphi^0) + a(U^0, \varphi^0) \right\} \mathrm{d}t \int_{I_m} (f, \varphi^0)\, \mathrm{d}t \quad \forall \varphi^0 \in P_0(I_m; V_h)^d.$$

With the nodal basis $\{\varphi_1, \ldots, \varphi_N\}$ of the spatial finite element space V_h, the linear-in-time trial function $U = \{U^0, U^1\}$ on I_m can be written as

$$U^j(x,t)|_{I_m} = \sum_{k=1}^N \left\{ y_{m-1,k}^j + k_m^{-1}(t - t_{m-1})(y_{m,k}^j - y_{m-1,k}^j) \right\} \varphi_k(x), \quad j \in \{0,1\}.$$

Here, $y_{m,k}^j$ and $y_{m-1,k}^j$ are the nodal values of U^j at the time levels t_{m-1} and t_m, where $y_{m-1,k}^j$ is known from the previous time step. Inserting this trial function into the above equations and testing successively with the basis functions φ_i, $i = 1, \ldots, N$, we obtain the following fully discrete equations for the first component,

$$\sum_{k=1}^N (y_{m,k}^0 - y_{m-1,k}^0) m(\varphi_k, \varphi_i) - \frac{1}{2} k_m \sum_{k=1}^N y_{m-1,k}^1 m(\varphi_k, \varphi_i) -$$
$$\frac{1}{2} k_m \sum_{k=1}^N y_{m,k}^1 m(\varphi_k, \varphi_i) = 0, \quad i = 1, \ldots, N,$$

and analogously for the second component,

$$\sum_{k=1}^N (y_{m,k}^1 - y_{m-1,k}^1) m(\varphi_k, \varphi_i) + \frac{1}{2} k_m \sum_{k=1}^N y_{m-1,k}^0 a(\varphi_k, \varphi_i) +$$
$$\frac{1}{2} k_m \sum_{k=1}^N y_{m,k}^0 a(\varphi_k, \varphi_i) = \int_{I_m} (f, \varphi_i)\, \mathrm{d}t, \quad i = 1, \ldots, N.$$

With the above notation for the mass and stiffness matrices, $M = (m(\varphi_k, \varphi_i))_{i,k=1}^N$ and $K = (a(\varphi_k, \varphi_i))_{i,k=1}^N$ respectively, and the force vector $F = (f, \varphi_i)_{i=1}^N$, this can be written in a more compact way as

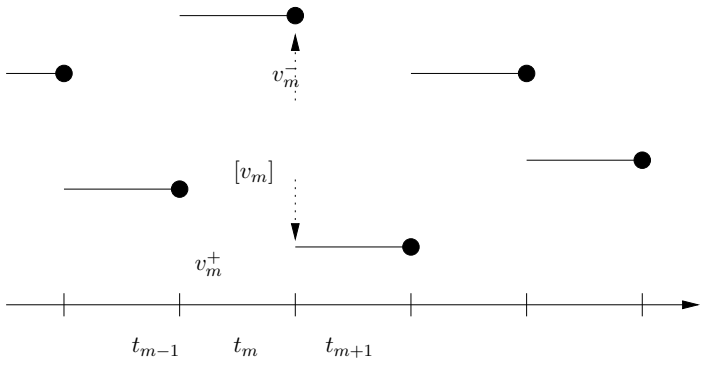
$$My_m^0 - \tfrac{1}{2}k_m My_m^1 = My_{m-1}^0 + \tfrac{1}{2}k_m My_{m-1}^1, \quad My_m^1 + \tfrac{1}{2}k_m Ky_m^0 = My_{m-1}^1 - \tfrac{1}{2}k_m Ky_{m-1}^0 + \int_{I_m} F(t)\mathrm{d}t. \tag{3.25}$$

From these equations it is evident that the cG(1)/cG(1) method can be regarded as a time-stepping scheme, which for zero forcing, i.e. $f \equiv 0$, coincides with the Crank-Nicolson scheme (trapezoidal rule) applied to the spatially semi-discrete variational system (3.6). In the case of non-zero forcing this equivalence is only modulo the evaluation of the time-integral on the right-hand side by the trapezoidal rule, i.e. up to a term of higher order $\mathcal{O}(k^3)$.

3.3.3. The "discontinuous-in-time" Galerkin (dG(r)) schemes. An alternative to the choice of continuous-in-time trial functions is the "discontinuous" Galerkin method that uses trial as well as test functions that may be discontinuous across time points t_m. Continuity of the solution is then enforced in the variational sense for the limit $k \to 0$. We introduce the notation

$$v_m^{\pm} = \lim_{s \searrow 0} v(x, t_m \pm s), \quad [v_m] = v_m^+ - v_m^-,$$

for the one-sided limits of piecewise continuous functions and the corresponding "jumps" at the discrete time points t_m (see Fig. 3.7).



F i g. 3.7. Left/right limit and jump in the lowest-order dG(0) method

The starting point for formulating the dG(r) time discretization is again the mixed variational formulation (3.17). We will use the time-discrete spaces

$$S_k^{r,\mathrm{d}}(I;V) = \big\{p \in L^2(I;V)^d \;\big|\; p|_{I_m} \in P_r(I_m;V)^d,\ m = 1, \ldots, M\big\},$$

and in abbreviated notation $\mathcal{V}_k := S_k^{r,\mathrm{d}}(I;V)$. Then, the time-discrete variational problem seeks a pair $\{u_k^0, u_k^1\} \in \mathcal{V}_k \times \mathcal{V}_k$ satisfying

$$m(\!(\partial_t u_k^0, \varphi_k^1)\!) + \sum_{m=1}^{M-1} m([u_{k,m}^0], \varphi_{k,m}^{1,+}) - m(\!(u_k^1, \varphi_k^1)\!) + m(u_{k,0}^{0,+}, \varphi_{k,0}^{1,+}) = m(u_0^0, \varphi_{k,0}^{1,+}),$$

$$m(\!(\partial_t u_k^1, \varphi_k^0)\!) + \sum_{m=1}^{M-1} m([u_{k,m}^1], \varphi_{k,m}^{0,+}) + a(\!(u_k^0, \varphi_k^0)\!) + m(u_{k,0}^{1,+}, \varphi_{k,0}^{0,+}) = m(u_0^1, \varphi_{k,0}^{0,+}) + (\!(f, \varphi_k^0)\!), \tag{3.26}$$

for all test pairs $\{\varphi^1,\varphi^0\} \in \mathcal{W}_k \times \mathcal{W}_k$, where in this particular case we can take $\mathcal{W}_k := \mathcal{V}_k$. Again by adding up the two equations (3.26), we obtain a compact expression for the semi-discrete equations: Find $u_k \in \mathcal{V}_k \times \mathcal{V}_k$ satisfying

$$A(u_k,\varphi) = F(\varphi) \quad \forall \varphi = \{\varphi^1,\varphi^0\} \in \mathcal{V}_k \times \mathcal{V}_k, \tag{3.27}$$

with the bilinear form and force term defined, respectively, as follows:

$$A(u_k,\varphi) := m(\!(\partial_t u_k^0,\varphi_k^1)\!) + \sum_{m=1}^{M-1} m([u_{k,m}^0],\varphi_{k,m}^{1,+}) - m(\!(u_k^1,\varphi_k^1)\!) + m(u_{k,0}^{0,+},\varphi_{k,0}^{1,+}) +$$
$$m(\!(\partial_t u_k^1,\varphi_k^0)\!) + \sum_{m=1}^{M-1} m([u_{k,m}^1],\varphi_{k,m}^{0,+}) + a(\!(u_k^0,\varphi_k^0)\!) + m(u_{k,0}^{1,+},\varphi_{k,0}^{0,+}),$$
$$F(\varphi) := m(u_0^0,\varphi_{k,0}^{1,+}) + m(u_0^1,\varphi_{k,0}^{0,+}) + (\!(f,\varphi_k^0)\!).$$

We note that the discretization (3.27) again is strongly consistent with the continuous problem in mixed formulation (2.6), i.e., the exact solution $\hat{u} := \{u^0,u^1\} = \{u,\partial_t u\} \in \hat{\mathcal{V}}^0 \times \hat{\mathcal{V}}^1$ also automatically satisfies

$$A(\hat{u},\varphi) = F(\varphi), \quad \varphi \in \mathcal{W}_k \times \mathcal{W}_k, \tag{3.28}$$

and therefore the following "Galerkin orthogonality" relation holds:

$$A(\hat{u}-u_k,\varphi) = 0, \quad \varphi \in \mathcal{W}_k \times \mathcal{W}_k. \tag{3.29}$$

To discretize equations (3.26) in space, we again replace the continuous space V in the definition of the above trial and test spaces by the finite element space V_h defined above. Since the trial as well as the test functions in the dG(r) method may be discontinuous in time it is natural to allow the spatial finite element spaces to be different on each of the subintervals I_m, which is indicated by V_h^m using the additional superscript m. Accordingly, we introduce the fully discrete function space

$$\mathcal{V}_{hk} := \big\{\varphi \in S_k^{r,\mathrm{d}}(I;V) \;\big|\; \varphi|_{I_m} \in P_r(I_m;V_h^m)^d,\ m=1,\dots,M\big\}.$$

Referring to Remark 3.8, in each time slab $\bar{\Omega}\times I_m$ the "trial" as well as the "test" functions are defined on a common spatial mesh $\mathbb{T}_h^m$, which corresponds to that from the right end point t_m.

Then, the resulting fully discrete problem seeks a pair $\{U^0,U^1\} \in \mathcal{V}_{hk} \times \mathcal{V}_{hk}$ satisfying

$$m(\!(\partial_t U_k^0,\varphi_k^1)\!) + \sum_{m=1}^{M-1} m([U_{k,m}^0],\varphi_{k,m}^{1,+}) - m(\!(U_k^1,\varphi_k^1)\!) + m(U_{k,0}^{0,+},\varphi_{k,0}^{1,+}) = m(U_0^0,\varphi_{k,0}^{1,+}),$$
$$m(\!(\partial_t U_k^1,\varphi_k^0)\!) + \sum_{m=1}^{M-1} m([U_{k,m}^1],\varphi_{k,m}^{0,+}) + a(\!(U_k^0,\varphi_k^0)\!) + m(U_{k,0}^{1,+},\varphi_{k,0}^{0,+}) = m(U_0^1,\varphi_{k,0}^{0,+}) + (\!(f,\varphi_k^0)\!), \tag{3.30}$$

for all test pairs $\{\varphi^1,\varphi^0\} \in \mathcal{W}_{hk} \times \mathcal{W}_{hk}$ where again we can take $\mathcal{W}_{hk} := \mathcal{V}_{hk}$. As above, these equations can be written in compact form: Find $U = \{U^0,U^1\} \in \mathcal{V}_{hk} \times \mathcal{V}_{hk}$ satisfying

$$A(U,\varphi) = F(\varphi) \quad \forall \varphi = \{\varphi^1,\varphi^0\} \in \mathcal{V}_{hk} \times \mathcal{V}_{hk}. \tag{3.31}$$

It can be shown that the fully discrete problems written in the form (3.30) or (3.31) possess unique solutions. For the lowest-order case $r = 0$ this follows from the equivalence of this particular Galerkin method to the well-known backward Euler scheme, which will be established below.

The full discretization of the wave equation represented by (3.31) is denoted as the cG(s)/dG(r) method where "dG" stands for "discontinuous Galerkin" (i.e. trial and test functions discontinuous in time) and $s, r \in \mathbb{N}$ refer to the local polynomial degrees of the trial functions in space and time, respectively. Below, we will focus on the simplest version of this method, namely the cG(1)/dG(0) method, which uses trial and test functions which are continuous piecewise linear/n-linear in space and discontinuous piecewise constant in time. This discretization is of second order in space but only of first order in time and admits *a priori* error estimates of the form

$$\sup_{t\in I} \left\{ \|(u^0 - U^0)(t)\| + \|(u^1 - U^1)(t)\| \right\} = \mathcal{O}(h^2 + k), \tag{3.32}$$

provided that the continuous solution of problem (2.1) is sufficiently smooth (see Johnson [24] and Hughes & Hulbert [20]).

Remark 3.9. Following Remark 3.8, one may think that the dG(r) method is better suited to choosing different meshes in different time steps since both trial and test functions are now entirely localized to individual time intervals, and functions defined on different meshes do no longer overlap in time. However, this is not so: System (3.30) calls for the integration of the jump terms $[U^i_{k,m}], i = 0, 1$ against test functions $\varphi^{i,+}_{k,m}$. Such terms also combine functions defined on different meshes if the mesh changes at time instant t_m, resulting in the same difficulties encountered with the cG(r) method.

3.3.4. Relation between the cG(1)/dG(0) and backward Euler schemes. The trial and test functions in the cG(1)/dG(0) method are piecewise constant in time. Hence setting $U^i(t_m) = U^i_{m-1}$, we have that $U^i|_{I_m} = U^i_m$, and further $\dot{U}^i \equiv 0$ for $i \in \{0, 1\}$. This implies for $m = 1, \ldots, M$ the identities

$$U^{i,-}_m = U^i_{m-1}, \quad U^{i,+}_m = U^i_m, \quad [U^i_m] = U^i_m - U^i_{m-1}.$$

Here, we set $U^{i,-}_0$ equal to the corresponding initial value u^i_0. With these identities, (3.30) takes on the following form for $m = 1, \ldots, M$:

$$m(U^0_m, \varphi^1) - \int_{I_m} m(U^1_m, \varphi^1)\,\mathrm{d}t = m(U^0_{m-1}, \varphi^1),$$

$$m(U^1_m, \varphi^0) + \int_{I_m} a(U^0_m, \varphi^0)\,\mathrm{d}t = \int_{I_m} (f, \varphi^0)\,\mathrm{d}t + m(U^1_{m-1}, \varphi^0), \tag{3.33}$$

for all $\varphi \in P_0(I_m; V_h)^d \times P_0(I_m; V_h)^d$. Hence with the above notation for the mass and stiffness matrices, $M = (m(\varphi_k, \varphi_i))^N_{i,k=1}$ and $K = (a(\varphi_k, \varphi_i))^N_{i,k=1}$ respectively, and the force vector $F = (f, \varphi_i)^N_{i=1}$ this can be written more compactly as

$$MU^0_m - k_m MU^1_m = MU^0_{m-1}, \quad MU^1_m + k_m KU^0_m = MU^1_{m-1} + \int_{I_m} F(t)\,\mathrm{d}t. \tag{3.34}$$

From these equations it is evident that the cG(1)/dG(0) method can be regarded as a time stepping scheme, which for zero forcing, i.e. $f \equiv 0$, coincides with the backward Euler scheme applied to the spatially semi-discrete variational system (3.6). In the case of non-zero forcing this equivalence is only modulo the evaluation of the time-integral on the right-hand side by the box rule, i.e. up to a term of higher order $\mathcal{O}(k^2)$.

Remark 3.10. As mentioned in Remark 3.5, the Newmark scheme for $\gamma = 2\beta = 1$ is algebraically equivalent to a combination of the Crank-Nicolson and the backward Euler scheme applied separately to u^0 and u^1. Hence, thanks to the equivalence of the Crank-Nicolson scheme to the cG(1) scheme and that of the backward Euler scheme to the dG(0) scheme, for this particular choice of parameters the Newmark scheme turns out to be equivalent to a (Petrov-)Galerkin method in time.

3.4. Comparison of the different discretization methods. Table 3.1 gives an overview of the theoretical convergence behavior of the different discretization methods introduced above, measured in terms of the "end-time" error norm $\|(u^0 - U^0)(T)\| + \|(u^1 - U^1)(T)\|$. This is well confirmed by the results shown in Fig. 3.8, using numerical tests for a spatially two-dimensional model problem with exact solution $u(x, y, t) = \sin(\pi t)\cos(\frac{1}{2}\pi x)\cos(\frac{1}{2}\pi y)$ on the space-time region $(-1, 1)^2 \times (0, 1]$.

Table 3.1. **Order of convergence of the different methods with respect to the error norm $\|(u^0 - U^0)(T)\| + \|(u^1 - U^1)(T)\|$**

Method	cG(1)/dG(0) backward Euler	cG(1)/cG(1) Crank-Nicolson	Newmark $\gamma \neq 0.5$	Newmark $\gamma = 0.5$
Order	$\mathcal{O}(h^2 + k)$	$\mathcal{O}(h^2 + k^2)$	$\mathcal{O}(h^2 + k)$	$\mathcal{O}(h^2 + k^2)$

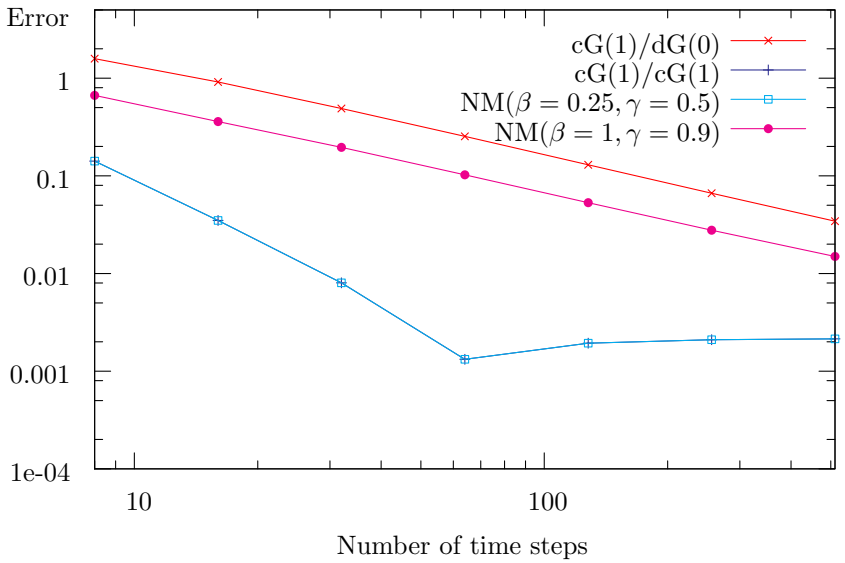


Fig. 3.8. Comparison of orders of convergence of the different time-stepping schemes on spatial meshes with 4,096 cells: linear convergence of cG(1)/dG(0) and Newmark ($\beta = 1, \gamma = 0.9$) schemes, quadratic convergence of cG(1)/cG(1) and Newmark ($\beta = 0.25, \gamma = 0.5$) scheme. The curves for the cG(1)/cG(1) and the Newmark ($\beta = 0.25, \gamma = 0.5$) scheme coincide, as expected. At error level 0.001 the spatial discretization error becomes dominant

4. A posteriori error estimation and mesh adaptation in space and time

4.1. The general framework. The main subject of this article is the a posteriori error control and step-size adaptation in the Galerkin discretization of the wave equation. Because of the superior approximation properties of the cG(1)/cG(1) method for the wave equation (second-order accuracy, energy conservation, unconditional stability, etc.) compared to the cG(1)/dG(0) method, this method will be discussed further in detail in the context of adaptivity. The cG(1)/dG(0) method will only briefly be treated in remarks. In view of the algebraic equivalence of the cG(1)/cG(1) method to the Crank-Nicolson and the trapezoidal Newmark schemes the results described below directly carry over to these particular finite difference schemes. Higher-order cG(s)/cG(r) methods can be analyzed by the same arguments with analogous results.

In practice, one is often not interested in the calculated solution itself but rather in derived quantities that can be computed from it. We will here assume that a single scalar is the goal of the numerical simulation. For example, this may be a certain norm of the solution, e. g. the global L^2- or energy norm at the final time T, the energy at a given time point, the energy flow through a curve Γ within the given domain Ω, or even simply a point value of the solution. Each of these quantities of interest can be thought to be obtained by applying a certain "output functional" $J(\cdot)$ to the solution. In the general case of nonlinear output functionals one has to carry out a linearization, which is described in detail in Becker & Rannacher [7] within the "Dual Weighted Residual" (DWR) method for a posteriori error estimation and mesh-size adaptation in the Galerkin finite element approximation of general nonlinear variational equations. In the following, for notational simplicity, we will only consider *linear* output functionals, which are given in the form $J(\varphi) := j^0(\varphi^0) + j^1(\varphi^1)$. The goal is the computation of $J(u)$, but only $J(U)$ is available; due to the assumed linearity, we can then estimate the error in this quantity by considering $J(u) - J(U) = J(u - U)$.

Note that this general formalism can also be used for certain special cases such as the L^2 error at the end-time,

$$J(u - U) := \|(u^0 - U^0)(T)\|.$$

Though the latter functional is *nonlinear* it can be fitted into the present framework by assuming that a sufficiently good approximation $\hat{e}^0$ to $(u^0 - U^0)(T)$ is known (for example obtained by extrapolation from preceding refinement levels) and then setting

$$J(\varphi) := \big(\hat{e}^0, \varphi^0(T)\big)/\|\hat{e}^0\|.$$

For the derivation of a posteriori error estimates for the fully discrete solution $U := U_{kh}$ we recall the following useful abstract theorem, which is a generalization of a similar result in Becker & Rannacher [7]; see Schmich & Vexler [33] and Meidner [29]. Though the examples presented below in Section 5 are *linear*, i.e., they involve *linear* operators and *linear* output functionals, the abstract theory in this section covers the most general *nonlinear* situation, in order to prepare for subsequent work.

Proposition 4.1. *Let X be a function space and $L : X \to \mathbb{R}$ a three times Gâteaux differentiable functional. We seek a stationary point $\hat{x}$ of L in a subspace of ("continuous") solutions $\hat{X} \subset X$, i. e., we seek $\hat{x} \in \hat{X}$ that satisfies*

$$L'(\hat{x})(\varphi) = 0 \quad \forall \varphi \in \hat{X}. \tag{4.1}$$

This equation is approximated by a Galerkin method in a finite dimensional ("discrete") subspace $X_d \subset X$, where we do not necessarily assume that $X_d \subset \hat{X}$. The approximation yields a stationary point $x_d \in X_d$ that satisfies

$$L'(x_d)(\varphi) = 0 \quad \forall \varphi \in X_d. \tag{4.2}$$

If the stationary point $\hat{x}$ of the continuous problem in addition satisfies

$$L'(\hat{x})(x_d) = 0, \tag{4.3}$$

then, we can represent the error in the form

$$L(\hat{x}) - L(x_d) = \tfrac{1}{2} L'(x_d)(\hat{x} - \varphi) + R, \tag{4.4}$$

where $\varphi \in X_d$ can be arbitrarily chosen and the cubic remainder R is given in terms of the error $e := \hat{x} - x_d$ as

$$R = \frac{1}{2} \int_0^1 L'''(x_d + se)(e,e,e)\, s(s-1)\, \mathrm{d}s.$$

Proof. By the fundamental theorem of calculus, we have

$$L(\hat{x}) - L(x_d) = \int_0^1 L'(x_d + se)(e)\, \mathrm{d}s.$$

Replacing the integral by the trapezoidal rule plus corresponding remainder term and using the above assumptions yields

$$L(\hat{x}) - L(x_d) = \frac{1}{2}\big(L'(x_d)(e) + L'(\hat{x})(e)\big) + \frac{1}{2} \int_0^1 L'''(x_d + se)(e,e,e)\, s(s-1)\, \mathrm{d}s =$$

$$\frac{1}{2} L'(x_d)(\hat{x} - \varphi) + \frac{1}{2} \int_0^1 L'''(x_d + se)(e,e,e)\, s(s-1)\, \mathrm{d}s,$$

for arbitrary $\varphi \in X_d$. □

Remark 4.1. The somewhat complicated setting of Proposition 4.1 is motivated by situations such as the cG(s)/dG(r) method, in which the variational form on the discrete level does not fit the well-posed formulation of the continuous problem, i.e., the continuous solution space $\hat{X}$ is a strict subspace of the space X underlying the approximation, $X_d \subset X$. In the standard situation, $X_d \subset X = \hat{X}$ and condition (4.3) is automatically satisfied. However, here $X_d \not\subset \hat{X}$ and condition (4.3) requires that $\hat{x}$ is not only a stationary point with respect to all (smooth) test functions in $\hat{X}$, but also with respect to the additional (discrete) "test function" x_d. That $\hat{x}$ satisfies this additional condition depends on its higher degree of smoothness and the particular structural properties of the variational formulation, i.e. the functional $J(\cdot)$ used in the Galerkin approximation.

In the next step, we apply the results of Proposition 4.1 to the general Galerkin or Petrov-Galerkin approximation of variational equations such as occurring in the context of the Galerkin discretization of the wave equation. Let E, E^* be two function spaces and

$\hat{E} \subset E$ a proper subspace. Furthermore, let $A(\cdot)(\cdot), F(\cdot)$ be generic semi-linear and linear forms that we will later identify with those used in the continuous and discontinuous Galerkin methods (3.23) and (3.31). We consider the task of computing a functional value $J(\hat{u})$ from the solution $\hat{u} \in \hat{E}$ of the variational problem

$$A(\hat{u})(\varphi) = F(\varphi) \quad \forall \varphi \in E^*. \tag{4.5}$$

Here, the functional $J : E \to \mathbb{R}$, the semi-linear form $A : E \times E^* \to \mathbb{R}$ and the linear right-hand side $F : E \to \mathbb{R}$ are assumed to be three times Gâteaux differentiable. This problem is approximated by a Galerkin or Petrov-Galerkin method in subspaces $E_d \subset E$, $E_d^* \subset E^*$ resulting in an approximate solution $u_d \in E_d$, satisfying

$$A(u_d)(\varphi) = F(\varphi) \quad \forall \varphi \in E_d^*, \tag{4.6}$$

and the corresponding approximate functional value $J(u_d)$. The solvability (not necessarily unique) of problems (4.5) and (4.6) is assumed. We want to estimate the error $J(\hat{u}) - J(u_d)$. To this end, we introduce the Lagrangian functional

$$\mathfrak{L}(u, z) := J(u) + F(z) - A(u)(z),$$

for arguments $\{u, z\} \in E \times E^*$. A stationary point $\{\hat{u}, \hat{z}\} \in \hat{E} \times E^*$ of $\mathfrak{L}(\cdot, \cdot)$ on $\hat{E} \times E^*$ is determined by the equation

$$\mathfrak{L}'(\hat{u}, \hat{z})(\psi, \varphi) = 0 \quad \forall \{\psi, \varphi\} \in \hat{E} \times E^*, \tag{4.7}$$

or equivalently by the system of equations

$$A'(\hat{u})(\psi, \hat{z}) = J'(\hat{u})(\psi) \quad \psi \in \hat{E}, \tag{4.8}$$

$$A(\hat{u})(\varphi) = F(\varphi) \quad \varphi \in E^*. \tag{4.9}$$

The second of these two equations is just the given "state" equation (4.5) and the first one is the so-called "dual" (or "adjoint") equation governed by the given goal functional $J(\cdot)$. Correspondingly, a "discrete" stationary point $\{u_d, z_d\} \in E_d \times E_d^*$ of $\mathfrak{L}(\cdot, \cdot)$ on $E_d \times E_d^*$ is determined by the equation

$$\mathfrak{L}'(u_d, z_d)(\psi, \varphi) = 0 \quad \forall \{\psi, \varphi\} \in E_d \times E_d^*, \tag{4.10}$$

or equivalently by the system of equations

$$A'(u_d)(\psi, z_d) = J'(u_d)(\psi) \quad \psi \in E_d, \tag{4.11}$$

$$A(u_d)(\varphi) = F(\varphi) \quad \varphi \in E_d^*. \tag{4.12}$$

Clearly, for stationary points $\{\hat{u}, \hat{z}\} \in \hat{E} \times E^*$ and $\{u_d, z_d\} \in E_d \times E_d^*$, we have that

$$J(\hat{u}) - J(u_d) = \mathfrak{L}(\hat{u}, \hat{z}) - \mathfrak{L}(u_d, z_d). \tag{4.13}$$

Corollary 4.1. *With the above notation let $\{\hat{u}, \hat{z}\} \in \hat{E} \times E^*$ and $\{u_d, z_d\} \in E_d \times E_d^*$ be stationary points of $\mathfrak{L}$ on $\hat{E} \times E^*$ and on $E_d \times E_d^*$, respectively. If the condition*

$$J'(\hat{u})(u_d) - A'(\hat{u})(u_d, \hat{z}) = 0 \tag{4.14}$$

is satisfied, then we have the error representation

$$J(\hat{u}) - J(u_d) = \tfrac{1}{2}\mathfrak{L}'(u_d, z_d)(\hat{u} - \psi, \hat{z} - \varphi) + R, \tag{4.15}$$

for arbitrary $\psi \in E_d$, $\varphi \in E_d^*$ and a remainder R which is cubic in the errors $e := \hat{u} - u_d$ and $\varepsilon := \hat{z} - z_d$.

Proof. We embed the current situation into the framework of Proposition 4.1. To this end, we set $X := E\times E^*$, $\hat{X} := \hat{E}\times E^*$, $X_d := E_d\times E_d^*$ and $L(\hat{x}) := \mathcal{L}(\hat{u},\hat{z})$, $\hat{x} = \{\hat{u},\hat{z}\} \in \hat{E}\times E^*$. Since $z_d \in E_d^* \subset E^*$ and observing (4.14), we find that

$$L'(\hat{x})(x_d) = \left\{ \begin{array}{c} J'(\hat{u})(u_d) - A'(\hat{u})(u_d,\hat{z}) \\ F(z_d) - A(\hat{u})(z_d) \end{array} \right\} = 0,$$

i.e., condition (4.3) is satisfied. Hence, Proposition 4.1 yields the error representation

$$L(\hat{x}) - L(x_d) = \tfrac{1}{2}L'(x_d)(\hat{x}-\varphi) + R$$

for arbitrary $\varphi \in X_d$ and a remainder R which is cubic in the error $\hat{x} - x_d$. In view of (4.15) this implies the asserted representation. □

For the particular form of the Lagrangian functional $L(x) = \mathcal{L}(u,z) = J(u) + F(z) - A(u)(z)$, $x = \{u,z\} \in E \times E^*$ the error representation (4.15) takes the concrete form

$$J(\hat{u}) - J(u_d) = \tfrac{1}{2}\rho(u_d)(\hat{z}-\varphi) + \tfrac{1}{2}\rho^*(z_d)(\hat{u}-\psi) + R, \tag{4.16}$$

for arbitrary $\psi \in E_d$, $\varphi \in E_d^*$, with the primal and dual residuals

$$\rho(u_d)(\cdot) := F(\cdot) - A(u_d)(\cdot), \quad \rho^*(z_d)(\cdot) := J(\cdot) - A'(u_d)(\cdot, z_d).$$

Next, we specialize the discussion to *linear* problems such as those that are mainly considered in this paper.

Corollary 4.2. *Suppose the notation and assumptions as in Corollary* 4.1*, particularly* (4.14)*, hold. In the case of a linear variational problem with bilinear form $A(\cdot,\cdot)$ and linear goal functional $J(\cdot)$, we have the a posteriori error representation*

$$J(\hat{u} - u_d) = F(\hat{z}-\varphi) - A(u_d, \hat{z}-\varphi), \tag{4.17}$$

with arbitrary $\varphi \in E_d^*$, where $\hat{z} \in E^*$ is the solution of the dual problem

$$A(\psi,\hat{z}) = J(\psi) \quad \forall \psi \in \hat{E}. \tag{4.18}$$

Proof. In the case of linear problems the remainder R in the error representation (4.16) vanishes. Further, in view of (4.14), for $\varphi \in E_d^*$, $\psi \in E_d$, we have that

$$\rho(u_d)(\hat{z}-\varphi) = F(\hat{z}-\varphi) - A(u_d,\hat{z}-\varphi) = F(\hat{z}-z_d) - A(u_d,\hat{z}-z_d) =$$

$$A(\hat{u},\hat{z}-z_d) - A(u_d,\hat{z}-z_d) = A(\hat{u}-u_d,\hat{z}-z_d) = A(\hat{u}-u_d,z) - A(\hat{u}-u_d,z_d) =$$

$$J(\hat{u}-u_d) - A(\hat{u}-u_d,z_d) = J(\hat{u}-\psi) - A(\hat{u}-\psi,z_d) = \rho^*(z_d)(\hat{u}-\psi).$$

Consequently (4.16) reduces to the form $J(\hat{u}-u_d) = \rho(u_d)(\hat{z}-\varphi)$, for arbitrary $\varphi \in E_d^*$, which does not contain the unknown solution $\hat{u}$. This implies (4.17). □

Remark 4.2. The practical evaluation of the general nonlinear error representation (4.16) or its linear special case (4.17) requires the generation of approximations to the generally unknown (exact) "primal" and "dual" solutions $u \in E$ and $z \in E^*$. Strategies for this crucial process will be described below in the context of the different time and space discretizations.

4.2. A posteriori error estimation for the Galerkin methods.

4.2.1. The cG(1)/cG(1) method. We begin with the cG(1)/cG(1) method written in its compact form (3.23) for discretizing the linear problem (2.6). In order to apply the results of the preceding section, we identify the spaces

$$E := \mathcal{V}\times\mathcal{H}, \quad E^* := \mathcal{W}\times\mathcal{W}, \quad \hat{E} := \hat{\mathcal{V}}^0\times\hat{\mathcal{V}}^1, \quad E_d := \mathcal{V}_{hk}\times\mathcal{V}_{hk}, \quad E_d^* := \mathcal{W}_{hk}\times\mathcal{W}_{hk}.$$

Further, we suppose that the linear goal functional is given in the form $J(\varphi) := j^0(\varphi^0) + j^1(\varphi^1)$ with certain linear functionals j^0, j^1, such as described above. Then, from Corollary 4.2, we have the following error representation for the fully discrete approximation $U = U_{kh} \in \mathcal{V}_{kh}\times\mathcal{V}_{kh}$:

$$J(u-U) = \rho(U)(z-\varphi) = F(z-\varphi) - A(U, z-\varphi), \tag{4.19}$$

with arbitrary $\varphi \in \mathcal{W}_{kh}\times\mathcal{W}_{kh}$, where $z = \{z^1, z^0\} \in \mathcal{H}\times\mathcal{V}$ is the solution of the associated dual problem

$$A(\psi, z) = J(\psi) \quad \forall\psi = \{\psi^0, \psi^1\} \in \hat{\mathcal{V}}^0\times\hat{\mathcal{V}}^1, \tag{4.20}$$

the existence of which is assumed. However, in the present linear situation the solvability of the dual problem will become obvious. Indeed in Section 4, we will see that for many goal functionals considered in the numerical examples below, the dual problem has the structure of a wave equation such as (2.1) but running backward in time. Thus, using the same setting as used for the forward problem, $\hat{z} \in \hat{\mathcal{V}}$ if $J(\cdot)$ is regular enough. In this case the crucial condition (4.14) is satisfied.

Corollary 4.3. *For the approximation of problem* (3.17) *by the cG(1)/cG(1) method* (3.23)*, we have the following a posteriori error representation:*

$$J(u-U) = \eta_\omega(U) := \sum_{K\in\mathbb{T}_h^0} \Big\{ \big(\rho(u_0^0-U_0^0), z_0^1-\varphi_0^1\big)_K + \big(\rho(u_0^1-U_0^1), z_0^0-\varphi_0^0\big)_K \Big\} +$$

$$\sum_{m=1}^{M}\sum_{K\in\mathbb{T}_h^m} \Big\{ (R_0(U), z^1-\varphi^1)_{K\times I_m} + (R_1(U), z^0-\varphi^0)_{K\times I_m} + (r_1(U), z^0-\varphi^0)_{\partial K\times I_m} \Big\}, \tag{4.21}$$

for arbitrary $\{\varphi^0, \varphi^1\} \in \mathcal{W}_{kh}\times\mathcal{W}_{kh}$*. Here,* $\hat{z} = \{z^1, z^0\} \in \mathcal{H}\times\mathcal{V}$ *is the solution of the dual problem* (4.20) *and the residuals are defined by*

$$R_0(U) := \rho U^1 - \rho\partial_t U^0, \quad R_1(U) := f - \rho\partial_t U^1 - \mathcal{A}U^0, \quad r_1(U) := -\tfrac{1}{2}[\partial_n^{\mathcal{A}} U^0],$$

with $[\partial_n^{\mathcal{A}} U^0]$ *denoting the jump of* $\partial_n^{\mathcal{A}} U^0$ *across the cell interfaces.*

Proof. First, we note that, in view of the discussion in Section 4 below, condition (4.14) is satisfied in the present situation and now reads $A(U, \hat{z}) = J(U)$. Therefore, Corollary 4.2 is applicable. Recalling the definition of the bilinear form $A(\cdot,\cdot)$ and the functional $F(\cdot)$, the abstract a posteriori error representation (4.19) takes the following concrete form:

$$J(u-U) = m(u_0^0, z_0^1-\varphi_0^1) + m(u_0^1, z_0^0-\varphi_0^0) + (\!(f, z^0-\varphi^0)\!) - m(\!(\partial_t U^0, z^1-\varphi^1)\!) + m(\!(U^1, z^1-\varphi^1)\!) -$$

$$m(U_0^0, z_0^1-\varphi_0^1) - m(\!(\partial_t U^1, z^0-\varphi^0)\!) - a(\!(U^0, z^0-\varphi^0)\!) - m(U_0^1, z_0^0-\varphi_0^0),$$

with arbitrary pairs $\{\varphi^0,\varphi^1\}\in\mathcal{W}_{kh}\times\mathcal{W}_{kh}$. Reordering terms results in

$$J(u-U)=m(u_0^0-U_0^0,z_0^1-\varphi_0^1)+m(u_0^1-U_0^1,z_0^0-\varphi_0^0)+(\!(f,z^0-\varphi^0)\!)+$$

$$m(\!(U^1-\partial_tU^0,z^1-\varphi^1)\!)-m(\!(\partial_tU^1,z^0-\varphi^0)\!)-a(\!(U^0,z^0-\varphi^0)\!).$$

Cellwise integration by parts yields

$$a(\!(U^0,z^0-\varphi^0)\!)=\int_0^T a(U^0,z^0-\varphi^0)\,\mathrm{d}t=\int_0^T\sum_{K\in\mathbb{T}_h}\Big\{(\mathcal{A}U^0,z^0-\varphi^0)_K-(\partial_n^{\mathcal{A}}U^0,z^0-\varphi^0)_{\partial K}\Big\}\,\mathrm{d}t=$$

$$\int_0^T\sum_{K\in\mathbb{T}_h}\Big\{(\mathcal{A}U^0,z^0-\varphi^0)_K-\tfrac{1}{2}([\partial_n^{\mathcal{A}}U^0],z^0-\varphi^0)_{\partial K}\Big\}\,\mathrm{d}t.$$

Combining these identities and splitting up the contributions from the different time intervals I_m and spatial cells $K\in\mathbb{T}_h^m$, we obtain

$$J(u-U)=\sum_{K\in\mathbb{T}_h^0}(\rho(u_0^0-U_0^0),z_0^1-\varphi_0^1)_K+\sum_{K\in\mathbb{T}_h^0}(\rho(u_0^1-U_0^1),z_0^0-\varphi_0^0)_K+\sum_{m=1}^M\sum_{K\in\mathbb{T}_h^m}(\rho(U^1-\partial_tU^0),z^1-\varphi^1)_{K\times I_m}+$$

$$\sum_{m=1}^M\sum_{K\in\mathbb{T}_h^m}\Big\{(f-\rho\partial_tU^1-\mathcal{A}U^0,z^0-\varphi^0)_{K\times I_m}-\tfrac{1}{2}([\partial_n^{\mathcal{A}}U^0],z^0-\varphi^0)_{\partial K\times I_m}\Big\}.$$

From this, we conclude the asserted error representation (4.21). □

In the error representation (4.21), the error contributions by the spatial and the temporal discretization do not appear in separated form. Hence it cannot be taken as the basis for independent adaptation of the time step and spatial mesh. However, in order to separate these error components, we can utilize the "free" functions φ^i, $i=0,1$ by choosing them as the cellwise defined natural nodal interpolation $I_{hk}z^i\in\mathcal{W}_{hk}$, i.e., piecewise bi- or trilinear in space and constant in time. First, we introduce the local temporal L^2 projection $\overline{v}\in P_0(I_m)$ of a general function $v\in L^2(I_m)$ into $P_0(I_m)$ defined by

$$\int_{I_m}\overline{v}\,\mathrm{d}t=\int_{I_m}v\,\mathrm{d}t. \tag{4.22}$$

With this notation, we define the cellwise interpolation $I_{hk}z^i\in\mathcal{W}_{hk}$ of z^i by prescribing

$$I_{hk}z^i(a)=\bar{z}^i(a),\quad\text{for all nodal points } a \text{ of } \mathbb{T}_h^m, \tag{4.23}$$

where $\mathbb{T}_h^m$ is the spatial mesh used in the time slab $\bar{\Omega}\times I_m$. Then, with the corresponding spatial nodal interpolation I_h^m on $\mathbb{T}_h^m$, there holds $I_{hk}z^i|_{\bar{\Omega}\times I_m}=I_h^m\bar{z}^i$. Notice that by construction these interpolations $I_{hk}z^i$ are continuous in space but usually discontinuous in time. Using this construction in the error representation (4.21), we obtain the following result.

Corollary 4.4. *For the approximation of problem* (3.17) *by the cG(1)/cG(1) method* (3.23) *there holds the a posteriori error estimate*

$$|\eta_\omega(U)| \leqslant \sum_{K\in\mathbb{T}_h^0} \{\rho_{K,h}^{0,0}\omega_{K,h}^{1,0} + \rho_{K,h}^{1,0}\omega_{K,h}^{0,0}\} + \sum_{m=1}^{M}\sum_{K\in\mathbb{T}_h^m} \Big\{\rho_{K,h}^{0,m}\,\omega_{K,h}^{1,m} + \rho_{K,h}^{1,m}\,\omega_{K,h}^{0,m} + \rho_{\partial K,h}^{1,m}\,\omega_{\partial K,h}^{0,m}\Big\}+$$

$$\sum_{m=1}^{M}\sum_{K\in\mathbb{T}_h^m} \Big\{\rho_{K,k}^{0,m}\,\omega_{K,k}^{1,m} + \rho_{K,k}^{1,m}\,\omega_{K,k}^{0,m} + \rho_{\partial K,k}^{1,m}\,\omega_{\partial K,k}^{0,m}\Big\}. \tag{4.24}$$

With the notation of Corollary 4.3 the residual terms and weights are defined by:

1) spatial terms

$$\rho_{K,h}^{0,0} := \|\rho(u_0^0-U_0^0)\|_K, \quad \omega_{K,h}^{1,0} := \|z_0^1-(I_h^1 z^1)_0\|_K,$$

$$\rho_{K,h}^{1,0} := \|\rho(u_0^1-U_0^1)\|_K, \quad \omega_{K,h}^{0,0} := \|z_0^0-(I_h^1 z^0)_0\|_K,$$

$$\rho_{K,h}^{0,m} := \|R_0(U)\|_{K\times I_m}, \quad \omega_{K,h}^{1,m} := \|\bar z^1-I_h^m\bar z^1\|_{K\times I_m},$$

$$\rho_{K,h}^{1,m} := \|R_1(U)\|_{K\times I_m}, \quad \omega_{K,h}^{0,m} := \|\bar z^0-I_h^m\bar z^0\|_{K\times I_m},$$

$$\rho_{\partial K,h}^{1,m} := h_K^{-1/2}\|r_1(U)\|_{\partial K\times I_m}, \quad \omega_{\partial K,h}^{0,m} := h_K^{1/2}\|\bar z^0-I_h^m\bar z^0\|_{\partial K\times I_m},$$

2) temporal terms

$$\rho_{K,k}^{0,m} := \|R_0(U)-\overline{R_0(U)}\|_{K\times I_m}, \quad \omega_{K,k}^{1,m} := \|z^1-\bar z^1\|_{K\times I_m},$$

$$\rho_{K,k}^{1,m} := \|R_1(U)-\overline{R_1(U)}\|_{K\times I_m}, \quad \omega_{K,k}^{0,m} := \|z^0-\bar z^0\|_{K\times I_m},$$

$$\rho_{\partial K,k}^{1,m} := h_K^{-1/2}\|r_1(U)-\overline{r_1(U)}\|_{\partial K\times I_m}, \quad \omega_{\partial K,k}^{0,m} := h_K^{1/2}\|z^0-\bar z^0\|_{\partial K\times I_m}.$$

Proof. Taking $\varphi^i := I_{hk}z^i$, $i=0,1$, in the error representation (4.21), yields

$$\eta_\omega(U) := \sum_{K\in\mathbb{T}_h^0}\Big\{\big(\rho(u_0^0-U_0^0), z_0^1-(I_{hk}z^1)_0\big)_K + \big(\rho(u_0^1-U_0^1), z_0^0-(I_{hk}z^0)_0\big)_K\Big\}+$$

$$\sum_{m=1}^{M}\sum_{K\in\mathbb{T}_h^m}\Big\{(R_0(U), z^1-I_{hk}z^1)_{K\times I_m} + (R_1(U), z^0-I_{hk}z^0)_{K\times I_m} + (r_1(U), z^0-I_{hk}z^0)_{\partial K\times I_m}\Big\}.$$

Now, we additionally introduce the local time-averages $\bar z^i$ to obtain

$$\eta_\omega(U) := \sum_{K\in\mathbb{T}_h^0}\Big\{\big(\rho(u_0^0-U_0^0), z_0^1-(I_{hk}z^1)_0\big)_K + \big(\rho(u_0^1-U_0^1), z_0^0-(I_{hk}z^0)_0\big)_K\Big\}+$$

$$\sum_{m=1}^{M}\sum_{K\in\mathbb{T}_h^m}\Big\{(R_0(U), z^1-\bar z^1)_{K\times I_m} + (R_0(U), \bar z^1-I_{hk}z^1)_{K\times I_m}+$$

$$(R_1(U), z^0-\bar z^0)_{K\times I_m}+(R_1(U), \bar z^0-I_{hk}z^0)_{K\times I_m}+(r_1(U), z^0-\bar z^0)_{\partial K\times I_m}+(r_1(U), \bar z^0-I_{hk}z^0)_{\partial K\times I_m}\Big\}.$$

Then, using the projection property of $z^i - \bar{z}^i$ and the relation $I_{hk} z^i|_{\bar{\Omega}\times I_m} = I_h^m \bar{z}^i$ stated above, we arrive at

$$\eta_\omega(U) := \sum_{K\in\mathbb{T}_h^0} \Big\{ \big(\rho(u_0^0 - U_0^0), z_0^1 - (I_h^1 \bar{z}^1)_0\big)_K + \big(\rho(u_0^1 - U_0^1), z_0^0 - (I_h^1 \bar{z}^0)_0\big)_K \Big\} +$$

$$\sum_{m=1}^{M} \sum_{K\in\mathbb{T}_h^m} \Big\{ (R_0(U) - \overline{R_0(U)}, z^1 - \bar{z}^1)_{K\times I_m} + (R_0(U), \bar{z}^1 - I_h^m \bar{z}^1)_{K\times I_m} + (R_1(U) - \overline{R_1(U)}, z^0 - \bar{z}^0)_{K\times I_m} +$$

$$(R_1(U), \bar{z}^0 - I_h^m \bar{z}^0)_{K\times I_m} + (r_1(U) - \overline{r_1(U)}, z^0 - \bar{z}^0)_{\partial K\times I_m} + (r_1(U), \bar{z}^0 - I_h^m \bar{z}^0)_{\partial K\times I_m} \Big\}.$$

From this identity, we obtain the asserted error estimate (4.24) by applying cellwise the Hölder inequality to all the scalar products. □

Remark 4.3. We note that in the error estimate given in Corollary 4.4, the effect of the space discretization is separated from that of the time discretization, i.e., on each space-time cell $K \times I_m$ the respective indicators can be used to control the spatial mesh width h_K and the time step k_m. The different cell residual terms contain information about different aspects of the quality of the discretization:

- $\rho_{K,h}^{0,0}$ and $\rho_{K,h}^{1,0}$ measure the spatial accuracy in approximating the initial data;
- $\rho_{K,h}^{0,m}$ and $\rho_{K,h}^{1,m}$ measure the spatial and $\rho_{K,k}^{0,m}$ and $\rho_{K,k}^{1,m}$ the temporal accuracy in representing the equations $\partial_t u^0 = u^1$ and $\partial_t u^1 + \mathcal{A} u^0 = f$, respectively;
- $\rho_{\partial K,h}^{0,m}$ and $\rho_{\partial K,k}^{0,m}$ measure the spatial "smoothness" of the discrete solution U_{hk} depending on the spatial and time discretization, respectively.

Remark 4.4. In controlling the discretization by the cG(1)/cG(1) method, we follow two different goals. First, we need to accurately estimate the actual errors (in terms of the goal functional) on the generated meshes for getting a stopping criterion of the adaptation process. Second, we need effective (non-negative) "error indicators" on each of the space-time mesh cells $K\times\bar{I}_m$ for steering the adaptation process. The first goal is achieved by the error estimator $\eta_\omega(U)$ defined in Corollary 4.3, which is to be evaluated directly without further estimation using the strategies described in Section 4, below. In fact, the subtraction of the arbitrary function $\varphi \in \mathcal{W}_{hk} \times \mathcal{W}_{hk}$ may be suppressed since it has no effect on the value of the estimator $\eta_\omega(U)$, due to Galerkin orthogonality (3.21). The second goal is achieved by the error estimate (4.24) of Corollary 4.4. We emphasize that the use of this error "estimate" for deriving a stopping criterion may result in strong overestimation of the true error, since possible global cancellation effects of the residuals are not captured. Therefore, in the numerical examples discussed in Section 5 below, the mesh refinement is controlled by an estimate such as (4.24) while the effectivity index I_{eff}, which measures the accuracy of the error estimation, is determined using the error representation (4.21).

Sometimes, the target functional one is interested in is sufficiently global such that its domain of influence (which is given by the support of the dual solution) is more or less the whole domain. Then one does not gain much from the effort of numerically approximating the dual solution and one can get cheaper error indicators than the one above by using analytical a priori estimates for it. This kind of analysis is well known in the derivation of error estimates in global norms for the Laplace equation. We refer to Johnson [24] for

corresponding analysis of the discretization of the wave equation by the *discontinuous*-in-time Galerkin finite element method.

In our numerical examples below, we will explore this issue by comparing the grids obtained by the weighted error estimator $\eta_\omega(U)$ derived above with those resulting from the use of one of these "traditional" error indicators. Without further justification, we select as a "baseline" a rather simple estimator proposed by Kelly & al. [28] in an entirely different context, namely the Laplace equation:

$$\eta_E^{\rm red}(U) := \Big(\sum_{m=1}^{M} \sum_{K\in\mathbb{T}_h^m} h_K (\rho_{\partial K,h}^{0,m})^2 \Big)^{1/2}. \tag{4.25}$$

It only measures the spatial smoothness of the computed solution U and neglects contributions by the time discretization. A more complete estimator that still avoids the evaluation of a dual solution is given by

$$\eta_E(U)^2 := \sum_{K\in\mathbb{T}_h^0} h_K^2 \Big\{ (\rho_{K,h}^{0,0})^2 + (\rho_{K,h}^{1,0})^2 \Big\} + \sum_{m=1}^{M} \sum_{K\in\mathbb{T}_h^m} h_K^2 \Big\{ (\rho_{K,h}^{0,m})^2 + (\rho_{K,h}^{1,m})^2 + (\rho_{\partial K,h}^{0,m})^2 \Big\} +$$

$$\sum_{m=1}^{M} \sum_{K\in\mathbb{T}_h^m} (k_m^2 + h_K^2) \Big\{ (\rho_{K,k}^{0,m})^2 + (\rho_{K,k}^{1,m})^2 + (\rho_{\partial K,k}^{0,m})^2 \Big\}. \tag{4.26}$$

It involves all residuals also occurring in the estimate (4.24). Such estimators are typically referred to as "energy error indicators" since they were originally derived as indicators for the energy norm of the error. Clearly such heuristic error indicators may be useful for mesh adaptation but will hardly yield good quantitative estimates of the error, particularly in cases with heterogeneous data.

4.2.2. The cG(1)/dG(0) method. The derivation of error representation formulas such as those shown in the previous section is trivially extended to the case of the cG(1)/dG(0) method. Let us here state without proof such an extension, analogous to that given for the cG(1)/cG(1) method in Corollary 4.3, and a resulting a posteriori error estimate analogous to that in Corollary 4.4. Starting from the abstract equations (4.19) and (4.20) and observing that in this case $\mathcal{W}_{kh} = \mathcal{V}_{kh}$, we obtain the following result.

Corollary 4.5. *For the approximation of problem* (3.17) *by the cG(1)/dG(0) method* (3.31)*, we have the following a posteriori error representation:*

$$J(u-U) = \eta_\omega(U) := \sum_{K\in\mathbb{T}_h^0} \Big\{ \big(\rho(u_0^0 - U_0^{0,+}), z_0^1 - \varphi_0^{1,+}\big)_K + \big(\rho(u_0^1 - U_0^{1,+}), z_0^0 - \varphi_0^{0,+}\big)_K \Big\} +$$

$$\sum_{m=1}^{M} \sum_{K\in\mathbb{T}_h^m} \Big\{ (R_0(U), z^1 - \varphi^1)_K + (R_1(U), z^0 - \varphi^0)_K + (r_1(U), z^0 - \varphi^0)_{\partial K} \Big\} -$$

$$\sum_{m=1}^{M-1} \sum_{K\in\mathbb{T}_h^m} \Big\{ (\rho[U_m^1], z_m^{0,+} - \varphi_m^{0,+})_K + (\rho[U_m^0], z_m^{1,+} - \varphi_m^{1,+})_K \Big\}, \tag{4.27}$$

with arbitrary $\{\varphi^0, \varphi^1\} \in \mathcal{W}_{kh} \times \mathcal{W}_{kh}$. *Here,* $\{z^1, z^0\} \in \mathcal{H} \times \mathcal{V}$ *is the solution of the dual problem* (4.20). $R_0(U), R_1(U)$ *and* $r_1(U)$ *are defined as before and* $[U_m^i]$ *is the jump of* $U^i(t)$ *at time instant* t_m *as indicated in Fig.* 3.7.

The proof follows the same line of argument as that of Corollary 4.3 and is therefore omitted. From this error representation we can then derive an a posteriori error estimate analogous to the one derived before in Corollary 4.4 for the cG(1)/cG(1) method. We omit the details since the cG(1)/dG(0) method is not used in the test examples in Section 5, below.

4.2.3. The dual problem. All error representation formulas derived above contain the solution $\{z^0, z^1\}$ of a dual problem (4.20). This problem can be given an intuitive interpretation which we will outline here for the cG(1)/cG(1) method. Recalling the definition of the bilinear form $A(\cdot,\cdot)$ and the functional $J(\cdot)$ the dual problem reads more explicitly as

$$m(\!(\partial_t\psi^0, z^1)\!)+m(\psi_0^0, z_0^1)+a(\!(\psi^0, z^0)\!)+m(\psi_0^1, z_0^0)-m(\!(\psi^1, z^1)\!)+m(\!(\partial_t\psi^1, z^0)\!) = j^0(\psi^0)+j^1(\psi^1), \tag{4.28}$$

or equivalently as the following system of equations:

$$\begin{aligned} m(\!(\partial_t\psi^0, z^1)\!) + m(\psi_0^0, z_0^1) + a(\!(\psi^0, z^0)\!) &= j^0(\psi^0),\\ m(\!(\partial_t\psi^1, z^0)\!) + m(\psi_0^1, z_0^0) - m(\!(\psi^1, z^1)\!) &= j^1(\psi^1), \end{aligned} \tag{4.29}$$

for all $\{\psi^0, \psi^1\} \in \hat{\mathcal{V}}^0\times\hat{\mathcal{V}}^1$. Assuming sufficient regularity of $\{z^0, z^1\}$ and integrating by parts in time, we obtain

$$\begin{aligned} -m(\!(\psi^0, \partial_t z^1)\!) + m(\psi^0, z^1)\Big|_{t=0}^{t=T} + m(\psi_0^0, z_0^1) + a(\!(\psi^0, z^0)\!) &= j^0(\psi^0),\\ -m(\!(\psi^1, \partial_t z^0)\!) + m(\psi^1, z^0)\Big|_{t=0}^{t=T} + m(\psi_0^1, z_0^0) - m(\!(\psi^1, z^1)\!) &= j^1(\psi^1), \end{aligned}$$

and, consequently, the dual system

$$\begin{aligned} -m(\!(\psi^0, \partial_t z^1)\!) + m(\psi^0(T), z^1(T)) + a(\!(\psi^0, z^0)\!) &= j^0(\psi^0),\\ -m(\!(\psi^1, \partial_t z^0)\!) + m(\psi^1(T), z^0(T)) - m(\!(\psi^1, z^1)\!) &= j^1(\psi^1), \end{aligned} \tag{4.30}$$

for all $\{\psi^0, \psi^1\} \in \hat{\mathcal{V}}^0 \times \hat{\mathcal{V}}^1$. This variational problem can be expressed as a wave equation running backward in time with initial conditions (or "terminal conditions", depending on the viewpoint, as they are posed at $t = T$) and right-hand side depending on the particular choice of the (sufficiently regular) functionals j^0 and j^1. In turn, according to the discussion in Section 2, this wave equation possesses a unique solution $z \in \hat{\mathcal{V}}$ or $\{z^0, z^1\} \in \hat{\mathcal{V}}^0\times\hat{\mathcal{V}}^1$ for its equivalent "mixed" formulation. Following the above argument backwards this solution also satisfies (4.29) and (4.28) and therefore is "the" dual solution of the problem corresponding to the chosen goal functional $J(\cdot)$.

Intuitively, the dual solution transports back in time information about how important a particular space-time point is for the evaluation of the goal functional. This will become particularly clear in Example 5.3. Let us consider two particular examples to give these ideas a more concrete form:

Example 4.1. The first example concerns the estimation of the end-time L^2-norm error. As mentioned above, we can approximate $J(u-U) = \|(u-u_k)(T)\|$ if we choose

$$j^0(\psi^0) := \frac{(\psi^0(T), \hat{e}^0)}{\|\hat{e}^0\|}, \quad j^1(\psi^1) := 0,$$

in (4.30), where $\hat{e}^0 \approx (u^0 - U^0)(T)$. We then obtain the system

$$-\rho\partial_t z^1 + \mathcal{A}z^0 = 0, \quad \text{in } \Omega \times [0,T),$$

$$-\rho\partial_t z^0 - \rho z^1 = 0, \quad \text{in } \Omega \times [0,T),$$

with the "initial" conditions

$$z^0(T) = 0, \quad z^1(T) = \frac{\hat{e}^0}{\|\hat{e}^0\|}, \tag{4.31}$$

and the usual boundary conditions $z^0|_{\partial\Omega_D} = 0$, $\partial_n^{\mathcal{A}} z^0|_{\partial\Omega_N} = 0$. Clearly, this system is equivalent to the (backward in time) wave propagation problem

$$\rho\partial_t^2 z + \mathcal{A}z = 0, \quad \text{in } \Omega \times [0,T),$$

$$z|_{t=T} = 0, \quad \partial_t z|_{t=T} = \frac{\hat{e}^0}{\|\hat{e}^0\|}, \quad \text{in } \Omega,$$

$$z|_{\partial\Omega_D} = 0, \quad \partial_n^{\mathcal{A}} z|_{\partial\Omega_N} = 0, \quad \text{on } [0,T). \tag{4.32}$$

In this case the data satisfy $f \equiv 0$ and $z^0(T) = 0$, $z^1(T) \in H$, so that $z \in \hat{\mathcal{V}}$.

Example 4.2. The second example concerns a weighted space-time average over a subdomain $\Omega_0 \subset \Omega$,

$$J(u) = \int_0^T \int_{\Omega_0} u(x,t)\omega(x,t)\,\mathrm{d}x\,\mathrm{d}t,$$

where ω is a smooth, non-negative weighting function. With the choice

$$j^0(\psi^0) := \int_0^T \int_{\Omega_0} \psi^0(x,t)\omega(x,t)\,\mathrm{d}x\,\mathrm{d}t, \quad j^1(\psi^1) := 0,$$

in (4.30) and denoting by χ_{Ω_0} the characteristic function of Ω_0, we obtain the system

$$-\rho\partial_t z^1 + \mathcal{A}z^0 = \omega\chi_{\Omega_0}, \quad \text{in } \Omega \times [0,T),$$

$$-\rho\partial_t z^0 - \rho z^1 = 0, \quad \text{in } \Omega \times [0,T),$$

with homogeneous initial and boundary conditions $z^0(T) = z^1(T) = 0$ and $z^0|_{\partial\Omega_D} = 0$, $\partial_n^{\mathcal{A}} z^0|_{\partial\Omega_N} = 0$, respectively. Again, this system is equivalent to a (backward in time) wave propagation problem analogous to (4.32) with data satisfying $f \in L^2(I;H)$ and $z^0(T) = z^1(T) = 0$, so that again $z \in \hat{\mathcal{V}}$. In the related case of a functional of the form

$$J(\psi) = \int_0^T \int_{\Gamma} \psi^0(s,t)\omega(s,t)\,\mathrm{d}s\,\mathrm{d}t,$$

where Γ is a part of the spatial boundary $\partial\Omega$, we only have $f \in L^2(I;V^*)$, but this still suffices to guarantee that $z \in \hat{\mathcal{V}}$.

If the goal functional J is less regular for initial values $z^0(T) \notin V$, $z^1(T) \notin H$, or for a forcing term $f \notin L^2(I;V^*)$, involving, for example, spatial point values such as

$$J(u) = \int_{t_1}^{t_2} u(x_0,t)\, t \,\mathrm{d}t, \quad x_0 \in \bar{\Omega},$$

(see Example 5.4) the dual solution z lacks smoothness so that the theory developed above may not be directly applicable. In this case, one may appropriately regularize the functional by introducing

$$J_\varepsilon(u) := \int_{t_1}^{t_2} \left(\frac{1}{|B_\varepsilon(x_0)|} \int_{B_\varepsilon(x_0)} u(x_0,t)\,\mathrm{d}x \right) t \,\mathrm{d}t,$$

where $B_\varepsilon(x_0)$ is a spatial ball centered at x_0 with radius $\varepsilon \approx \mathrm{TOL}$. However, this regularization is usually only necessary formally on the continuous level for making the abstract theory applicable. On the discrete level, i.e. in the practical realization of the adaptive scheme, it may frequently be possible to work with the original "singular" form of J; see Bangerth & Rannacher [6] for a more detailed discussion of this issue.

4.3. Practical aspects. The a posteriori error estimates for the cG(s)/cG(r) and the cG(s)/dG(r) methods derived above are not immediately practically applicable: First, they still contain the continuous solution of a dual problem that is, in general, equally difficult to obtain as the continuous solution of the forward problem; its numerical approximation is therefore necessary for every practical method. Second, we need to define how we want to use the resulting cell-wise error estimators to refine the spatial and temporal meshes. We will discuss these issues in the following subsections.

4.3.1. Evaluation of the a posteriori error estimates. For the use of the above a posteriori error representations and estimates, we have to evaluate the weights $\omega_{K,h}^m$ and $\omega_{K,k}^m$. This requires the construction of approximations to the dual solution $z = \{z^0, z^1\}$ or more precisely to the local "interpolation" errors $(z^i - I_{hk} z^i)|_{K\times I_m}$, $(z^i - \bar{z}^i)|_{K\times I_m}$. These approximations are used in the error representations (4.21) for the cG(1)/cG(1) method and (4.27) for the cG(1)/dG(0) method resulting in approximate error representations denoted by

$$J(u-U) \approx \tilde{\eta}_\omega(U). \tag{4.33}$$

A variety of techniques for the evaluation of these weights have been discussed in the literature. Among those, the solution of the dual problem globally by a higher-order method, say the cG(2)/cG(2) method, is not feasible in practice as it is clearly too expensive, particularly in three dimensions. On the other hand, is has often turned out to be sufficient to apply local high-order post-processing based on the "discrete" dual solution computed by the same method as used for the primal problem, for instance by the cG(1)/cG(1) method (Crank-Nicolson scheme). For a detailed discussion of this approach and several of its variants, we refer to Bangerth & Rannacher [6] and the literature cited therein. The basic idea is to compute a discrete dual solution Z_{hk} on the current (or a slightly finer) space-time mesh and construct from that the desired approximations by one of the following strategies.

Approximation by higher-order local interpolation. Let Z_{hk} be an approximation to the dual solution z computed by the cG(1)/cG(1) method (Crank-Nicolson scheme in time) or the cG(1)/dG(0) methods (backward Euler scheme in time) on the current space-time

mesh (possibly with smaller time steps $\frac{1}{2}k_m$). The temporal mesh $\{I_m,\, m=1,\dots,M\}$ and the spatial meshes $\mathbb{T}_h^m$ are grouped into 2-patches and 2×2-patches (in two dimensions), respectively. From the nodal values of Z_{hk}, we construct locally a higher-order (in the present case $(n+1)$-quadratic) interpolation $\tilde I_{hk}^{(2)} Z_{hk}$ on each time slab $\bar\Omega\times I_m$ (see Figs. 4.1 and 4.2), which is then used in the approximation

$$(z^i - I_{hk} z^i)|_{K\times I_m} \approx (\tilde I_{hk}^{(2)} Z_{hk}^i - I_{hk} Z_{hk}^i)|_{K\times I_m}. \tag{4.34}$$

Since in both methods, cG(1)/cG(1) and cG(1)/dG(0), the test functions are piecewise *constant* in time, it may seem sufficient to use only *linear* interpolation in time of the nodal values of Z_{hk} in constructing the approximation $\tilde I_{hk}^{(2)} Z_{hk}$. However, practical experience shows that this simple approximation may lead to strong underestimation of the true error on coarser meshes.

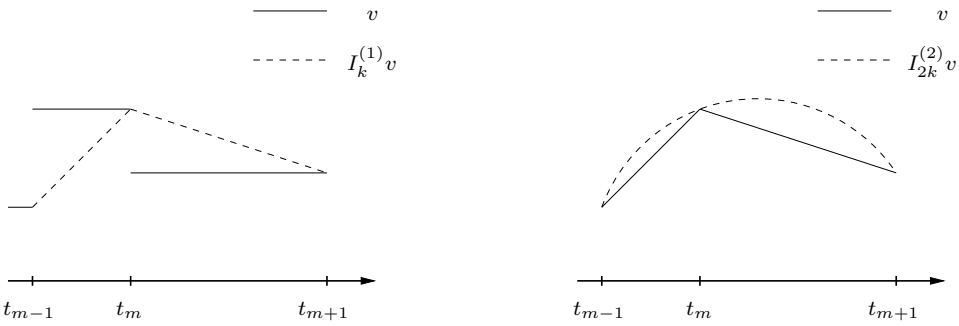


Fig. 4.1. Local post-processing in time by higher-order patchwise interpolation: "linear" (left) or "quadratic" (right) interpolation of computed "constant" or "linear" nodal values

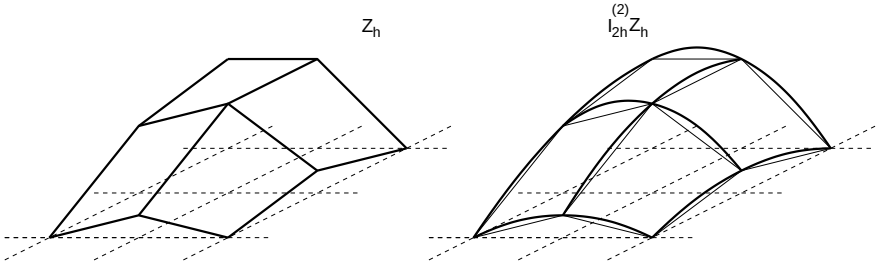


Fig. 4.2. Local post-processing in space by higher-order patchwise interpolation: "biquadratic" interpolation of computed "bilinear" nodal values

The observed success of this relatively cheap post-processing approach seems to be largely based on super-approximation effects, which can be expected on essentially uniform or at least very structured meshes. However, its rigorous theoretical justification is still missing and it may appear questionable whether the distance of the functions Z_{hk}^i and $\bar Z_{hk}^i$ is large enough (compared to the size of the spatial error) to justify the proposed approximation.

Use of interpolation estimates and approximation of higher-order derivatives. Alternatively, one may use interpolation estimates of the form

$$\|z^i - I_{hk} z^i\|_{K\times I_m} \leqslant c_I\{h_K^2 \|\nabla^2 z^i\|_{K\times I_m} + k_m \|\partial_t z^i\|_{K\times I_m} + k_m^2 \|\partial_t^2 z^i\|_{K\times I_m}\},$$

$$\|z^i - \bar z^i\|_{K\times I_m} \leqslant c_I k_m \|\partial_t z^i\|_{K\times I_m},$$

in order to reduce the estimation of the weights $\|z^i - I_{hk} z^i\|_{K\times I_m}$ and $\|z^i - \bar z^i\|_{K\times I_m}$ to terms only involving derivatives of z^i. These derivatives are then approximated by corresponding

difference quotients in space and time of the computed discrete dual solution Z_{hk} on the current space-time meshes.

This method seems computationally cheaper than that using higher-order patchwise interpolation described above. However, it involves generally only approximately known "interpolation constants" c_I and cannot capture local oscillations of the dual solution. Therefore, it is not advisable if real *error estimation* is required, but it is sufficiently accurate for steering the local mesh adaptation.

Remark 4.5. The two local approximation strategies described above seem questionable for equations with highly oscillatory solutions – such as the wave equation – if the region of evaluation of the target functional $J(\cdot)$ extends over more than one wave length in space or over more than one period in time. For this reason, in the corresponding examples in Section 5 below, globally higher-order methods or finer meshes have been used for generating approximations of the weighting terms. However, this is obviously not the way to go if computational resources are limited. Better alternatives will have to be developed for this purpose in the future.

4.3.2. Strategies for time-step and mesh-size adaptation. Adaptive finite element methods based on error estimates have a two-fold goal: They want to guarantee that a computation satisfies a prescribed error tolerance TOL, and they want to achieve the first goal with the least amount of work by choosing meshes adaptively in the most efficient way. In practice, the second goal is admittedly often more important: error estimates may not be accurate enough for hard guarantees, or computational resources may not be adequate to actually achieve practically desirable error tolerances and we will consequently have to be content with the best that is possible under the circumstances. However, in any case it is important that we have effective ways to let local error estimates guide us in deciding which cells or time steps to refine and/or which to coarsen.

To do so, let us start from the approximate a posteriori error estimate (4.33):

$$|J(e)| \approx |\tilde{\eta}_\omega(U)| \leqslant \sum_{m=1}^{M} \sum_{K\in\mathbb{T}_h^m} \left\{ \rho_{K,h}^m \, \tilde{\omega}_{K,h}^m + \rho_{K,k}^m \, \tilde{\omega}_{K,k}^m \right\}. \tag{4.35}$$

From this, we obtain error "indicators" for each cell K and time step m,

$$\eta_{K,h}^m := \rho_{K,h}^m \, \tilde{\omega}_{K,h}^m, \quad \eta_{K,k}^m := \rho_{K,k}^m \, \tilde{\omega}_{K,k}^m, \quad \eta_k^m := \sum_{K\in\mathbb{T}_h^m} \eta_{K,k}^m.$$

Here, $\eta_{K,h}^m$ and $\eta_{K,k}^m$ are meant to indicate the local contributions to the global error due to spatial and temporal discretization, respectively. η_k^m is the spatial discretization error of all cells in time interval I_m.

Let M be the number of time steps, N_m the numbers of cells of mesh $\mathbb{T}_h^m$, and $N = \sum_{m=1}^{M} N_m$ be the total number of space-time cells. We base our refinement strategy on the principle that the most efficient strategy to achieve a certain tolerance TOL is to choose meshes adaptively in such a way that each of the N cells contributes a roughly equal amount $\eta_{K,h}^m + \eta_{K,k}^m \approx \mathrm{TOL}/N$ to the global error (see Bangerth & Rannacher [6]).

Since the time step can not be chosen individually for different cells but needs to remain fixed globally, algorithms can only provide approximate solutions to this goal. A strategy for this is to aim at the following distribution of errors that allots roughly half of the total error budget to spatial and temporal discretization errors, and splits the spatial error budget among time steps proportionally to the lengths of time intervals:

1. *Adaptation in time.* Choose the time step k_m so that

$$\frac{\alpha}{2}\frac{\mathrm{TOL}}{M} \leqslant \eta_k^m \leqslant \frac{1}{2}\frac{\mathrm{TOL}}{M}.$$

2. *Adaptation in space.* Choose h_K such that

$$\frac{\beta}{2}\frac{k_m}{T}\frac{\mathrm{TOL}}{N_m} \leqslant \eta_{K,h}^m \leqslant \frac{1}{2}\frac{k_m}{T}\frac{\mathrm{TOL}}{N_m}.$$

A practical choice of tuning parameters is $\alpha = \beta = \frac{1}{4}$.

In actual implementations, it is often not feasible to immediately choose k_m, h_K in such a way that these inequalities are satisfied. Rather, they are iteratively achieved. To this end, one starts with solving the discrete problem on a coarse space and time mesh. Error indicators are then evaluated for each cell and time interval. Those time intervals and cells that do not satisfy the upper error bound are then refined, while those that do not satisfy the lower bound may be coarsened. Alternatively, a fixed fraction (say 25%) of those time intervals with the largest error indicator η_k^m may be refined, and similarly on each time interval a fixed fraction of cells with the largest indicators $\eta_{K,h}^m$ will be refined, while a separate fraction of time intervals and cells with the lowest indicators may be coarsened. In either case, the process repeats until these cycles yield time step and mesh sizes that satisfy the above error bounds. At this point, there holds

$$\tilde{\eta}_\omega(U) \leqslant \sum_{m=1}^{M}\sum_{K\in\mathbb{T}_h^m}\eta_{K,h}^m + \sum_{m=1}^{M}\eta_k^m \leqslant \sum_{m=1}^{M}\sum_{K\in\mathbb{T}_h^m}\frac{1}{2}\frac{k_m}{T}\frac{\mathrm{TOL}}{N_m} + \sum_{m=1}^{M}\frac{1}{2}\frac{\mathrm{TOL}}{M} =$$

$$\frac{\mathrm{TOL}}{2}\sum_{m=1}^{M}\frac{k_m}{T}\sum_{K\in\mathbb{T}_h^m}\frac{1}{N_m} + \frac{TOL}{2}\sum_{m=1}^{M}\frac{1}{M} \;=\; TOL, \tag{4.36}$$

i.e., the adaptation process has reached the prescribed goal and is stopped.

Alternatively, if no fixed error tolerance TOL but rather a maximum number of cells is prescribed, the goal of the adaptation process is to reach a maximum of accuracy under this constraint. This can be achieved by the process described above using a decaying sequence of tolerances $\mathrm{TOL}_k \to 0$ $(k \to \infty)$.

5. Numerical examples

All the test examples presented in this section have been calculated using a program based on the Open Source finite element library `deal.II`, see Bangerth & al. [4] and the project website `http://www.dealii.org/`.

5.1. Adaptation in time. The first two tests concern the adaptation of the time discretization by the DWR approach in order to reach a certain error tolerance TOL for a pre-chosen spatial discretization of sufficiently high accuracy. Hence in all error estimates the contribution by the spatial discretization is neglected, and we will here only consider spatially one-dimensional examples. The error measure is the natural "energy norm"

$$J(e_k) = \|e_k^0(T)\| + \|e_k^1(T)\|, \qquad e_k^i := u^i - u_k^i, \;\; i = 1, 2,$$

at the end-time T, which is *local* in time but *global* in space. In view of the discussion in Section 4 the corresponding error functional is taken as

$$J(\varphi) := \frac{(\varphi^0(T), \hat{e}_k^0(T))}{\|\hat{e}_k^0(T)\|} + \frac{(\varphi^1(T), \hat{e}_k^1(T))}{\|\hat{e}_k^1(T)\|},$$

which for $\varphi := e_k = \{e_k^0, e_k^1\}$ returns the desired error quantity if $\hat{e}_k^i = e_k^i$. To this end, the a priori unknown quantities $\hat{e}_k^i(T)$, $i = 0, 1$, in this functional $J(\cdot)$ are successively chosen within the refinement process by extrapolating the discrete solution from preceding coarser meshes. To start this approximation on coarse meshes, we use the result from the computation on an auxiliary finer mesh. The quality of the error estimation is measured in terms of the "effectivity index"

$$I_{\text{eff}} := \left| \frac{J(e_k)}{\tilde{\eta}_\omega(u_k)} \right|,$$

where $\tilde{\eta}_\omega(u_k)$ is the approximate error estimator (4.33) (here only applied for the time discretization) evaluated using the technique of "local high-order interpolation" for approximating the dual solution as outlined in Section 4. If the error estimator yields a good approximation of the actual error, I_{eff} should be close to 1.

Example 5.1. The first test problem is the acoustic wave equation with a right-hand side independent of the spatial variable,

$$\partial_t^2 u(x,t) - \partial_x^2 u(x,t) = 10 \cdot e^{-100(t-9.0)^2},$$

$$u(x,0) = 0, \quad \partial_t u(x,0) = x - x^2,$$

on the space-time region

$$\Omega \times I = (0,1) \times (0,10).$$

We fix the spatial grid at 256 elements. Furthermore, we start with a uniform time grid of 16 time intervals which are to be both globally and locally refined. The reference solutions are computed on a highly refined time grid of $2^{19} = 524\,288$ time intervals.

First, we consider global uniform refinement. Table 5.1 shows the obtained values of the time-discretization error $J(e_k)$, the error estimator $\eta(u_k)$ and the corresponding effectivity index I_{eff}. Then, we consider local time-step adaptation based on the a posteriori error estimate derived in Section 4. The "fixed-fraction" strategy is used for step-size adaptation with results shown in Table 5.2. This data is also visualized in Figs. 5.1 and 5.2. A closer look at the distribution of the time steps shows that the local refinement mainly occurs at the time period around and after $t = 9$. This reflects the fact that only at $t = 9$ energy is fed into the system by the increase in the right-hand side. However, according to the energy conservation property of both the continuous problem and the cG(1)/cG(1) approximation (or equivalently the Crank-Nicolson scheme) there is no energy dissipation in the system afterwards, i.e., significant re-coarsening of the time grid can not be expected.

Table 5.1. **Example 5.1: Time-discretization error $J(e_k)$, error estimator $\tilde{\eta}_\omega(u_k)$ and effectivity index I_{eff} for increasing number of time steps under uniform refinement**

Number of time steps	$J(e_k)$	$\tilde{\eta}_\omega(u_k)$	I_{eff}
32	$6.79 \cdot 10^{-1}$	$3.22 \cdot 10^{-1}$	2.11
64	$4.39 \cdot 10^{-1}$	$3.55 \cdot 10^{-1}$	1.24
128	$2.20 \cdot 10^{-1}$	$2.19 \cdot 10^{-1}$	1.00
256	$6.98 \cdot 10^{-2}$	$7.48 \cdot 10^{-2}$	0.93
512	$9.03 \cdot 10^{-3}$	$9.72 \cdot 10^{-3}$	0.93
1024	$1.18 \cdot 10^{-3}$	$1.07 \cdot 10^{-3}$	1.10
2048	$3.12 \cdot 10^{-4}$	$2.94 \cdot 10^{-4}$	1.06
4096	$8.84 \cdot 10^{-5}$	$8.32 \cdot 10^{-5}$	1.06

Table 5.2. **Example 5.1: Quantities as in Table 5.1 but for local refinement (only a few steps shown)**

Number of time steps	$J(e_k)$	$\tilde{\eta}_\omega(u_k)$	I_{eff}
31	$6.86 \cdot 10^{-1}$	$6.18 \cdot 10^{-1}$	1.11
58	$2.15 \cdot 10^{-1}$	$2.42 \cdot 10^{-1}$	0.89
112	$4.29 \cdot 10^{-2}$	$4.40 \cdot 10^{-2}$	0.98
218	$1.33 \cdot 10^{-2}$	$1.48 \cdot 10^{-2}$	0.90
425	$2.87 \cdot 10^{-3}$	$3.08 \cdot 10^{-3}$	0.93
828	$7.62 \cdot 10^{-4}$	$8.41 \cdot 10^{-4}$	0.91
1616	$2.96 \cdot 10^{-4}$	$3.30 \cdot 10^{-4}$	0.90
3156	$4.56 \cdot 10^{-5}$	$4.74 \cdot 10^{-5}$	0.96

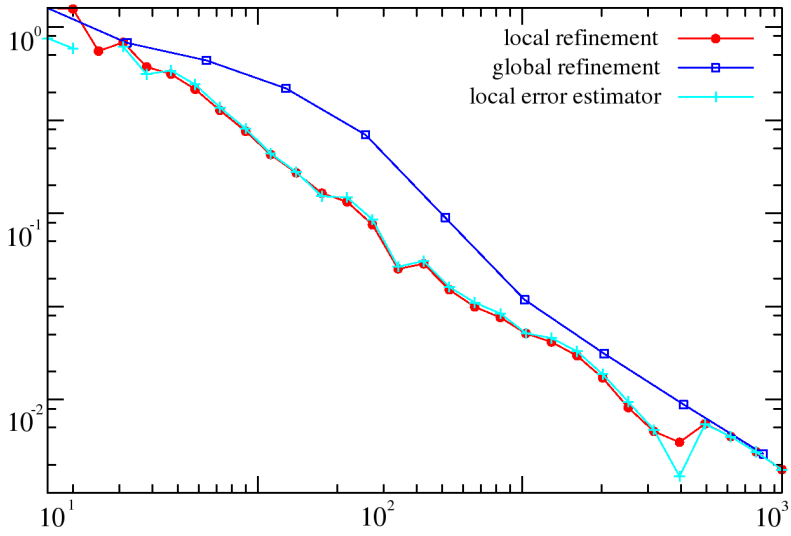


Fig. 5.1. Example 5.1: Error for different numbers of time steps with local and global refinement

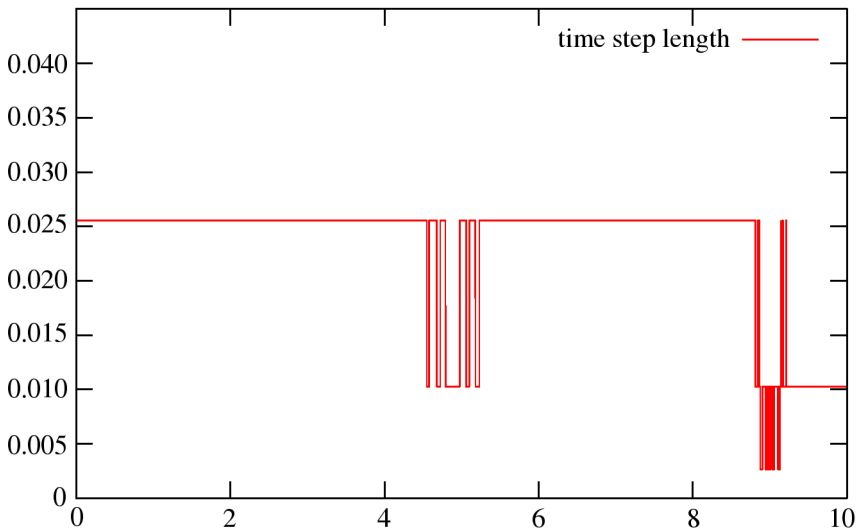


Fig. 5.2. Example 5.1: Local time-step size after several steps of local refinement

Example 5.2. The second test problem is the acoustic wave equation with a homogeneous right-hand side $f \equiv 0$ and the exact solution $u(x,t) = \sin(\pi x)\sin(\pi t)$. The space-time domain is the same as in Example 5.1. Compared to the previous example, we replace the refinement strategy by one where we refine all those time intervals where the local temporal error is larger than the mean value of the error over all time intervals. The results of global as well as local refinement in this case are shown in Fig. 5.3. It can be seen that the error estimator shows the same behavior as global grid refinement. This results from the fact that the temporal error in this example is equally distributed over all time steps as the solution is solely driven by the initial conditions; local refinement on the basis of local error indicators can therefore not be expected to improve the accuracy any more than global refinement. Indeed, after several refinement cycles, all time intervals are equally long. However, the effectivity index is close to 1.

Table 5.3. **Example 5.2: Discretization error in time $J(e_k)$, error estimator $\tilde{\eta}_\omega(u_k)$ and effectivity index I_{eff} for increasing number of time steps with global uniform refinement**

Number of time steps	$J(e_k)$	$\tilde{\eta}_\omega(u_k)$	I_{eff}
16	0.271	–	–
32	3.246	3.607	0.89
64	0.445	0.459	0.96
128	0.117	0.116	1.01
256	0.031	0.031	1.00
512	0.008	0.008	1.00
1024	0.002	0.002	1.00
2048	0.000	0.000	1.00

Table 5.4. **Example 5.2: Quantities as in Table 5.3 but for local refinement (only a few steps shown)**

Number of time steps	$J(e_k)$	$\tilde{\eta}_\omega(u_k)$	I_{eff}
15	1.389	–	–
29	4.278	4.612	0.92
53	1.168	1.284	0.91
98	0.259	0.257	1.00
193	0.075	0.074	1.00
384	0.019	0.019	1.00
768	0.005	0.005	1.00
1536	0.001	0.001	0.99

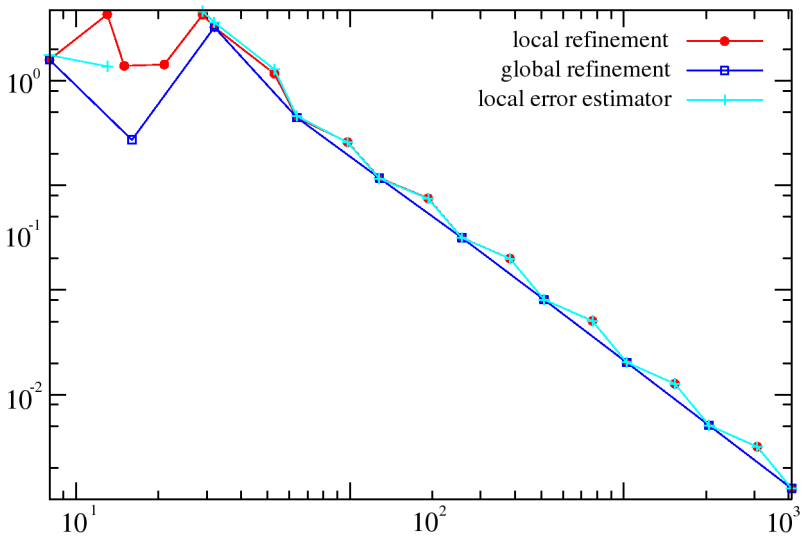


Fig. 5.3. Example 5.2: Error for different numbers of time steps with local and global refinement

5.2. Adaptation in space. The next three examples illustrate the performance of the DWR approach for spatial mesh adaptation. To this end in all examples the data of the

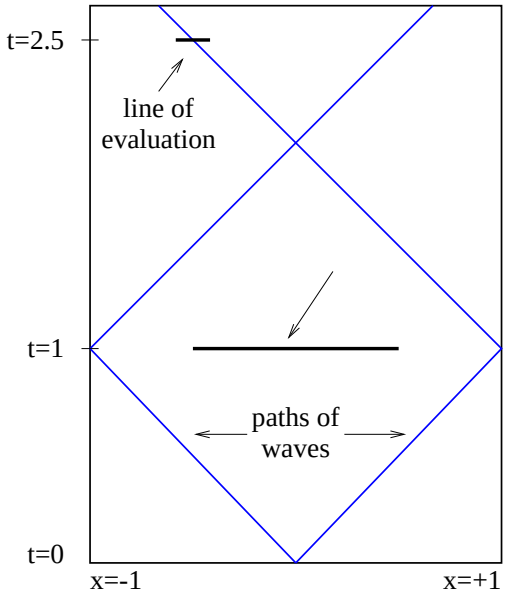
problem (right-hand side and diffusion coefficient) are chosen such that the exact solution has a complex dynamic spatial behavior. Here, we choose the time step sizes so that they satisfy a local CFL condition; the temporal errors are therefore small on all but the smallest cells and their contributions in the error estimates will consequently be neglected.

Example 5.3. We begin with an example in one space dimension. Let in (2.2) be $d = 1$, $a \equiv 1$, and $\Omega = (-1, 1)$. We choose initial and boundary conditions as shown in Fig. 5.4, with half-width $s = 0.1$ and end-time $T = 2.7$. This choice of initial conditions leads to two "peaks" traveling to the left and right and being reflected at the boundaries.

$$\begin{aligned} u(x,0) &= e^{-|x_s|^2}(1-|x_s|^2)\Theta(1-|x_s|), \\ \partial_t u(x,0) &= 0, \\ \partial_x u(1) = 0, &\quad \partial_x u(-1) = 0, \end{aligned}$$

with $x_s = x/s$ and the jump function

$$\Theta(y) = \begin{cases} 0 & \text{for } y < 0, \\ 1 & \text{for } y \geqslant 0. \end{cases}$$



F i g. 5.4. Example 5.3: Configuration

We are now interested in the evaluation of only one branch of the solution, for example the one initially traveling to the right. Here, we choose to localize the measurements at $t = 2.5$ and around $x = -0.5$ and use as goal quantity

$$J(u) = \int_{-0.6}^{-0.4} u^0(x, 2.5)\, \mathrm{d}x. \tag{5.1}$$

Note that the solution's two peaks are centered around $x = \pm 0.5$ at $t = 2.5$, with diameter $2s = 0.2$ as in the initial distribution. The solutions u, z of the primal and dual problems are shown in Fig. 5.5 on the left. As can easily be seen, the integral kernel of the functional $J(\cdot)$, i.e., the characteristic function of $[-0.6, 0.4] \times \{2.5\}$, serves as source term for the dual solution. The dual solution therefore is discontinuous in time due to the singular integral kernel.

The resulting space-time grid after three refinement cycles is also shown in Fig. 5.5. As can be seen, the error estimator does not only track just one branch as would be the obvious thing to do, but also takes into account errors occurring in the whole space-time domain as long as the laws of wave propagation allow them to affect the goal functional $J(\cdot)$. It is therefore clearly more efficient than almost any choice of a priori refining the mesh by hand. Also note that for $t > 2.5$, the dual solution is zero and consequently the mesh is coarsened in each refinement cycle. Consequently, the solution is hardly resolved at these times, in accordance with the fact that it then does not matter any more for our goal.

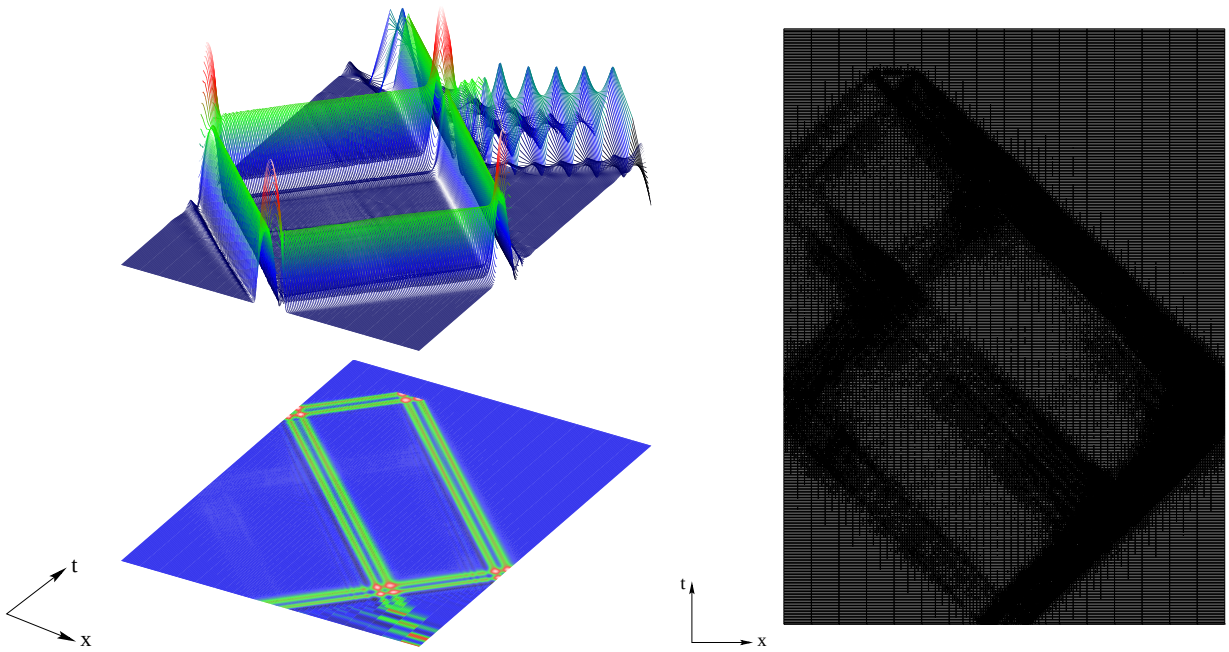


Fig. 5.5. Example 5.3: Primal solution u and dual solution z (left). (Notice that the dual solution is computed backward in time.) Resulting space-time grid after three cycles of refinement (right)

Example 5.4. Next, let us consider a more realistic example in two space dimensions that could mimic the propagation of waves in a layered medium such as a simplified model of Earth. Let $\Omega = (-1,1)^2 \subset \mathbb{R}^2$ and the initial values in (2.1) be

$$u_0^0(x) = e^{-|x_s|^2}(1-|x_s|^2)\,\Theta(1-|x_s|), \qquad u_0^1 = 0,$$

with x_s and the jump function $\Theta(\cdot)$ as defined in Example 5.3, and $s = 0.01$. We choose the elasticity coefficient discontinuous, $a = 1$ for $y < 0.2$, and $a = 9$ for $y \geqslant 0.2$. A typical wave pattern is shown in Fig. 5.6.

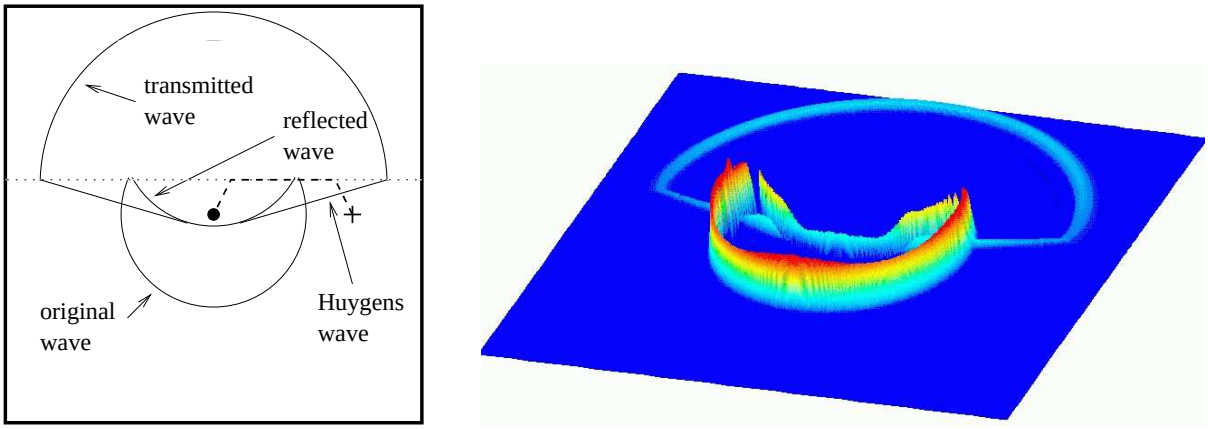


Fig. 5.6. Example 5.4: Layout of the wave pattern after some time, a bullet and a cross indicate the positions of source and receiver, respectively (left). The dotted line indicates the discontinuity in the coefficient, while the thick dashed line denotes the path of least action. Plot of the solution at $t = 0.45$ (right)

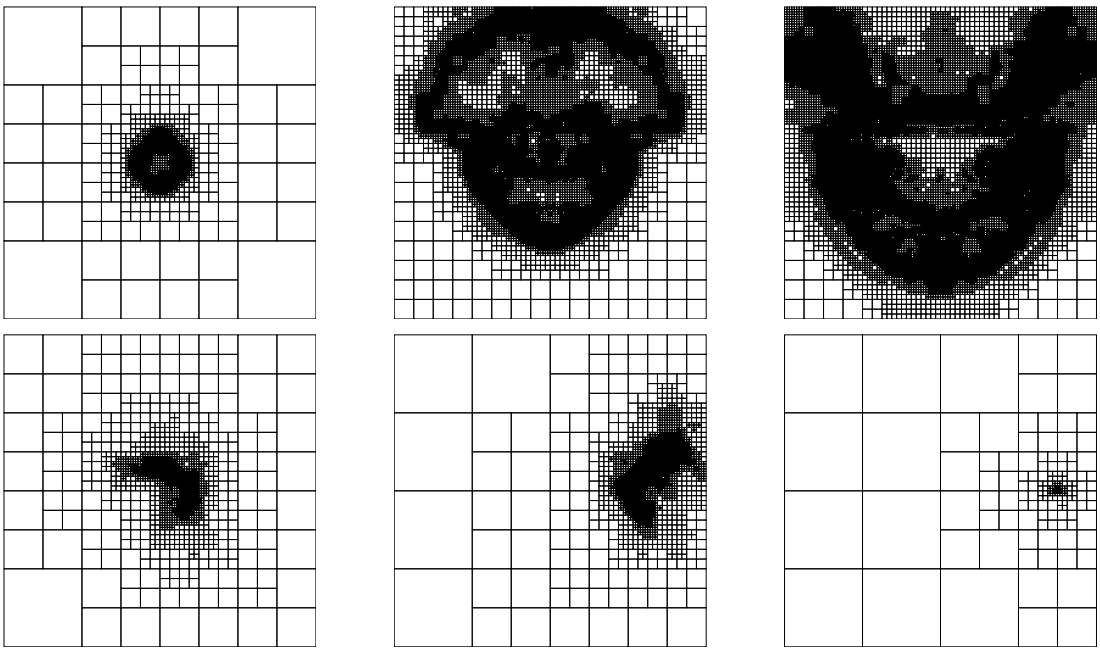
In geophysics and seismics it is an important task to accurately model the signal arrival time at a given point. In our case, let us assume that we are interested in the situation at the point $x_0 = (0.75, 0)$. As shown in the layout (see Fig. 5.6) the three first waves arriving at this point are the Huygens' wave, the direct wave, and the one reflected from

the discontinuity. The first one travels into the medium of higher wave velocity, travels some distance parallel to the discontinuity and then back towards the point of measurement. Among all waves it is the one which has the least action along its path and is therefore called Huygens' wave. From extrapolation of computed data, we estimate its arrival time to be approximately $\tau_H \approx 0.618$, while the arrival times of the other ones are $\tau_d \approx 0.75$ for the direct wave and $\tau_r \approx 0.85$ for the reflected wave. A quantity related to the arrival time is

$$J(u) = \int_{t_1}^{t_2} u(x_0, t)\, t \, \mathrm{d}t,$$

with a time interval $[t_1, t_2]$ suitably chosen around the signal and such that it does not include other signals. This interval is usually chosen in accordance with experimental data. We take $t_1 = 0.55$ and $t_2 = 0.68$, to catch the first wave only. Accordingly, we choose $T = t_2$, to stop the computation at the first possible time – although we could also extend T with the effect that automatic refinement would coarsen meshes after t_2 to a single cell as in the previous example. In this case the goal functional $J(\cdot)$ is not regular enough to guarantee that the corresponding dual solution satisfies $\hat{z} \in \hat{\mathcal{V}}$. This complication may be solved by "regularization" as discussed in Section 4.

In Fig. 5.7, we show the computational grids at times $t = 0.15$, $t = 0.45$, and $t = T$, as generated by refinement by the heuristic energy error indicator (4.25) and by the weighted estimator (DWR method). It is readily seen that the latter only tracks that part of the wave field that travels to the right. A closer look at a more complete sequence of grids than shown here reveals that the most refined parts of the grids indeed track the path of least action (the dashed line in Fig. 5.6) which marks the path of the first signal to arrive at the receiver. The first grid shown is at a time where the wave to arrive first is still traveling upward, while in the second it is already traveling downward again. These complicated features of wave propagation are clearly reflected in the grids.



F i g. 5.7. Example 5.4: Grids at times $t = 0.15$, $t = 0.45$ and $t = T$, with refinement by the simplified energy error indicator (4.25) (rop row). Grids produced by the DWR method (four cycles of refinement and coarsening) (bottom row)

In Figure 5.8 the convergence of $J(U_{hk})$ towards the inferred value $J(u) \approx 0.618$ is shown. Since the grids only tracked the interesting part of the wave it is not surprising that it accomplishes the same accuracy with a significantly lower number of space-time cells than the grids refined with the simplified energy error indicator (4.25). Note that the dip in each curve is due to the error, $J(U_{hk}) - 0.618$, changing its sign, which happens to bring $J(U_{hk})$ close to the exact value. Leaving aside these two data points, the grids as refined by the DWR method show a higher order of convergence than the grids as refined by the heuristic approach. It should be mentioned that refinement by the two methods starts from the same grid, but that in the first step the DWR methods coarsens more cells than it refines, which leads to an overall decrease of space-time cells.

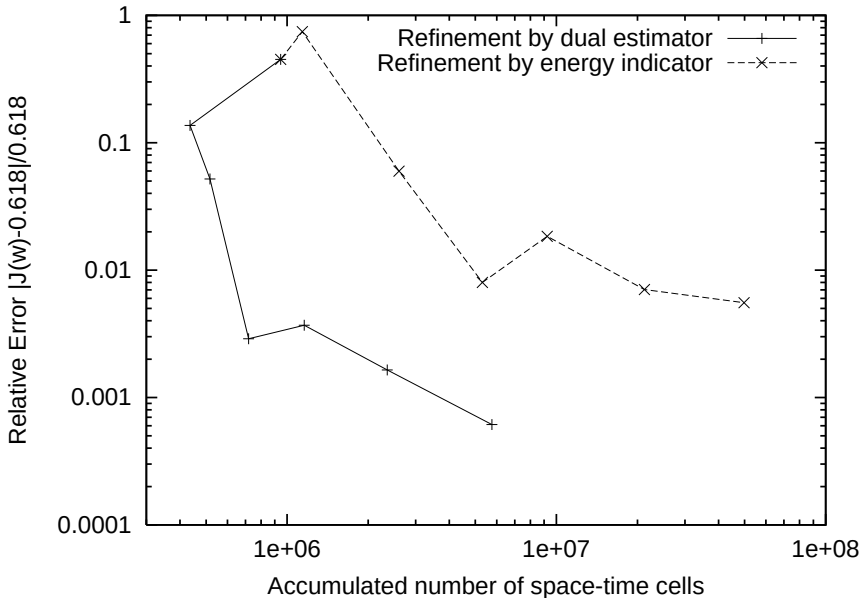


Fig. 5.8. Example 5.4: Convergence of the functional $J(U)$ to the exact value $J(u) \approx 0.618$ (The dips in the curves are due to sign changes of the error)

Example 5.5. The last example demonstrates the performance of the DWR method for computing the propagation of an outward traveling wave on $\Omega = (-1, 1)^2$ with a strongly heterogeneous coefficient as is frequently found in many engineering and earth sciences applications. Layout of the domain and structure of the coefficient are shown in Fig. 5.9. We choose initial conditions as in the previous example but with $s = 0.02$, and boundary conditions as follows:

$$n \cdot \{a \nabla u\} = 0 \quad \text{on } y = 1, \quad u = 0 \quad \text{on } \partial\Omega \backslash \{y = 1\}.$$

The region of origin of the wave field is significantly smaller than shown in Fig. 5.9.

Notice that the lowest frequency in this initial wave field has wavelength $\lambda = 4s$; hence taking the common minimum ten grid points per wavelength would yield 62,500 cells already for the largest wavelength, rendering uniformly refined grids unable to produce high accuracy for such cases. If we consider this example as a model of propagation of seismic waves in a faulted region of rock, then we would be interested in recording seismograms at the surface,

here chosen as the top line Γ of the domain. A corresponding functional output is

$$J(u) = \int_0^T \int_\Gamma u(s,t)\,\omega(s,t)\,\mathrm{ds}\,\mathrm{dt},$$

with a weight factor $\omega(s,t) = \sin(3\pi s)\sin(5\pi t/T)$, and end-time $T = 2$. The frequency of oscillation of this weight is chosen to match the frequencies in the wave field to obtain good resolution of changes.

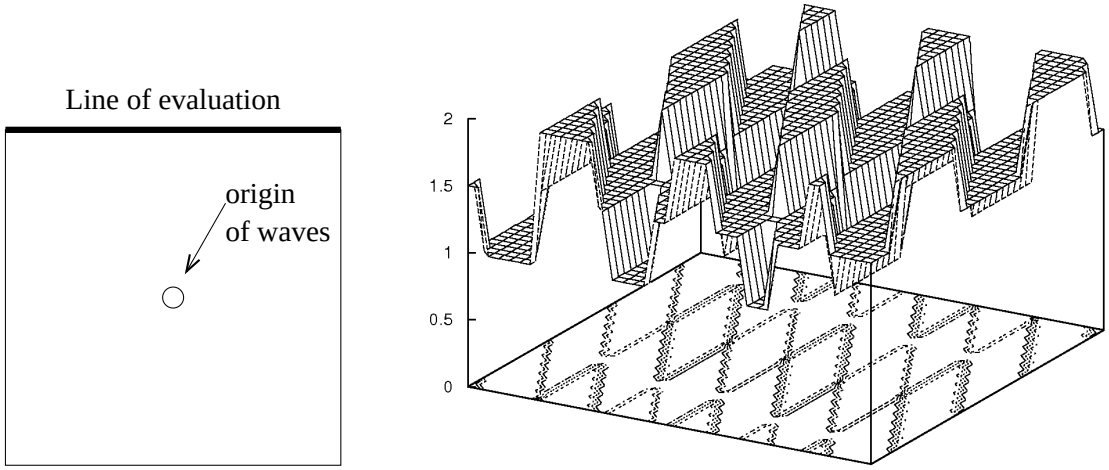


Fig. 5.9. Example 5.5: Layout of the domain (left) and structure of the coefficient $a(x)$ (right)

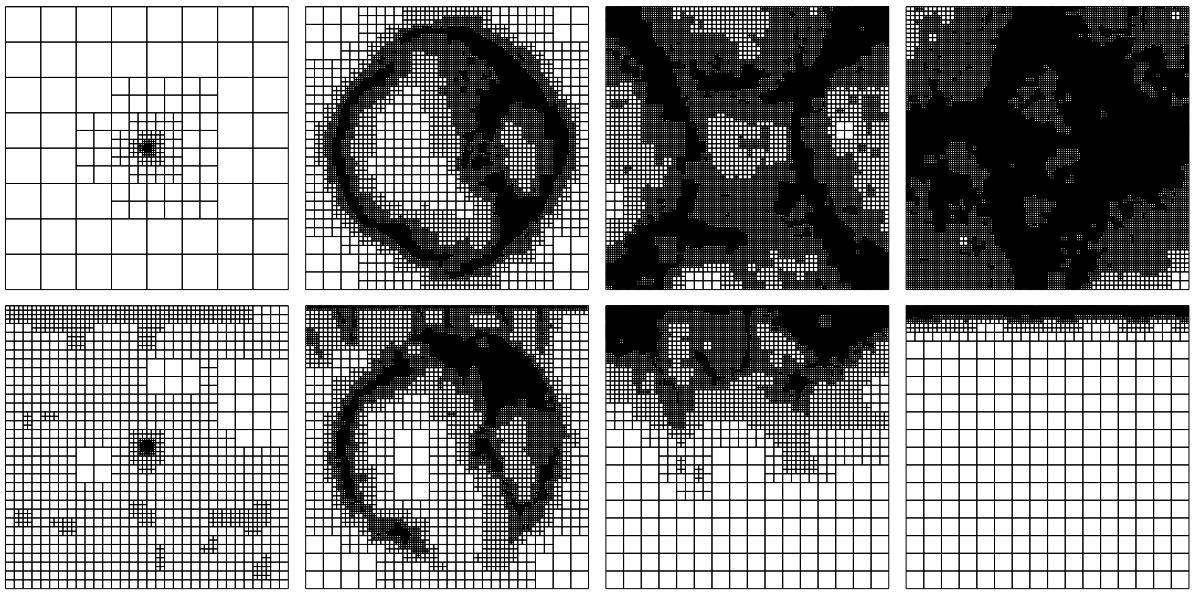
Remark 5.1. The evaluation of the weighted *a posteriori* error estimate of the DWR method requires a careful approximation of the adjoint solution z. Therefore, in this example, we have used a higher-order method (bi-quadratic elements) for solving the space-time adjoint problem, though this does not seem feasible for complex higher-dimensional problems.

Table 5.5. **Example 5.5: Results obtained by adaptation of spatial discretization using the DWR method (reference value $J(u) \approx -4.515 \cdot 10^{-6}$, M = number time steps, N = average number of mesh cells)**

Weighted estimator		Heuristic indicator	
$N \times M$	$J(U)$	$N \times M$	$J(U)$
327 789	$-2.085 \cdot 10^{-6}$	327 789	$-2.085 \cdot 10^{-6}$
920 380	$-4.630 \cdot 10^{-6}$	920 380	$-4.630 \cdot 10^{-6}$
2 403 759	$-4.286 \cdot 10^{-6}$	2 403 759	$-4.286 \cdot 10^{-6}$
1 918 696	$-4.177 \cdot 10^{-6}$	5 640 223	$-4.385 \cdot 10^{-6}$
2 975 119	$-4.438 \cdot 10^{-6}$	10 189 837	$-4.463 \cdot 10^{-6}$
6 203 497	$-4.524 \cdot 10^{-6}$	17 912 981	$-4.521 \cdot 10^{-6}$
		41 991 779	$-4.517 \cdot 10^{-6}$

In Fig. 5.10, we show the grids resulting from refinement by the DWR method compared with the heuristic energy-error-based method. Both initially resolve the wave field quite well, including reflections from discontinuities in the coefficient. On the other hand, the DWR refinement indicator additionally takes into account that the lower parts of the domain lie outside the domain of influence of the target functional if we truncate the time domain at

$T = 2$; this domain of influence constricts to the top as we approach the final time, as is reflected by the produced grids. The meshes obtained in this way are obviously much more economical, without degrading the accuracy in approximating the quantity of interest.



F i g. 5.10. Example 5.5: Grids produced by the energy-error indicator (top row) and by the weighted estimator (bottom row) at times $t = 0$, $\frac{2}{3}$, $\frac{4}{3}$, 2

6. Summary

In this paper, we have discussed several discretization methods for the wave equation including a posteriori error control and mesh-size adaptation for computing physically meaningful quantities. Adjustment of spatial and temporal mesh sizes is based on an a posteriori representation of the exact error with respect to an arbitrary functional of the solution, which includes the local residuals of the numerical solution and local weights derived from the solution of a dual problem associated with the quantity of interest. This approach, the "Dual Weighted Residual (DWR)" method, fundamentally relies on the Galerkin character of the underlying space-time discretization. Therefore it is important to note that also common finite difference time-stepping schemes such as certain variants of the Crank-Nicolson and the (trapezoidal) Newmark scheme fit into this framework as they are algebraically equivalent to certain lower-order "continuous" or "discontinuous" Galerkin schemes. Consequently, the error estimation techniques outlined here are also applicable to these established and well-understood methods.

It has been demonstrated that meshes generated with the aid of refinement criteria derived by the DWR approach are significantly superior to meshes obtained by a simplified refinement indicator which does not include information on the quantity of interest. The superiority has been demonstrated by several examples of one and two dimensional wave propagation, including high-frequency waves and discontinuous coefficients. In particular, it has been shown that refinement based on the error representation is able to track where information comes from, thus leading to highly localized mesh refinement if the target functional is localized. In general, the smaller the region of evaluation of the target functional is, the larger are the savings of the DWR approach compared to global refinement and to

more traditional approaches of adaptivity. Furthermore, by this approach quantitative error control is feasible.

Finally, several aspects of the underlying mechanisms have been discussed, particularly alternative ways of evaluating the a posteriori error representation formula. Good mesh refinement criteria that include localized information about the target functional can be obtained by solving the dual problem to the same accuracy as the primal one. This may double the computational cost compared to the pure forward solution, but usually reduces the computing work by at least an order of magnitude compared to simple ad hoc approaches to adaptivity, due to the more economical meshes produced, and can therefore allow the numerical treatment of problems for which sufficient accuracy would otherwise not be achievable.

References

1. L. Bales and I. Lasiecka, *Continuous finite elements in space and time for the nonhomogeneous wave equation*, Computers Math. Applic. **27** (1994), pp. 91–102.

2. W. Bangerth, *Adaptive Finite-Elemente-Methoden zur Lösung der Wellengleichung mit Anwendung in der Physik der Sonne*, Diploma thesis, University of Heidelberg, 1998.

3. W. Bangerth and R. Rannacher, *Finite element approximation of the acoustic wave equation: Error control and mesh adaptivity*, East-West J. Numer. Math. **7** (1999), pp. 263–282.

4. W. Bangerth, R. Hartmann, and G. Kanschat, *deal.II – a General Purpose Object Oriented Finite Element Library*, ACM Trans. Math. Softw. **33** (2007), pp. 24/1–27.

5. W. Bangerth and R. Rannacher, *Adaptive finite element techniques for the acoustic wave equation*, J. Comput. Acoustics **9** (2001), pp. 575–591.

6. W. Bangerth and R. Rannacher, *Adaptive Finite Element Methods for Differential Equations*, Birkhäuser, Basel, 2003.

7. R. Becker and R. Rannacher, *An optimal control approach to a posteriori error estimation in finite element methods*, Acta Numerica 2000 (A. Iserles, ed.), pp. 1–102, Cambridge University Press, 2001.

8. C. Bernardi and E. Süli, *Time and space adaptivity for the second-order wave equation*, Math. Models Methods Appl. Sci. **15** (2005), pp. 199–225.

9. S. C. Brenner and R. L. Scott, *The Mathematical Theory of Finite Element Methods*, Springer, Berlin-Heidelberg-New York, 1994.

10. C. E. Castro, M. Käser, and E. F. Toro, *Space-time adaptive numerical methods for geophysical applications*, Roy. Soc. Phil. Trans. A. **376** (2009), pp. 4613–4631.

11. P. Ciarlet, *Finite Element Methods for Elliptic Problems*, North-Holland, Amsterdam, 1978.

12. M. Dumbser, M. Käser, and E. F. Toro, *An arbitrary high order discontinuous Galerkin method for elastic waves on unstructured meshes V: local time stepping and p-adaptivity*, Geophys. J. Int. **171** (2007), pp. 695–717.

13. D. A. French and T. E. Peterson, *A continuous space-time finite element method for the wave equation*, Math. Comput. **65** (1996), pp. 491–506.

14. M. Geiger, *Vergleich dreier Zeitschrittverfahren zur numerischen Lösung der akustischen Wellengleichung*, Diploma thesis, University of Heidelberg, 2008.

15. G. L. Goudreau and R. L. Taylor, *Evaluation of numerical integration methods in elastodynamics*, Comp. Meth. Appl. Mech. Eng. **2** (1972), pp. 69–97.

16. C. Großmann and H.-G. Roos, *Numerische Behandlung partieller Differentialgleichungen*, Teubner, 2005.

17. M. Grote, A. Schneebeli, and D. Schötzau, *Discontinuous Galerkin finite element method for the wave equation*, SIAM J. Numer. Anal. **44** (2006), pp. 2408–2431.

18. R. Hartmann, *A posteriori Fehlerschätzung und adaptive Schrittweiten- und Ortsgittersteuerung bei Galerkin-Verfahren für die Wärmeleitungsgleichung*, Diploma thesis, University of Heidelberg, 1998.

19. T. J. R. Hughes, *The Finite Element Method*, Dover Publications, 2000.

20. T. J. R. Hughes and G. Hulbert, *Space-time finite element methods for elastodynamics: formulations and error estimates*, Comput. Methods Appl. Mech. Engrg. **66** (1988), pp. 339–363.

21. T. J. R. Hughes and G. Hulbert, *Space-time finite element methods for second-order hyperbolic equations*, Comput. Methods Appl. Mech. Engrg. **84** (1990), pp. 327–348.

22. G. Hulbert, *Time finite element methods for structural dynamics*, Inter. J. Numer. Methods Engrg. (1992), pp. 307–331.

23. C. Johnson, *Numerical Solution of Partial Differential Equations by the Finite Element Method*, Cambridge University Press, Cambridge, 1987.

24. C. Johnson, *Discontinuous Galerkin finite element methods for second order hyperbolic problems*, Comput. Method. Appl. Mech. Engrg. **107** (1993), pp. 117–129.

25. J. L. Lions, *Optimal Control of Systems Governed by Partial Differential Equations*, Springer, Berlin-Heidelberg-New York, 1971.

26. J. L. Lions and E. Magenes, *Problèmes aux Limites non Homogènes et Applications, 1,2,3*, Dunod, Paris, 1968.

27. X. Li and N.-E. Wiberg, *Implementation and adaptivity of a space-time finite element method for structural dynamics*, Comput. Methods Appl. Mech. Engrg. **156** (1998), pp. 211–229.

28. D. W. Kelly, J. R. Gago, O. C. Zienkiewicz, and I. Babuska, *A posteriori error analysis and adaptive processes in the finite element method*, Int. J. Numer. Methods Engrg. **156** (1998), pp. 211–229.

29. D. Meidner, *Adaptive Space-Time Finite Element Methods for Optimization Problems Governed by Nonlinear Parabolic Systems*, Dissertation, University of Heidelberg, 2008.

30. N. M. Newmark, *A method of computation for structural dynamics*, J. Eng. Mech. Div. ASCE. **85** (1959), pp. 67–94.

31. I. Romero and L. M. Lacoma, *A methodology for the formulation of error estimators for time integration in linear solid and structural dynamics*, Int. J. Numer. Methods Eng. **66** (2006), pp. 635–660.

32. M. Schemann and F. A. Bornemann, *An adaptive Rothe method for the wave equation*, Comput. Visual. Sci. **1** (1998), pp. 137–144.

33. M. Schmich and B. Vexler, *Adaptivity with dynamic meshes for space-time finite element discretizations of parabolic equations*, SIAM J. Sci. Comput. **30** (2008), pp. 369–393.

34. K. Schweizerhof, J. Neumann, and S. Kizio, *On time integration error estimation and adaptive time stepping in structural dynamics*, Proc. in Applied Mathematics and Mechanics. **4** (2004), pp. 35–38.

35. E. Süli and C. Wilkins, *Adaptive finite element methods for the damped wave equation*, Report no. 96/23, Oxford University Computing Laboratory, 1996.

36. E. L. Wilson, *Static and dynamic analysis of structures*, Computers and Structures, Inc., 2000.

37. N.-E. Wiberg and X. Li, *Adaptive finite element procedures for linear and non-linear dynamics*, Int. J. Numer. Methods Eng. **4** (1998), pp. 1781–1802.

38. J. Wloka, *Partial Differential Equations*, Cambridge University Press, Cambridge-London-New York-New Rochelle-Melbourne-Sydney, 1987.

39. W. L. Wood, *A unified set of single step algorithms. Part II: Theory*, Int. J. Numer. Meth. Eng. **20** (1984), pp. 2303–2309.

40. W. L. Wood, *Practical Time-stepping Schemes*, Clarendon Press, 1990.

41. O. C. Zienkiewicz, W. L. Wood, and N. W. Hine, *A unified set of single step algorithms. Part I: General formulation and applications*, Int. J. Numer. Meth. Eng. **20** (1984), pp. 1529–1552.