Lecture Notes

Topics in Finite Elements Part I Introduction and Classical Methods for the Wave Equation

Marco Zank
Fakultät für Mathematik, Universität Wien,
Oskar-Morgenstern-Platz 1, 1090 Wien

Draft from 2020-09-08, 14:39:36

Contents

The model problem is the scalar wave equation

$$\left.\begin{array}{rcll} \partial_{tt}u(x,t) - \Delta_x u(x,t) & = & f(x,t) & \text{for } (x,t) \in Q = \Omega \times (0,T), \\ u(x,t) & = & 0 & \text{for } (x,t) \in \Sigma = \partial\Omega \times [0,T], \\ u(x,0) & = & u_0(x) & \text{for } x \in \Omega, \\ \partial_t u(x,0) & = & u_1(x) & \text{for } x \in \Omega, \end{array}\right\} \tag{1}$$

where

- a function $u\colon \overline{Q} \to \mathbb{R}$ is sought,
- the function $f\colon Q \to \mathbb{R}$ is a given right-hand side,
- the functions $u_0, u_1\colon \Omega \to \mathbb{R}$ are given initial conditions,
- $\Omega \subset \mathbb{R}^d$, $d = 1,2,3$, is a given bounded Lipschitz domain with boundary $\partial\Omega$,
- $T > 0$ is a given terminal time.

To approximate the solution u of (1), the classical approaches are time stepping schemes or finite difference methods in time together with finite element methods in space. An alternative is to discretise (1) without separating the temporal and spatial variables, which leads to the so-called space-time methods.

1 Variational Setting of the Wave Equation

This section is based on [13, Chapter 23 and 24]. The variational formulation of (1) is given as follows:

Find $u \in L^2(0,T;H_0^1(\Omega))$ with

- $\partial_t u \in L^2(0,T;L^2(\Omega))$,
- $\partial_{tt} u \in L^2(0,T;[H_0^1(\Omega)]')$,
- $u(\cdot,0) = u_0$ in $H_0^1(\Omega)$ and
- $\partial_t u(\cdot,0) = u_1$ in $L^2(\Omega)$

such that

$$\langle \partial_{tt}u(\cdot,t), v\rangle_\Omega + \langle \nabla_x u(\cdot,t), \nabla_x v\rangle_{L^2(\Omega)} = \langle f(\cdot,t), v\rangle_{L^2(\Omega)} \tag{2}$$

for almost all $t \in (0,T)$ and all $v \in H_0^1(\Omega)$, where $f \in L^2(0,T;L^2(\Omega))$, $u_0 \in H_0^1(\Omega)$ and $u_1 \in L^2(\Omega)$ are the given right-hand side and the given initial conditions.

Here, for $s \geq 0$, the usual Sobolev spaces of real-valued functions $H^s(\Omega)$, $H_0^s(\Omega)$ are endowed with the Sobolev-Slobodeckij inner product $\langle\cdot,\cdot\rangle_{H^s(\Omega)}$ and the norm $\|\cdot\|_{H^s(\Omega)}$. For the closed subspace $H_0^1(\Omega) \subset H^1(\Omega)$, the inner product

$$\langle w, v\rangle_{H_0^1(\Omega)} := \langle \nabla_x w, \nabla_x v\rangle_{L^2(\Omega)} = \int_\Omega \nabla_x w(x) \cdot \nabla_x v(x) \mathrm{d}x, \quad w, v \in H_0^1(\Omega),$$

and the induced norm

$$|w|_{H^1(\Omega)} := \|w\|_{H^1_0(\Omega)} = \sqrt{\langle w, w\rangle_{H^1_0(\Omega)}} = \sqrt{\int_\Omega |\nabla_x w(x)|^2 \,\mathrm{d}x}, \quad w \in H^1_0(\Omega),$$

are considered due to the Poincaré inequality

$$\forall v \in H^1_0(\Omega): \quad \|v\|_{L^2(\Omega)} \le C_P \|\nabla_x v\|_{L^2(\Omega)} \tag{3}$$

with a constant $C_P > 0$ independent of v. The dual space $[H^1_0(\Omega)]'$ is a Hilbert space characterised as the completion of $L^2(\Omega)$ with respect to the Hilbertian norm

$$\|g\|_{[H^1_0(\Omega)]'} := \sup_{0 \neq v \in H^1_0(\Omega)} \frac{|\langle g, v\rangle_\Omega|}{|v|_{H^1(\Omega)}},$$

where $\langle \cdot, \cdot \rangle_\Omega$ denotes the duality pairing as extension of the inner product in $L^2(\Omega)$.

In the following subsection, a short introduction to vector-valued functions and their function spaces is given.

1.1 Short Introduction to Vector-Valued Functions

In this subsection, let $(X, \langle \cdot, \cdot \rangle_X)$ be a real Hilbert space and $T \in (0, \infty)$.

Definition 1.1. *Let $m \in \mathbb{N}_0$ and $1 \le p < \infty$ be given.*

- *Define*

  $$C^m([0,T];X) := \left\{ v\colon [0,T] \to X \colon \; v^{(k)} \text{ is continuous on } [0,T] \text{ for all } k = 0, \dots, m \right\}$$

  with the norm

  $$\|v\|_{C^m([0,T];X)} := \sum_{k=0}^m \max_{0 \le t \le T} \left\| v^{(k)}(t) \right\|_X,$$

  where $v^{(0)} := v$ and

  $$\partial_t v(t) := v^{(1)}(t) := \lim_{s \to t} \frac{v(s) - v(t)}{s - t} \text{ in } X \quad \text{for each } t \in (0,T).$$

  In addition, set $C([0,T];X) := C^0([0,T];X)$.

- *Define*

  $$L^p(0,T;X) := \left\{ v\colon (0,T) \to X \colon \; v \text{ is measurable and } \|v\|_{L^p(0,T;X)} < \infty \right\}$$

  with the norm

  $$\|v\|_{L^p(0,T;X)} := \left(\int_0^T \|v(t)\|_X^p \mathrm{d}t \right)^{1/p},$$

  where v is measurable if $t \mapsto \langle v(t), w \rangle_X$ is measurable for all $w \in X$, see Bochner measurable, Bochner integral.

Lemma 1.2. *Let $m \in \mathbb{N}_0$ and $1 \le p < \infty$ be given.*

1. *The space $C^m([0,T];X)$ with the norm $\|\cdot\|_{C^m([0,T];X)}$ is a real Banach space.*
2. *The space $L^p(0,T;X)$ with the norm $\|\cdot\|_{L^p(0,T;X)}$ is a real Banach space in the case, where one identifies functions that are equal almost everywhere on $(0,T)$.*
3. *The space $L^2(0,T;X)$ is a real Hilbert space with the inner product*

$$\langle u,v\rangle_{L^2(0,T;X)} := \int_0^T \langle u(t),v(t)\rangle_X \mathrm{d}t \quad \text{for } u,v \in L^2(0,T;X).$$

4. *The space $C^m([0,T];X)$ is dense in $L^p(0,T;X)$ and the embedding*

$$C^m([0,T];X) \hookrightarrow L^p(0,T;X)$$

is continuous.

5. *Let $v \in L^1(0,T;X)$ be given. Then, there exists a unique element $u \in X$ such that*

$$\forall w \in X\colon \quad \int_0^T \langle v(t),w\rangle_X \mathrm{d}t = \langle u,w\rangle_X.$$

Set $\int_0^T v(t)\mathrm{d}t := u \in X$.

6. *The spaces $L^2(0,T;L^2(\Omega))$ and $L^2(Q)$ are isometric, i.e.*

$$L^2(0,T;L^2(\Omega)) \cong L^2(Q).$$

7. *The spaces $C([0,T];C(\overline{\Omega}))$ and $C(\overline{\Omega}\times[0,T])$ are isometric, i.e.*

$$C([0,T];C(\overline{\Omega})) \cong C(\overline{\Omega}\times[0,T]) = C(\overline{Q}).$$

Proof. See [13, Chapter 23], [1, Lemma 8.22, page 276]. For 7., see [8, page 50] or [3, page 288]. □

Definition 1.3. *Let $(Y,\langle\cdot,\cdot\rangle_Y)$ be another real Hilbert space. Further, let $u \in L^1(0,T;X)$ and $w \in L^1(0,T;Y)$ be given. For $n \in \mathbb{N}$, the function w is called the n-th generalised derivative of the function u on $(0,T)$ if*

$$\forall \varphi \in C_0^\infty(0,T)\colon \quad \underbrace{\int_0^T \varphi^{(n)}(t)u(t)\mathrm{d}t}_{\in X} = (-1)^n \underbrace{\int_0^T \varphi(t)w(t)\mathrm{d}t}_{\in Y}. \tag{4}$$

Set $\partial_t^n u := u^{(n)} := w$ and $\partial_t u := u^{(1)}$, $\partial_{tt} u := u^{(2)}$.

Note that the integrals in (4) exist since $\left\|\varphi^{(n)}(t)u(t)\right\|_X \le c\|u(t)\|_X$ for all $t \in (0,T)$ with a constant $c > 0$.

Lemma 1.4. *The following properties of the generalised derivative hold true:*

- *The generalised derivative is unique in* $L^1(0,T;Y)$.
- *Let* $X = Y$ *and* $u \in C^n([0,T];X)$ *for* $n \in \mathbb{N}$. *Then, the classical derivative and the generalised derivative coincide on* $(0,T)$.

Proof. See [13, Chapter 23]. $\square$

1.2 Existence and Uniqueness

In the variational formulation (2), ∂_{tt} is the distributional derivative on $(0,T)$, i.e. equality (2) means that

$$\int_0^T \langle u(\cdot,t), v\rangle_{L^2(\Omega)} \frac{\mathrm{d}^2\varphi}{\mathrm{d}t^2}(t)\mathrm{d}t + \int_0^T \langle \nabla_x u(\cdot,t), \nabla_x v\rangle_{L^2(\Omega)}\varphi(t)\mathrm{d}t = \int_0^T \langle f(\cdot,t), v\rangle_{L^2(\Omega)}\varphi(t)\mathrm{d}t$$

for all $\varphi \in C_0^\infty(0,T)$, where the equalities

$$\int_0^T \langle u(\cdot,t), v\rangle_{L^2(\Omega)} \frac{\mathrm{d}^2\varphi}{\mathrm{d}t^2}(t)\mathrm{d}t = \int_0^T \frac{\mathrm{d}^2}{\mathrm{d}t^2}\left[\langle u(\cdot,t), v\rangle_{L^2(\Omega)}\right]\varphi(t)\mathrm{d}t = \int_0^T \langle \partial_{tt}u(\cdot,t), v\rangle_\Omega \varphi(t)\mathrm{d}t$$

are used, see [13, Proposition 23.20]. The variational formulation in (2) is examined in many books, for example, [5, Théorème 8.1, Chapitre 3, page 287, and Théorème 8.2, Chapitre 3, page 296], [4, Theorem 4.2, Chapter IV, page 167], [11, Satz 29.1, Kapitel V, page 422], [13, Section 24.1, Chapter 24, page 453], [2, Mathematical Example 1, Chapter XVIII, page 581] or [9, Theorem 12.4, page 227]. In these books, the following existence and uniqueness result is proven.

Theorem 1.5. *For given* $f \in L^2(0,T;L^2(\Omega))$, $u_0 \in H_0^1(\Omega)$ *and* $u_1 \in L^2(\Omega)$, *a unique solution* u *of the variational formulation* (2) *exists. This solution* u *satisfies*

$$\begin{aligned} u &\in L^2(0,T;H_0^1(\Omega)) \cap C([0,T];H_0^1(\Omega)), \\ \partial_t u &\in L^2(0,T;L^2(\Omega)) \cap C([0,T];L^2(\Omega)), \\ \partial_{tt} u &\in L^2(0,T;[H_0^1(\Omega)]'), \end{aligned}$$

the stability estimate

$$\sqrt{\|u\|^2_{L^2(0,T;H_0^1(\Omega))} + \|\partial_t u\|^2_{L^2(0,T;L^2(\Omega))}} \le c\left(|u_0|_{H^1(\Omega)} + \|u_1\|_{L^2(\Omega)} + \|f\|_{L^2(0,T;L^2(\Omega))}\right)$$

with a constant $c > 0$, *and the energy equality*

$$\mathcal{E}(t_2) - \mathcal{E}(t_1) = \int_{t_1}^{t_2} \langle f(\cdot,s), \partial_t u(\cdot,s)\rangle_{L^2(\Omega)}\mathrm{d}s \qquad \text{with } 0 \le t_1 < t_2 \le T,$$

where the total energy of the solution u *is given by*

$$\mathcal{E}(t) := \frac{1}{2}\|\partial_t u(\cdot,t)\|^2_{L^2(\Omega)} + \frac{1}{2}\|\nabla_x u(\cdot,t)\|^2_{L^2(\Omega)}, \quad t \in [0,T].$$

Proof. For the existence and uniqueness, see the books [5, 11, 13, 2, 4, 9] as mentioned above. For the regularity and energy equality, see

- [9, Korollar 12.6, page 231],
- [5, Lemme 8.3, page 298],
- [2, Lemma 7, page 578].

□

Corollary 1.6. *For $f = 0$, the energy equality is*

$$\mathcal{E}(t) \stackrel{!}{=} \mathcal{E}(0) = \frac{1}{2}\|u_1\|^2_{L^2(\Omega)} + \frac{1}{2}\|\nabla_x u_0\|^2_{L^2(\Omega)} \quad \textit{for all } t \in (0, T],$$

i.e. a conservation of the total energy holds true.

In the remainder of the lecture, write

$$v(t) := v(\cdot, t) \quad \text{for } t \in [0, T],$$

i.e.

$$v(t)(x) = v(x, t) \quad \text{for } (x, t) \in \Omega \times [0, T],$$

where $v\colon \Omega \times [0, T] \to \mathbb{R}$ is a given function.

2 Classical Methods: Leapfrog Method and Crank-Nicolson Method

In the standard approach, the discretisation schemes for time-dependent partial differential equations, e.g., the wave equation, are based on semi-discretisations:

- <u>Method of lines:</u>
  1. discretisation in space
  2. discretisation in time

  After the first discretisation, a system of ordinary differential equations has to be solved.

- <u>Rothe's method:</u>
  1. discretisation in time
  2. discretisation in space

  After the first discretisation, a boundary value problem has to be solved for every time step.

In this lecture, the method of lines is used, i.e. the first discretisation is done with respect to the spatial variable.

2.1 Finite Element Spaces for the Spatial Variable

Let the bounded Lipschitz domain $\Omega \subset \mathbb{R}^d$ be an interval $\Omega = (0, L)$ for $d = 1$, or polygonal for $d = 2$, or polyhedral for $d = 3$. For this situation, discretisations in space are introduced as follows. The spatial domain Ω is decomposed as

$$\overline{\Omega} = \bigcup_{\ell=1}^{N} \overline{\omega}_\ell$$

with N spatial elements $\omega_\ell \subset \mathbb{R}^d$, i.e.

$$\mathcal{T} := \{\omega_\ell\}_{\ell=1}^{N}$$

is an admissible decomposition or mesh of Ω. Here, the spatial elements ω_ℓ are intervals for $d = 1$, triangles for $d = 2$ and tetrahedra for $d = 3$. The local mesh sizes are given as

$$h_\ell := \left(\int_{\omega_\ell} \mathrm{d}x \right)^{1/d} \qquad \text{for } \ell = 1, \ldots, N.$$

In addition,

$$h := h_{\max}(\mathcal{T}) := \max_{\ell=1,\ldots,N} h_\ell \qquad \text{and} \qquad h_{\min} := h_{\min}(\mathcal{T}) := \min_{\ell=1,\ldots,N} h_\ell$$

are the global mesh size and minimal local mesh size. Furthermore, $\tilde{M}$ is the number of vertices $\{x_i\}_{i=1}^{\tilde{M}}$ of the decomposition $\mathcal{T}$. In the following, a sequence

$$(\mathcal{T}_\nu)_{\nu\in\mathbb{N}} := \{\mathcal{T}_\nu \colon \ \nu \in \mathbb{N}\}$$

of decompositions of Ω is considered. The sequence $(\mathcal{T}_\nu)_{\nu\in\mathbb{N}}$ of decompositions of Ω is called *shape-regular*, if a constant $c_F > 0$ exists such that

$$\forall \nu \in \mathbb{N} \colon \ \forall \omega \in \mathcal{T}_\nu \colon \qquad \sup_{x,y\in\omega} |x - y| \leq c_F \, \mathrm{r}_\omega,$$

where r_ω is the radius of the largest ball that can be inscribed in the spatial element $\omega \in \mathcal{T}_\nu$. The sequence $(\mathcal{T}_\nu)_{\nu\in\mathbb{N}}$ of decompositions of Ω is called *globally quasi-uniform*, if a constant $c_G \geq 1$ exists such that

$$\forall \nu \in \mathbb{N} \colon \quad \frac{h_{\max}(\mathcal{T}_\nu)}{h_{\min}(\mathcal{T}_\nu)} \leq c_G.$$

Remark 2.1. *In the literature, a shape-regular sequence of decomposition is also called regular or quasi-uniform. In that case, a globally quasi-uniform sequence is called uniform. See, e.g., [1, Definition 9.26 and Definition 9.33] and [10, page 205].*

In the whole work, the sequence $(\mathcal{T}_\nu)_{\nu\in\mathbb{N}}$ of decompositions of Ω is assumed to be admissible and shape-regular.

The space

$$S_h^1(\Omega) = \operatorname{span}\{\psi_i\}_{i=1}^{\tilde{M}} \subset H^1(\Omega)$$

is the space of piecewise linear, continuous functions on intervals ($d = 1$), triangles ($d = 2$), tetrahedra ($d = 3$), where the functions ψ_i are the usual nodal basis functions satisfying $\psi_i(x_k) = \delta_{ik}$ for $i, k = 1, \ldots, \tilde{M}$. In addition, the subspace $S_{h,0}^1(\Omega) \subset S_h^1(\Omega)$ satisfies the homogeneous Dirichlet boundary condition, i.e.

$$S_{h,0}^1(\Omega) = S_h^1(\Omega) \cap H_0^1(\Omega).$$

After an ordering of the vertices $\{x_i\}_{i=1}^{\tilde{M}}$ in interior vertices $\{x_i\}_{i=1}^{M} \subset \Omega$ and boundary vertices $\{x_i\}_{i=M+1}^{\tilde{M}} \subset \partial\Omega$, this $H_0^1(\Omega)$ conforming subspace is written as

$$S_{h,0}^1(\Omega) = \operatorname{span}\{\psi_i\}_{i=1}^{M}.$$

A function $U_h \in S_{h,0}^1(\Omega)$ admits the representation

$$U_h(x) = \sum_{i=1}^{M} U_i \psi_i(x)$$

for $x \in \overline{\Omega}$ with the coefficients $U_i \in \mathbb{R}$. In the remainder of this work, $M_h \in \mathbb{R}^{M\times M}$ and $A_h \in \mathbb{R}^{M\times M}$ denote mass and stiffness matrices defined via

$$M_h[i,j] = \langle \psi_j, \psi_i \rangle_{L^2(\Omega)} \tag{5}$$

for $i, j = 1, \ldots, M$, and

$$A_h[i,j] = \langle \nabla_x \psi_j, \nabla_x \psi_i \rangle_{L^2(\Omega)} \tag{6}$$

for $i, j = 1, \ldots, M$.

Lemma 2.2 (Inverse Inequality)**.** *Let the sequence $(\mathcal{T}_\nu)_{\nu\in\mathbb{N}}$ of decompositions of Ω be admissible, shape-regular and globally quasi-uniform. Then, the inverse inequality*

$$\forall v_h \in S_h^1(\Omega): \quad |v_h|_{H^1(\Omega)}^2 \leq c_{\mathrm{inv}}\, h^{-2} \|v_h\|_{L^2(\Omega)}^2$$

holds true with a constant $c_{\mathrm{inv}} > 0$.

Proof. See previous courses or [10, Lemma 9.8, page 217]. □

Definition 2.3 (Ritz Projection)**.** *For a symmetric, continuous and elliptic bilinear form*

$$a(\cdot,\cdot)\colon H_0^1(\Omega) \times H_0^1(\Omega) \to \mathbb{R},$$

the Ritz projection $R_h\colon H_0^1(\Omega) \to S_{h,0}^1(\Omega)$ is defined as unique solution of the variational formulation to find $R_h v \in S_{h,0}^1(\Omega)$ such that

$$\forall w_h \in S_{h,0}^1(\Omega): \quad a(R_h v, w_h) = a(v, w_h)$$

for a function $v \in H_0^1(\Omega)$.

Note that the Ritz projection is the orthogonal projection of $H_0^1(\Omega)$ onto $S_{h,0}^1(\Omega)$ with respect to the energy inner product $a(\cdot,\cdot)$. In this lecture, the special case

$$a(v,w) := \langle \nabla_x v, \nabla_x w \rangle_{L^2(\Omega)}, \quad v, w \in H_0^1(\Omega),$$

is considered, i.e. the Ritz projection R_h is the $H_0^1(\Omega)$ projection $Q_h^1 \colon H_0^1(\Omega) \to S_{h,0}^1(\Omega)$ defined by

$$\forall w_h \in S_{h,0}^1(\Omega)\colon \quad \left\langle \nabla_x Q_h^1 v, \nabla_x w_h \right\rangle_{L^2(\Omega)} = \langle \nabla_x v, \nabla_x w_h \rangle_{L^2(\Omega)} \tag{7}$$

for a function $v \in H_0^1(\Omega)$.

Lemma 2.4. *Let the sequence $(\mathcal{T}_\nu)_{\nu\in\mathbb{N}}$ of decompositions of Ω be admissible and shape-regular. Then,*

- *the stability estimate*

$$\forall v \in H_0^1(\Omega)\colon \quad \left|Q_h^1 v\right|_{H^1(\Omega)} \le |v|_{H^1(\Omega)},$$

- *the $H^1(\Omega)$ error estimate*

$$\forall v \in H_0^1(\Omega) \cap H^s(\Omega)\colon \quad \left|Q_h^1 v - v\right|_{H^1(\Omega)} \le c\, h^{s-1} \|v\|_{H^s(\Omega)}$$

  for $s \in [1,2]$ and a constant $c > 0$, and

- *if Ω is sufficiently regular, e.g., Ω convex or $\partial\Omega$ smooth, the $L^2(\Omega)$ error estimate*

$$\forall v \in H_0^1(\Omega) \cap H^s(\Omega)\colon \quad \left\|Q_h^1 v - v\right\|_{L^2(\Omega)} \le c\, h^{s} \|v\|_{H^s(\Omega)}$$

  for $s \in [1,2]$ and a constant $c > 0$

hold true.

Proof. See previous courses, i.e. Céa's lemma, Aubin-Nitsche duality argument. □

Lemma 2.5. *Let $v \in C^m([0,T]; H_0^1(\Omega))$ be a given function for some $m \in \mathbb{N}_0$. Then, the relations*

$$t \mapsto Q_h^1 v(\cdot,t) \in C^m([0,T]; S_{h,0}^1(\Omega))$$

and

$$\partial_t^j Q_h^1 v = Q_h^1 \partial_t^j v \quad \textit{for } j = 1,\dots,m$$

hold true.

Proof. Use an orthonormal basis of $S_{h,0}^1(\Omega)$ and the definition of the derivative. □

2.2 Finite Difference Method for the Temporal Variable

Let $N_t \in \mathbb{N}$, $N_t \geq 2$, be the time discretisation parameter. Set the constant time step size

$$\tau = \frac{T}{N_t} \in \left(0, \frac{T}{2}\right]$$

and the time steps

$$t^k := k\tau \quad \text{for } k = 0, \dots, N_t.$$

Definition 2.6 (Finite Difference Operators)**.** *Let the constant time step size $\tau \in \left(0, \frac{T}{2}\right]$ and the corresponding time steps t^k be given.*

- *For a function $v\colon \Omega \times [0, T] \to \mathbb{R}$, define*
  - *the continuous first-order forward operator*

$$\mathrm{D}_\tau^+ v(x,t) := \frac{v(x, t+\tau) - v(x,t)}{\tau}, \quad x \in \Omega,\ t \in [0, T-\tau],$$

  - *the continuous first-order backward operator*

$$\mathrm{D}_\tau^- v(x,t) := \frac{v(x,t) - v(x, t-\tau)}{\tau}, \quad x \in \Omega,\ t \in [\tau, T],$$

  - *the continuous second-order central difference operator*

$$\mathrm{D}_\tau^2 v(x,t) := \frac{v(x, t+\tau) - 2v(x,t) + v(x, t-\tau)}{\tau^2}, \quad x \in \Omega,\ t \in [\tau, T-\tau],$$

  - *the function $v^k\colon \Omega \to \mathbb{R}$ by*

$$v^k(x) := v(x, t^k), \quad x \in \Omega,\ k \in \{0, 1, \dots, N_t\}. \tag{8}$$

- *For a sequence of functions $V^k\colon \Omega \to \mathbb{R}$, $k = 0, \dots, N_t$, define*
  - *the discrete first-order forward operator*

$$\mathrm{D}_\tau^+ V^k(x) := \frac{V^{k+1}(x) - V^k(x)}{\tau}, \quad x \in \Omega,\ k \in \{0, 1, \dots, N_t - 1\},$$

  - *the discrete first-order backward operator*

$$\mathrm{D}_\tau^- V^k(x) := \frac{V^k(x) - V^{k-1}(x)}{\tau}, \quad x \in \Omega,\ k \in \{1, \dots, N_t\},$$

  - *the discrete second-order central difference operator*

$$\mathrm{D}_\tau^2 V^k(x) := \frac{V^{k+1}(x) - 2V^k(x) + V^{k-1}(x)}{\tau^2}, \quad x \in \Omega,\ k \in \{1, 2, \dots, N_t - 1\}.$$

Note that by using the notation (8), the equalities

$$\begin{aligned} \mathrm{D}_\tau^+ v(x,t^k) &= (\mathrm{D}_\tau^+ v)^k(x) = \mathrm{D}_\tau^+ v^k(x), && x\in\Omega,\ k\in\{0,1,\dots,N_t-1\},\\ \mathrm{D}_\tau^- v(x,t^k) &= (\mathrm{D}_\tau^- v)^k(x) = \mathrm{D}_\tau^- v^k(x), && x\in\Omega,\ k\in\{1,\dots,N_t\}, \end{aligned}$$

and

$$\mathrm{D}_\tau^2 v(x,t^k) = (\mathrm{D}_\tau^2 v)^k(x) = \mathrm{D}_\tau^2 v^k(x), \quad x\in\Omega,\ k\in\{1,2,\dots,N_t-1\},$$

hold true for a function $v\colon \Omega\times[0,T]\to\mathbb{R}$ and the corresponding sequence $v^0, v^1, \dots, v^{N_t}$ defined by $v^k(\cdot) = v(\cdot, t^k)$, $k = 0,1,\dots,N_t$.

Lemma 2.7. *Let the constant time step size $\tau \in \left(0, \frac{T}{2}\right]$ and a function $v \in C^2([0,T]; L^2(\Omega))$ be given. Then, for $x\in\Omega$ and $t\in[0,T-\tau]$, a number $\Theta(x,t,\tau)\in(0,1)$ exists such that*

$$\partial_t v(x,t) = \mathrm{D}_\tau^+ v(x,t) - \frac{\tau}{2}\partial_{tt} v(x, t+\tau\,\Theta(x,t,\tau)).$$

Analogously, for $x\in\Omega$ and $t\in[\tau,T]$, a number $\Theta(x,t,\tau)\in(0,1)$ exists such that

$$\partial_t v(x,t) = \mathrm{D}_\tau^- v(x,t) + \frac{\tau}{2}\partial_{tt} v(x, t-\tau\,\Theta(x,t,\tau)).$$

Proof. Frist, for fixed $x\in\Omega$ and $t\in[0,T-\tau]$, Taylor's theorem gives

$$v(x,t+\tau) = v(x,t) + \partial_t v(x,t)\,\tau + \frac{\partial_{tt} v(x,t+\tau\Theta(x,t,\tau))}{2}\tau^2$$

with $\Theta(x,t,\tau)\in(0,1)$ depending on x, t and τ. Hence, with

$$\mathrm{D}_\tau^+ v(x,t) = \frac{v(x,t+\tau)-v(x,t)}{\tau} = \partial_t v(x,t) + \frac{\partial_{tt} v(x,t+\tau\Theta(x,t,\tau))}{2}\tau,$$

the first assertion follows.

Second, for fixed $x\in\Omega$ and $t\in[\tau,T]$, Taylor's theorem gives

$$v(x,t-\tau) = v(x,t) - \partial_t v(x,t)\,\tau + \frac{\partial_{tt} v(x,t-\tau\Theta(x,t,\tau))}{2}\tau^2$$

with $\Theta(x,t,\tau)\in(0,1)$ depending on x, t and τ. Hence, with

$$\mathrm{D}_\tau^- v(x,t) = \frac{v(x,t)-v(x,t-\tau)}{\tau} = \partial_t v(x,t) - \frac{\partial_{tt} v(x,t-\tau\Theta(x,t,\tau))}{2}\tau,$$

the second assertion follows. □

Lemma 2.8. *Let the constant time step size $\tau \in \left(0, \frac{T}{2}\right]$ and a function $v \in C^4([0,T]; L^2(\Omega))$ be given. Then, for $x\in\Omega$ and $t\in[\tau,T-\tau]$, a number $\Theta(x,t,\tau)\in[-1,1]$ exists such that*

$$\partial_{tt} v(x,t) = \mathrm{D}_\tau^2 v(x,t) - \frac{\tau^2}{12}\partial_t^4 v(x, t+\tau\,\Theta(x,t,\tau)).$$

Proof. For fixed $x \in \Omega$ and $t \in [\tau, T-\tau]$, Taylor's theorem gives

$$v(x,t+\tau) = v(x,t) + \partial_t v(x,t)\,\tau + \frac{\partial_{tt} v(x,t)}{2}\tau^2 + \frac{\partial_t^3 v(x,t)}{6}\tau^3 + \frac{\partial_t^4 v(x,t+\tau\hat{\Theta})}{24}\tau^4,$$
$$v(x,t-\tau) = v(x,t) - \partial_t v(x,t)\,\tau + \frac{\partial_{tt} v(x,t)}{2}\tau^2 - \frac{\partial_t^3 v(x,t)}{6}\tau^3 + \frac{\partial_t^4 v(x,t-\tau\tilde{\Theta})}{24}\tau^4$$

with $\hat{\Theta}, \tilde{\Theta} \in (0,1)$ depending on x, t and τ. Adding these equations up yields

$$\partial_{tt} v(x,t)\tau^2 = \underbrace{v(x,t+\tau) - 2v(x,t) + v(x,t-\tau)}_{=\tau^2 \mathrm{D}_\tau^2 v(x,t)} - \frac{\tau^4}{24}[\partial_t^4 v(x,t+\tau\hat{\Theta}) + \partial_t^4 v(x,t-\tau\tilde{\Theta})].$$

By the intermediate value theorem, $s \mapsto \partial_t^4 v(x,t+\tau s)$ must assume every value between $\partial_t^4 v(x,t+\tau\hat{\Theta})$ and $\partial_t^4 v(x,t-\tau\tilde{\Theta})$ on the interval $(-\tilde{\Theta}, \hat{\Theta}) \subset (-1,1)$, including also the average of these two numbers, i.e. an element $\Theta(x,t,\tau) \in [-1,1]$ exists such that

$$\partial_t^4 v(x,t+\tau\Theta(x,t,\tau)) = \frac{\partial_t^4 v(x,t+\tau\hat{\Theta}) + \partial_t^4 v(x,t-\tau\tilde{\Theta})}{2}.$$

Hence, the assertion follows. □

2.3 Leapfrog Method

This subsection is based on [1, Subsection 9.5.2] and [6, Chapter 3]. In the remainder of this subsection, let $f \in C([0,T]; L^2(\Omega))$ be a given right-hand side of (1), let the sequence $(\mathcal{T}_\nu)_{\nu\in\mathbb{N}}$ of decompositions of Ω be admissible, shape-regular and globally quasi-uniform, and let the constant time step size $\tau \in \left(0, \frac{T}{2}\right]$ and the corresponding time steps t^k be given. Introduce for $x \in \Omega$ and $k \in \{0, 1, \dots, N_t\}$ the approximations

$$U^k(x) := \sum_{i=1}^{M} \alpha_i^k \psi_i(x) \approx u(x,t^k), \tag{9}$$

where $\alpha_i^k \in \mathbb{R}$ are the unknown coefficients of $U^k \in S_{h,0}^1(\Omega)$. For $k = 1, \dots, N_t - 1$, using the approximation (9) in the variational formulation (2) together with the discrete second-order central difference operator $\mathrm{D}_\tau^2 U^k(x) \approx \partial_{tt} u(x,t^k)$, $x \in \Omega$, gives the leapfrog method.

Leapfrog Method

Find functions $U^k \in S_{h,0}^1(\Omega)$ for $k = 0, 1, \dots, N_t$ such that

$$U^0 = u_{0,h}, \quad U^1 = u_{1,h} \tag{10}$$

and for $k = 1, 2, \dots, N_t - 1$,

$$\forall v_h \in S^1_{h,0}(\Omega): \quad \underbrace{\frac{1}{\tau^2}\left\langle U^{k+1} - 2U^k + U^{k-1}, v_h \right\rangle_{L^2(\Omega)}}_{=\left\langle \mathrm{D}^2_\tau U^k, v_h \right\rangle_{L^2(\Omega)}} + \left\langle \nabla_x U^k, \nabla_x v_h \right\rangle_{L^2(\Omega)} = \left\langle f^k, v_h \right\rangle_{L^2(\Omega)} \tag{11}$$

with $f^k(x) := f(x, t^k)$ for $x \in \Omega$, where

$$u(\cdot, 0) = u_0 \approx u_{0,h} \in S^1_{h,0}(\Omega) \quad \text{and } u(\cdot, t^1) \approx u_{1,h} \in S^1_{h,0}(\Omega)$$

are approximations using the initial conditions u_0, u_1 and the right-hand side f, which will be defined later. In the following, it is shown that the leapfrog method can be realised as a two-step method. With (9) and $v_h = \psi_j$, $j = 1, \dots, M$, in (11), the discrete variational formulation (11) is equivalent to solving the linear system

$$\frac{1}{\tau^2} M_h(\underline{\alpha}^{k+1} - 2\underline{\alpha}^k + \underline{\alpha}^{k-1}) + A_h \underline{\alpha}^k = \begin{pmatrix} \left\langle f^k, \psi_1 \right\rangle_{L^2(\Omega)} \\ \left\langle f^k, \psi_2 \right\rangle_{L^2(\Omega)} \\ \vdots \\ \left\langle f^k, \psi_M \right\rangle_{L^2(\Omega)} \end{pmatrix} =: \underline{f}^k, \quad k = 1, \dots, N_t - 1, \tag{12}$$

where the unknown coefficient vectors are given by

$$\underline{\alpha}^k := \begin{pmatrix} \alpha^k_1 \\ \alpha^k_2 \\ \vdots \\ \alpha^k_M \end{pmatrix} \in \mathbb{R}^M, \quad k = 0, \dots, N_t.$$

Here, M_h and A_h are the mass and stiffness matrices given in (5) and (6). Rewriting the linear system (12) gives

$$M_h \underline{\alpha}^{k+1} = \tau^2(\underline{f}^k - A_h \underline{\alpha}^k) + M_h(2\underline{\alpha}^k - \underline{\alpha}^{k-1}), \quad k = 1, \dots, N_t - 1, \tag{13}$$

i.e. in every time step, a linear system with the mass matrix has to solved for $\underline{\alpha}^{k+1}$, when $\underline{\alpha}^k$, $\underline{\alpha}^{k-1}$ are already computed. Since the mass matrix is positive definite, the linear system (13) is uniquely solvable. Hence, the discrete variational formulation (11) has a unique solution $U^{k+1} \in S^1_{h,0}(\Omega)$ for $k = 1, \dots, N_t - 1$, when U^k, U^{k-1} are known.

In the following, the leapfrog method (10), (11) is analysed. Therefore,

- an energy conservation,
- the stability and
- error estimates

are proven.

Definition 2.9. *Let $U^k \in S^1_{h,0}(\Omega)$ be the solution of* (10), (11) *for the time step t^k, $k = 0, 1, \ldots, N_t$. The discrete energy of the leapfrog method* (10), (11) *is defined by*

$$\mathcal{E}^k_{\mathrm{LF}} := \frac{1}{2}\left\|\mathrm{D}^+_\tau U^k\right\|^2_{L^2(\Omega)} + \frac{1}{2}\left\langle \nabla_x U^k, \nabla_x U^{k+1}\right\rangle_{L^2(\Omega)}$$

for $k = 0, 1, \ldots, N_t - 1$.

Note that $\mathcal{E}^k_{\mathrm{LF}}$ is in general not a non-negative quantity. For the stability analysis, it is desirable to determine a condition for which the discrete energy is non-negative. This leads to the so-called CFL condition (Courant, Friedrichs, Lewy).

Lemma 2.10. *Let the sequence $(\mathcal{T}_\nu)_{\nu\in\mathbb{N}}$ of decompositions of Ω be admissible, shape-regular and globally quasi-uniform, and let the constant time step size $\tau \in \left(0, \frac{T}{2}\right]$ be given. Assume that $\lambda \in (0, 2)$ exists such that the CFL condition*

$$c_{\mathrm{inv}} \frac{\tau^2}{h^2} \leq 2 - \lambda \tag{14}$$

is satisfied, where $c_{\mathrm{inv}} > 0$ is the constant of the inverse inequality (Lemma 2.2). Then, the inequality

$$\mathcal{E}^k_{\mathrm{LF}} \geq \frac{\lambda}{4}\left\|\mathrm{D}^+_\tau U^k\right\|^2_{L^2(\Omega)} + \frac{1}{4}\left|U^{k+1}\right|^2_{H^1(\Omega)} + \frac{1}{4}\left|U^k\right|^2_{H^1(\Omega)}$$

holds true for $k = 0, \ldots, N_t - 1$.

Proof. Let $k \in \{0, \ldots, N_t - 1\}$ be fixed. The second part of $\mathcal{E}^k_{\mathrm{LF}}$ fulfils

$$\begin{aligned}
2\left\langle \nabla_x U^k, \nabla_x U^{k+1}\right\rangle_{L^2(\Omega)} &= \left\langle \nabla_x U^{k+1}, \nabla_x U^{k+1}\right\rangle_{L^2(\Omega)} + \left\langle \nabla_x U^k, \nabla_x U^k\right\rangle_{L^2(\Omega)} \\
&\quad - \left\langle \nabla_x U^{k+1} - \nabla_x U^k, \nabla_x U^{k+1} - \nabla_x U^k\right\rangle_{L^2(\Omega)} \\
&= \left|U^{k+1}\right|^2_{H^1(\Omega)} + \left|U^k\right|^2_{H^1(\Omega)} - \tau^2\left|\mathrm{D}^+_\tau U^k\right|^2_{H^1(\Omega)} \\
&\geq \left|U^{k+1}\right|^2_{H^1(\Omega)} + \left|U^k\right|^2_{H^1(\Omega)} - c_{\mathrm{inv}}\frac{\tau^2}{h^2}\left\|\mathrm{D}^+_\tau U^k\right\|^2_{L^2(\Omega)},
\end{aligned}$$

where the inverse inequality (Lemma 2.2) is used since $\mathrm{D}^+_\tau U^k \in S^1_h(\Omega)$. Inserting the last inequality in the definition of $\mathcal{E}^k_{\mathrm{LF}}$ gives

$$\mathcal{E}^k_{\mathrm{LF}} \geq \frac{1}{2}\underbrace{\left(1 - c_{\mathrm{inv}}\frac{\tau^2}{2h^2}\right)}_{\geq 1 - \frac{1}{2}(2-\lambda) = \frac{\lambda}{2}}\left\|\mathrm{D}^+_\tau U^k\right\|^2_{L^2(\Omega)} + \frac{1}{4}\left|U^{k+1}\right|^2_{H^1(\Omega)} + \frac{1}{4}\left|U^k\right|^2_{H^1(\Omega)}$$

and hence, the assertion follows. □

Remark 2.11. *In the situation of Lemma 2.10, using the inequality*

$$2\left\langle \nabla_x U^k, \nabla_x U^{k+1} \right\rangle_{L^2(\Omega)} = 2\left|U^{k+1/2}\right|_{H^1(\Omega)}^2 - \frac{\tau^2}{2}\left|\mathrm{D}_\tau^+ U^k\right|_{H^1(\Omega)}^2$$
$$\geq 2\left|U^{k+1/2}\right|_{H^1(\Omega)}^2 - c_{\mathrm{inv}}\frac{\tau^2}{2h^2}\left\|\mathrm{D}_\tau^+ U^k\right\|_{L^2(\Omega)}^2$$

for $k = 0, \ldots, N_t - 1$ *in the proof of Lemma 2.10 with the average*

$$U^{k+1/2} := \frac{1}{2}\left(U^{k+1} + U^k\right)$$

leads to the inequality

$$\mathcal{E}_{\mathrm{LF}}^k \geq \frac{\widetilde{\lambda}}{8}\left\|\mathrm{D}_\tau^+ U^k\right\|_{L^2(\Omega)}^2 + \frac{1}{2}\left|U^{k+1/2}\right|_{H^1(\Omega)}^2$$

for $k = 0, \ldots, N_t - 1$, *if a constant* $\widetilde{\lambda} \in (0, 4)$ *exists such that the weaker CFL condition*

$$c_{\mathrm{inv}}\frac{\tau^2}{h^2} \leq 4 - \widetilde{\lambda} \tag{15}$$

is satisfied. Note that

$$\left|U^{k+1/2}\right|_{H^1(\Omega)} = \left|\frac{U^{k+1} + U^k}{2}\right|_{H^1(\Omega)} \leq \frac{1}{2}\left|U^{k+1}\right|_{H^1(\Omega)} + \frac{1}{2}\left|U^k\right|_{H^1(\Omega)}$$

for $k = 0, \ldots, N_t - 1$.

Lemma 2.10 states that the discrete $\mathcal{E}_{\mathrm{LF}}^k$ is non-negative, providing the CFL condition (14). Next, an energy conservation is proven.

Theorem 2.12 (Energy Conservation of the Leapfrog Method)**.** *Let* $f \in C([0, T]; L^2(\Omega))$ *be a given right-hand side of* (1), *let the sequence* $(\mathcal{T}_\nu)_{\nu\in\mathbb{N}}$ *of decompositions of* Ω *be admissible, shape-regular and globally quasi-uniform, and let the constant time step size* $\tau \in \left(0, \frac{T}{2}\right]$ *be given. Assume that* $\lambda \in (0, 2)$ *exists such that the CFL condition* (14) *holds true. Then,*

- *for the right-hand side* $f = 0$, *the leapfrog method* (10), (11) *conserves energy in the sense that*
$$\mathcal{E}_{\mathrm{LF}}^k = \mathcal{E}_{\mathrm{LF}}^0 \qquad \text{for } k = 1, \ldots, N_t - 1,$$
- *for the right-hand sides* $f \neq 0$, *the relation*
$$\sqrt{\mathcal{E}_{\mathrm{LF}}^k} \leq \sqrt{\mathcal{E}_{\mathrm{LF}}^0} + \sum_{j=1}^{k} \frac{\tau}{\sqrt{\lambda}}\left\|f^j\right\|_{L^2(\Omega)} \qquad \text{for } k = 1, \ldots, N_t - 1$$

  holds true, where $f^j(x) = f(x, t^j)$, $x \in \Omega$.

Proof. Let $k \in \{1, \ldots, N_t - 1\}$ be fixed. For $j \in \{1, \ldots, k\}$, choosing

$$v_h = U^{j+1} - U^{j-1} = (U^{j+1} - U^j) + (U^j - U^{j-1}) \in S^1_{h,0}(\Omega)$$

in the variational formulation (11) yields

$$\begin{aligned}
&\langle f^j, U^{j+1} - U^{j-1} \rangle_{L^2(\Omega)} \\
&\qquad = \frac{1}{\tau^2} \langle U^{j+1} - 2U^j + U^{j-1}, U^{j+1} - U^{j-1} \rangle_{L^2(\Omega)} \\
&\qquad\quad + \langle \nabla_x U^j, \nabla_x U^{j+1} - \nabla_x U^{j-1} \rangle_{L^2(\Omega)} \\
&\qquad = \frac{1}{\tau^2} \langle (U^{j+1} - U^j) - (U^j - U^{j-1}), (U^{j+1} - U^j) + (U^j - U^{j-1}) \rangle_{L^2(\Omega)} \\
&\qquad\quad + \langle \nabla_x U^j, \nabla_x U^{j+1} - \nabla_x U^{j-1} \rangle_{L^2(\Omega)} \\
&\qquad = \left\| \frac{U^{j+1} - U^j}{\tau} \right\|^2_{L^2(\Omega)} - \left\| \frac{U^j - U^{j-1}}{\tau} \right\|^2_{L^2(\Omega)} \\
&\qquad\quad + \langle \nabla_x U^j, \nabla_x U^{j+1} \rangle_{L^2(\Omega)} - \langle \nabla_x U^j, \nabla_x U^{j-1} \rangle_{L^2(\Omega)} \\
&\qquad = 2(\mathcal{E}^j_{\mathrm{LF}} - \mathcal{E}^{j-1}_{\mathrm{LF}}).
\end{aligned} \tag{16}$$

Hence, in the case $f = 0$, the equality

$$\mathcal{E}^k_{\mathrm{LF}} = \mathcal{E}^0_{\mathrm{LF}} \qquad \text{for } k = 1, \ldots, N_t - 1$$

follows by induction on j from (16).

In the case $f \neq 0$, the relation (16) gives

$$\begin{aligned}
2(\mathcal{E}^j_{\mathrm{LF}} - \mathcal{E}^{j-1}_{\mathrm{LF}}) &= \langle f^j, U^{j+1} - U^{j-1} \rangle_{L^2(\Omega)} \\
&= \langle f^j, U^{j+1} - U^j \rangle_{L^2(\Omega)} + \langle f^j, U^j - U^{j-1} \rangle_{L^2(\Omega)} \\
&= \tau \langle f^j, \mathrm{D}^+_\tau U^j \rangle_{L^2(\Omega)} + \tau \langle f^j, \mathrm{D}^+_\tau U^{j-1} \rangle_{L^2(\Omega)} \\
&\leq \tau \| f^j \|_{L^2(\Omega)} \left(\| \mathrm{D}^+_\tau U^j \|_{L^2(\Omega)} + \| \mathrm{D}^+_\tau U^{j-1} \|_{L^2(\Omega)} \right) \\
&\leq \tau \| f^j \|_{L^2(\Omega)} \left(\sqrt{\frac{4}{\lambda} \mathcal{E}^j_{\mathrm{LF}}} + \sqrt{\frac{4}{\lambda} \mathcal{E}^{j-1}_{\mathrm{LF}}} \right),
\end{aligned} \tag{17}$$

where the Cauchy-Schwarz inequality and Lemma 2.10, i.e the CFL condition (14), are used. If $\mathcal{E}^j_{\mathrm{LF}} > 0$, then

$$\sqrt{\mathcal{E}^j_{\mathrm{LF}}} - \sqrt{\mathcal{E}^{j-1}_{\mathrm{LF}}} = \frac{\mathcal{E}^j_{\mathrm{LF}} - \mathcal{E}^{j-1}_{\mathrm{LF}}}{\sqrt{\mathcal{E}^j_{\mathrm{LF}}} + \sqrt{\mathcal{E}^{j-1}_{\mathrm{LF}}}} \overset{(17)}{\leq} \frac{\tau}{\sqrt{\lambda}} \| f^j \|_{L^2(\Omega)}. \tag{18}$$

If $\mathcal{E}^j_{\mathrm{LF}} = 0$, then the estimate (18) still holds true since $\mathcal{E}^{j-1}_{\mathrm{LF}} \geq 0$. Hence, summing (18) over $j = 1, \ldots, k$ gives

$$\sqrt{\mathcal{E}^k_{\mathrm{LF}}} - \sqrt{\mathcal{E}^0_{\mathrm{LF}}} = \sum_{j=1}^k \left(\sqrt{\mathcal{E}^j_{\mathrm{LF}}} - \sqrt{\mathcal{E}^{j-1}_{\mathrm{LF}}} \right) \leq \sum_{j=1}^k \frac{\tau}{\sqrt{\lambda}} \| f^j \|_{L^2(\Omega)},$$

where a telescoping sum occurs. $\square$

Remark 2.13. *An energy conservation analogous to Theorem 2.12 holds true, if the weaker CFL condition* (15) *is satisfied.*

Theorem 2.14 (Stability of the Leapfrog Method)**.** *Let $f \in C([0,T]; L^2(\Omega))$ be a given right-hand side of* (1)*, let the sequence $(\mathcal{T}_\nu)_{\nu\in\mathbb{N}}$ of decompositions of Ω be admissible, shape-regular and globally quasi-uniform, and let the constant time step size $\tau \in \left(0, \frac{T}{2}\right]$ be given. Assume that $\lambda \in (0,2)$ exists such that the CFL condition* (14) *holds true. Then, the leapfrog method* (10), (11) *is stable in the sense that*

$$\left\|\mathrm{D}_\tau^+ U^k\right\|_{L^2(\Omega)} + \left|U^{k+1}\right|_{H^1(\Omega)} \leq \left(\frac{1}{\sqrt{\lambda}} + 1\right) \left[\sqrt{2}\left\|\mathrm{D}_\tau^+ U^0\right\|_{L^2(\Omega)} + \left|U^0\right|_{H^1(\Omega)} + \left|U^1\right|_{H^1(\Omega)} + \sum_{j=1}^{k} \frac{2\tau}{\sqrt{\lambda}} \left\|f^j\right\|_{L^2(\Omega)}\right]$$

for $k = 1, \ldots, N_t - 1$.

Proof. Let $k \in \{1, \ldots, N_t - 1\}$ be fixed. With the inequality

$$\begin{aligned}
2\langle\nabla_x v, \nabla_x w\rangle_{L^2(\Omega)} &= \langle\nabla_x v, \nabla_x v\rangle_{L^2(\Omega)} + \langle\nabla_x w, \nabla_x w\rangle_{L^2(\Omega)} - \langle\nabla_x(v-w), \nabla_x(v-w)\rangle_{L^2(\Omega)} \\
&= |v|_{H^1(\Omega)}^2 + |w|_{H^1(\Omega)}^2 - |v-w|_{H^1(\Omega)}^2 \\
&\leq |v|_{H^1(\Omega)}^2 + |w|_{H^1(\Omega)}^2
\end{aligned}$$

for all $v, w \in H^1(\Omega)$, it follows that

$$\begin{aligned}
\mathcal{E}_{\mathrm{LF}}^0 &= \frac{1}{2}\left\|\mathrm{D}_\tau^+ U^0\right\|_{L^2(\Omega)}^2 + \frac{1}{2}\left\langle\nabla_x U^0, \nabla_x U^1\right\rangle_{L^2(\Omega)} \\
&\leq \frac{1}{2}\left\|\mathrm{D}_\tau^+ U^0\right\|_{L^2(\Omega)}^2 + \frac{1}{4}\left|U^0\right|_{H^1(\Omega)}^2 + \frac{1}{4}\left|U^1\right|_{H^1(\Omega)}^2 .
\end{aligned}$$

Using Lemma 2.10, the energy conservation (Theorem 2.12) and the last estimate yields

$$\begin{aligned}
\left\|\mathrm{D}_\tau^+ U^k\right\|_{L^2(\Omega)} + \left|U^{k+1}\right|_{H^1(\Omega)} &\leq \sqrt{\frac{4}{\lambda}\mathcal{E}_{\mathrm{LF}}^k} + \sqrt{4\mathcal{E}_{\mathrm{LF}}^k} \\
&= 2\left(\frac{1}{\sqrt{\lambda}} + 1\right)\sqrt{\mathcal{E}_{\mathrm{LF}}^k} \\
&\leq 2\left(\frac{1}{\sqrt{\lambda}} + 1\right)\left(\sqrt{\mathcal{E}_{\mathrm{LF}}^0} + \sum_{j=1}^{k}\frac{\tau}{\sqrt{\lambda}}\left\|f^j\right\|_{L^2(\Omega)}\right) \\
&\leq \left(\frac{1}{\sqrt{\lambda}} + 1\right)\left[\sqrt{2\left\|\mathrm{D}_\tau^+ U^0\right\|_{L^2(\Omega)}^2 + \left|U^0\right|_{H^1(\Omega)}^2 + \left|U^1\right|_{H^1(\Omega)}^2} + \sum_{j=1}^{k}\frac{2\tau}{\sqrt{\lambda}}\left\|f^j\right\|_{L^2(\Omega)}\right].
\end{aligned}$$

So, the assertion follows by using $(a+b+c)^2 = a^2 + b^2 + c^2 + 2ab + 2ac + 2bc$. $\square$

Remark 2.15. *A stability result analogous to Theorem 2.14 in the sense that*

$$\left\|\mathrm{D}_\tau^+ U^k\right\|_{L^2(\Omega)} + \left|U^{k+1/2}\right|_{H^1(\Omega)}, \quad k = 1, \ldots, N_t - 1,$$

is estimated holds true, if the weaker CFL condition (15) *is satisfied.*

Remark 2.16. *For $f \in C([0,T]; L^2(\Omega))$, the estimate*

$$\sum_{j=1}^{k} \tau \left\|f^j\right\|_{L^2(\Omega)} \leq \max_{i=1,\ldots,k} \left\|f^i\right\|_{L^2(\Omega)} \sum_{j=1}^{k} \tau \leq T \|f\|_{C([0,T];L^2(\Omega))} \tag{19}$$

is valid for all $k = 1, \ldots, N_t$. Hence, Theorem 2.12 states that the discrete energy $\mathcal{E}_{\mathrm{LF}}^k$ is bounded for any $k \in \{1, \ldots, N_t - 1\}$ by $\mathcal{E}_{\mathrm{LF}}^0$, i.e. by U^0, U^1 and the right-hand side f. Analogously, Theorem 2.14 gives the boundedness of the approximations $U^2, U^3, \ldots, U^{N_t}$ by U^0, U^1 and the right-hand side f. Note that U^0 and U^1 are determined by approximations (10) *of the initial conditions u_0, u_1 and the right-hand side f.*

Theorem 2.17 (Error Estimates of the Leapfrog Method)**.** *Let $f \in C([0,T]; L^2(\Omega))$ be a given right-hand side of* (1), *let the sequence $(\mathcal{T}_\nu)_{\nu\in\mathbb{N}}$ of decompositions of Ω be admissible, shape-regular and globally quasi-uniform, and let the constant time step size $\tau \in \left(0, \frac{T}{2}\right]$ be given. Assume that $\lambda \in (0,2)$ exists such that the CFL condition* (14) *holds true. In addition, let the unique solution u of Theorem 1.5 satisfy $u \in C^4([0,T]; H_0^1(\Omega))$.*

Then, the leapfrog method (10), (11) *fulfils*

- *the $L^2(\Omega)$ error estimate*

$$\begin{aligned}\max_{k=2,\ldots,N_t} \left\|U^k - u(\cdot,t^k)\right\|_{L^2(\Omega)} &\leq C_P \left(\frac{1}{\sqrt{\lambda}} + 1\right) \Big[\sqrt{2}\left\|\mathrm{D}_\tau^+(U^0 - Q_h^1 u(\cdot,t^0))\right\|_{L^2(\Omega)} \\ &+ \left|U^0 - Q_h^1 u_0\right|_{H^1(\Omega)} + \left|U^1 - Q_h^1 u(\cdot,t^1)\right|_{H^1(\Omega)} + \frac{2T}{\sqrt{\lambda}}\left\|Q_h^1 \partial_{tt} u - \partial_{tt} u\right\|_{C([0,T];L^2(\Omega))} \\ &+ \frac{\tau^2 C_P T}{2\sqrt{\lambda}}\left\|\partial_t^4 u\right\|_{C([0,T];H_0^1(\Omega))}\Big] + \left\|Q_h^1 u - u\right\|_{C([0,T];L^2(\Omega))},\end{aligned}$$

- *the $H_0^1(\Omega)$ error estimate*

$$\begin{aligned}\max_{k=2,\ldots,N_t} \left|U^k - u(\cdot,t^k)\right|_{H^1(\Omega)} &\leq \left(\frac{1}{\sqrt{\lambda}} + 1\right) \Big[\sqrt{2}\left\|\mathrm{D}_\tau^+(U^0 - Q_h^1 u(\cdot,t^0))\right\|_{L^2(\Omega)} \\ &+ \left|U^0 - Q_h^1 u_0\right|_{H^1(\Omega)} + \left|U^1 - Q_h^1 u(\cdot,t^1)\right|_{H^1(\Omega)} + \frac{2T}{\sqrt{\lambda}}\left\|Q_h^1 \partial_{tt} u - \partial_{tt} u\right\|_{C([0,T];L^2(\Omega))} \\ &+ \frac{\tau^2 C_P T}{2\sqrt{\lambda}}\left\|\partial_t^4 u\right\|_{C([0,T];H_0^1(\Omega))}\Big] + \left\|Q_h^1 u - u\right\|_{C([0,T];H_0^1(\Omega))},\end{aligned}$$

- *the $L^2(\Omega)$ error estimate of the discrete time derivative*

$$\max_{k=1,\dots,N_t-1} \left\|\mathrm{D}_\tau^+(U^k - u(\cdot,t^k))\right\|_{L^2(\Omega)} \le \left(\frac{1}{\sqrt{\lambda}} + 1\right) \Big[\sqrt{2}\left\|\mathrm{D}_\tau^+(U^0 - Q_h^1 u(\cdot,t^0))\right\|_{L^2(\Omega)}$$
$$+ \left|U^0 - Q_h^1 u_0\right|_{H^1(\Omega)} + \left|U^1 - Q_h^1 u(\cdot,t^1)\right|_{H^1(\Omega)} + \frac{2T}{\sqrt{\lambda}}\left\|Q_h^1 \partial_{tt} u - \partial_{tt} u\right\|_{C([0,T];L^2(\Omega))}$$
$$+ \frac{\tau^2 C_P T}{2\sqrt{\lambda}}\left\|\partial_t^4 u\right\|_{C([0,T];H_0^1(\Omega))}\Big]$$
$$+ \left\|Q_h^1 \partial_t u - \partial_t u\right\|_{C([0,T];L^2(\Omega))} + \frac{\tau}{2}\left\|Q_h^1 \partial_{tt} u - \partial_{tt} u\right\|_{C([0,T];L^2(\Omega))},$$

where $Q_h^1\colon H_0^1(\Omega) \to S_{h,0}^1(\Omega)$ is the $H_0^1(\Omega)$ projection (7).

Proof. For $k \in \{0,\dots,N_t\}$ and $x \in \Omega$, the pointwise error is written as

$$U^k(x) - u(x,t^k) = \underbrace{U^k(x) - Q_h^1 u(x,t^k)}_{=:\theta^k(x)} + \underbrace{Q_h^1 u(x,t^k) - u(x,t^k)}_{=:\rho(x,t^k)=:\rho^k(x)} = \theta^k(x) + \rho^k(x).$$

The quantities $\theta^k \in S_{h,0}^1(\Omega)$ and $\rho^k \in H_0^1(\Omega)$ are treated separately.

First, for $k \in \{1,\dots,N_t-1\}$, the finite element function $\theta^k \in S_{h,0}^1(\Omega)$ is estimated. Therefore, with the notation (8), the leapfrog method (11), the $H_0^1(\Omega)$ projection (7) and the variational formulation (2), it follows that

$$\begin{aligned}
&\left\langle \mathrm{D}_\tau^2\theta^k, v_h\right\rangle_{L^2(\Omega)} + \left\langle \nabla_x\theta^k, \nabla_x v_h\right\rangle_{L^2(\Omega)} \\
&= \underbrace{\left\langle \mathrm{D}_\tau^2 U^k, v_h\right\rangle_{L^2(\Omega)} + \left\langle \nabla_x U^k, \nabla_x v_h\right\rangle_{L^2(\Omega)}}_{\overset{(11)}{=}\left\langle f^k, v_h\right\rangle_{L^2(\Omega)}} - \left\langle \mathrm{D}_\tau^2 (Q_h^1 u)^k, v_h\right\rangle_{L^2(\Omega)} - \underbrace{\left\langle \nabla_x (Q_h^1 u)^k, \nabla_x v_h\right\rangle_{L^2(\Omega)}}_{\overset{(7)}{=}\left\langle \nabla_x u^k, \nabla_x v_h\right\rangle_{L^2(\Omega)}} \\
&= \left\langle (\partial_{tt}u)^k, v_h\right\rangle_{L^2(\Omega)} + \left\langle \nabla_x u^k, \nabla_x v_h\right\rangle_{L^2(\Omega)} - \left\langle \mathrm{D}_\tau^2 (Q_h^1 u)^k, v_h\right\rangle_{L^2(\Omega)} - \left\langle \nabla_x u^k, \nabla_x v_h\right\rangle_{L^2(\Omega)} \\
&= \left\langle (\partial_{tt}u)^k, v_h\right\rangle_{L^2(\Omega)} \underbrace{- \left\langle \mathrm{D}_\tau^2 u^k, v_h\right\rangle_{L^2(\Omega)} + \left\langle \mathrm{D}_\tau^2 u^k, v_h\right\rangle_{L^2(\Omega)}}_{=0} - \left\langle \mathrm{D}_\tau^2 (Q_h^1 u)^k, v_h\right\rangle_{L^2(\Omega)} \\
&= \left\langle (\partial_{tt}u)^k - \mathrm{D}_\tau^2 u^k - \mathrm{D}_\tau^2 \rho^k, v_h\right\rangle_{L^2(\Omega)}
\end{aligned}$$

for all $v_h \in S_{h,0}^1(\Omega)$ and for $k = 1,\dots,N_t-1$. Hence, the functions $\theta^k \in S_{h,0}^1(\Omega)$, $k = 0,1,\dots,N_t$, are the solutions of the leapfrog method (10), (11) with

$$\theta^0 = U^0 - Q_h^1 u(\cdot,0), \quad \theta^1 = U^1 - Q_h^1 u(\cdot,t^1)$$

for the right-hand side $g \in C([0,T];L^2(\Omega))$ with

$$g(x,t) := \begin{cases} \partial_{tt}u(x,\tau) - \mathrm{D}_\tau^2 u(x,\tau) - \mathrm{D}_\tau^2\rho(x,\tau), & (x,t) \in \Omega\times[0,\tau), \\ \partial_{tt}u(x,t) - \mathrm{D}_\tau^2 u(x,t) - \mathrm{D}_\tau^2\rho(x,t), & (x,t) \in \Omega\times[\tau,T-\tau], \\ \partial_{tt}u(x,T-\tau) - \mathrm{D}_\tau^2 u(x,T-\tau) - \mathrm{D}_\tau^2\rho(x,T-\tau), & (x,t) \in \Omega\times(T-\tau,T], \end{cases}$$

where $\rho = Q_h^1 u - u$. So, the stability of the leapfrog method (Theorem 2.14) for θ^k gives the estimate

$$\begin{aligned}\left\|\mathrm{D}_\tau^+\theta^k\right\|_{L^2(\Omega)} + \left|\theta^{k+1}\right|_{H^1(\Omega)} \leq \left(\frac{1}{\sqrt{\lambda}}+1\right)\Bigg[\sqrt{2}\left\|\mathrm{D}_\tau^+\theta^0\right\|_{L^2(\Omega)} + \left|\theta^0\right|_{H^1(\Omega)} + \left|\theta^1\right|_{H^1(\Omega)} \\ + \sum_{j=1}^{k}\frac{2\tau}{\sqrt{\lambda}}\left\|(\partial_{tt}u)^j - \mathrm{D}_\tau^2 u^j - \mathrm{D}_\tau^2\rho^j\right\|_{L^2(\Omega)}\Bigg] \quad (20)\end{aligned}$$

for $k = 1, \dots, N_t - 1$. For the last term in (20), the triangle inequality, Lemma 2.8 with $\Theta_1(x, t^j, \tau)$, $\Theta_2(x, t^j, \tau) \in [-1, 1]$ and the Poincaré inequality (3) with constant $C_P > 0$ yield

$$\begin{aligned}&\left\|(\partial_{tt}u)^j - \mathrm{D}_\tau^2 u^j - \mathrm{D}_\tau^2\rho^j\right\|_{L^2(\Omega)} \\ &\leq \left\|(\partial_{tt}u)^j - \mathrm{D}_\tau^2 u^j\right\|_{L^2(\Omega)} + \left\|\mathrm{D}_\tau^2\rho^j\right\|_{L^2(\Omega)} \\ &= \left\|-\frac{\tau^2}{12}\partial_t^4 u(\cdot, t^j + \tau\,\Theta_1(\cdot, t^j, \tau))\right\|_{L^2(\Omega)} + \left\|\mathrm{D}_\tau^2\rho(\cdot, t^j)\right\|_{L^2(\Omega)} \\ &\leq \frac{\tau^2}{12}\left\|\partial_t^4 u\right\|_{C([0,T];L^2(\Omega))} + \left\|\partial_{tt}\rho(\cdot, t^j) + \frac{\tau^2}{12}\partial_t^4\rho(\cdot, t^j + \tau\,\Theta_2(\cdot, t^j, \tau))\right\|_{L^2(\Omega)} \\ &\leq \frac{\tau^2}{12}\underbrace{\left\|\partial_t^4 u\right\|_{C([0,T];L^2(\Omega))}}_{\leq C_P\left\|\partial_t^4 u\right\|_{C([0,T];H_0^1(\Omega))}} + \left\|\partial_{tt}\rho\right\|_{C([0,T];L^2(\Omega))} + \frac{\tau^2}{12}\max_{0\leq t\leq T}\underbrace{\left\|\partial_t^4\rho(\cdot, t)\right\|_{L^2(\Omega)}}_{\leq C_P\left|\partial_t^4\rho(\cdot,t)\right|_{H^1(\Omega)}}\end{aligned}$$

and further, Lemma 2.5 and the $H_0^1(\Omega)$ stability of Q_h^1 (Lemma 2.4) give

$$\begin{aligned}\left\|(\partial_{tt}u)^j - \mathrm{D}_\tau^2 u^j - \mathrm{D}_\tau^2\rho^j\right\|_{L^2(\Omega)} \leq \frac{\tau^2}{12}C_P\left\|\partial_t^4 u\right\|_{C([0,T];H_0^1(\Omega))} + \left\|\partial_{tt}u - Q_h^1\partial_{tt}u\right\|_{C([0,T];L^2(\Omega))} \\ + \frac{\tau^2}{12}C_P\max_{0\leq t\leq T}\underbrace{\left|Q_h^1\partial_t^4 u(\cdot, t) - \partial_t^4 u(\cdot, t)\right|_{H^1(\Omega)}}_{\leq 2\left|\partial_t^4 u(\cdot,t)\right|_{H^1(\Omega)}}\end{aligned}$$

for $j = 1, \dots, k$. With the last estimate and the inequality (19), it follows for (20) that

$$\begin{aligned}\left\|\mathrm{D}_\tau^+\theta^k\right\|_{L^2(\Omega)} + \left|\theta^{k+1}\right|_{H^1(\Omega)} \leq \left(\frac{1}{\sqrt{\lambda}}+1\right)\Bigg[\sqrt{2}\left\|\mathrm{D}_\tau^+\theta^0\right\|_{L^2(\Omega)} + \left|\theta^0\right|_{H^1(\Omega)} + \left|\theta^1\right|_{H^1(\Omega)} \\ + \frac{2T}{\sqrt{\lambda}}\left\|Q_h^1\partial_{tt}u - \partial_{tt}u\right\|_{C([0,T];L^2(\Omega))} + \frac{\tau^2 C_P T}{2\sqrt{\lambda}}\left\|\partial_t^4 u\right\|_{C([0,T];H_0^1(\Omega))}\Bigg] \quad (21)\end{aligned}$$

for $k = 1, \dots, N_t - 1$.

Second, for $k \in \{0, \dots, N_t - 1\}$, the function $\mathrm{D}_\tau^+\rho^k \in H_0^1(\Omega)$ is estimated. With Lemma 2.7 with $\Theta(x, t^k, \tau) \in (0, 1)$ and Lemma 2.5, one concludes that

$$\begin{aligned}\left\|\mathrm{D}_\tau^+\rho^k\right\|_{L^2(\Omega)} = \left\|\mathrm{D}_\tau^+\rho(\cdot, t^k)\right\|_{L^2(\Omega)} = \left\|\partial_t\rho(\cdot, t^k) + \frac{\tau}{2}\partial_{tt}\rho(\cdot, t^k + \tau\Theta(\cdot, t^k, \tau))\right\|_{L^2(\Omega)} \\ \leq \left\|Q_h^1\partial_t u - \partial_t u\right\|_{C([0,T];L^2(\Omega))} + \frac{\tau}{2}\left\|Q_h^1\partial_{tt}u - \partial_{tt}u\right\|_{C([0,T];L^2(\Omega))} \quad (22)\end{aligned}$$

for $k \in \{0, \dots, N_t - 1\}$.

Hence, the assertion follows due to

$$\begin{aligned}
\left\|U^{k+1} - u(\cdot, t^{k+1})\right\|_{L^2(\Omega)} &\le \left\|\theta^{k+1}\right\|_{L^2(\Omega)} + \left\|\rho^{k+1}\right\|_{L^2(\Omega)} \le C_P \left|\theta^{k+1}\right|_{H^1(\Omega)} + \left\|\rho^{k+1}\right\|_{L^2(\Omega)}, \\
\left|U^{k+1} - u(\cdot, t^{k+1})\right|_{H^1(\Omega)} &\le \left|\theta^{k+1}\right|_{H^1(\Omega)} + \left|\rho^{k+1}\right|_{H^1(\Omega)}, \\
\left\|\mathrm{D}_\tau^+ (U^k - u(\cdot, t^k))\right\|_{L^2(\Omega)} &\le \left\|\mathrm{D}_\tau^+ \theta^k\right\|_{L^2(\Omega)} + \left\|\mathrm{D}_\tau^+ \rho^k\right\|_{L^2(\Omega)}
\end{aligned}$$

for $k = 1, \dots, N_t - 1$ with the help of (21), (22) and the Poincaré inequality (3). □

Remark 2.18. *Error estimates analogous to Theorem 2.17 in the norms*

$$\begin{aligned}
&\max_{k=1,\dots,N_t-1} \left\|U^{k+1/2} - u(\cdot, t^{k+1/2})\right\|_{L^2(\Omega)}, \\
&\max_{k=1,\dots,N_t-1} \left|U^{k+1/2} - u(\cdot, t^{k+1/2})\right|_{H^1(\Omega)}, \\
&\max_{k=1,\dots,N_t-1} \left\|\mathrm{D}_\tau^+ (U^k - u(\cdot, t^k))\right\|_{L^2(\Omega)}
\end{aligned}$$

hold true with

$$t^{k+1/2} = \frac{t^{k+1} + t^k}{2} = t^k + \frac{\tau}{2}, \quad k = 0, \dots, N_t - 1,$$

if the weaker CFL condition (15) *is satisfied, see also Theorem 2.26 and its proof.*

The Theorem 2.17 motivates to choose

$$U^0 = u_{0,h} = Q_h^1 u_0 \quad \text{and} \quad U^1 = u_{1,h} = Q_h^1 \left(u_0 + \tau u_1 + \frac{\tau^2}{2} (\Delta_x u_0 + f(\cdot, 0)) \right) \tag{23}$$

in (10) for $u_0, u_1 \in H_0^1(\Omega)$ with $\Delta_x u_0 \in H_0^1(\Omega)$ and $f \in C([0,T]; L^2(\Omega))$ with $f(\cdot, 0) \in H_0^1(\Omega)$. Using the representation (9), the initialisation (23) can be realised as

$$\underline{\alpha}^0 = A_h^{-1} \underline{f}_{u_0} \quad \text{and} \quad \underline{\alpha}^1 = \underline{\alpha}^0 + A_h^{-1} \left(\tau \underline{f}_{u_1} + \frac{\tau^2}{2} \underline{f}_{\Delta_x u_0} + \frac{\tau^2}{2} \underline{f}_f \right)$$

with the stiffness matrix A_h given in (6) and the vectors

$$\begin{aligned}
\underline{f}_{u_0}[i] &= \langle \nabla_x u_0, \nabla_x \psi_i \rangle_{L^2(\Omega)}, \\
\underline{f}_{u_1}[i] &= \langle \nabla_x u_1, \nabla_x \psi_i \rangle_{L^2(\Omega)}, \\
\underline{f}_{\Delta_x u_0}[i] &= \langle \nabla_x \Delta_x u_0, \nabla_x \psi_i \rangle_{L^2(\Omega)}, \\
\underline{f}_f[i] &= \langle \nabla_x f(\cdot, 0), \nabla_x \psi_i \rangle_{L^2(\Omega)}
\end{aligned}$$

for $i = 1, \dots, M$.

Lemma 2.19. *Let the sequence $(\mathcal{T}_\nu)_{\nu\in\mathbb{N}}$ of decompositions of Ω be admissible, shape-regular and not necessarily globally quasi-uniform, and let the constant time step size $\tau \in \left(0, \frac{T}{2}\right]$ be given. Let $u_0, u_1 \in H_0^1(\Omega)$ with $\Delta_x u_0 \in H_0^1(\Omega)$, $f \in C([0,T]; L^2(\Omega))$ with $f(\cdot,0) \in H_0^1(\Omega)$ be the given initial conditions and right-hand side of the wave equation* (1)*. In addition, assume that the unique solution u of Theorem 1.5 fulfils $u \in C^3([0,T]; H_0^1(\Omega))$. Then, the initialisation* (23) *leads to*

$$\sqrt{2}\left\|\mathrm{D}_\tau^+(U^0 - Q_h^1 u(\cdot,t^0))\right\|_{L^2(\Omega)} + \left|U^0 - Q_h^1 u_0\right|_{H^1(\Omega)} + \left|U^1 - Q_h^1 u(\cdot,t^1)\right|_{H^1(\Omega)} \leq \frac{\tau^2}{6}\left(\tau + \sqrt{2}C_P\right)\left|\partial_t^3 u\right|_{C([0,T];H_0^1(\Omega))}$$

with the constant $C_P > 0$ of the Poincaré inequality (3).

Proof. First, $|U^0 - Q_h^1 u_0|_{H^1(\Omega)} = 0$ holds true.

Second, the assumptions on u and f yield $\Delta_x u \overset{(2)}{=} \partial_{tt} u - f \in C([0,T]; L^2(\Omega))$ and so

$$\partial_{tt} u(\cdot,0) = \lim_{t\to 0^+} \partial_{tt} u(\cdot,t) \overset{(2)}{=} \lim_{t\to 0^+} [\Delta_x u(\cdot,t) + f(\cdot,t)] = \underbrace{\Delta_x u(\cdot,0)}_{=\Delta_x u_0} + f(\cdot,0) \quad \text{in } L^2(\Omega).$$

Hence, Taylor's theorem gives

$$u(x,t^1) = u(x,0) + \partial_t u(x,0)\,\tau + \underbrace{\partial_{tt} u(x,0)}_{=\Delta_x u_0(x) + f(x,0)} \cdot \frac{\tau^2}{2} + \partial_t^3 u(x,\Theta(x,t^1))\frac{\tau^3}{6}, \quad x \in \Omega,$$

with $\Theta(x,t^1) \in (0,t^1)$, and thus,

$$\begin{aligned}\left|U^1 - Q_h^1 u(\cdot,t^1)\right|_{H^1(\Omega)} &= \Bigg|Q_h^1\left(u_0 + \tau u_1 + \frac{\tau^2}{2}(\Delta_x u_0 + f(\cdot,0))\right) \\ &\qquad - Q_h^1\left(u_0 + u_1\,\tau + (\Delta_x u_0 + f(\cdot,0))\frac{\tau^2}{2} + \partial_t^3 u(\cdot,\Theta(\cdot,t^1))\frac{\tau^3}{6}\right)\Bigg|_{H^1(\Omega)} \\ &= \frac{\tau^3}{6}\left|Q_h^1 \partial_t^3 u(\cdot,\Theta(\cdot,t^1))\right|_{H^1(\Omega)} \\ &\leq \frac{\tau^3}{6}\left|\partial_t^3 u\right|_{C([0,T];H_0^1(\Omega))}\end{aligned}$$

folllows, where the H_0^1 stability (Lemma 2.4) is used.

Third, the last estimate and the Poincaré inequality (3) yield

$$\begin{aligned}\left\|\mathrm{D}_\tau^+(U^0 - Q_h^1 u(\cdot,t^0))\right\|_{L^2(\Omega)} = \frac{1}{\tau}\left\|U^1 - Q_h^1 u(\cdot,t^1)\right\|_{L^2(\Omega)} &\leq \frac{C_P}{\tau}\left|U^1 - Q_h^1 u(\cdot,t^1)\right|_{H^1(\Omega)} \\ &\leq \frac{C_P\tau^2}{6}\left|\partial_t^3 u\right|_{C([0,T];H_0^1(\Omega))}.\end{aligned}$$

Hence, the assertion is proven. □

Corollary 2.20. *Let the assumptions of Theorem 2.17 and Lemma 2.19 be satisfied. In addition, assume that the solution u is sufficiently smooth and the domain $\Omega \subset \mathbb{R}^d$ is sufficiently regular, e.g., convex or smooth. Then, the error estimates*

$$\max_{k=0,\dots,N_t} \left\| U^k - u(\cdot, t^k) \right\|_{L^2(\Omega)} \leq c \cdot (h^2 + \tau^2),$$
$$\max_{k=0,\dots,N_t} \left| U^k - u(\cdot, t^k) \right|_{H^1(\Omega)} \leq c \cdot (h + \tau^2),$$
$$\max_{k=0,\dots,N_t-1} \left\| \mathrm{D}_\tau^+ (U^k - u(\cdot, t^k)) \right\|_{L^2(\Omega)} \leq c \cdot (h^2 + \tau^2)$$

hold true with a constant $c > 0$ independent of u and the mesh sizes h, τ.

Remark 2.21. *The implementation of the initialisation procedure* (23) *requires spatial derivatives of the initial conditions u_0, u_1 and the right-hand side f. However, nodal interpolations or*

$$\underline{\alpha}^0 = M_h^{-1} \underline{\hat{f}}_{u_0} \quad \text{and} \quad \underline{\alpha}^1 = \underline{\alpha}^0 + M_h^{-1} \left(\tau \underline{\hat{f}}_{u_1} + \frac{\tau^2}{2} \underline{f}^0 - \frac{\tau^2}{2} A_h \underline{\alpha}^0 \right)$$

with the stiffness matrix A_h and mass matrix M_h given in (6), (5) *and the vectors*

$$\underline{\hat{f}}_{u_0}[i] = \langle u_0, \psi_i \rangle_{L^2(\Omega)}, \quad \underline{\hat{f}}_{u_1}[i] = \langle u_1, \psi_i \rangle_{L^2(\Omega)}, \quad \underline{f}^0[i] = \langle f(\cdot, 0), \psi_i \rangle_{L^2(\Omega)}, \qquad i = 1, \dots, M,$$

are easy to implement and are often sufficiently accurate in practice.

As numerical examples, the domain and the time interval

$$\Omega = (0,1) \quad \text{and} \quad (0,T) = (0,11)$$

with the exact solutions

$$u_A(x,t) = \sin(2\pi t)\sin(2\pi x) + \cos(\pi t)\sin(\pi x) \qquad \text{for } (x,t) \in [0,1] \times [0,11],$$
$$u_B(x,t) = x(1-x)\sin(\pi(3x - t)) \qquad \text{for } (x,t) \in [0,1] \times [0,11]$$

and the corresponding right-hand sides f_A, f_B and initial conditions $u_{A,0}$, $u_{A,1}$, $u_{B,0}$, $u_{B,1}$ are considered. Note that $f_A = 0$ and the total energy

$$\mathcal{E}_A(t) = \mathcal{E}_A(0) = \frac{1}{2} \|u_{A,1}\|^2_{L^2(\Omega)} + \frac{1}{2} \|\nabla_x u_{A,0}\|^2_{L^2(\Omega)} = \frac{5\pi^2}{4} \approx 12.337, \quad t \in [0, 11],$$

is conserved, see Corollary 1.6. The uniform spatial mesh size and the constant time step size are given by

$$h = \frac{1}{N}, \quad N \in \mathbb{N}, \qquad \text{and} \qquad \tau = \frac{11}{N_t}, \quad N_t \in \mathbb{N}.$$

In this case, the CFL condition (14) and the weaker CFL condition (15) read as

$$\frac{\tau}{h} \le \sqrt{\frac{1}{6} - \frac{\lambda}{12}} < \frac{1}{\sqrt{6}} \approx 0.4082 \qquad \text{for } \lambda \in (0,2), \tag{24}$$

$$\frac{\tau}{h} \le \sqrt{\frac{1}{3} - \frac{\widetilde{\lambda}}{12}} < \frac{1}{\sqrt{3}} \approx 0.5774 \qquad \text{for } \widetilde{\lambda} \in (0,4), \tag{25}$$

since $c_{\mathrm{inv}} = 12$, see the derivation of [10, (9.19), page 217].

In Table 1, Table 2, Table 3 and Table 4, the errors of Corollary 2.20 for the solutions u_A, u_B are given for the leapfrog method (10), (11) with initialisation (23) when a uniform refinement strategy is applied for the temporal and spatial discretisations, i.e. N and N_t are doubled in each refinement step. The appearing integrals to compute the initialisation (23) and the related right-hand sides are calculated by using high-order quadrature rules. In Table 1 and Table 3, the CFL condition (24) is satisfied with

$$\frac{\tau}{h} = 0.4$$

and hence, the leapfrog method is stable and the error estimates of Corollary 2.20 are confirmed. Additionally, in Figure 1 the total energy $\mathcal{E}_A(t) = \frac{5\pi^2}{4} \approx 12.337$ and the discrete energy $\mathcal{E}_{\mathrm{LF}}^k$ of the leapfrog method for the solution u_A are depicted for the refinement level 0 and level 3 with $\frac{\tau}{h} = 0.4$, where the discrete energy conservation of the leapfrog method (Theorem 2.12) is verified. In Table 2 and Table 4, the CFL conditions (24) and (25) are violated with

$$\frac{\tau}{h} = \frac{22}{38} \approx 0.5789.$$

So, Table 2 and Table 4 show that the leapfrog method is not stable in these cases.

h	τ	$\max_k \lVert\cdot\rVert_{L^2(\Omega)}$	eoc	$\max_k \lvert\cdot\rvert_{H^1(\Omega)}$	eoc	$\max_k \lVert \mathrm{D}_\tau^+(\cdot)\rVert_{L^2(\Omega)}$	eoc
0.50000	0.20000	1.33e+00	-	5.62e+00	-	5.88e+00	-
0.25000	0.10000	1.24e+00	0.08	7.89e+00	-0.4	8.01e+00	-0.4
0.12500	0.05000	1.17e+00	0.08	7.41e+00	0.08	7.34e+00	0.11
0.06250	0.02500	3.57e-01	1.63	2.25e+00	1.64	2.20e+00	1.65
0.03125	0.01250	9.09e-02	1.93	5.74e-01	1.93	5.59e-01	1.94
0.01562	0.00625	2.28e-02	1.97	1.46e-01	1.95	1.40e-01	1.97
0.00781	0.00313	5.69e-03	1.99	6.40e-02	1.19	3.50e-02	1.99
0.00391	0.00156	1.42e-03	1.99	3.20e-02	1.00	8.75e-03	1.99
0.00195	0.00078	3.56e-04	2.00	1.60e-02	1.00	2.19e-03	2.00
0.00098	0.00039	8.89e-05	2.00	7.99e-03	1.00	5.47e-04	2.00

Table 1: Numerical results of the leapfrog method (10), (11), (23) with uniform refinement for the solution u_A, where the CFL condition (24) is satisfied with $\frac{\tau}{h} = 0.4$.

h	τ	$\max_k \lVert\cdot\rVert_{L^2(\Omega)}$	eoc	$\max_k \lvert\cdot\rvert_{H^1(\Omega)}$	eoc	$\max_k \lVert \mathrm{D}_\tau^+(\cdot)\rVert_{L^2(\Omega)}$	eoc
0.50000	0.28947	1.34e+00	-	5.66e+00	-	5.65e+00	-
0.25000	0.14474	1.44e+00	-0.1	8.46e+00	-0.5	8.29e+00	-0.4
0.12500	0.07237	1.27e+00	0.16	8.04e+00	0.07	7.88e+00	0.07
0.06250	0.03618	4.11e-01	1.56	2.59e+00	1.56	2.53e+00	1.56
0.03125	0.01809	1.05e-01	1.93	6.60e-01	1.93	6.44e-01	1.93
0.01562	0.00905	3.21e+47	-	7.11e+49	-	6.69e+49	-
0.00781	0.00452	1.14e+134	-	5.04e+136	-	4.69e+136	-

Table 2: Numerical results of the leapfrog method (10), (11), (23) with uniform refinement for the solution u_A, where the CFL conditions (24) and (25) are violated with $\frac{\tau}{h} = \frac{22}{38} \approx 0.5789$.

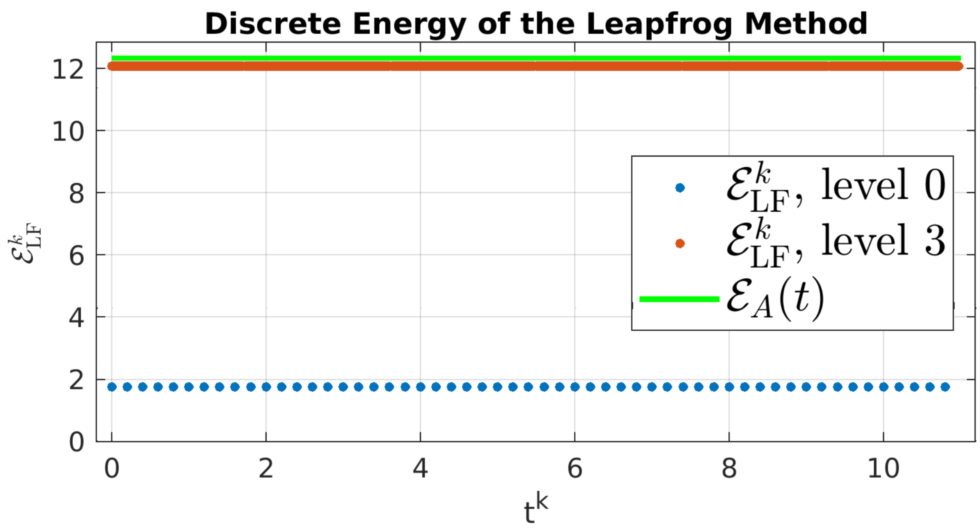

Figure 1: Total energy $\mathcal{E}_A(t) = \frac{5\pi^2}{4} \approx 12.337$ and discrete energy $\mathcal{E}^k_{\mathrm{LF}}$ of the leapfrog method (10), (11), (23) with uniform refinement for the solution u_A, where the CFL condition (24) is satisfied with $\frac{\tau}{h} = 0.4$.

h	τ	$\max_k \|\cdot\|_{L^2(\Omega)}$	eoc	$\max_k \lvert\cdot\rvert_{H^1(\Omega)}$	eoc	$\max_k \|\mathrm{D}_\tau^+(\cdot)\|_{L^2(\Omega)}$	eoc
0.50000	0.20000	9.03e-01	-	3.01e+00	-	2.93e+00	-
0.25000	0.10000	1.29e-01	2.17	1.20e+00	1.03	4.12e-01	2.19
0.12500	0.05000	2.39e-02	2.19	5.23e-01	1.08	7.83e-02	2.16
0.06250	0.02500	6.06e-03	1.88	2.68e-01	0.92	2.13e-02	1.79
0.03125	0.01250	1.52e-03	1.95	1.35e-01	0.97	5.24e-03	1.97
0.01562	0.00625	3.79e-04	1.98	6.74e-02	0.99	1.28e-03	2.01
0.00781	0.00313	9.46e-05	1.99	3.37e-02	0.99	3.20e-04	1.99
0.00391	0.00156	2.37e-05	1.99	1.69e-02	1.00	8.01e-05	1.99
0.00195	0.00078	5.92e-06	2.00	8.43e-03	1.00	2.00e-05	2.00
0.00098	0.00039	1.48e-06	2.00	4.22e-03	1.00	5.01e-06	2.00

Table 3: Numerical results of the leapfrog method (10), (11), (23) with uniform refinement for the solution u_B, where the CFL condition (24) is satisfied with $\frac{\tau}{h} = 0.4$.

h	τ	$\max_k \|\cdot\|_{L^2(\Omega)}$	eoc	$\max_k \lvert\cdot\rvert_{H^1(\Omega)}$	eoc	$\max_k \|\mathrm{D}_\tau^+(\cdot)\|_{L^2(\Omega)}$	eoc
0.50000	0.28947	7.69e-01	-	2.59e+00	-	2.49e+00	-
0.25000	0.14474	1.29e-01	1.99	1.18e+00	0.87	4.15e-01	2.00
0.12500	0.07237	2.41e-02	2.18	5.23e-01	1.06	8.01e-02	2.14
0.06250	0.03618	6.17e-03	1.87	2.68e-01	0.92	2.14e-02	1.81
0.03125	0.01809	1.54e-03	1.95	1.35e-01	0.97	5.32e-03	1.96
0.01562	0.00905	1.50e+55	-	3.33e+57	-	3.13e+57	-
0.00781	0.00452	2.46e+139	-	1.09e+142	-	1.01e+142	-

Table 4: Numerical results of the leapfrog method (10), (11), (23) with uniform refinement for the solution u_B, where the CFL conditions (24) and (25) are violated with $\frac{\tau}{h} = \frac{22}{38} \approx 0.5789$.

2.4 Crank-Nicolson Method or Newmark Method

This subsection is based on [1, Subsection 9.5.2] and [6, Chapter 3]. The Crank-Nicolson method considered in this subsection is also known as Newmark method with Newmark parameters $\beta = \frac{1}{4}$, $\gamma = \frac{1}{2}$, see the classical reference [7, (8.6-4), (8.6-5), (8.6-6), page 205]. In the remainder of this subsection, let $f \in C([0,T]; L^2(\Omega))$ be a given right-hand side of (1), let the sequence $(\mathcal{T}_\nu)_{\nu\in\mathbb{N}}$ of decompositions of Ω be admissible, shape-regular and not necessarily globally quasi-uniform, and let the constant time step size $\tau \in \left(0, \frac{T}{2}\right]$ and the corresponding time steps t^k be given. Introduce for $x \in \Omega$ and $k \in \{0, 1, \ldots, N_t\}$ the approximations

$$U^k(x) := \sum_{i=1}^{M} \alpha_i^k \psi_i(x) \approx u(x, t^k), \tag{26}$$

where $\alpha_i^k \in \mathbb{R}$ are the unknown coefficients of $U^k \in S^1_{h,0}(\Omega)$. For $k = 1, \ldots, N_t - 1$, replacing

$$\nabla_x u(x, t^k) \approx \frac{1}{4}\left(\nabla_x U^{k+1}(x) + 2\nabla_x U^k(x) + \nabla_x U^{k-1}(x)\right), \quad x \in \Omega,$$

and

$$f(x, t^k) \approx \frac{1}{4}\left(f(x, t^{k+1}) + 2f(x, t^k) + f(x, t^{k-1})\right), \quad x \in \Omega,$$

in the variational formulation (2) together with the discrete second-order central difference operator $\mathrm{D}^2_\tau U^k(x) \approx \partial_{tt} u(x, t^k)$, $x \in \Omega$, gives the Crank-Nicolson method.

Crank-Nicolson Method
Find functions $U^k \in S^1_{h,0}(\Omega)$ for $k = 0, 1, \ldots, N_t$ such that

$$U^0 = u_{0,h}, \quad U^1 = u_{1,h} \tag{27}$$

and for $k = 1, 2, \ldots, N_t - 1$,

$$\begin{aligned} \forall v_h \in S^1_{h,0}(\Omega)\colon \quad & \underbrace{\frac{1}{\tau^2}\left\langle U^{k+1} - 2U^k + U^{k-1}, v_h\right\rangle_{L^2(\Omega)}}_{=\left\langle \mathrm{D}^2_\tau U^k, v_h\right\rangle_{L^2(\Omega)}} \\ & + \frac{1}{4}\left\langle \nabla_x U^{k+1} + 2\nabla_x U^k + \nabla_x U^{k-1}, \nabla_x v_h\right\rangle_{L^2(\Omega)} = \frac{1}{4}\left\langle f^{k+1} + 2f^k + f^{k-1}, v_h\right\rangle_{L^2(\Omega)} \end{aligned} \tag{28}$$

with $f^k(x) := f(x, t^k)$ for $x \in \Omega$, where

$$u(\cdot, 0) = u_0 \approx u_{0,h} \in S^1_{h,0}(\Omega) \quad \text{and} \quad u(\cdot, t^1) \approx u_{1,h} \in S^1_{h,0}(\Omega)$$

are approximations using the initial conditions u_0, u_1 and the right-hand side f, which will be defined later. In the following, it is shown that the Crank-Nicolson method can be realised as a two-step method. With (26) and $v_h = \psi_j$, $j = 1, \dots, M$, in (28), the discrete variational formulation (28) is equivalent to solving the linear system

$$\frac{1}{\tau^2} M_h(\underline{\alpha}^{k+1} - 2\underline{\alpha}^k + \underline{\alpha}^{k-1}) + \frac{1}{4} A_h(\underline{\alpha}^{k+1} + 2\underline{\alpha}^k + \underline{\alpha}^{k-1}) = \frac{1}{4}\left(\underline{f}^{k+1} + 2\underline{f}^k + \underline{f}^{k-1}\right) \tag{29}$$

for $k = 1, \dots, N_t - 1$, where the unknown coefficient vectors are given by

$$\underline{\alpha}^k := \begin{pmatrix} \alpha_1^k \\ \alpha_2^k \\ \vdots \\ \alpha_M^k \end{pmatrix} \in \mathbb{R}^M, \quad k = 0, \dots, N_t.$$

Here, M_h and A_h are the mass and stiffness matrices given in (5) and (6), and the vectors of the right-hand side are defined by

$$\underline{f}^k := \begin{pmatrix} \left\langle f^k, \psi_1 \right\rangle_{L^2(\Omega)} \\ \left\langle f^k, \psi_2 \right\rangle_{L^2(\Omega)} \\ \vdots \\ \left\langle f^k, \psi_M \right\rangle_{L^2(\Omega)} \end{pmatrix}, \quad k = 0, \dots, N_t.$$

Rewriting the linear system (29) gives

$$\left(M_h + \frac{\tau^2}{4} A_h\right) \underline{\alpha}^{k+1} = \frac{\tau^2}{4}\left(\underline{f}^{k+1} + 2\underline{f}^k + \underline{f}^{k-1}\right) - \frac{\tau^2}{4} A_h(2\underline{\alpha}^k + \underline{\alpha}^{k-1}) + M_h(2\underline{\alpha}^k - \underline{\alpha}^{k-1}) \tag{30}$$

for $k = 1, \dots, N_t - 1$, i.e. in every time step, a linear system with the system matrix $M_h + \frac{\tau^2}{4} A_h$ has to solved for $\underline{\alpha}^{k+1}$, when $\underline{\alpha}^k$, $\underline{\alpha}^{k-1}$ are already computed. Since the mass and stiffness matrices are positive definite, the linear system (30) is uniquely solvable. Hence, the discrete variational formulation (28) has a unique solution $U^{k+1} \in S_{h,0}^1(\Omega)$ for $k = 1, \dots, N_t - 1$, when U^k, U^{k-1} are known.

In the following, the Crank-Nicolson method (27), (28) is analysed. Therefore,

- an energy conservation,
- the stability and
- error estimates

are proven.

Definition 2.22. *Let $U^k \in S_{h,0}^1(\Omega)$ be the solution of (27), (28) for the time step t^k, $k = 0, 1, \dots, N_t$. The discrete energy of the Crank-Nicolson method (27), (28) is defined by*

$$\mathcal{E}_{\mathrm{CN}}^k := \frac{1}{2}\left\| \mathrm{D}_\tau^+ U^k \right\|_{L^2(\Omega)}^2 + \frac{1}{2} \underbrace{\left\langle \nabla_x U^{k+1/2}, \nabla_x U^{k+1/2} \right\rangle_{L^2(\Omega)}}_{= \left| U^{k+1/2} \right|_{H^1(\Omega)}^2}$$

for $k = 0, 1, \ldots, N_t - 1$ *with the average*

$$U^{k+1/2} := \frac{1}{2}\left(U^{k+1} + U^k\right).$$

Note that $\mathcal{E}_{\mathrm{CN}}^k$ is in general a non-negative quantity, i.e. no CFL condition (Courant, Friedrichs, Lewy) is required such that $\mathcal{E}_{\mathrm{CN}}^k \geq 0$. Next, an energy conservation is proven.

Theorem 2.23 (Energy Conservation of the Crank-Nicolson Method)**.** *Let the function* $f \in C([0,T]; L^2(\Omega))$ *be a given right-hand side of* (1)*, let the sequence* $(\mathcal{T}_\nu)_{\nu\in\mathbb{N}}$ *of decompositions of* Ω *be admissible, shape-regular and not necessarily globally quasi-uniform, and let the constant time step size* $\tau \in \left(0, \frac{T}{2}\right]$ *be given. Then,*

- *for the right-hand side* $f = 0$*, the Crank-Nicolson method* (27), (28) *conserves energy in the sense that*

$$\mathcal{E}_{\mathrm{CN}}^k = \mathcal{E}_{\mathrm{CN}}^0 \qquad \text{for } k = 1, \ldots, N_t - 1,$$

- *for the right-hand sides* $f \neq 0$*, the relation*

$$\sqrt{\mathcal{E}_{\mathrm{CN}}^k} \leq \sqrt{\mathcal{E}_{\mathrm{CN}}^0} + \sum_{j=1}^{k} \frac{\tau}{4\sqrt{2}} \left\| f^{j+1} + 2f^j + f^{j-1} \right\|_{L^2(\Omega)} \qquad \text{for } k = 1, \ldots, N_t - 1$$

  holds true, where $f^j(x) = f(x, t^j)$*,* $x \in \Omega$*.*

Proof. Let $k \in \{1, \ldots, N_t - 1\}$ be fixed. For $j \in \{1, \ldots, k\}$, choosing

$$v_h = U^{j+1} - U^{j-1} = (U^{j+1} - U^j) + (U^j - U^{j-1}) = (U^{j+1} + U^j) - (U^j + U^{j-1}) \in S_{h,0}^1(\Omega)$$

in the variational formulation (28) yields

$$\begin{aligned}
&\frac{1}{4}\left\langle f^{j+1} + 2f^j + f^{j-1}, U^{j+1} - U^{j-1} \right\rangle_{L^2(\Omega)} \\
&\quad = \frac{1}{\tau^2}\left\langle U^{j+1} - 2U^j + U^{j-1}, U^{j+1} - U^{j-1} \right\rangle_{L^2(\Omega)} \\
&\quad\quad + \frac{1}{4}\left\langle \nabla_x U^{j+1} + 2\nabla_x U^j + \nabla_x U^{j-1}, \nabla_x U^{j+1} - \nabla_x U^{j-1} \right\rangle_{L^2(\Omega)} \\
&\quad = \frac{1}{\tau^2}\left\langle (U^{j+1} - U^j) - (U^j - U^{j-1}), (U^{j+1} - U^j) + (U^j - U^{j-1}) \right\rangle_{L^2(\Omega)} \\
&\quad\quad + \frac{1}{4}\left\langle \nabla_x(U^{j+1} + U^j) + \nabla_x(U^j + U^{j-1}), \nabla_x(U^{j+1} + U^j) - \nabla_x(U^j + U^{j-1}) \right\rangle_{L^2(\Omega)} \\
&\quad = \left\| \frac{U^{j+1} - U^j}{\tau} \right\|_{L^2(\Omega)}^2 - \left\| \frac{U^j - U^{j-1}}{\tau} \right\|_{L^2(\Omega)}^2 + \left| \frac{U^{j+1} + U^j}{2} \right|_{H^1(\Omega)}^2 - \left| \frac{U^j + U^{j-1}}{2} \right|_{H^1(\Omega)}^2 \\
&\quad = 2(\mathcal{E}_{\mathrm{CN}}^j - \mathcal{E}_{\mathrm{CN}}^{j-1}).
\end{aligned} \tag{31}$$

Hence, in the case $f = 0$, the equality

$$\mathcal{E}_{\mathrm{CN}}^k = \mathcal{E}_{\mathrm{CN}}^0 \qquad \text{for } k = 1, \ldots, N_t - 1$$

follows by induction on j from (31).

In the case $f \neq 0$, the relation (31) gives

$$\begin{aligned}
2(\mathcal{E}_{\mathrm{CN}}^{j} - \mathcal{E}_{\mathrm{CN}}^{j-1}) =& \frac{1}{4}\left\langle f^{j+1} + 2f^{j} + f^{j-1}, U^{j+1} - U^{j-1}\right\rangle_{L^2(\Omega)} \\
=& \frac{1}{4}\left\langle f^{j+1} + 2f^{j} + f^{j-1}, U^{j+1} - U^{j}\right\rangle_{L^2(\Omega)} \\
&+ \frac{1}{4}\left\langle f^{j+1} + 2f^{j} + f^{j-1}, U^{j} - U^{j-1}\right\rangle_{L^2(\Omega)} \\
=& \frac{\tau}{4}\left\langle f^{j+1} + 2f^{j} + f^{j-1}, \mathrm{D}_\tau^+ U^{j}\right\rangle_{L^2(\Omega)} + \frac{\tau}{4}\left\langle f^{j+1} + 2f^{j} + f^{j-1}, \mathrm{D}_\tau^+ U^{j-1}\right\rangle_{L^2(\Omega)} \\
\leq& \frac{\tau}{4}\left\| f^{j+1} + 2f^{j} + f^{j-1}\right\|_{L^2(\Omega)} \left(\left\|\mathrm{D}_\tau^+ U^{j}\right\|_{L^2(\Omega)} + \left\|\mathrm{D}_\tau^+ U^{j-1}\right\|_{L^2(\Omega)}\right) \\
\leq& \frac{\tau}{4}\left\| f^{j+1} + 2f^{j} + f^{j-1}\right\|_{L^2(\Omega)} \left(\sqrt{2\mathcal{E}_{\mathrm{CN}}^{j}} + \sqrt{2\mathcal{E}_{\mathrm{CN}}^{j-1}}\right),
\end{aligned} \tag{32}$$

where the Cauchy-Schwarz inequality is used. If $\mathcal{E}_{\mathrm{CN}}^{j} > 0$, then

$$\sqrt{\mathcal{E}_{\mathrm{CN}}^{j}} - \sqrt{\mathcal{E}_{\mathrm{CN}}^{j-1}} = \frac{\mathcal{E}_{\mathrm{CN}}^{j} - \mathcal{E}_{\mathrm{CN}}^{j-1}}{\sqrt{\mathcal{E}_{\mathrm{CN}}^{j}} + \sqrt{\mathcal{E}_{\mathrm{CN}}^{j-1}}} \overset{(32)}{\leq} \frac{\tau}{4\sqrt{2}}\left\| f^{j+1} + 2f^{j} + f^{j-1}\right\|_{L^2(\Omega)}. \tag{33}$$

If $\mathcal{E}_{\mathrm{CN}}^{j} = 0$, then the estimate (33) still holds true since $\mathcal{E}_{\mathrm{CN}}^{j-1} \geq 0$. Hence, summing (33) over $j = 1, \ldots, k$ gives

$$\sqrt{\mathcal{E}_{\mathrm{CN}}^{k}} - \sqrt{\mathcal{E}_{\mathrm{CN}}^{0}} = \sum_{j=1}^{k}\left(\sqrt{\mathcal{E}_{\mathrm{CN}}^{j}} - \sqrt{\mathcal{E}_{\mathrm{CN}}^{j-1}}\right) \leq \sum_{j=1}^{k} \frac{\tau}{4\sqrt{2}}\left\| f^{j+1} + 2f^{j} + f^{j-1}\right\|_{L^2(\Omega)},$$

where a telescoping sum occurs. □

Theorem 2.24 (Stability of the Crank-Nicolson Method)**.** *Let $f \in C([0,T]; L^2(\Omega))$ be a given right-hand side of* (1)*, let the sequence $(\mathcal{T}_\nu)_{\nu\in\mathbb{N}}$ of decompositions of Ω be admissible, shape-regular and not necessarily globally quasi-uniform, and let the constant time step size $\tau \in \left(0, \frac{T}{2}\right]$ be given. Then, the Crank-Nicolson method* (27)*,* (28) *is stable in the sense that*

$$\begin{aligned}
\left\|\mathrm{D}_\tau^+ U^{k}\right\|_{L^2(\Omega)} + \left|U^{k+1/2}\right|_{H^1(\Omega)} \leq 2\left\|\mathrm{D}_\tau^+ U^{0}\right\|_{L^2(\Omega)} + 2\left|U^{1/2}\right|_{H^1(\Omega)} \\
+ \sum_{j=1}^{k} \frac{\tau}{2}\left\| f^{j+1} + 2f^{j} + f^{j-1}\right\|_{L^2(\Omega)}
\end{aligned}$$

for $k = 1, \ldots, N_t - 1$.

Proof. Let $k \in \{1, \ldots, N_t - 1\}$ be fixed. Using the definition of the discrete energy and the energy conservation (Theorem 2.23) yields

$$
\begin{aligned}
\left\|\mathrm{D}_\tau^+ U^k\right\|_{L^2(\Omega)} + \left|U^{k+1/2}\right|_{H^1(\Omega)} &\le \sqrt{2\mathcal{E}_{\mathrm{CN}}^k} + \sqrt{2\mathcal{E}_{\mathrm{CN}}^k} \\
&= 2\sqrt{2}\sqrt{\mathcal{E}_{\mathrm{CN}}^k} \\
&\le 2\sqrt{2}\left(\sqrt{\mathcal{E}_{\mathrm{CN}}^0} + \sum_{j=1}^{k} \frac{\tau}{4\sqrt{2}}\left\|f^{j+1} + 2f^j + f^{j-1}\right\|_{L^2(\Omega)}\right) \\
&= 2\sqrt{2}\left[\sqrt{\frac{1}{2}\|\mathrm{D}_\tau^+ U^0\|_{L^2(\Omega)}^2 + \frac{1}{2}\left|U^{1/2}\right|_{H^1(\Omega)}^2}\right. \\
&\qquad \left. + \sum_{j=1}^{k} \frac{\tau}{4\sqrt{2}}\left\|f^{j+1} + 2f^j + f^{j-1}\right\|_{L^2(\Omega)}\right].
\end{aligned}
$$

So, the assertion follows by using the binomial formula $(a+b)^2 = a^2 + 2ab + b^2$. □

Remark 2.25. *For $f \in C([0,T]; L^2(\Omega))$, the estimate*

$$
\sum_{j=1}^{k} \tau\left\|f^{j+1} + 2f^j + f^{j-1}\right\|_{L^2(\Omega)} \le 4 \max_{i=0,\dots,k+1}\left\|f^i\right\|_{L^2(\Omega)} \sum_{j=1}^{k} \tau \le 4T\|f\|_{C([0,T];L^2(\Omega))}
$$

is valid for all $k = 1, \dots, N_t - 1$. Hence, Theorem 2.23 states that the discrete energy $\mathcal{E}_{\mathrm{CN}}^k$ is bounded for any $k \in \{1, \dots, N_t - 1\}$ by $\mathcal{E}_{\mathrm{CN}}^0$, i.e. by U^0, U^1 and the right-hand side f. Analogously, Theorem 2.24 gives the boundedness of the approximations $U^2, U^3, \dots, U^{N_t}$ by U^0, U^1 and the right-hand side f. Note that U^0 and U^1 are determined by approximations (27) *of the initial conditions u_0, u_1 and the right-hand side f.*

Theorem 2.26 (Error Estimates of the Crank-Nicolson Method)**.** *Let $f \in C([0,T]; L^2(\Omega))$ be a given right-hand side of* (1)*, let the sequence $(\mathcal{T}_\nu)_{\nu\in\mathbb{N}}$ of decompositions of Ω be admissible, shape-regular and not necessarily globally quasi-uniform, and let the constant time step size $\tau \in \left(0, \frac{T}{2}\right]$ be given. In addition, let the unique solution u of Theorem 1.5 satisfy $u \in C^4([0,T]; H_0^1(\Omega))$. Then, the Crank-Nicolson method* (27), (28) *fulfils*

- *the $L^2(\Omega)$ error estimate*

$$
\begin{aligned}
&\max_{k=1,\dots,N_t-1}\left\|U^{k+1/2} - u(\cdot, t^{k+1/2})\right\|_{L^2(\Omega)} \le C_P\left[2\left\|\mathrm{D}_\tau^+(U^0 - Q_h^1 u(\cdot, t^0))\right\|_{L^2(\Omega)}\right. \\
&\quad + \left|U^0 - Q_h^1 u_0\right|_{H^1(\Omega)} + \left|U^1 - Q_h^1 u(\cdot, t^1)\right|_{H^1(\Omega)} + \tau^2 C_P T\left\|\partial_t^4 u\right\|_{C([0,T];H_0^1(\Omega))} \\
&\quad \left. + 2T\left\|\partial_{tt} u - Q_h^1 \partial_{tt} u\right\|_{C([0,T];L^2(\Omega))}\right] + \left\|Q_h^1 u - u\right\|_{C([0,T];L^2(\Omega))} + \frac{\tau^2}{8}\|\partial_{tt} u\|_{C([0,T];L^2(\Omega))},
\end{aligned}
$$

- *the $H_0^1(\Omega)$ error estimate*

$$\max_{k=1,\dots,N_t-1} \left|U^{k+1/2} - u(\cdot,t^{k+1/2})\right|_{H^1(\Omega)} \le 2\left\|\mathrm{D}_\tau^+(U^0 - Q_h^1 u(\cdot,t^0))\right\|_{L^2(\Omega)}$$
$$+ \left|U^0 - Q_h^1 u_0\right|_{H^1(\Omega)} + \left|U^1 - Q_h^1 u(\cdot,t^1)\right|_{H^1(\Omega)} + \tau^2 C_P T \left\|\partial_t^4 u\right\|_{C([0,T];H_0^1(\Omega))}$$
$$+ 2T\left\|\partial_{tt} u - Q_h^1 \partial_{tt} u\right\|_{C([0,T];L^2(\Omega))} + \left\|Q_h^1 u - u\right\|_{C([0,T];H_0^1(\Omega))} + \frac{\tau^2}{8}\|\partial_{tt} u\|_{C([0,T];H_0^1(\Omega))},$$

- *the $L^2(\Omega)$ error estimate of the discrete time derivative*

$$\max_{k=1,\dots,N_t-1} \left\|\mathrm{D}_\tau^+(U^k - u(\cdot,t^k))\right\|_{L^2(\Omega)} \le 2\left\|\mathrm{D}_\tau^+(U^0 - Q_h^1 u(\cdot,t^0))\right\|_{L^2(\Omega)}$$
$$+ \left|U^0 - Q_h^1 u_0\right|_{H^1(\Omega)} + \left|U^1 - Q_h^1 u(\cdot,t^1)\right|_{H^1(\Omega)}$$
$$+ \tau^2 C_P T\left\|\partial_t^4 u\right\|_{C([0,T];H_0^1(\Omega))} + 2T\left\|\partial_{tt} u - Q_h^1\partial_{tt} u\right\|_{C([0,T];L^2(\Omega))}$$
$$+ \left\|Q_h^1 \partial_t u - \partial_t u\right\|_{C([0,T];L^2(\Omega))} + \frac{\tau}{2}\left\|Q_h^1 \partial_{tt} u - \partial_{tt} u\right\|_{C([0,T];L^2(\Omega))},$$

where $Q_h^1 \colon H_0^1(\Omega) \to S_{h,0}^1(\Omega)$ is the $H_0^1(\Omega)$ projection (7) *and*

$$t^{k+1/2} = \frac{t^{k+1} + t^k}{2} = t^k + \frac{\tau}{2}, \quad k = 0,\dots,N_t - 1.$$

Proof. For $k \in \{0,\dots,N_t\}$ and $x \in \Omega$, the pointwise error is written as

$$U^k(x) - u(x,t^k) = \underbrace{U^k(x) - Q_h^1 u(x,t^k)}_{=:\theta^k(x)} + \underbrace{Q_h^1 u(x,t^k) - u(x,t^k)}_{=:\rho(x,t^k)=:\rho^k(x)} = \theta^k(x) + \rho^k(x).$$

With this notation, the averaged pointwise error is

$$\begin{aligned}
U^{k+1/2}(x) - \frac{u(x,t^{k+1}) + u(x,t^k)}{2} &= \frac{U^{k+1}(x) + U^k(x)}{2} - \frac{u(x,t^{k+1}) + u(x,t^k)}{2} \\
&= \frac{1}{2}\left[(U^{k+1}(x) - u(x,t^{k+1})) + (U^k(x) - u(x,t^k))\right] \\
&= \frac{1}{2}\left[(\theta^{k+1}(x) + \rho^{k+1}(x)) + (\theta^k(x) + \rho^k(x))\right] \\
&= \theta^{k+1/2}(x) + \rho^{k+1/2}(x)
\end{aligned}$$

for $k \in \{0,\dots,N_t - 1\}$ and $x \in \Omega$. The quantities $\theta^k \in S_{h,0}^1(\Omega)$ and $\rho^k \in H_0^1(\Omega)$ are treated separately.

First, for $k \in \{1,\dots,N_t - 1\}$, the finite element function $\theta^k \in S_{h,0}^1(\Omega)$ is estimated. Therefore, with the notation (8), the Crank-Nicolson method (28), the $H_0^1(\Omega)$ projection

(7) and the variational formulation (2), it follows that

$$\begin{aligned}
&\left\langle \mathrm{D}_\tau^2\theta^k, v_h\right\rangle_{L^2(\Omega)} + \frac{1}{4}\left\langle \nabla_x\theta^{k+1} + 2\nabla_x\theta^k + \nabla_x\theta^{k-1}, \nabla_x v_h\right\rangle_{L^2(\Omega)} \\
&= \underbrace{\left\langle \mathrm{D}_\tau^2 U^k, v_h\right\rangle_{L^2(\Omega)} + \frac{1}{4}\left\langle \nabla_x U^{k+1} + 2\nabla_x U^k + \nabla_x U^{k-1}, \nabla_x v_h\right\rangle_{L^2(\Omega)}}_{\overset{(28)}{=}\frac{1}{4}\left\langle f^{k+1}+2f^k+f^{k-1},v_h\right\rangle_{L^2(\Omega)}} \\
&\quad - \left\langle \mathrm{D}_\tau^2(Q_h^1 u)^k, v_h\right\rangle_{L^2(\Omega)} - \underbrace{\frac{1}{4}\left\langle \nabla_x(Q_h^1 u)^{k+1} + 2\nabla_x(Q_h^1 u)^k + \nabla_x(Q_h^1 u)^{k-1}, \nabla_x v_h\right\rangle_{L^2(\Omega)}}_{\overset{(7)}{=}\frac{1}{4}\left\langle \nabla_x u^{k+1}+2\nabla_x u^k+\nabla_x u^{k-1},\nabla_x v_h\right\rangle_{L^2(\Omega)}} \\
&= \frac{1}{4}\left\langle f^{k+1} + 2f^k + f^{k-1}, v_h\right\rangle_{L^2(\Omega)} - \left\langle \mathrm{D}_\tau^2(Q_h^1 u)^k, v_h\right\rangle_{L^2(\Omega)} \\
&\quad - \underbrace{\frac{1}{4}\left\langle \nabla_x u^{k+1} + 2\nabla_x u^k + \nabla_x u^{k-1}, \nabla_x v_h\right\rangle_{L^2(\Omega)}}_{\overset{(2)}{=}\frac{1}{4}\left\langle f^{k+1}+2f^k+f^{k-1},v_h\right\rangle_{L^2(\Omega)}-\frac{1}{4}\left\langle(\partial_{tt}u)^{k+1}+2(\partial_{tt}u)^k+(\partial_{tt}u)^{k-1},v_h\right\rangle_{L^2(\Omega)}}
\end{aligned}$$

and further,

$$\begin{aligned}
&\left\langle \mathrm{D}_\tau^2\theta^k, v_h\right\rangle_{L^2(\Omega)} + \frac{1}{4}\left\langle \nabla_x\theta^{k+1} + 2\nabla_x\theta^k + \nabla_x\theta^{k-1}, \nabla_x v_h\right\rangle_{L^2(\Omega)} \\
&\qquad = \frac{1}{4}\left\langle (\partial_{tt}u)^{k+1} + 2(\partial_{tt}u)^k + (\partial_{tt}u)^{k-1}, v_h\right\rangle_{L^2(\Omega)} - \left\langle \mathrm{D}_\tau^2(Q_h^1 u)^k, v_h\right\rangle_{L^2(\Omega)} \\
&\qquad \quad \underbrace{+\left\langle \mathrm{D}_\tau^2 u^k, v_h\right\rangle_{L^2(\Omega)} - \left\langle \mathrm{D}_\tau^2 u^k, v_h\right\rangle_{L^2(\Omega)}}_{=0} \\
&\qquad = \left\langle (\partial_{tt}u)^k - \mathrm{D}_\tau^2 u^k - \mathrm{D}_\tau^2\rho^k + \frac{1}{4}\left[(\partial_{tt}u)^{k+1} - 2(\partial_{tt}u)^k + (\partial_{tt}u)^{k-1}\right], v_h\right\rangle_{L^2(\Omega)} \\
&\qquad = \left\langle (\partial_{tt}u)^k - \mathrm{D}_\tau^2 u^k - \mathrm{D}_\tau^2\rho^k + \frac{\tau}{4}\left[\mathrm{D}_\tau^+(\partial_{tt}u)^k - \mathrm{D}_\tau^-(\partial_{tt}u)^k\right], v_h\right\rangle_{L^2(\Omega)}
\end{aligned}$$

for all $v_h \in S_{h,0}^1(\Omega)$ and for $k = 1,\ldots,N_t - 1$. Hence, the functions $\theta^k \in S_{h,0}^1(\Omega)$, $k = 0,1,\ldots,N_t$, are the solutions of the Crank-Nicolson method (27), (28) with

$$\theta^0 = U^0 - Q_h^1 u(\cdot,0), \quad \theta^1 = U^1 - Q_h^1 u(\cdot,t^1)$$

for the right-hand side $g \in C([0,T];L^2(\Omega))$ with

$$g(x,t) := \partial_{tt}u(x,t) - \mathrm{D}_\tau^2 u(x,t) - \mathrm{D}_\tau^2\rho(x,t) + \frac{\tau}{4}\left[\mathrm{D}_\tau^+\partial_{tt}u(x,t) - \mathrm{D}_\tau^-\partial_{tt}u(x,t)\right]$$

for $(x,t) \in \Omega \times [\tau, T-\tau]$,

$$g(x,t) := \partial_{tt}u(x,\tau) - \mathrm{D}_\tau^2 u(x,\tau) - \mathrm{D}_\tau^2\rho(x,\tau) + \frac{\tau}{4}\left[\mathrm{D}_\tau^+\partial_{tt}u(x,\tau) - \mathrm{D}_\tau^-\partial_{tt}u(x,\tau)\right]$$

for $(x,t) \in \Omega \times [0,\tau)$ and

$$g(x,t) := \partial_{tt}u(x,T-\tau) - \mathrm{D}_\tau^2 u(x,T-\tau) - \mathrm{D}_\tau^2 \rho(x,T-\tau) \\ + \frac{\tau}{4}\left[\mathrm{D}_\tau^+ \partial_{tt}u(x,T-\tau) - \mathrm{D}_\tau^- \partial_{tt}u(x,T-\tau)\right]$$

for $(x,t) \in \Omega \times (T-\tau, T]$, where $\rho = Q_h^1 u - u$. So, the stability of the Crank-Nicolson method (Theorem 2.24) for θ^k gives the estimate

$$\left\|\mathrm{D}_\tau^+\theta^k\right\|_{L^2(\Omega)} + \left|\theta^{k+1/2}\right|_{H^1(\Omega)} \le 2\left\|\mathrm{D}_\tau^+\theta^0\right\|_{L^2(\Omega)} + 2\left|\theta^{1/2}\right|_{H^1(\Omega)} \\ + \underbrace{\sum_{j=1}^{k} \frac{\tau}{2}\left\|g^{j+1} + 2g^j + g^{j-1}\right\|_{L^2(\Omega)}}_{\le \max_{j=0,\ldots,k+1}\|g^j\|_{L^2(\Omega)} \sum_{j=1}^k 2\tau} \tag{34}$$

for $k = 1, \ldots, N_t - 1$. For the last term in (34), the triangle inequality, Lemma 2.8 with $\Theta_1(x,t^j,\tau)$, $\Theta_2(x,t^j,\tau) \in [-1,1]$, Lemma 2.7 with $\Theta_3(x,t^j,\tau)$, $\Theta_4(x,t^j,\tau) \in (0,1)$ and the Poincaré inequality (3) with constant $C_P > 0$ yield

$$\begin{aligned}
&\left\|g^j\right\|_{L^2(\Omega)} = \left\|(\partial_{tt}u)^j - \mathrm{D}_\tau^2 u^j - \mathrm{D}_\tau^2\rho^j + \frac{\tau}{4}\left[\mathrm{D}_\tau^+(\partial_{tt}u)^j - \mathrm{D}_\tau^-(\partial_{tt}u)^j\right]\right\|_{L^2(\Omega)} \\
&\le \left\|(\partial_{tt}u)^j - \mathrm{D}_\tau^2 u^j\right\|_{L^2(\Omega)} + \left\|\mathrm{D}_\tau^2\rho^j\right\|_{L^2(\Omega)} + \frac{\tau}{4}\left\|\mathrm{D}_\tau^+(\partial_{tt}u)^j - \mathrm{D}_\tau^-(\partial_{tt}u)^j\right\|_{L^2(\Omega)} \\
&= \left\|-\frac{\tau^2}{12}\partial_t^4 u(\cdot, t^j + \tau\,\Theta_1(\cdot,t^j,\tau))\right\|_{L^2(\Omega)} + \left\|\mathrm{D}_\tau^2\rho(\cdot,t^j)\right\|_{L^2(\Omega)} \\
&+ \frac{\tau}{4}\left\|\partial_t^3 u(\cdot,t^j) + \frac{\tau}{2}\partial_t^4 u(\cdot,t^j + \tau\,\Theta_3(\cdot,t^j,\tau)) - \partial_t^3 u(\cdot,t^j) + \frac{\tau}{2}\partial_t^4 u(\cdot,t^j - \tau\,\Theta_4(\cdot,t^j,\tau))\right\|_{L^2(\Omega)} \\
&\le \frac{\tau^2}{3}\left\|\partial_t^4 u\right\|_{C([0,T];L^2(\Omega))} + \left\|\partial_{tt}\rho(\cdot,t^j) + \frac{\tau^2}{12}\partial_t^4\rho(\cdot,t^j + \tau\,\Theta_2(\cdot,t^j,\tau))\right\|_{L^2(\Omega)} \\
&\le \frac{\tau^2}{3}\left\|\partial_t^4 u\right\|_{C([0,T];L^2(\Omega))} + \left\|\partial_{tt}\rho\right\|_{C([0,T];L^2(\Omega))} + \frac{\tau^2}{12}\max_{0\le t\le T}\underbrace{\left\|\partial_t^4\rho(\cdot,t)\right\|_{L^2(\Omega)}}_{\le C_P\left|\partial_t^4\rho(\cdot,t)\right|_{H^1(\Omega)}}
\end{aligned}$$

and further, Lemma 2.5, the $H_0^1(\Omega)$ stability of Q_h^1 (Lemma 2.4) and the Poincaré inequality (3) with constant $C_P > 0$ give

$$\left\|g^j\right\|_{L^2(\Omega)} \le \frac{\tau^2}{3}\underbrace{\left\|\partial_t^4 u\right\|_{C([0,T];L^2(\Omega))}}_{\le C_P\left\|\partial_t^4 u\right\|_{C([0,T];H_0^1(\Omega))}} + \left\|\partial_{tt}u - Q_h^1\partial_{tt}u\right\|_{C([0,T];L^2(\Omega))} \\ + \frac{\tau^2}{12}C_P\max_{0\le t\le T}\underbrace{\left|Q_h^1\partial_t^4 u(\cdot,t) - \partial_t^4 u(\cdot,t)\right|_{H^1(\Omega)}}_{\le 2\left|\partial_t^4 u(\cdot,t)\right|_{H^1(\Omega)}} \tag{35}$$

for $j = 1, \ldots, N_t - 1$. Since the relations

$$g^0 = g^1 \quad \text{and} \quad g^{N_t-1} = g^{N_t}$$

hold true, the estimate (35) is also valid for $j = 0, N_t$, i.e. the estimate

$$\|g^j\|_{L^2(\Omega)} \leq \frac{\tau^2}{2} C_P \|\partial_t^4 u\|_{C([0,T];H_0^1(\Omega))} + \|\partial_{tt} u - Q_h^1 \partial_{tt} u\|_{C([0,T];L^2(\Omega))}$$

holds true for $j = 0, \ldots, N_t$. With the last estimate, it follows for (34) that

$$\begin{aligned}
\|\mathrm{D}_\tau^+ \theta^k\|_{L^2(\Omega)} + |\theta^{k+1/2}|_{H^1(\Omega)} &\leq 2\|\mathrm{D}_\tau^+ \theta^0\|_{L^2(\Omega)} + 2 \underbrace{|\theta^{1/2}|_{H^1(\Omega)}}_{=\left|\frac{\theta^1+\theta^0}{2}\right|_{H^1(\Omega)}} + \max_{j=0,\ldots,k+1} \|g^j\|_{L^2(\Omega)} \underbrace{\sum_{j=1}^{k} 2\tau}_{\leq 2T} \\
&\leq 2\|\mathrm{D}_\tau^+ \theta^0\|_{L^2(\Omega)} + |\theta^0|_{H^1(\Omega)} + |\theta^1|_{H^1(\Omega)} \\
&\quad + \tau^2 C_P T \|\partial_t^4 u\|_{C([0,T];H_0^1(\Omega))} + 2T \|\partial_{tt} u - Q_h^1 \partial_{tt} u\|_{C([0,T];L^2(\Omega))}
\end{aligned} \tag{36}$$

for $k = 1, \ldots, N_t - 1$.

Second, for $k \in \{0, \ldots, N_t - 1\}$, the terms with the function $\rho^{k+1/2} = \frac{\rho(\cdot,t^{k+1}) + \rho(\cdot,t^k)}{2} \in H_0^1(\Omega)$ are estimated by

$$\|\rho^{k+1/2}\|_{L^2(\Omega)} \leq \frac{1}{2}\|\rho^{k+1}\|_{L^2(\Omega)} + \frac{1}{2}\|\rho^k\|_{L^2(\Omega)} \leq \|Q_h^1 u - u\|_{C([0,T];L^2(\Omega))}$$

and analogously,

$$|\rho^{k+1/2}|_{H^1(\Omega)} \leq \|Q_h^1 u - u\|_{C([0,T];H_0^1(\Omega))}.$$

Third, for $k \in \{0, \ldots, N_t - 1\}$, the function $\mathrm{D}_\tau^+ \rho^k \in H_0^1(\Omega)$ is estimated. With Lemma 2.7 with $\Theta(x, t^k, \tau) \in (0, 1)$ and Lemma 2.5, one concludes that

$$\begin{aligned}
\|\mathrm{D}_\tau^+ \rho^k\|_{L^2(\Omega)} = \|\mathrm{D}_\tau^+ \rho(\cdot, t^k)\|_{L^2(\Omega)} &= \left\|\partial_t \rho(\cdot, t^k) + \frac{\tau}{2} \partial_{tt} \rho(\cdot, t^k + \tau\Theta(\cdot, t^k, \tau))\right\|_{L^2(\Omega)} \\
&\leq \|Q_h^1 \partial_t u - \partial_t u\|_{C([0,T];L^2(\Omega))} + \frac{\tau}{2}\|Q_h^1 \partial_{tt} u - \partial_{tt} u\|_{C([0,T];L^2(\Omega))}
\end{aligned} \tag{37}$$

for $k \in \{0, \ldots, N_t - 1\}$.

Fourth, for $k \in \{0, \ldots, N_t - 1\}$, Taylor' theorem yields

$$u(\cdot, t^k) = u(\cdot, t^{k+1/2}) - \partial_t u(\cdot, t^{k+1/2})\frac{\tau}{2} + \partial_{tt} u(\cdot, t^{k+1/2} - \frac{\tau}{2}\tilde{\Theta}(\cdot, t^{k+1/2}, \tau))\frac{\tau^2}{8}$$

and

$$u(\cdot, t^{k+1}) = u(\cdot, t^{k+1/2}) + \partial_t u(\cdot, t^{k+1/2})\frac{\tau}{2} + \partial_{tt} u(\cdot, t^{k+1/2} + \frac{\tau}{2}\hat{\Theta}(\cdot, t^{k+1/2}, \tau))\frac{\tau^2}{8}$$

with $\tilde{\Theta}(x,t^{k+1/2},\tau), \hat{\Theta}(x,t^{k+1/2},\tau) \in (0,1)$. Adding these equalities up gives

$$\begin{aligned}&\frac{u(\cdot,t^{k+1})+u(\cdot,t^k)}{2}\\&= u(\cdot,t^{k+1/2}) + \frac{\tau^2}{16}\left[\partial_{tt}u(\cdot,t^{k+1/2}-\frac{\tau}{2}\tilde{\Theta}(\cdot,t^{k+1/2},\tau)) + \partial_{tt}u(\cdot,t^{k+1/2}+\frac{\tau}{2}\hat{\Theta}(\cdot,t^{k+1/2},\tau))\right]\end{aligned}$$

and the triangle inequality states

$$\left\|U^{k+1/2}-u(\cdot,t^{k+1/2})\right\|_{L^2(\Omega)} \le \left\|U^{k+1/2}-\frac{u(\cdot,t^{k+1})+u(\cdot,t^k)}{2}\right\|_{L^2(\Omega)} + \frac{\tau^2}{8}\|\partial_{tt}u\|_{C([0,T];L^2(\Omega))} \tag{38}$$

for $k=0,\ldots,N_t-1$. Analogously, the estimate

$$\left|U^{k+1/2}-u(\cdot,t^{k+1/2})\right|_{H^1(\Omega)} \le \left|U^{k+1/2}-\frac{u(\cdot,t^{k+1})+u(\cdot,t^k)}{2}\right|_{H^1(\Omega)} + \frac{\tau^2}{8}\|\partial_{tt}u\|_{C([0,T];H_0^1(\Omega))} \tag{39}$$

holds true for $k=0,\ldots,N_t-1$.

Hence, the assertion follows due to

$$\begin{aligned}\left\|U^{k+1/2}-\frac{u(\cdot,t^{k+1})+u(\cdot,t^k)}{2}\right\|_{L^2(\Omega)} &\le \left\|\theta^{k+1/2}\right\|_{L^2(\Omega)} + \left\|\rho^{k+1/2}\right\|_{L^2(\Omega)}\\&\le C_P\left|\theta^{k+1/2}\right|_{H^1(\Omega)} + \left\|\rho^{k+1/2}\right\|_{L^2(\Omega)},\\ \left|U^{k+1/2}-\frac{u(\cdot,t^{k+1})+u(\cdot,t^k)}{2}\right|_{H^1(\Omega)} &\le \left|\theta^{k+1/2}\right|_{H^1(\Omega)} + \left|\rho^{k+1/2}\right|_{H^1(\Omega)},\\ \left\|\mathrm{D}_\tau^+(U^k-u(\cdot,t^k))\right\|_{L^2(\Omega)} &\le \left\|\mathrm{D}_\tau^+\theta^k\right\|_{L^2(\Omega)} + \left\|\mathrm{D}_\tau^+\rho^k\right\|_{L^2(\Omega)}\end{aligned}$$

for $k=1,\ldots,N_t-1$ with the help of (36), (37), the Poincaré inequality (3) and (38) and (39). □

The Theorem 2.26 motivates to choose the same initialisation procedure (23) as for the leapfrog method in Subsection 2.3, see also Remark 2.21. Hence, with Lemma 2.19, the following corollary holds true.

Corollary 2.27. *Let the assumptions of Theorem 2.26 and Lemma 2.19 be satisfied. In addition, assume that the solution u is sufficiently smooth and the domain $\Omega\subset\mathbb{R}^d$ is sufficiently regular, e.g., convex or smooth. Then, the error estimates*

$$\begin{aligned}\max_{k=0,\ldots,N_t-1}\left\|U^{k+1/2}-u(\cdot,t^{k+1/2})\right\|_{L^2(\Omega)} &\le c\cdot(h^2+\tau^2),\\ \max_{k=0,\ldots,N_t-1}\left|U^{k+1/2}-u(\cdot,t^{k+1/2})\right|_{H^1(\Omega)} &\le c\cdot(h+\tau^2),\\ \max_{k=0,\ldots,N_t-1}\left\|\mathrm{D}_\tau^+(U^k-u(\cdot,t^k))\right\|_{L^2(\Omega)} &\le c\cdot(h^2+\tau^2)\end{aligned}$$

hold true with a constant $c>0$ independent of u and the mesh sizes h, τ.

The same numerical examples as for the leapfrog method in Subsection 2.3 are investigated. Hence, the domain and the time interval

$$\Omega = (0,1) \quad \text{and} \quad (0,T) = (0,11)$$

with the exact solutions

$$\begin{aligned} u_A(x,t) &= \sin(2\pi t)\sin(2\pi x) + \cos(\pi t)\sin(\pi x) && \text{for } (x,t) \in [0,1]\times[0,11], \\ u_B(x,t) &= x(1-x)\sin(\pi(3x-t)) && \text{for } (x,t) \in [0,1]\times[0,11] \end{aligned}$$

and the corresponding right-hand sides f_A, f_B and initial conditions $u_{A,0}$, $u_{A,1}$, $u_{B,0}$, $u_{B,1}$ are considered. Note that $f_A = 0$ and the total energy

$$\mathcal{E}_A(t) = \mathcal{E}_A(0) = \frac{1}{2}\|u_{A,1}\|^2_{L^2(\Omega)} + \frac{1}{2}\|\nabla_x u_{A,0}\|^2_{L^2(\Omega)} = \frac{5\pi^2}{4} \approx 12.337, \quad t \in [0,11],$$

is conserved, see Corollary 1.6. The uniform spatial mesh size and the constant time step size are given by

$$h = \frac{1}{N}, \quad N \in \mathbb{N}, \qquad \text{and} \qquad \tau = \frac{11}{N_t}, \quad N_t \in \mathbb{N}.$$

In Table 5, Table 6, Table 7 and Table 8, the errors of Corollary 2.27 for the solutions u_A, u_B are given for the Crank-Nicolson method (27), (28) with initialisation (23) when a uniform refinement strategy is applied for the temporal and spatial discretisations, i.e. N and N_t are doubled in each refinement step. The appearing integrals to compute the initialisation (23) and the related right-hand sides are calculated by using high-order quadrature rules. Table 5 and Table 7 show that the Crank-Nicolson method is stable and the error estimates of Corollary 2.27 are confirmed, where the mesh sizes are chosen such that the ratio

$$\frac{\tau}{h} = 0.4$$

is fulfilled. Additionally, in Figure 2 the total energy $\mathcal{E}_A(t) = \frac{5\pi^2}{4} \approx 12.337$ and the discrete energy $\mathcal{E}^k_{\mathrm{CN}}$ of the Crank-Nicolson method for the solution u_A are depicted for the refinement level 0 and level 3 with $\frac{\tau}{h} = 0.4$, where the discrete energy conservation of the Crank-Nicolson method (Theorem 2.23) is verified. Table 6 and Table 8 show that the Crank-Nicolson method is also stable and the error estimates of Corollary 2.27 are also confirmed, when the mesh sizes are chosen such that the ratio

$$\frac{\tau}{h} = \frac{22}{38} \approx 0.5789$$

is fulfilled. Note that in the last case, the leapfrog method is not stable, see Table 2 and Table 4.

h	τ	$\max_k \|\cdot\|_{L^2(\Omega)}$	eoc	$\max_k \|\cdot\|_{H^1(\Omega)}$	eoc	$\max_k \|\mathrm{D}_\tau^+(\cdot)\|_{L^2(\Omega)}$	eoc
0.50000	0.20000	1.07e+00	-	4.55e+00	-	5.11e+00	-
0.25000	0.10000	1.25e+00	-0.2	7.90e+00	-0.6	8.05e+00	-0.5
0.12500	0.05000	7.46e-01	0.67	4.73e+00	0.67	4.69e+00	0.70
0.06250	0.02500	2.09e-01	1.75	1.32e+00	1.76	1.29e+00	1.78
0.03125	0.01250	5.32e-02	1.93	3.40e-01	1.91	3.27e-01	1.93
0.01562	0.00625	1.33e-02	1.97	1.28e-01	1.39	8.20e-02	1.97
0.00781	0.00313	3.34e-03	1.99	6.39e-02	0.99	2.05e-02	1.99
0.00391	0.00156	8.34e-04	1.99	3.20e-02	1.00	5.13e-03	1.99
0.00195	0.00078	2.09e-04	2.00	1.60e-02	1.00	1.28e-03	2.00
0.00098	0.00039	5.22e-05	2.00	7.99e-03	1.00	3.20e-04	2.00

Table 5: Numerical results of the Crank-Nicolson method (27), (28), (23) with uniform refinement for the solution u_A with the ratio $\frac{\tau}{h} = 0.4$.

h	τ	$\max_k \|\cdot\|_{L^2(\Omega)}$	eoc	$\max_k \|\cdot\|_{H^1(\Omega)}$	eoc	$\max_k \|\mathrm{D}_\tau^+(\cdot)\|_{L^2(\Omega)}$	eoc
0.50000	0.28947	6.02e-01	-	3.00e+00	-	4.02e+00	-
0.25000	0.14474	8.08e-01	-0.3	5.20e+00	-0.6	5.61e+00	-0.4
0.12500	0.07237	3.62e-01	1.04	2.31e+00	1.05	2.24e+00	1.19
0.06250	0.03618	1.01e-01	1.76	6.49e-01	1.74	6.21e-01	1.76
0.03125	0.01809	2.57e-02	1.92	2.55e-01	1.31	1.58e-01	1.93
0.01562	0.00905	6.46e-03	1.97	1.28e-01	0.99	3.97e-02	1.97
0.00781	0.00452	1.62e-03	1.99	6.39e-02	0.99	9.94e-03	1.99
0.00391	0.00226	4.05e-04	1.99	3.20e-02	1.00	2.49e-03	1.99
0.00195	0.00113	1.01e-04	2.00	1.60e-02	1.00	6.22e-04	2.00
0.00098	0.00057	2.53e-05	2.00	7.99e-03	1.00	1.55e-04	2.00

Table 6: Numerical results of the Crank-Nicolson method (27), (28), (23) with uniform refinement for the solution u_A with the ratio $\frac{\tau}{h} = \frac{22}{38} \approx 0.5789$.

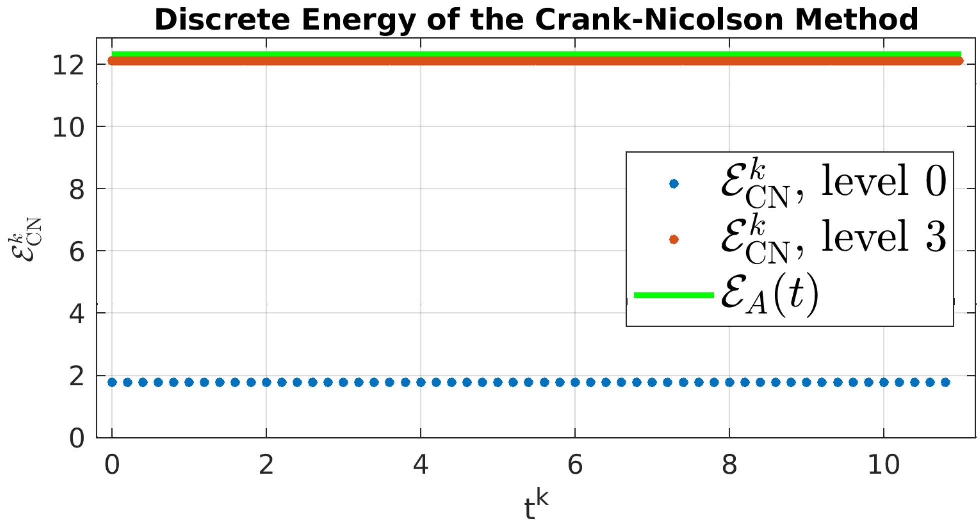

Figure 2: Total energy $\mathcal{E}_A(t) = \frac{5\pi^2}{4} \approx 12.337$ and discrete energy $\mathcal{E}^k_{\mathrm{CN}}$ of the Crank-Nicolson method (27), (28), (23) with uniform refinement for the solution u_A with ratio $\frac{\tau}{h} = 0.4$.

h	τ	$\max_k \|\cdot\|_{L^2(\Omega)}$	eoc	$\max_k \|\cdot\|_{H^1(\Omega)}$	eoc	$\max_k \|\mathrm{D}^+_\tau(\cdot)\|_{L^2(\Omega)}$	eoc
0.50000	0.20000	1.26e+00	-	4.18e+00	-	4.45e+00	-
0.25000	0.10000	1.23e-01	2.60	1.18e+00	1.42	3.99e-01	2.69
0.12500	0.05000	2.24e-02	2.21	5.21e-01	1.06	7.76e-02	2.13
0.06250	0.02500	5.88e-03	1.84	2.67e-01	0.92	2.08e-02	1.81
0.03125	0.01250	1.49e-03	1.94	1.35e-01	0.97	5.05e-03	2.00
0.01562	0.00625	3.71e-04	1.98	6.74e-02	0.99	1.27e-03	1.98
0.00781	0.00313	9.27e-05	1.99	3.37e-02	0.99	3.17e-04	1.99
0.00391	0.00156	2.32e-05	1.99	1.69e-02	1.00	7.92e-05	1.99
0.00195	0.00078	5.79e-06	2.00	8.43e-03	1.00	1.98e-05	2.00
0.00098	0.00039	1.45e-06	2.00	4.22e-03	1.00	4.95e-06	2.00

Table 7: Numerical results of the Crank-Nicolson method (27), (28), (23) with uniform refinement for the solution u_B with the ratio $\frac{\tau}{h} = 0.4$.

h	τ	$\max_k \|\cdot\|_{L^2(\Omega)}$	eoc	$\max_k \|\cdot\|_{H^1(\Omega)}$	eoc	$\max_k \|\mathrm{D}_\tau^+(\cdot)\|_{L^2(\Omega)}$	eoc
0.50000	0.28947	1.26e+00	-	4.28e+00	-	4.72e+00	-
0.25000	0.14474	1.17e-01	2.65	1.16e+00	1.45	3.78e-01	2.82
0.12500	0.07237	2.23e-02	2.15	5.19e-01	1.05	7.39e-02	2.12
0.06250	0.03618	5.83e-03	1.85	2.67e-01	0.91	2.01e-02	1.79
0.03125	0.01809	1.47e-03	1.94	1.35e-01	0.97	5.01e-03	1.95
0.01562	0.00905	3.68e-04	1.98	6.74e-02	0.99	1.26e-03	1.97
0.00781	0.00452	9.20e-05	1.99	3.37e-02	0.99	3.15e-04	1.99
0.00391	0.00226	2.30e-05	1.99	1.69e-02	1.00	7.87e-05	1.99
0.00195	0.00113	5.75e-06	2.00	8.43e-03	1.00	1.97e-05	2.00
0.00098	0.00057	1.44e-06	2.00	4.22e-03	1.00	4.92e-06	2.00

Table 8: Numerical results of the Crank-Nicolson method (27), (28), (23) with uniform refinement for the solution u_B with the ratio $\frac{\tau}{h} = \frac{22}{38} \approx 0.5789$.

3 Space-Time Methods

Space-time methods are characterised by interpreting the time direction as an additional spatial coordinate. Thus, the time-dependent problem is discretised without separating the temporal and spatial variables, i.e. a space-time discretisation. This ansatz may lead to structured decompositions, i.e. tensor-product approaches, see Figure 3, or unstructured decompositions, see Figure 4, of the space-time domain Q. In general, the main advantages of space-time methods are space-time adaptivity, space-time parallelisation and the treatment of moving boundaries. At a first glance, a disadvantage is that a global linear system must be solved at once. Therefore, fast solvers and preconditioning are essential, which are not investigated in this work. However, space-time approximation methods depend strongly on the space-time variational formulations on the continuous level.

As model problem, consider the homogeneous Dirichlet problem for the wave equation,

$$\left.\begin{array}{rcll} \partial_{tt}u(x,t) - \Delta_x u(x,t) &=& f(x,t) & \text{for } (x,t) \in Q = \Omega \times (0,T), \\ u(x,t) &=& 0 & \text{for } (x,t) \in \Sigma = \partial\Omega \times [0,T], \\ u(x,0) = \partial_t u(x,0) &=& 0 & \text{for } x \in \Omega, \end{array}\right\} \tag{40}$$

where $\Omega \subset \mathbb{R}^d$, $d = 1,2,3$, is a bounded domain with Lipschitz boundary $\partial\Omega$, $T > 0$ is a terminal time and f is a given right-hand side. For the space-time variational formulation of the wave equation (40), the Sobolev spaces

$$H^{1,1}_{0;0,}(Q) := H^1_{0,}(0,T;L^2(\Omega)) \cap L^2(0,T;H^1_0(\Omega)) \subset H^1(Q)$$

and

$$H^{1,1}_{0;,0}(Q) := H^1_{,0}(0,T;L^2(\Omega)) \cap L^2(0,T;H^1_0(\Omega)) \subset H^1(Q)$$

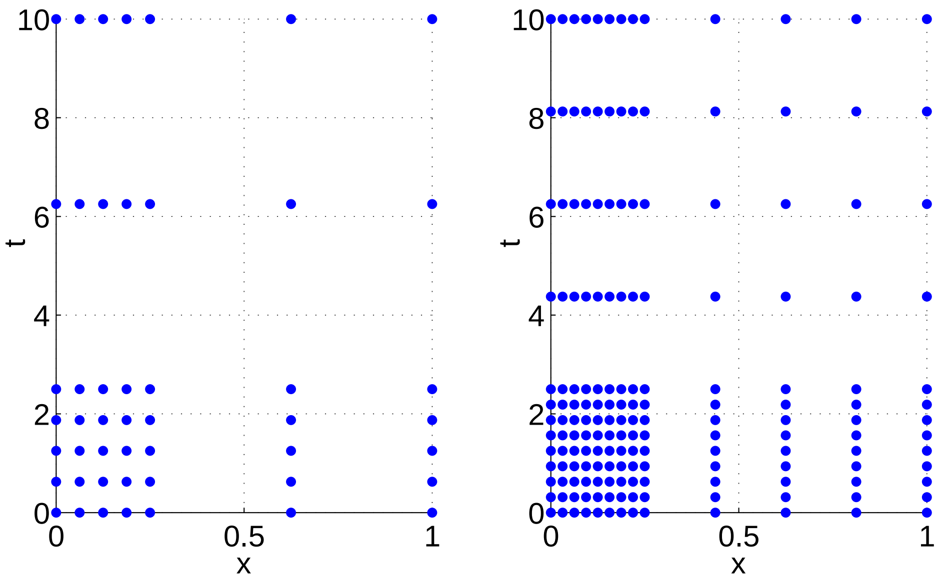

Figure 3: Nonuniform structured space-time meshes for a spatially one-dimensional domain: Starting mesh and the mesh after one uniform refinement step.

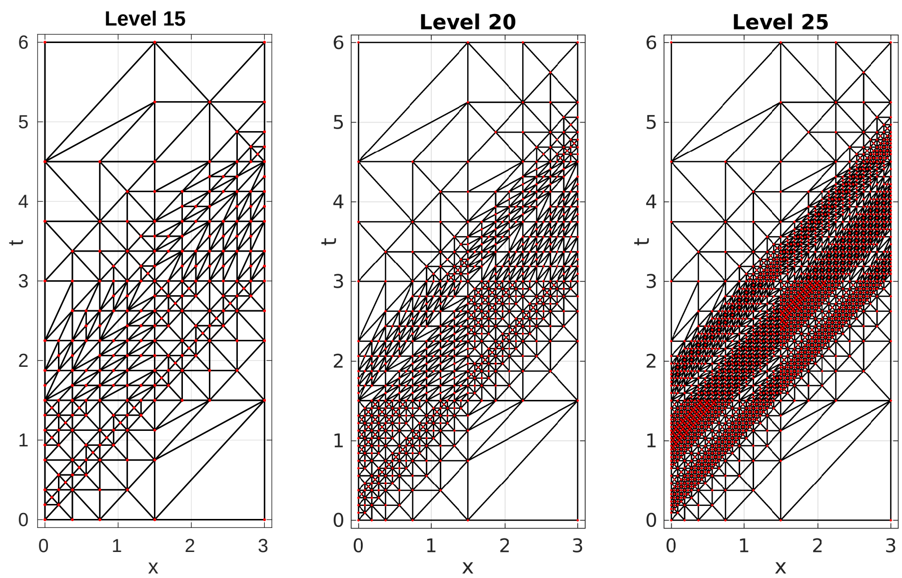

Figure 4: Unstructured space-time meshes for a spatially one-dimensional domain, resulting from an adaptive refinement strategy.

are endowed with the inner product

$$\langle u, v\rangle_{H^{1,1}_{0;}(Q)} := \int_0^T \int_\Omega \left(\partial_t u(x,t)\partial_t v(x,t) + \nabla_x u(x,t)\cdot\nabla_x v(x,t)\right) \mathrm{d}x\mathrm{d}t$$

and the induced norm

$$|u|_{H^1(Q)} := \sqrt{\langle u, u\rangle_{H^{1,1}_{0;}(Q)}} = \left\{ \int_0^T \int_\Omega \left(|\partial_t u(x,t)|^2 + \sum_{m=1}^d |\partial_{x_m} u(x,t)|^2 \right) \mathrm{d}x\mathrm{d}t \right\}^{1/2}$$

with the subspaces

$$H^1_{0,}(0,T;L^2(\Omega)) := \left\{ v \in H^1(0,T;L^2(\Omega)) \colon\ v(\cdot,0) = 0 \text{ in } L^2(\Omega) \right\} \subset C([0,T];L^2(\Omega))$$

and

$$H^1_{,0}(0,T;L^2(\Omega)) := \left\{ v \in H^1(0,T;L^2(\Omega)) \colon\ v(\cdot,T) = 0 \text{ in } L^2(\Omega) \right\} \subset C([0,T];L^2(\Omega)).$$

Note that in $H^{1,1}_{0;0,}(Q)$ and $H^{1,1}_{0;,0}(Q)$, the seminorm $|\cdot|_{H^1(Q)}$ is a to $\|\cdot\|_{H^1(Q)}$ equivalent norm due to the Poincaré inequality. The dual space $[H^{1,1}_{0;,0}(Q)]'$ is characterised as completion of $L^2(Q)$ with respect to the Hilbertian norm

$$\|f\|_{[H^{1,1}_{0;,0}(Q)]'} := \sup_{0\neq w\in H^{1,1}_{0;,0}(Q)} \frac{|\langle f, w\rangle_Q|}{|w|_{H^1(Q)}},$$

where $\langle\cdot,\cdot\rangle_Q$ denotes the duality pairing as extension of the inner product in $L^2(Q)$.

To derive a variational formulation in subspaces of $H^1(Q)$, take the second-order time derivative $\partial_{tt}u(x,t)$, multiply by a sufficiently smooth function $w(x,t)$ with $w(x,T) = 0$ and integrate via the space-time cylinder Q. Then, integration by parts gives

$$\begin{aligned}\int_0^T \int_\Omega \partial_{tt}u(x,t)\cdot w(x,t)\mathrm{d}x\mathrm{d}t = &-\int_0^T \int_\Omega \partial_t u(x,t)\cdot\partial_t w(x,t)\mathrm{d}x\mathrm{d}t \\ &+ \int_\Omega \partial_t u(x,T)\underbrace{w(x,T)}_{\stackrel{!}{=}0}\mathrm{d}x - \int_\Omega \underbrace{\partial_t u(x,0)}_{=0} w(x,0)\mathrm{d}x.\end{aligned}$$

So, one defines the bilinear form $a(\cdot,\cdot)\colon\ H^{1,1}_{0;0,}(Q)\times H^{1,1}_{0;,0}(Q)\to\mathbb{R}$,

$$a(u,w) := -\langle\partial_t u, \partial_t w\rangle_{L^2(Q)} + \langle\nabla_x u, \nabla_x w\rangle_{L^2(Q)}, \quad u\in H^{1,1}_{0;0,}(Q), \quad w\in H^{1,1}_{0;,0}(Q).$$

The boundedness of the bilinear form $a(\cdot,\cdot)$ is stated in the next lemma.

Lemma 3.1. *The bilinear form $a(\cdot,\cdot)\colon\ H^{1,1}_{0;0,}(Q)\times H^{1,1}_{0;,0}(Q)\to\mathbb{R}$ is bounded, i.e.*

$$|a(u,w)| \le |u|_{H^1(Q)}\,|w|_{H^1(Q)}$$

for all $u\in H^{1,1}_{0;0,}(Q)$, $w\in H^{1,1}_{0;,0}(Q)$.

Proof. The assertion follows immediately from the Cauchy-Schwarz inequality. □

The variational formulation of the wave equation (40) is to find $u \in H^{1,1}_{0;0,}(Q)$ such that

$$\forall w \in H^{1,1}_{0;,0}(Q): \quad a(u,w) = \langle f, w\rangle_{L^2(Q)}, \tag{41}$$

where $f \in L^2(Q)$ is given. Note that the initial condition $u(\cdot,0)=0$ is considered in the strong sense, whereas the initial condition $\partial_t u(\cdot,0)=0$ is incorporated in a weak sense.

Next, an existence and uniqueness result for the variational formulation (41) is given. Such a result is contained in [4, Chapter IV], see also [12].

Theorem 3.2. *Let $f \in L^2(Q)$ be given. Then, a unique solution $u \in H^{1,1}_{0;0,}(Q)$ of (41) exists, satisfying*

$$|u|_{H^1(Q)} \leq \frac{1}{\sqrt{2}} T \|f\|_{L^2(Q)}.$$

Proof. For the variational formulation (41), there exists a unique solution $u \in H^{1,1}_{0;0,}(Q)$, see [4, Chapter IV, Theorem 3.1, page 157, and Theorem 3.2, page 160]. The estimate follows by a Fourier series ansatz as in [4, Section 7, Chapter IV], see [12, Theorem 4.2.23, page 153]. □

The solution operator from Theorem 3.2 is not an isomorphism and hence, to derive from Theorem 3.2 an inf-sup condition, like

$$\forall u \in H^{1,1}_{0;0,}(Q): \qquad \sup_{0\neq w \in H^{1,1}_{0;,0}(Q)} \frac{|a(u,w)|}{|w|_{H^1(Q)}} \geq C_S \, |u|_{H^1(Q)} \tag{42}$$

with a constant $C_S > 0$, is not possible.

Theorem 3.3. *There does not exist a constant $C > 0$ such that each right-hand side $f \in L^2(Q)$ and the corresponding solution $u \in H^{1,1}_{0;0,}(Q)$ of (41) satisfy*

$$|u|_{H^1(Q)} \leq C \, \|f\|_{[H^{1,1}_{0;,0}(Q)]'}.$$

In particular, the inf-sup condition (42) does not hold true.

Proof. See [12, Theorem 4.2.24, page 155]. □

In the remainder of this section, conforming finite element discretisations for the variational formulation (41), resulting in Galerkin-Petrov schemes, are introduced and examined. For a discretization scheme, let the bounded Lipschitz domain $\Omega \subset \mathbb{R}^d$ be an interval $\Omega = (0, L)$ for $d = 1$, or polygonal for $d = 2$, or polyhedral for $d = 3$. For a tensor-product ansatz, admissible decompositions

$$\overline{Q} = \overline{\Omega} \times [0,T] = \bigcup_{i=1}^{N_x} \overline{\omega_i} \times \bigcup_{\ell=1}^{N_t} \overline{\tau_\ell}$$

are considered with $N := N_x \cdot N_t$ space-time elements, where the time intervals $\tau_\ell := (t_{\ell-1}, t_\ell) \subset \mathbb{R}$ with mesh sizes

$$h_{t,\ell} := t_\ell - t_{\ell-1} \quad \text{for } \ell = 1, \dots, N_t,$$

$h_t := h_{t,\max} := \max_\ell h_{t,\ell}$, $h_{t,\min} := \min_\ell h_{t,\ell}$, are defined via the decomposition

$$0 = t_0 < t_1 < t_2 < \cdots < t_{N_t - 1} < t_{N_t} = T$$

of the time interval $(0, T)$. For the spatial domain Ω, consider a shape-regular sequence $(\mathcal{T}_\nu)_{\nu \in \mathbb{N}}$ of admissible decompositions

$$\mathcal{T}_\nu := \{\omega_i \subset \mathbb{R}^d \colon i = 1, \dots, N_x\}$$

of Ω into finite elements $\omega_i \subset \mathbb{R}^d$ with mesh sizes

$$h_{x,i} := \left(\int_{\omega_i} \mathrm{d}x \right)^{1/d} \quad \text{for } i = 1, \dots, N_x,$$

$h_x := \max_i h_{x,i}$. The spatial elements ω_i are intervals for $d = 1$, triangles or quadrilaterals for $d = 2$, and tetrahedra or hexahedra for $d = 3$. Next, introduce the finite element space

$$Q_h^1(Q) := V_{h_x,0}^1(\Omega) \otimes S_{h_t}^1(0, T)$$

of piecewise multilinear, continuous functions, i.e.

$$V_{h_x,0}^1(\Omega) = \operatorname{span}\{\psi_j\}_{j=1}^{M_x} \subset H_0^1(\Omega), \quad S_{h_t}^1(0, T) = \operatorname{span}\{\varphi_\ell\}_{\ell=0}^{N_t} \subset H^1(0, T).$$

In fact, $V_{h_x,0}^1(\Omega)$ is either the space $S_{h_x}^1(\Omega) \cap H_0^1(\Omega)$ of piecewise linear, continuous functions on intervals ($d = 1$), triangles ($d = 2$), and tetrahedra ($d = 3$), or $V_{h_x,0}^1(\Omega)$ is the space $Q_{h_x}^1(\Omega) \cap H_0^1(\Omega)$ of piecewise linear/bilinear/trilinear, continuous functions on intervals ($d = 1$), quadrilaterals ($d = 2$), and hexahedra ($d = 3$).

The Galerkin-Petrov finite element discretisation of the variational formulation (41) is to find

$$u_h \in Q_h^1(Q) \cap H_{0;0,}^{1,1}(Q)$$

such that

$$a(u_h, w_h) = \langle f, w_h \rangle_{L^2(Q)} \tag{43}$$

for all $w_h \in Q_h^1(Q) \cap H_{0;,0}^{1,1}(Q)$. The approximate function u_h admits the representation

$$u_h(x,t) = \sum_{\ell=1}^{N_t} \sum_{j=1}^{M_x} u_j^\ell \psi_j(x) \varphi_\ell(t) = \sum_{j=1}^{M_x} U_{h_t,j}(t) \psi_j(x), \quad U_{h_t,j}(t) = \sum_{\ell=1}^{N_t} u_j^\ell \varphi_\ell(t)$$

for $(x,t) \in \overline{Q}$. After an appropriate ordering of the degrees of freedom, the discrete variational formulation (43) is equivalent to the global linear system

$$K_h \underline{u} = \underline{F}$$

with the system matrix

$$K_h = -A_{h_t} \otimes M_{h_x} + M_{h_t} \otimes A_{h_x} \in \mathbb{R}^{N_t \cdot M_x \times N_t \cdot M_x},$$

where $M_{h_x} \in \mathbb{R}^{M_x \times M_x}$ and $A_{h_x} \in \mathbb{R}^{M_x \times M_x}$ denote spatial mass and stiffness matrices given by

$$M_{h_x}[i,j] = \langle \psi_j, \psi_i \rangle_{L^2(\Omega)}, \quad A_{h_x}[i,j] = \langle \nabla_x \psi_j, \nabla_x \psi_i \rangle_{L^2(\Omega)}, \quad i,j = 1, \dots, M_x,$$

and the temporal stiffness matrix

$$A_{h_t} = \begin{pmatrix} \frac{-1}{h_{t,1}} & & & & & \\ \frac{1}{h_{t,1}} + \frac{1}{h_{t,2}} & \frac{-1}{h_{t,2}} & & & & \\ \frac{-1}{h_{t,2}} & \frac{1}{h_{t,2}} + \frac{1}{h_{t,3}} & \frac{-1}{h_{t,3}} & & & \\ & \ddots & \ddots & \ddots & & \\ & & \frac{-1}{h_{t,N_t-1}} & \frac{1}{h_{t,N_t-1}} + \frac{1}{h_{t,N_t}} & \frac{-1}{h_{t,N_t}} \end{pmatrix} \in \mathbb{R}^{N_t \times N_t},$$

the temporal mass matrix

$$M_{h_t} = \frac{1}{6} \begin{pmatrix} h_{t,1} & & & & \\ 2h_{t,1} + 2h_{t,2} & h_{t,2} & & & \\ h_{t,2} & 2h_{t,2} + 2h_{t,3} & h_{t,3} & & \\ & \ddots & \ddots & \ddots & \\ & & h_{t,N_t-1} & 2h_{t,N_t-1} + 2h_{t,N_t} & h_{t,N_t} \end{pmatrix} \in \mathbb{R}^{N_t \times N_t},$$

and with the corresponding vector $\underline{F} \in \mathbb{R}^{N_t \cdot M_x}$ of the right-hand side.

Remark 3.4. *Note that the Galerkin-Petrov finite element discretisation* (43) *can be realised as a two-step method.*

For a uniform time mesh size $h_t > 0$, stability of (43) in a certain sense can be proven, see [12, page 156-158], when the CFL condition

$$h_t < \sqrt{\frac{12}{c_{\text{inv}}}}\, h_x$$

is satisfied with the constant $c_{\text{inv}} > 0$ of the inverse inequality (Lemma 2.2) with respect to the spatial mesh, i.e. the sequence $(\mathcal{T}_\nu)_{\nu \in \mathbb{N}}$ of decompositions of Ω must be globally quasi-uniform.

In the particular case $d = 1$, it holds true that $c_{\text{inv}} = 12$, see the derivation of [10, (9.19), page 217] and therefore, stability follows for

$$h_t < h_x.$$

When $V^1_{h_x,0}(\Omega) \subset H^1_0(\Omega)$ is also of tensor-product structure, i.e.

$$V^1_{h_x,0}(\Omega) = \left(S^1_{h_{x_1}}(0, L_1) \otimes \ldots \otimes S^1_{h_{x_d}}(0, L_d) \right) \cap H^1_0(\Omega),$$

for example, when considering $\Omega = (0, L_1) \times \ldots \times (0, L_d) \subset \mathbb{R}^d$ with uniform mesh sizes $h_{x_1}, \ldots, h_{x_d}$, one concludes $c_{\text{inv}} = 12d$, and therefore, the stability condition

$$h_t < \frac{h_{x,\min}}{\sqrt{d}},$$

where $h_{x,\min} = \min\{h_{x_1}, \ldots, h_{x_d}\}$.

As a numerical example, consider for $d = 2$ the spatial domain $\Omega = (0, 1)^2$ with uniform discretisations with mesh sizes $h_x = h_{x_1} = h_{x_2}$ and the exact solution

$$u(x_1, x_2, t) = t^2 \sin(\pi x_1) \sin(\pi x_2) \quad \text{for } (x_1, x_2, t) \in Q = \Omega \times (0, T)$$

with different terminal times $T \in \left\{ \frac{7}{10}, \frac{1}{\sqrt{2}}, \frac{3}{4}, 1, 2 \right\}$. Then stability follows when choosing

$$\frac{h_t}{h_x} < \frac{1}{\sqrt{2}} \approx 0.7071068. \tag{44}$$

In Table 9, Table 10, Table 11, Table 12 and Table 13, the $L^2(Q)$ error, the $H^1(Q)$ error and the maximal and minimal singular values of the related system matrix K_h of the Galerkin-Petrov formulation (43) are given, where the observed convergence rates are as expected, provided the CFL condition (44) is satisfied, i.e. the CFL condition (44) seems to be sharp. Here, the number of the degrees of freedom is denoted by

$$\text{dof} = \dim Q_h^1(Q) \cap H^{1,1}_{0;0,}(Q) = \dim Q_h^1(Q) \cap H^{1,1}_{0;,0}(Q).$$

dof	h_x	h_t	$\|u - u_h\|_{L^2(Q)}$	eoc	$\|u - u_h\|_{H^1(Q)}$	eoc	$\sigma_{\max}(K_h)$	$\sigma_{\min}(K_h)$
2	0.500	0.3500	2.034e-02	-	4.204e-01	-	4.8e-01	4.7e-01
36	0.250	0.1750	4.737e-03	2.1	2.089e-01	1.0	9.4e-01	5.4e-02
392	0.125	0.0875	1.161e-03	2.0	1.040e-01	1.0	6.5e-01	7.0e-03
3600	0.062	0.0438	2.887e-04	2.0	5.193e-02	1.0	3.5e-01	1.2e-03
30752	0.031	0.0219	7.207e-05	2.0	2.595e-02	1.0	1.8e-01	2.3e-04
254016	0.016	0.0109	1.801e-05	2.0	1.298e-02	1.0	8.9e-02	3.6e-05

Table 9: Numerical results for the Galerkin-Petrov formulation (43) with uniform meshes for $Q = (0, 1)^2 \times (0, \frac{7}{10})$, satisfying the CFL condition (44).

dof	h_x	h_t	$\|u-u_h\|_{L^2(Q)}$	eoc	$\|u-u_h\|_{H^1(Q)}$	eoc	$\sigma_{\max}(K_h)$	$\sigma_{\min}(K_h)$
2	0.500	0.3536	2.094e-02	-	4.303e-01	-	4.7e-01	4.7e-01
36	0.250	0.1768	4.888e-03	2.1	2.141e-01	1.0	9.3e-01	5.2e-02
392	0.125	0.0884	1.198e-03	2.0	1.066e-01	1.0	6.4e-01	6.1e-03
3600	0.062	0.0442	2.981e-04	2.0	5.323e-02	1.0	3.4e-01	7.6e-04
30752	0.031	0.0221	7.444e-05	2.0	2.661e-02	1.0	1.8e-01	9.6e-05
254016	0.016	0.0110	1.860e-05	2.0	1.330e-02	1.0	8.8e-02	1.2e-05

Table 10: Numerical results for the Galerkin-Petrov formulation (43) with uniform meshes for $Q = (0,1)^2 \times (0, \frac{1}{\sqrt{2}})$ for the limit case of the CFL condition (44).

dof	h_x	h_t	$\|u-u_h\|_{L^2(Q)}$	eoc	$\|u-u_h\|_{H^1(Q)}$	eoc	$\sigma_{\max}(K_h)$	$\sigma_{\min}(K_h)$
2	0.500	0.3750	2.476e-02	-	4.937e-01	-	5.0e-01	4.3e-01
36	0.250	0.1875	5.862e-03	2.1	2.469e-01	1.0	8.7e-01	4.1e-02
392	0.125	0.0938	1.443e-03	2.0	1.231e-01	1.0	6.0e-01	2.7e-03
3600	0.062	0.0469	3.594e-04	2.0	6.153e-02	1.0	3.2e-01	4.5e-05
30752	0.031	0.0234	8.977e-05	2.0	3.076e-02	1.0	1.7e-01	3.5e-08
254016	0.016	0.0117	2.244e-05	2.0	1.538e-02	1.0	8.3e-02	5.7e-14

Table 11: Numerical results for the Galerkin-Petrov formulation (43) with uniform meshes for $Q = (0,1)^2 \times (0, \frac{3}{4})$, violating the CFL condition (44).

dof	h_x	h_t	$\|u-u_h\|_{L^2(Q)}$	eoc	$\|u-u_h\|_{H^1(Q)}$	eoc	$\sigma_{\max}(K_h)$	$\sigma_{\min}(K_h)$
2	0.500	0.5000	5.418e-02	-	9.782e-01	-	7.2e-01	2.7e-01
36	0.250	0.2500	1.353e-02	2.0	4.986e-01	1.0	7.2e-01	1.5e-02
392	0.125	0.1250	3.381e-03	2.0	2.502e-01	1.0	4.6e-01	1.7e-04
3600	0.062	0.0625	8.453e-04	2.0	1.252e-01	1.0	2.4e-01	6.6e-08
30752	0.031	0.0312	2.113e-04	2.0	6.263e-02	1.0	1.2e-01	2.8e-14
254016	0.016	0.0156	8.621e+08	-41.9	4.635e+11	-42.8	6.2e-02	≈ 0

Table 12: Numerical results for the Galerkin-Petrov formulation (43) with uniform meshes for $Q = (0,1)^2 \times (0,1)$, violating the CFL condition (44).

dof	h_x	h_t	$\|u-u_h\|_{L^2(Q)}$	eoc	$\|u-u_h\|_{H^1(Q)}$	eoc	$\sigma_{\max}(K_h)$	$\sigma_{\min}(K_h)$
2	0.500	1.0000	2.777e-01	-	5.638e+00	-	1.7e+00	1.8e-01
36	0.250	0.5000	7.355e-02	1.9	2.798e+00	1.0	1.5e+00	6.9e-03
392	0.125	0.2500	1.863e-02	2.0	1.404e+00	1.0	9.3e-01	2.4e-05
3600	0.062	0.1250	4.732e-03	2.0	7.028e-01	1.0	4.9e-01	7.0e-10
30752	0.031	0.0625	7.852e-01	-7.4	1.796e+02	-8.0	2.5e-01	≈ 0
254016	0.016	0.0312	1.710e+21	-70.9	7.642e+23	-71.8	1.2e-01	≈ 0

Table 13: Numerical results for the Galerkin-Petrov formulation (43) with uniform meshes for $Q = (0,1)^2 \times (0,2)$, violating the CFL condition (44).

References

[1] Arendt, W., and Urban, K. *Partielle Differenzialgleichungen. Eine Einführung in analytische und numerische Methoden.*, 2nd edition ed. Berlin: Springer Spektrum, 2018.

[2] Dautray, R., and Lions, J.-L. *Mathematical analysis and numerical methods for science and technology. Vol. 5.* Springer-Verlag, Berlin, 1992. Evolution problems. I, With the collaboration of Michel Artola, Michel Cessenat and Hélène Lanchon, Translated from the French by Alan Craig.

[3] Köthe, G. *Topological vector spaces. II.*, vol. 237. Springer, Berlin, 1979.

[4] Ladyzhenskaya, O. A. *The boundary value problems of mathematical physics*, vol. 49 of *Applied Mathematical Sciences.* Springer-Verlag, New York, 1985.

[5] Lions, J.-L., and Magenes, E. *Problèmes aux limites non homogènes et applications. Vol. 1.* Travaux et Recherches Mathématiques, No. 17. Dunod, Paris, 1968.

[6] Longva, A. B. *Finite element solutions to the wave equation in non-convex domains - A relaxation of the CFL condition in the presence of local mesh refinement.* Master thesis, NTNU, Trondheim, 2017.

[7] Raviart, P.-A., and Thomas, J. M. *Introduction à l'analyse numérique des équations aux dérivées partielles.* 1983.

[8] Ryan, R. A. *Introduction to tensor products of Banach spaces.* London: Springer, 2002.

[9] Schweizer, B. *Partielle Differentialgleichungen. Eine anwendungsorientierte Einführung.* Heidelberg: Springer Spektrum, 2018.

[10] Steinbach, O. *Numerical approximation methods for elliptic boundary value problems. Finite and boundary elements.* New York, NY: Springer, 2008.

[11] Wloka, J. *Partielle Differentialgleichungen.* B. G. Teubner, Stuttgart, 1982.

[12] Zank, M. *Inf–sup stable space–time methods for time–dependent partial differential equations.* volume 36 of Monographic Series TU Graz: Computation in Engineering and Science. Feb 2020.

[13] Zeidler, E. *Nonlinear functional analysis and its applications. II/A.* Springer-Verlag, New York, 1990.