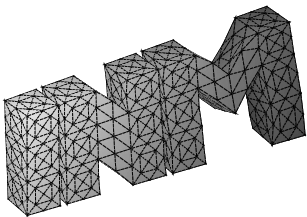

Space–time finite element methods
for parabolic problems

O. Steinbach



**Berichte aus dem
Institut für Numerische Mathematik**

Technische Universität Graz

Space–time finite element methods
for parabolic problems

O. Steinbach

**Berichte aus dem
Institut für Numerische Mathematik**

Bericht 2015/1

Technische Universität Graz
Institut für Numerische Mathematik
Steyrergasse 30
A 8010 Graz

WWW: <http://www.numerik.math.tu-graz.ac.at>

© Alle Rechte vorbehalten. Nachdruck nur mit Genehmigung des Autors.

Space–time finite element methods for parabolic problems

Olaf Steinbach

Institut für Numerische Mathematik, TU Graz,
Steyrergasse 30, 8010 Graz, Austria

`o.steinbach@tugraz.at`

Abstract

We propose and analyze a space–time finite element method for the numerical solution of parabolic evolution equations. This approach allows the use of general and unstructured space–time finite elements which do not require any tensor product structure. The stability of the numerical scheme is based on a stability condition which holds for standard finite element spaces. We also provide related a priori error estimates which are confirmed by numerical experiments.

1 Introduction

Standard approaches to the numerical solution of parabolic evolution problems are based on semi–discretizations, either in space resulting in the method of lines [20], or in time yielding the Rothe method [11]. In both cases, the underlying structure is some tensor product of space and time variables. The sequential solution of the resulting linear systems complicates parallel solution algorithms, and adaptive refinement strategies which are done simultaneously in space and time. Due to the increasing computing capacities, an overall solution of a space–time discretized problem may become an alternative to think about. Applications of space–time finite element methods have already a long history [10]. However, equivalent parabolic operator equations were first analyzed in [17], see also [12, 21] for a discussion of the involved stability condition. Note that in all cases, space–time tensor product wavelet bases are used. Rather general space–time finite elements are already used in many applications, and also in the context of adaptive space–time mesh generation algorithms, see, e.g., [2, 13, 14, 15, 19], and the references given therein. Usually, discontinuous Galerkin methods are used, at least for time discretizations, which increase the number of degrees of freedom, and which involve certain parameters to be chosen to ensure stability of the numerical scheme. Therefore we are interested in the application of a standard finite element approach for the solution of evolution problems.

To give a motivation of our space–time finite element approach, we consider the Dirichlet boundary value problem for the heat equation in one spatial dimension,

$$\left. \begin{aligned} c_H \partial_t u(x, t) - \partial_{xx} u(x, t) &= 0 && \text{for } x \in (0, 1), t \in (0, T], \\ u(0, t) = u(1, t) &= 0, && \text{for } t \in (0, T], \\ u(x, 0) &= u_0(x) && \text{for } x \in (0, 1). \end{aligned} \right\} \quad (1.1)$$

We assume the compatibility condition $u_0(0) = u_0(1) = 0$, and $c_H > 0$ is a given heat capacity constant. It is well known that the solution of (1.1) allows the series representation

$$u(x, t) = \sum_{k=1}^{\infty} u_k e^{-(k\pi)^2 t / c_H} \sin k\pi x, \quad u_k = 2 \int_0^1 u_0(x) \sin k\pi x dx. \quad (1.2)$$

In particular for

$$u_0(x) = \begin{cases} 256x^2(1-4x)^2 & \text{for } x \in [0, \frac{1}{4}], \\ 0 & \text{for } x \in (\frac{1}{4}, 1] \end{cases} \quad (1.3)$$

we obtain

$$u_k = 1024 \frac{(192 - k^2 \pi^2)(1 - \cos \frac{k\pi}{4}) - 24k\pi \sin \frac{k\pi}{4}}{(k\pi)^5}.$$

The solution of the initial boundary value problem (1.1) with $c_H = 100$, $T = 2$, and subject to the initial condition (1.3) is depicted in Fig. 1. Now we may ask for a piecewise linear interpolation for this particular solution. For this we consider a sequence of uniform and admissible triangulations Q_h of refinement level L with mesh size h of the space–time domain $Q := (0, 1) \times (0, 2)$, i.e.

$$\overline{Q} = [0, 1] \times [0, 2] = \bigcup_{\ell=1}^N \overline{q}_\ell =: Q_h.$$

The triangular space–time finite elements q_ℓ are assumed to be shape–regular, see also Fig. 1. The related set of nodes is denoted by $\{(x_k, t_k)\}_{k=1}^M$. With respect to the triangulation Q_h we now define the piecewise linear nodal interpolation

$$I_h u(x, t) = \sum_{k=1}^M u(x_k, t_k) \varphi_k(x, t),$$

where $\varphi_k(x, t)$ are the usual two–dimensional continuous and piecewise linear basis functions. Since the solution (1.2) of the initial boundary value problem (1.1) turns out to be smooth, we can apply standard interpolation error estimates [4, 5, 18] to obtain

$$\|u - I_h u\|_{L_2(Q)} \leq c h^2 |u|_{H^2(Q)}. \quad (1.4)$$

Since the solution (1.2) is known, we can compute the interpolation error (1.4) for a sequence of nested finite element spaces. Hence we can estimate the order of convergence

(eoc) which confirms second order convergence, see Table 1. Starting from a coarse finite element mesh we can also drive an adaptive algorithm where we refine only those finite elements q_ℓ where the local error satisfies

$$\|u - I_h u\|_{L_2(q_\ell)} \geq \frac{1}{2} \max_{k=1, \dots, N} \|u - I_h u\|_{L_2(q_k)}.$$

To end up with an admissible triangulation, we apply a red–green–blue refinement strategy [22]. To obtain a comparable error, the number of nodes, i.e., the number of degrees of freedom, is approximately a fourth as for the case when using a uniform refinement strategy. The final mesh is given in Fig. 1. We observe that this adaptive mesh is completely unstructured, both in space and time. In Table 1 we present the interpolation errors for a sequence of space–time finite element meshes up to the seventh refinement level L . Obviously, second order convergence is confirmed. While in the case of a given function, i.e. of the given solution (1.2), we can easily determine the piecewise linear interpolations, we now may ask to find the optimal space–time finite element approximation by solving the initial boundary value problem (1.1) by using a space–time finite element method.

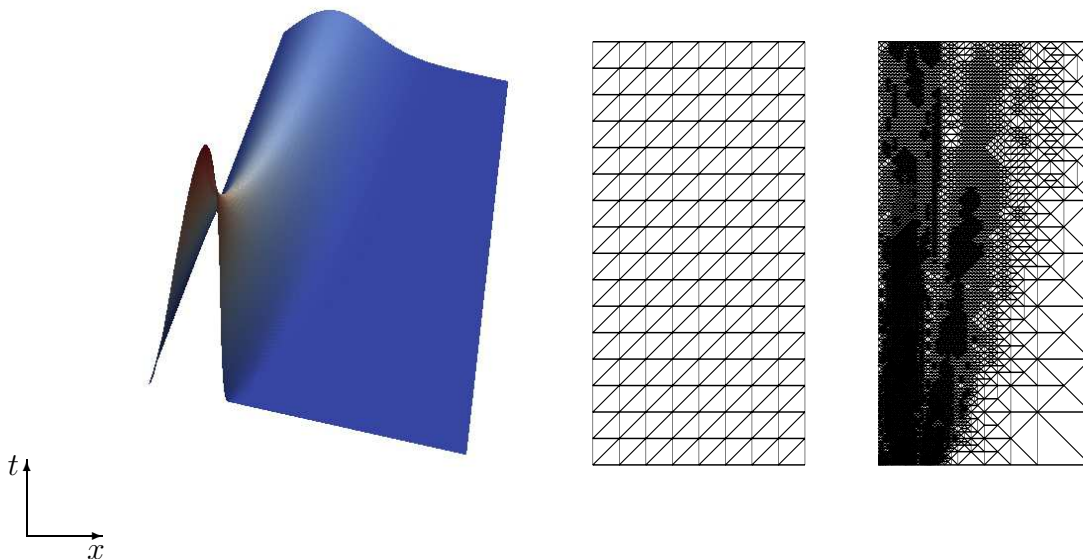


Figure 1: Solution, uniform mesh ($L = 3$), and adaptive mesh ($L = 15$).

In this paper we present a stability and a priori error analysis for a standard Galerkin finite element method for the numerical solution of parabolic evolution equations. In Sect. 2 we describe the model problem of a linear heat equation with Dirichlet boundary conditions, its Galerkin–Petrov variational formulation, and we prove a related stability condition which ensures unique solvability. The finite element discretisation is described and analysed in Sect. 3, where the main results are formulated in Theorem 3.1 on the

L	N	M	$\ u - I_h u\ _{L_2(Q)}$	eoc
Uniform refinement				
3	256	153	2.929 -2	
4	1024	561	1.081 -2	1.44
5	4096	2145	2.750 -3	1.97
6	16384	8385	6.870 -4	2.00
7	65536	33153	1.718 -4	2.00
Adaptive refinement				
15	19083	9645	1.973 -4	

Table 1: Interpolation error of the space–time solution.

stability of the proposed scheme, and in Theorem 3.3 on the a priori error estimate. For the particular case of a spatial one–dimensional problem we present a more detailed analysis on the interpolation error in anisotropic norms in Sect. 4. Numerical examples given in Sect. 5 confirm all theoretical results, and we also comment on a comparison with a space–time discontinuous Galerkin finite element approach as considered in [14]. We end this paper by some concluding remarks in Sect. 6, where we point out some open problems in the a priori error analysis, and we comment on challenging topics concerning a posteriori error estimation techniques and adaptive mesh refinement strategies, efficient preconditioned and parallel solution methods, space–time boundary element methods and their coupling with space–time finite element methods, and on applications including nonlinearities.

2 Variational formulations

As model problem we consider the Dirichlet problem for the heat equation,

$$\begin{aligned}
c_H \partial_t u(x, t) - \operatorname{div}_x [A(x, t) \nabla_x u(x, t)] &= f(x, t) & \text{for } (x, t) \in Q := \Omega \times (0, T), \\
u(x, t) &= 0 & \text{for } (x, t) \in \Sigma := \Gamma \times (0, T), \\
u(x, 0) &= u_0(x) & \text{for } x \in \Omega,
\end{aligned} \tag{2.1}$$

where $\Omega \subset \mathbb{R}^n$, $n = 1, 2, 3$, is a bounded domain with, for $n = 2, 3$, Lipschitz boundary $\Gamma := \partial\Omega$. Moreover, the coefficient matrix $A(x, t) = [A(x, t)]^\top \in \mathbb{R}^{n \times n}$ is assumed to be positive definite uniform in $(x, t) \in Q$; $c_H > 0$ is a given heat capacity constant, and we assume $u_0 \in H_0^1(\Omega)$. Finally, $f \in L_2(0, T; H^{-1}(\Omega))$ is a given source term.

The variational formulation of (2.1) is to find $u \in L_2(0, T; H_0^1(\Omega)) \cap H^1(0, T, H^{-1}(\Omega))$, $u(x, 0) = u_0(x)$ for $x \in \Omega$, such that

$$\begin{aligned}
c_H \int_0^T \int_\Omega \partial_t u(x, t) v(x, t) \, dx dt + \int_0^T \int_\Omega [A(x, t) \nabla_x u(x, t)] \cdot \nabla_x v(x, t) \, dx dt & \tag{2.2} \\
= \int_0^T \int_\Omega f(x, t) v(x, t) \, dx dt &
\end{aligned}$$

is satisfied for all $v \in L_2(0, T; H_0^1(\Omega))$, $v(x, 0) = 0$ for $x \in \Omega$.

Before we discuss the unique solvability of the variational problem (2.2), let us first consider the quasi-static elliptic Dirichlet boundary value problem, with $c_H > 0$ being the heat conductivity, for all $t \in (0, T)$, and for given $\phi(x, t)$,

$$-\operatorname{div}_x[A(x, t)\nabla_x w(x, t)] = c\phi(x, t) \quad \text{for } x \in \Omega, \quad w(x, t) = 0 \quad \text{for } x \in \Gamma. \quad (2.3)$$

The related variational problem is to find $w \in L_2(0, T; H_0^1(\Omega))$ such that

$$\int_0^T \int_{\Omega} [A(x, t)\nabla_x w(x, t)] \cdot \nabla_x v(x, t) \, dx dt = c_H \int_0^T \int_{\Omega} \phi(x, t)v(x, t) \, dx dt \quad (2.4)$$

is satisfied for all $v \in L_2(0, T; H_0^1(\Omega))$, where $\phi \in L_2(0, T; H^{-1}(\Omega))$ is given. Since the coefficient matrix $A(x, t)$ is assumed to be uniform positive definite,

$$\|v\|_{L_2(0, T; H_0^1(\Omega))} := \sqrt{\int_0^T \int_{\Omega} [A(x, t)\nabla_x v(x, t)] \cdot \nabla_x v(x, t) \, dx dt}$$

defines an equivalent norm in $L_2(0, T; H_0^1(\Omega))$. Therefore, unique solvability of the variational problem (2.4) follows and we conclude

$$\begin{aligned} \|w\|_{L_2(0, T; H_0^1(\Omega))}^2 &= \int_0^T \int_{\Omega} [A(x, t)\nabla_x w(x, t)] \cdot \nabla_x w(x, t) \, dx dt \\ &= c_H \int_0^T \int_{\Omega} \phi(x, t)w(x, t) \, dx dt \leq \|\phi\|_{L_2(0, T; H^{-1}(\Omega))} \|w\|_{L_2(0, T; H_0^1(\Omega))} \end{aligned}$$

where we have used the dual norm, with the duality pairing $\langle \cdot, \cdot \rangle_Q$,

$$\|\phi\|_{L_2(0, T; H^{-1}(\Omega))} := \sup_{0 \neq v \in L_2(0, T; H_0^1(\Omega))} \frac{c_H \langle \phi, v \rangle_Q}{\|v\|_{L_2(0, T; H_0^1(\Omega))}}. \quad (2.5)$$

In particular, we have obtained the upper estimate

$$\|w\|_{L_2(0, T; H_0^1(\Omega))} \leq \|\phi\|_{L_2(0, T; H^{-1}(\Omega))}.$$

The solution of the variational formulation (2.4) implies for any given $\phi \in L_2(0, T; H^{-1}(\Omega))$ a unique $w \in L_2(0, T; H_0^1(\Omega))$, which defines the Newton potential

$$N\phi := w, \quad N : L_2(0, T; H^{-1}(\Omega)) \rightarrow L_2(0, T; H_0^1(\Omega)).$$

By using

$$\begin{aligned} \|\phi\|_{L_2(0, T; H^{-1}(\Omega))} &= \sup_{0 \neq v \in L_2(0, T; H_0^1(\Omega))} \frac{c_H \langle \phi, v \rangle_Q}{\|v\|_{L_2(0, T; H_0^1(\Omega))}} \\ &= \sup_{0 \neq v \in L_2(0, T; H_0^1(\Omega))} \frac{1}{\|v\|_{L_2(0, T; H_0^1(\Omega))}} \int_0^T \int_{\Omega} [A(x, t)\nabla_x w(x, t)] \cdot \nabla_x v(x, t) \, dx dt \\ &\leq \|w\|_{L_2(0, T; H_0^1(\Omega))} \end{aligned}$$

we then conclude

$$\|w\|_{L_2(0,T;H_0^1(\Omega))} = \|N\phi\|_{L_2(0,T;H_0^1(\Omega))} = \|\phi\|_{L_2(0,T;H^{-1}(\Omega))}. \quad (2.6)$$

In fact, this gives

$$\begin{aligned} \|\phi\|_{L_2(0,T;H^{-1}(\Omega))}^2 &= \|w\|_{L_2(0,T;H_0^1(\Omega))}^2 \\ &= \int_0^T \int_{\Omega} [A(x,t)\nabla_x w(x,t)] \cdot \nabla_x w(x,t) \, dxdt \\ &= c_H \int_0^T \int_{\Omega} \phi(x,t)w(x,t) \, dxdt = c_H \langle \phi, w \rangle_Q, \end{aligned}$$

i.e.

$$\|\phi\|_{L_2(0,T;H^{-1}(\Omega))}^2 = c_H \langle \phi, N\phi \rangle_Q \quad \text{for all } \phi \in L_2(0,T;H^{-1}(\Omega)). \quad (2.7)$$

Now we are in a position to prove boundedness of the bilinear form

$$a(u,v) := c_H \int_0^T \int_{\Omega} \partial_t u(x,t)v(x,t) \, dxdt + \int_0^T \int_{\Omega} [A(x,t)\nabla_x u(x,t)] \cdot \nabla_x v(x,t) \, dxdt. \quad (2.8)$$

By using duality, Cauchy–Schwarz and Hölder inequality we have

$$\begin{aligned} a(u,v) &= c_H \int_0^T \int_{\Omega} \partial_t u(x,t)v(x,t) \, dxdt + \int_0^T \int_{\Omega} [A(x,t)\nabla_x u(x,t)] \cdot \nabla_x v(x,t) \, dxdt \\ &\leq \left[\|\partial_t u\|_{L_2(0,T;H^{-1}(\Omega))} + \|u\|_{L_2(0,T;H_0^1(\Omega))} \right] \|v\|_{L_2(0,T;H_0^1(\Omega))} \\ &\leq \sqrt{2} \sqrt{\|\partial_t u\|_{L_2(0,T;H^{-1}(\Omega))}^2 + \|u\|_{L_2(0,T;H_0^1(\Omega))}^2} \|v\|_{L_2(0,T;H_0^1(\Omega))} \\ &= \sqrt{2} \|u\|_{L_2(0,T;H_0^1(\Omega)) \cap H^1(0,T;H^{-1}(\Omega))} \|v\|_{L_2(0,T;H_0^1(\Omega))} \end{aligned} \quad (2.9)$$

for all $u \in L_2(0,T;H_0^1(\Omega)) \cap H^1(0,T;H^{-1}(\Omega))$ and $v \in L_2(0,T;H_0^1(\Omega))$ where we have used the related energy norm, i.e.

$$\begin{aligned} \|u\|_{L_2(0,T;H_0^1(\Omega)) \cap H^1(0,T;H^{-1}(\Omega))}^2 &= \|\partial_t u\|_{L_2(0,T;H^{-1}(\Omega))}^2 + \|u\|_{L_2(0,T;H_0^1(\Omega))}^2 \\ &= \|N\partial_t u\|_{L_2(0,T;H_0^1(\Omega))}^2 + \|u\|_{L_2(0,T;H_0^1(\Omega))}^2. \end{aligned} \quad (2.10)$$

Since the Galerkin–Petrov variational formulation (2.2) is based on the use of different test and ansatz spaces, no concept of ellipticity is applicable to ensure unique solvability of (2.2). Instead, a suitable stability condition has to be established. For this we will make use of the trivial inclusion

$$L_2(0,T;H_0^1(\Omega)) \cap H^1(0,T;H^{-1}(\Omega)) \subset L_2(0,T;H_0^1(\Omega)). \quad (2.11)$$

Theorem 2.1 For all $u \in L_2(0, T; H_0^1(\Omega)) \cap H^1(0, T; H^{-1}(\Omega))$ satisfying $u(x, 0) = 0$ for $x \in \Omega$ there holds the stability condition

$$\frac{1}{2\sqrt{2}} \|u\|_{L_2(0, T; H_0^1(\Omega)) \cap H^1(0, T; H^{-1}(\Omega))} \leq \sup_{0 \neq v \in L_2(0, T; H_0^1(\Omega))} \frac{a(u, v)}{\|v\|_{L_2(0, T; H_0^1(\Omega))}}. \quad (2.12)$$

Proof. For any given $u \in L_2(0, T; H_0^1(\Omega)) \cap H^1(0, T; H^{-1}(\Omega))$, $u(x, 0) = 0$ for $x \in \Omega$, we consider the particular choice $v = u + N\partial_t u \in L_2(0, T; H_0^1(\Omega))$ to obtain

$$a(u, u + N\partial_t u) = a(u, u) + a(u, N\partial_t u).$$

Due to the assumption $u(x, 0) = 0$ for $x \in \Omega$ we further conclude

$$\begin{aligned} a(u, u) &= c_H \int_0^T \int_{\Omega} \partial_t u(x, t) u(x, t) \, dx dt + \int_0^T \int_{\Omega} [A(x, t) \nabla_x u(x, t)] \cdot \nabla_x u(x, t) \, dx dt \\ &= \frac{1}{2} c_H \int_0^T \frac{d}{dt} \|u(t)\|_{L_2(\Omega)}^2 \, dt + \|u\|_{L_2(0, T; H_0^1(\Omega))}^2 \\ &= \frac{1}{2} c_H \|u(T)\|_{L_2(\Omega)}^2 + \|u\|_{L_2(0, T; H_0^1(\Omega))}^2 \\ &\geq \|u\|_{L_2(0, T; H_0^1(\Omega))}^2. \end{aligned} \quad (2.13)$$

For the second term we obtain, by using (2.7), the Cauchy–Schwarz inequality, and (2.6),

$$\begin{aligned} a(u, N\partial_t u) &= c_H \langle \partial_t u, N\partial_t u \rangle_Q + \langle A \nabla_x u, \nabla_x N\partial_t u \rangle_{L_2(Q)} \\ &\geq \|\partial_t u\|_{L_2(0, T; H^{-1}(\Omega))}^2 - \|u\|_{L_2(0, T; H_0^1(\Omega))} \|N\partial_t u\|_{L_2(0, T; H_0^1(\Omega))} \\ &\geq \|\partial_t u\|_{L_2(0, T; H^{-1}(\Omega))}^2 - \frac{1}{2} \left[\|u\|_{L_2(0, T; H_0^1(\Omega))}^2 + \|N\partial_t u\|_{L_2(0, T; H_0^1(\Omega))}^2 \right] \\ &= \frac{1}{2} \|\partial_t u\|_{L_2(0, T; H^{-1}(\Omega))}^2 - \frac{1}{2} \|u\|_{L_2(0, T; H_0^1(\Omega))}^2. \end{aligned}$$

Hence we have

$$\begin{aligned} a(u, u + N_t \partial_t u) &\geq \frac{1}{2} \left[\|\partial_t u\|_{L_2(0, T; H^{-1}(\Omega))}^2 + \|u\|_{L_2(0, T; H_0^1(\Omega))}^2 \right] \\ &= \frac{1}{2} \|u\|_{L_2(0, T; H_0^1(\Omega)) \cap H^1(0, T; H^{-1}(\Omega))}^2. \end{aligned} \quad (2.14)$$

With

$$\begin{aligned} \|u + N\partial_t u\|_{L_2(0, T; H_0^1(\Omega))}^2 &\leq 2 \left[\|u\|_{L_2(0, T; H_0^1(\Omega))}^2 + \|N\partial_t u\|_{L_2(0, T; H_0^1(\Omega))}^2 \right] \\ &= 2 \left[\|u\|_{L_2(0, T; H_0^1(\Omega))}^2 + \|\partial_t u\|_{L_2(0, T; H^{-1}(\Omega))}^2 \right] \\ &= 2 \|u\|_{L_2(0, T; H_0^1(\Omega)) \cap H^1(0, T; H^{-1}(\Omega))}^2 \end{aligned}$$

we finally conclude

$$a(u, u + N_t \partial_t u) \geq \frac{1}{2\sqrt{2}} \|u\|_{L_2(0, T; H_0^1(\Omega)) \cap H^1(0, T; H^{-1}(\Omega))} \|u + N\partial_t u\|_{L_2(0, T; H_0^1(\Omega))}$$

from which the stability condition (2.12) follows. \blacksquare

Remark 2.1 *Instead of the energy norm (2.10) we may also use the alternative norm [21]*

$$\|u\|_{L_2(0,T;H_0^1(\Omega))\cap H^1(0,T;H^{-1}(\Omega)),c_H}^2 = \|N\partial_t u\|_{L_2(0,T;H_0^1(\Omega))}^2 + \|u\|_{L_2(0,T;H_0^1(\Omega))}^2 + c_H \|u(T)\|_{L_2(\Omega)}^2.$$

Then we can prove the improved stability estimate

$$\|u\|_{L_2(0,T;H_0^1(\Omega))\cap H^1(0,T;H^{-1}(\Omega)),c_H} \leq \sup_{0 \neq v \in L_2(0,T;H_0^1(\Omega))} \frac{a(u,v)}{\|v\|_{L_2(0,T;H_0^1(\Omega))}}$$

for all $u \in L_2(0,T;H_0^1(\Omega)) \cap H^1(0,T;H^{-1}(\Omega))$ satisfying $u(x,0) = 0$ for $x \in \Omega$. However, to avoid the additional term, here we will use the energy norm (2.10) and the stability estimate (2.12).

Since the initial condition in (2.1) is seen as a Dirichlet condition in the space–time domain $Q = \Omega \times (0, T)$, we consider the splitting $u(x, t) = \bar{u}(x, t) + \bar{u}_0(x, t)$ for $(x, t) \in Q$ where $\bar{u}_0 \in L_2(0, T; H_0^1(\Omega)) \cap H^1(0, T; H^{-1}(\Omega))$ is some extension of the given initial datum $u_0 \in H_0^1(\Omega)$. Hence we have to find $\bar{u} \in X$ such that

$$a(\bar{u}, v) = \langle f, v \rangle_Q - a(\bar{u}_0, v) \tag{2.15}$$

is satisfied for all $v \in Y$, where we have used the function spaces

$$X := \left\{ v \in L_2(0, T; H_0^1(\Omega)) \cap H^1(0, T; H^{-1}(\Omega)) : v(x, 0) = 0 \text{ for } x \in \Omega \right\}$$

and $Y := L_2(0, T; H_0^1(\Omega))$. Obviously, we have $X \subset Y$. Now we are in a position to state unique solvability of the Galerkin–Petrov variational formulation (2.15), see, e.g., [3] or [18, Theorem 3.7].

Corollary 2.2 *Let $\bar{u}_0 \in L_2(0, T; H_0^1(\Omega)) \cap H^1(0, T; H^{-1}(\Omega))$ be some given extension of the given initial datum $u_0 \in H_0^1(\Omega)$, and assume $f \in L_2(0, T; H^{-1}(\Omega))$. Since the bilinear form $a(\cdot, \cdot)$ as given in (2.8) is bounded satisfying (2.9), and satisfies the stability condition (2.12), there exists a unique solution $\bar{u} \in X$ of the variational formulation (2.15) satisfying*

$$\|\bar{u}\|_{L_2(0,T;H_0^1(\Omega))\cap H^1(0,T;H^{-1}(\Omega))} \leq 2\sqrt{2} \left[\|f\|_{L_2(0,T;H^{-1}(\Omega))} + \sqrt{2} \|\bar{u}_0\|_{L_2(0,T;H_0^1(\Omega))\cap H^1(0,T;H^{-1}(\Omega))} \right].$$

3 Finite element discretisation

Let $X_h \subset X$ and $Y_h \subset Y$ be some finite dimensional spaces which will be specified later. As in the continuous case $X \subset Y$ we assume the inclusion $X_h \subset Y_h$. The Galerkin–Petrov discretisation of the variational problem (2.15) is to find $\bar{u}_h \in X_h$ such that

$$a(\bar{u}_h, v_h) = \langle f, v_h \rangle_Q - a(\bar{u}_0, v_h) \quad \text{for all } v_h \in Y_h. \tag{3.1}$$

Before we discuss the unique solvability of the variational formulation (3.1) as well as related a priori error estimates, let us first consider the Galerkin discretization of the variational formulation (2.4) to find $w_h \in Y_h$ such that

$$\int_0^T \int_{\Omega} [A(x, t) \nabla_x w_h(x, t)] \cdot \nabla_x v_h(x, t) dx dt = c_H \int_0^T \int_{\Omega} \phi(x, t) v_h(x, t) dx dt \quad (3.2)$$

is satisfied for all $v_h \in Y_h$. By using standard arguments we conclude unique solvability of (3.2), in particular we may define the approximate Newton potential

$$N_h \phi := w_h, \quad N_h : L_2(0, T; H^{-1}(\Omega)) \rightarrow Y_h \subset L_2(0, T; H_0^1(\Omega)).$$

By using (3.2) and (2.4) we have, for a conformal discretization $Y_h \subset Y$,

$$\begin{aligned} \|w_h\|_{L_2(0, T; H_0^1(\Omega))}^2 &= \int_0^T \int_{\Omega} [A(x) \nabla_x w_h(x, t)] \cdot \nabla_x w_h(x, t) dx dt \\ &= c_H \int_0^T \int_{\Omega} \phi(x, t) w_h(x, t) dx dt \\ &= \int_0^T \int_{\Omega} [A(x) \nabla_x w(x, t)] \cdot \nabla_x w_h(x, t) dx dt \\ &\leq \|w\|_{L_2(0, T; H_0^1(\Omega))} \|w_h\|_{L_2(0, T; H_0^1(\Omega))}, \end{aligned}$$

i.e.

$$\|N_h \phi\|_{L_2(0, T; H_0^1(\Omega))} \leq \|N \phi\|_{L_2(0, T; H_0^1(\Omega))} \quad \text{for all } \phi \in L_2(0, T; H_0^1(\Omega)). \quad (3.3)$$

Instead of the energy norm (2.10) we will use the mesh dependent energy norm

$$\|u\|_{X_h}^2 = \|N_h \partial_t u\|_{L_2(0, T; H_0^1(\Omega))}^2 + \|u\|_{L_2(0, T; H_0^1(\Omega))}^2 \quad (3.4)$$

Due to (3.3) we have the upper estimate

$$\|u\|_{X_h} \leq \|u\|_X \quad \text{for all } u \in X. \quad (3.5)$$

As in the continuous case we are now able to prove a discrete stability condition.

Theorem 3.1 *Assume $X_h \subset X$, $Y_h \subset Y$, and $X_h \subset Y_h$. Then there holds the discrete stability condition*

$$\frac{1}{2\sqrt{2}} \|u_h\|_{X_h} \leq \sup_{0 \neq v_h \in Y_h} \frac{a(u_h, v_h)}{\|v_h\|_{L_2(0, T; H_0^1(\Omega))}} \quad \text{for all } u_h \in X_h \quad (3.6)$$

Proof. The proof of Theorem 3.1 follows as in the continuous case, see Theorem 2.1. For $u_h \in X_h \subset X$ we define $N_h \partial_t u_h := w_h \in Y_h$ as the unique solution of the Galerkin variational formulation (3.2), i.e.

$$\int_0^T \int_{\Omega} [A(x, t) \nabla_x w_h(x, t)] \cdot \nabla_x v_h(x, t) dx dt = c_H \int_0^T \int_{\Omega} \partial_t u_h(x, t) v_h(x, t) dx dt, \quad v_h \in Y_h.$$

Due to $X_h \subset Y_h$ we have $u_h + w_h \in Y_h$ satisfying

$$\|u_h + w_h\|_{L_2(0,T;H_0^1(\Omega))}^2 \leq 2 \left[\|u_h\|_{L_2(0,T;H_0^1(\Omega))}^2 + \|w_h\|_{L_2(0,T;H_0^1(\Omega))}^2 \right] = 2 \|u_h\|_{X_h}^2.$$

Moreover, we have, by using $u_h(0) = 0$,

$$\begin{aligned} a(u_h, u_h) &= c_H \int_0^T \int_{\Omega} \partial_t u_h(x, t) u_h(x, t) \, dx dt + \int_0^T \int_{\Omega} [A(x, t) \nabla_x u_h(x, t)] \cdot \nabla_x u_h(x, t) \, dx dt \\ &= \frac{1}{2} c_H \|u_h(T)\|_{L_2(\Omega)}^2 + \|u_h\|_{L_2(0,T;H_0^1(\Omega))}^2 \\ &\geq \|u_h\|_{L_2(0,T;H_0^1(\Omega))}^2 \end{aligned}$$

as well as

$$\begin{aligned} a(u_h, w_h) &= c_H \int_0^T \int_{\Omega} \partial_t u_h(x, t) w_h(x, t) \, dx dt + \int_0^T \int_{\Omega} [A(x, t) \nabla_x u_h(x, t)] \cdot \nabla_x w_h(x, t) \, dx dt \\ &= \int_0^T \int_{\Omega} [A(x, t) \nabla_x w_h(x, t)] \cdot \nabla_x w_h(x, t) \, dx dt + \int_0^T \int_{\Omega} [A(x, t) \nabla_x u_h(x, t)] \cdot \nabla_x w_h(x, t) \, dx dt \\ &\geq \|w_h\|_{L_2(0,T;H_0^1(\Omega))}^2 - \|u_h\|_{L_2(0,T;H_0^1(\Omega))} \|w_h\|_{L_2(0,T;H_0^1(\Omega))} \\ &\geq \|w_h\|_{L_2(0,T;H_0^1(\Omega))}^2 - \frac{1}{2} \left[\|u_h\|_{L_2(0,T;H_0^1(\Omega))}^2 + \|w_h\|_{L_2(0,T;H_0^1(\Omega))}^2 \right] \\ &= \frac{1}{2} \|w_h\|_{L_2(0,T;H_0^1(\Omega))}^2 - \frac{1}{2} \|u_h\|_{L_2(0,T;H_0^1(\Omega))}^2. \end{aligned}$$

Hence,

$$\begin{aligned} a(u_h, u_h + w_h) &\geq \frac{1}{2} \left[\|u_h\|_{L_2(0,T;H_0^1(\Omega))}^2 + \|w_h\|_{L_2(0,T;H_0^1(\Omega))}^2 \right] = \frac{1}{2} \|u_h\|_{X_h}^2 \\ &\geq \frac{1}{2\sqrt{2}} \|u_h\|_{X_h} \|u_h + w_h\|_{L_2(0,T;H_0^1(\Omega))}, \end{aligned}$$

which implies the discrete stability condition (3.6). ■

The discrete stability condition (3.6) implies unique solvability of the Galerkin–Petrov variational formulation (3.1). By combining the Galerkin variational formulation (3.1) with (2.15) and by using the inclusion $Y_h \subset Y$ we also conclude the Galerkin orthogonality

$$a(\bar{u} - \bar{u}_h, v_h) = 0 \quad \text{for all } v_h \in Y_h. \quad (3.7)$$

With this we can derive the following quasi-optimal error estimate.

Theorem 3.2 *Let $\bar{u} \in X$ and $\bar{u}_h \in X_h$ be the unique solutions of the variational formulations (2.15) and (3.1), respectively. Then there holds the a priori error estimate*

$$\|\bar{u} - \bar{u}_h\|_{X_h} \leq 5 \inf_{z_h \in X_h} \|\bar{u} - z_h\|_X. \quad (3.8)$$

Proof. From the stability condition (3.6), by using the Galerkin orthogonality (3.7) and (2.9), we conclude, for arbitrary $z_h \in X_h$,

$$\begin{aligned} \frac{1}{2\sqrt{2}} \|\bar{u}_h - z_h\|_{X_h} &\leq \sup_{0 \neq v_h \in Y_h} \frac{a(\bar{u}_h - z_h, v_h)}{\|v_h\|_{L_2(0,T;H_0^1(\Omega))}} \\ &= \sup_{0 \neq v_h \in Y_h} \frac{a(\bar{u}_h - \bar{u}, v_h) + a(\bar{u} - z_h, v_h)}{\|v_h\|_{L_2(0,T;H_0^1(\Omega))}} \\ &= \sup_{0 \neq v_h \in Y_h} \frac{a(\bar{u} - z_h, v_h)}{\|v_h\|_{L_2(0,T;H_0^1(\Omega))}} \leq \sqrt{2} \|\bar{u} - z_h\|_X, \end{aligned}$$

and the assertion follows from the triangle inequality. \blacksquare

Now we are in a position to introduce the space–time finite element spaces $X_h \subset Y_h$. For the space–time domain $Q = \Omega \times (0, T) \subset \mathbb{R}^{n+1}$ we consider a sequence of admissible decompositions Q_h into shape regular simplicial finite elements q_ℓ which are triangles ($n = 1$), tetrahedra ($n = 2$), or pentatops ($n = 3$, see, e.g., [15]), i.e.

$$Q_h = \bigcup_{\ell=1}^N \bar{q}_\ell.$$

By $\{(x_k, t_k)\}_{k=1}^M$ we denote the related set of nodes $(x_k, t_k) \in \mathbb{R}^{n+1}$. As usual we introduce a reference element $q \subset \mathbb{R}^{n+1}$ which allows to describe a finite element q_ℓ by using a local parametrization,

$$\begin{pmatrix} x \\ t \end{pmatrix} = \begin{pmatrix} x_{k_1} \\ t_{k_1} \end{pmatrix} + J_\ell \begin{pmatrix} \xi \\ \tau \end{pmatrix} \quad \text{for } \begin{pmatrix} \xi \\ \tau \end{pmatrix} \in q,$$

where $q \subset \mathbb{R}^{n+1}$ are the usual reference elements. Hence we can introduce the volume Δ_ℓ of the finite element q_ℓ ,

$$\Delta_\ell = \int_{q_\ell} dx dt = \det J_\ell \int_q d\xi d\tau = |q| \det J_\ell,$$

and the local mesh width

$$h_\ell = \Delta_\ell^{1/(n+1)}, \quad h := \max_{\ell=1, \dots, N} h_\ell.$$

Note that

$$|q| = \begin{cases} \frac{1}{2} & \text{for } n = 1, \\ \frac{1}{6} & \text{for } n = 2, \\ \frac{1}{24} & \text{for } n = 3. \end{cases}$$

With respect to the admissible decomposition Q_h we define the space–time finite element space $S_h^1(Q_h) = \text{span}\{\varphi_k\}_{k=1}^M$ of piecewise linear and continuous basis functions φ_k . Now we are able to introduce the finite element spaces

$$X_h = S_h^1(Q_h) \cap X, \quad Y_h = S_h^1(Q_h) \cap Y,$$

which obviously satisfy the assumptions of Theorem 3.1. Based on Theorem 3.2 we conclude the following a priori error estimate.

Theorem 3.3 *Let $\bar{u} \in X$ and $\bar{u}_h \in X_h = S_h^1(Q_h) \cap X$ be the unique solutions of the variational formulations (2.15) and (3.1), respectively. Assume $\bar{u} \in H^2(Q)$. Then there holds the energy error estimate*

$$\|\bar{u} - \bar{u}_h\|_{L_2(0,T;H_0^1(\Omega))} \leq c h |\bar{u}|_{H^2(Q)}. \quad (3.9)$$

Proof. From the definition of the related norms and by using Theorem 3.1 we first have

$$\|\bar{u} - \bar{u}_h\|_{L_2(0,T;H_0^1(\Omega))} \leq \|\bar{u} - \bar{u}_h\|_{X_h} \leq 5 \inf_{z_h \in X_h} \|\bar{u} - z_h\|_X,$$

and it remains to define a suitable approximation z_h of \bar{u} to prove a related approximation property. In fact, let $z_h = P_h \bar{u} \in X_h$ be the unique solution of the Galerkin variational formulation

$$\begin{aligned} \int_0^T \int_{\Omega} \left[\partial_t P_h \bar{u}(x, t) \partial_t v_h(x, t) + [A(x, t) \nabla_x P_h \bar{u}(x, t)] \cdot \nabla_x v_h(x, t) \right] dx dt \\ = \int_0^T \int_{\Omega} \left[\partial_t \bar{u}(x, t) \partial_t v_h(x, t) + [A(x, t) \nabla_x \bar{u}(x, t)] \cdot \nabla_x v_h(x, t) \right] dx dt \end{aligned}$$

for all $v_h \in X_h$. When using the related energy norm, note that $v(x, 0) = 0$ for $x \in \Omega$ and $v(x, t) = 0$ for $x \in \Gamma$, i.e.

$$\|v\|_{H^1(Q)}^2 := \int_0^T \int_{\Omega} \left[|\partial_t v(x, t)|^2 + [A(x, t) \nabla_x v(x, t)] \cdot \nabla_x v(x, t) \right] dx dt,$$

and by using standard arguments, i.e. Cea's lemma and standard approximation error estimates, we obtain

$$\|\bar{u} - P_h \bar{u}_h\|_{L_2(0,T;H_0^1(\Omega))} \leq \|\bar{u} - P_h \bar{u}_h\|_{H^1(Q)} \leq \inf_{v_h \in X_h} \|\bar{u} - v_h\|_{H^1(Q_h)} \leq c h |\bar{u}|_{H^2(Q)}.$$

Note that for $n \leq 2$ we can use standard nodal interpolation error estimates, while for $n = 3$ we have to use projection type error estimates.

Moreover,

$$\begin{aligned} \|\partial_t(\bar{u} - P_h \bar{u})\|_{L_2(0,T;H^{-1}(\Omega))} &= \sup_{0 \neq v \in L_2(0,T;H_0^1(\Omega))} \frac{c_H \langle \partial_t(\bar{u} - P_h \bar{u}), v \rangle_Q}{\|v\|_{L_2(0,T;H_0^1(\Omega))}} \\ &\leq c_H \sup_{0 \neq v \in L_2(0,T;H_0^1(\Omega))} \frac{\|\partial_t(\bar{u} - P_h \bar{u})\|_{L_2(Q)} \|v\|_{L_2(Q)}}{\|v\|_{L_2(0,T;H_0^1(\Omega))}} \\ &\leq c_H \|\partial_t(\bar{u} - P_h \bar{u})\|_{L_2(Q)} \leq c_H \|\bar{u} - P_h \bar{u}\|_{H^1(Q)} \leq c h \|\bar{u}\|_{H^2(Q)}. \end{aligned}$$

By taking into account the norm definition (2.10) the assertion follows from the previous estimates. \blacksquare

Corollary 3.4 *Let $\bar{u} \in X$ and $\bar{u}_h \in X_h = S_h^1(Q_h) \cap X$ be the unique solutions of the variational formulations (2.15) and (3.1), respectively. We now assume $\bar{u} \in H^s(Q)$ for some $s \in [1, 2]$. Then there holds the energy error estimate*

$$\|\bar{u} - \bar{u}_h\|_{L_2(0,T;H_0^1(\Omega))} \leq c h^{s-1} |\bar{u}|_{H^s(Q)}. \quad (3.10)$$

Proof. The assertion follows from the trivial error estimate, i.e. $s = 1$,

$$\|\bar{u} - P_h \bar{u}_h\|_{L_2(0,T;H_0^1(\Omega))} \leq \|\bar{u} - P_h \bar{u}_h\|_{H^1(Q)} \leq \|\bar{u}\|_{H^1(Q)},$$

by using the energy error estimate (3.9), i.e. $s = 2$, and an interpolation argument for $s \in (1, 2)$. \blacksquare

Remark 3.1 *Instead of space–time finite element spaces $X_h \subset Y_h$ spanned by piecewise linear basis functions φ_k one may use continuous basis functions of some polynomial degree $p > 1$ locally to define $X_h \subset Y_h$. Then, when assuming $\bar{u} \in H^s(Q)$ for some $s \in [1, p + 1]$, one obtains the error estimate*

$$\|\bar{u} - \bar{u}_h\|_{L_2(0,T;H_0^1(\Omega))} \leq c h^{s-1} |\bar{u}|_{H^s(Q)}, \quad s \in [1, p + 1]. \quad (3.11)$$

4 Space–time interpolation error estimates

As we will see in the numerical results, in particular of Example 5.1, the error estimates (3.9) and (3.11) are optimal in the case of a regular solution. However, in the particular case of solutions with some singular behavior, see Examples 5.2 and 5.3, the error estimate (3.10) turns out to be not optimal, or not applicable. This is mainly due to the unified handling of the spatial and temporal derivatives when proving the error estimate (3.9). In what follows we present a more detailed analysis of the interpolation error in the energy norm. To simplify the presentation, let us consider the spatial one–dimensional case only, where all space–time finite elements q_ℓ are a right–angled isosceles triangles, as used in Fig. 1. Without loss of generality we consider the space–time finite elements q_h^1 and q_h^2 which are given by

$$q_h^1 = \left\{ (x, t) \in \mathbb{R}^2 : 0 < x < h, 0 < t < h - x \right\},$$

and

$$q_h^2 = \left\{ (x, t) \in \mathbb{R}^2 : 0 < x < h, h - x < t < h \right\}.$$

The related linear interpolations $I_h^1 u$ and $I_h^2 u$ of a given function u then read

$$I_h^1 u(x, t) = u(0, 0) \frac{h - x - t}{h} + u(h, 0) \frac{x}{h} + u(0, h) \frac{t}{h},$$

and

$$I_h^2 u(x, t) = u(0, h) \frac{h - x}{h} + u(h, 0) \frac{h - t}{h} + u(h, h) \frac{x + t - h}{h}.$$

For the local interpolation error we then can prove the following result.

Lemma 4.1 *Let q_h^1 and q_h^2 be two right-angled isosceles triangles such that $\bar{q}_h^1 \cup \bar{q}_h^2 = [0, h]^2$. Let $u(x, t)$, $(x, t) \in [0, h]^2$, be a given function such that $u_x \in H^1((0, h)^2)$, and let $I_h u$ be the piecewise linear interpolation. Then there holds the error estimate*

$$\|\partial_x(u - I_h u)\|_{L_2((0, h)^2)}^2 \leq \frac{4}{3} h^2 \int_0^h \int_0^h [u_{xx}(\eta, t)]^2 d\eta dt + \frac{2}{3} h^2 \int_0^h \int_0^h [u_{xt}(s, \tau)]^2 d\tau ds. \quad (4.1)$$

Proof. To estimate the local error

$$\|\partial_x(u - I_h^1 u)\|_{L_2(q_h^1)}^2 = \int_0^h \int_0^{h-x} [\partial_x(u(x, t) - I_h^1 u(x, t))]^2 dx dt$$

we first compute

$$\begin{aligned} \partial_x(u(x, t) - I_h^1 u(x, t)) &= u_x(x, t) - \frac{1}{h} [u(h, 0) - u(0, 0)] \\ &= \frac{1}{h} \int_0^h [u_x(x, t) - u_x(s, 0)] ds \\ &= \frac{1}{h} \int_0^h [u_x(x, t) - u_x(s, t) + u_x(s, t) - u_x(s, 0)] ds \\ &= \frac{1}{h} \int_0^h \left[\int_s^x u_{xx}(\eta, t) d\eta + \int_0^t u_{xt}(s, \tau) d\tau \right] ds. \end{aligned}$$

Hence, by using the Cauchy–Schwarz inequality twice,

$$\begin{aligned} \left[\partial_x(u(x, t) - I_h^1 u(x, t)) \right]^2 &= \frac{1}{h^2} \left[\int_0^h \left(\int_s^x u_{xx}(\eta, t) d\eta + \int_0^t u_{xt}(s, \tau) d\tau \right) ds \right]^2 \\ &\leq \frac{1}{h} \int_0^h \left(\int_s^x u_{xx}(\eta, t) d\eta + \int_0^t u_{xt}(s, \tau) d\tau \right)^2 ds \\ &\leq \frac{2}{h} \int_0^h \left(|x - s| \left| \int_s^x [u_{xx}(\eta, t)]^2 d\eta \right| + t \int_0^t [u_{xt}(s, \tau)]^2 d\tau \right) ds \\ &\leq \frac{2}{h} \int_0^h |x - s| ds \int_0^h [u_{xx}(\eta, t)]^2 d\eta + \frac{2}{h} t \int_0^h \int_0^h [u_{xt}(s, \tau)]^2 d\tau ds, \end{aligned}$$

and therefore,

$$\begin{aligned}
& \int_0^h \int_0^{h-x} \left[\partial_x(u(x, t) - I_h^1 u(x, t)) \right]^2 dt dx \\
& \leq \frac{2}{h} \int_0^h \int_0^h |x - s| ds \int_0^{h-x} \int_0^h [u_{xx}(\eta, t)]^2 d\eta dt dx \\
& \quad + \frac{2}{h} \int_0^h \int_0^{h-x} t dt dx \int_0^h \int_0^h [u_{xt}(s, \tau)]^2 d\tau ds \\
& \leq \frac{2}{h} \int_0^h \int_0^h |x - s| ds dx \int_0^h \int_0^h [u_{xx}(\eta, t)]^2 d\eta dt \\
& \quad + \frac{1}{h} \int_0^h (h - x)^2 dx \int_0^h \int_0^h [u_{xt}(s, \tau)]^2 d\tau ds \\
& \leq \frac{2}{3} h^2 \int_0^h \int_0^h [u_{xx}(\eta, t)]^2 d\eta dt + \frac{1}{3} h^2 \int_0^h \int_0^h [u_{xt}(s, \tau)]^2 d\tau ds
\end{aligned}$$

follows. In the same way we have

$$\partial_x(u(x, t) - I_h^2 u(x, t)) = \frac{1}{h} \int_0^h \left[\int_s^x u_{xx}(\eta, t) d\eta - \int_t^h u_{xt}(s, \tau) d\tau \right] ds,$$

and therefore

$$\begin{aligned}
& \left[\partial_x(u(x, t) - I_h^2 u(x, t)) \right]^2 \\
& \leq \frac{2}{h} \int_0^h |x - s| ds \int_0^h [u_{xx}(\eta, t)]^2 d\eta + \frac{2}{h} (h - t) \int_0^h \int_0^h [u_{xt}(s, \tau)]^2 d\tau ds
\end{aligned}$$

as well as

$$\begin{aligned}
& \int_0^h \int_{h-x}^h \left[\partial_x(u(x, t) - I_h^2 u(x, t)) \right]^2 dt dx \\
& \leq \frac{2}{3} h^2 \int_0^h \int_0^h [u_{xx}(\eta, t)]^2 d\eta dt + \frac{1}{3} h^2 \int_0^h \int_0^h [u_{xt}(s, \tau)]^2 d\tau ds
\end{aligned}$$

follows. Summing up gives the desired result. ■

As an immediate consequence of Lemma 4.1 we can formulate the following interpolation error estimate.

Corollary 4.2 *Let $Q_h \subset \mathbb{R}^2$ be a space-time finite element mesh of right-angled isosceles triangles. Let u be a given function such that $\partial_x u \in H^1(Q)$ is satisfied, and let $I_h u$ be the continuous piecewise linear interpolation. Then there holds the error estimate*

$$\|u - I_h u\|_{L_2(0, T; H_0^1(\Omega))} \leq c h |\partial_x u|_{H^1(Q)}. \quad (4.2)$$

Although the interpolation error estimate (4.2) is not sufficient to establish an improved error estimate for the finite element solution \bar{u}_h of the Galerkin–Petrov variational formulation (3.1), it already indicates the expected order of convergence. However, since the second error term, i.e. $\|\partial_t(\bar{u} - z_h)\|_{L_2(0,T;H^{-1}(\Omega))}$, is measured in a Sobolev norm with a negative index, no interpolation error estimates are available. Instead, appropriate projection operators should be used to gain an additional order of convergence due to the Aubin–Nitsche trick. But since we have to use the same element $z_h \in X_h$ within the a priori error estimate (3.9) we have to bound the projection error in different norms by the interpolation error.

Although we have shown the improved interpolation error estimate (4.2) only in the particular case of right–angled isosceles triangular elements, we claim that this error estimate remains true for more general situations, i.e. for unstructured meshes as used in Example 5.2, and for higher spatial dimensions. Since the related proofs are much more technical, we do not consider them here.

To close this section, let us recall that the error estimate (3.10) already covers the most general situation when assuming some regularity of the solution in the space–time domain which can be ensured in many cases.

5 Numerical results

In what follows we provide some numerical results to confirm the theoretical estimates as given in Sect. 3 and 4. As in [14] we consider the initial boundary value problem (2.1) for the spatial domain $\Omega = (0, 1)$, the final time $T = 1$, the heat capacity $c_H = 1$, and $A(x, t) = I$. The discretization is done by considering a uniform and structured triangulation of the space–time domain $Q = (0, 1)^2$ as shown in Fig. 2 (left), and by using a more unstructured mesh as shown in Fig. 2 (right), and by using either first ($p = 1$) or second ($p = 2$) order continuous finite elements. Note that the structured mesh corresponds to the situation as considered in Lemma 4.1. The linear system is solved by a GMRES without preconditioning where it is sufficient to consider $Y_h = X_h$ to build up the linear system, i.e. we neglect the basis functions with nonzero initial conditions.

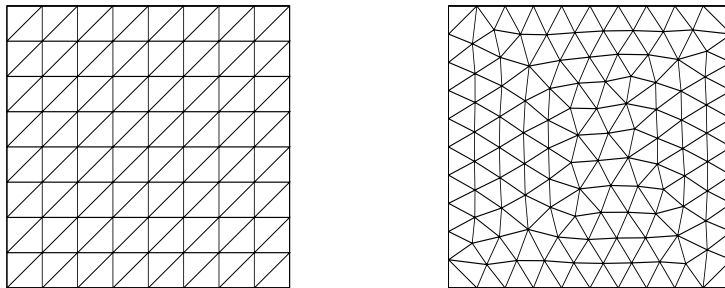


Figure 2: Structured mesh ($L = 3$), and unstructured mesh ($L = 0$).

Example 5.1 As in [14, Example 2.3.1] we first consider the exact solution to be

$$u(x, t) = \cos \pi t \sin \pi x \quad \text{for } (x, t) \in Q$$

and we determine the initial datum u_0 and the given volume density f accordingly. Since the solution u is smooth we expect first and second order convergence when using piecewise linear ($p = 1$) and piecewise quadratic ($p = 2$) continuous finite element approximations, see the error estimate (3.11). This is confirmed by the numerical results as given in Table 2.

		$p = 1$			$p = 2$		
L	N	dof	$\ u - u_h\ _{L_2(0,T;H_0^1(\Omega))}$	eoc	dof	$\ u - u_h\ _{L_2(0,T;H_0^1(\Omega))}$	eoc
2	32	12	5.960 -1		56	8.556 -2	
3	128	56	3.056 -1	0.964	240	2.172 -2	1.978
4	512	240	1.538 -1	0.991	992	5.456 -3	1.993
5	2048	992	7.705 -2	0.997	4032	1.366 -3	1.998
6	8192	4032	3.855 -2	0.999	16256	3.417 -4	1.999

Table 2: Numerical results in the case of a regular solution, structured mesh.

Note that the absolute errors and the estimated orders of convergence (eoc) almost coincide with the numerical results as given in [14, Table 2.1] in the case of a discontinuous Galerkin finite element approach where a related discrete energy norm is used which is not equivalent to the mesh dependent norm (3.4). It is worth to mention that there is a reduced order of convergence when using piecewise quadratic but continuous finite element approximations within the DG formulation [14, Table 2.7] which does not appear within our approach. This is probably due to the chosen energy norm of the DG approach which may not reflect the underlying function spaces in a proper way.

While the results of Table 2 correspond to a globally uniform mesh, see Fig. 2 (left), we may also use a more unstructured mesh of the space-time domain as shown in Fig. 2 (right). The related results are shown in Table 3, and again they confirm the theoretical findings.

		$p = 1$			$p = 2$		
L	N	dof	$\ u - u_h\ _{L_2(0,T;H_0^1(\Omega))}$	eoc	dof	$\ u - u_h\ _{L_2(0,T;H_0^1(\Omega))}$	eoc
0	216	98	1.562 -1		412	9.618 -3	
1	864	412	7.856 -2	0.992	1688	2.452 -3	1.972
2	3456	1688	3.938 -2	0.996	6832	6.225 -4	1.978
3	13824	6832	1.971 -2	0.999	27488	1.571 -4	1.986

Table 3: Numerical results in the case of a regular solution, unstructured mesh.

Example 5.2 As a second example we consider the solution

$$u(x, t) = (1 - t)^\alpha \sin \pi x \quad \text{for } (x, t) \in Q = (0, 1)^2, \quad \alpha = 0.75,$$

which has a line singularity at $t = 1$, i.e. $u \in H^{5/4-\varepsilon}(Q)$ for any $\varepsilon > 0$. According to the error estimate (3.10) one would only expect an order of convergence of almost 0.25. However, due to the improved interpolation error estimate (4.2) we can expect linear convergence which is confirmed by the numerical results given in Table 4 in the case of the structured mesh, and in Table 5 when using the more unstructured mesh. Note that when using second order elements ($p = 2$) we observe less than linear convergence, but the order of convergence is much better than expected from the error estimate (3.11). The understanding of this behavior requires a more detailed analysis to be done.

		$p = 1$			$p = 2$		
L	N	dof	$\ u - u_h\ _{L_2(0,T;H_0^1(\Omega))}$	eoc	dof	$\ u - u_h\ _{L_2(0,T;H_0^1(\Omega))}$	eoc
2	32	12	3.763 -1		56	4.553 -2	
3	128	56	1.942 -1	0.954	240	1.404 -2	1.697
4	512	240	9.864 -2	0.977	992	5.601 -3	1.326
5	2048	992	4.971 -2	0.989	4032	2.826 -3	0.987
6	8192	4032	2.498 -2	0.993	16256	1.581 -3	0.838

Table 4: Numerical results in the case of a singular solution ($\alpha = 0.75$), structured mesh.

		$p = 1$			$p = 2$		
L	N	dof	$\ u - u_h\ _{L_2(0,T;H_0^1(\Omega))}$	eoc	dof	$\ u - u_h\ _{L_2(0,T;H_0^1(\Omega))}$	eoc
0	216	98	1.171 -1		412	8.486 -3	
1	864	412	5.985 -2	0.968	1688	3.985 -3	1.091
2	3456	1688	3.031 -2	0.982	6832	2.170 -3	0.877
3	13824	6832	1.530 -2	0.986	27488	1.246 -3	0.800

Table 5: Numerical results in the case of a singular solution ($\alpha = 0.75$), unstructured mesh.

Example 5.3 As a last example we consider the solution [14, Example 2.3.3]

$$u(x, t) = \sqrt{1 - t} \sin \pi x \quad \text{for } (x, t) \in Q$$

with $u \in H^{1-\varepsilon}(Q)$, $\varepsilon > 0$. For the numerical results in the case of a piecewise linear continuous approximation see Table 6. We observe an almost linear order of convergence. While in the case of discontinuous piecewise linear approximations a convergence order of 0.5 is observed, see [14, Table 2.4], there is no convergence when using piecewise linear continuous approximations within the DG approach, see [14, Table 2.9]. The different observations are again due to the use of different energy norms within the different approaches and they may indicate that the chosen energy norm of the DG approach is not the optimal one.

L	N	dof	$\ u - u_h\ _{L_2(0,T;H_0^1(\Omega))}$	eoc
2	32	12	4.095 -1	
3	128	56	2.194 -1	0.900
4	512	240	1.175 -1	0.901
5	2048	992	6.351 -2	0.888
6	8192	4032	3.528 -2	0.848

Table 6: Numerical results in the case of a singular solution ($\alpha = 0.5$), structured mesh.

6 Concluding remarks

In this paper we have provided the stability and a priori error analysis of conforming space–time finite element methods for the solution of parabolic evolution equations. In contrast to other space–time approaches, which rely on space–time tensor product ansatz spaces, see, e.g., [17, 21], our approach is based on the use of arbitrary admissible finite element meshes of the space–time domain. Under rather mild assumptions on the regularity of the solution we have derived quasi–optimal error estimates which do not depend on the space dimension. In particular situations of non–smooth input data, less regular solutions may appear, where the above mentioned a priori error estimates may be either not optimal, or can not be applied. To get a better understanding of these phenomena we have presented a more detailed analysis of the piecewise linear interpolation error in the particular situation of right–angled isosceles triangular elements. It is obvious that additional work is required to improve these results, in particular to handle more general cases. To verify the theoretical results we have provided simple numerical examples for spatial one–dimensional domains. However, this paper extends previous work [14, 15] where space–time discontinuous Galerkin finite element methods were used even for three–dimensional spatial domains. In fact, using the results of [15] it will be possible to realize the proposed approach using continuous space–time finite elements also for three–dimensional spatial domains, this will be one topic of future work.

While in the present implementation of the continuous space–time finite element method we have used the GMRES algorithm without any preconditioner as an iterative solver, appropriate preconditioners and a parallel implementation are mandatory. Note that for the space–time discontinuous Galerkin approach related multigrid methods are analyzed and successfully implemented on state of the art parallel computer architectures showing optimal scaling properties [8, 9]. Although the multigrid approach may be also applied to continuous space–time formulations, probably there are other preconditioning strategies such as geometric and algebraic multilevel techniques, or domain decomposition methods applicable. Obviously, this will be a second topic of future research.

The main advantage of the proposed space–time finite element approach is the possibility to handle general finite element meshes which are adaptive simultaneously in space and time. To drive such an adaptive mesh refinement, appropriate a posteriori error estimators are required which are well established for elliptic boundary value problems, see, e.g., [22].

In the context of time-dependent problems, a posteriori error estimators are available as well, see, e.g., [1, 7, 16], but in most cases, adaptivity is considered either in space, or in time. So the design and analysis of appropriate space-time a posteriori error estimators is a third topic of future work.

The proposed strategy of rather general space-time approximations can be applied also to the numerical solution of boundary integral equations for parabolic partial differential equations. Then, the boundary $\Sigma = \Gamma \times (0, T)$ of the space-time domain $Q = \Omega \times (0, T)$ has to be decomposed into general space-time boundary elements. Note that the theory of boundary integral equations in particular for the heat equation is well established [6]. Although space-time tensor products are used in [6] for the approximation, the theory already covers the general case of arbitrary space-time boundary elements. However, the implementation of general space-time boundary element methods is more involved due to the fundamental solution which is non-local. In any case, space-time boundary element methods and their coupling with space-time finite element methods will be interesting topics for future work as well.

Applications of the proposed space-time finite element approach are numerous, not only to handle evolution equations of parabolic type, but also hyperbolic equations with the wave equation as the most simple example. While space-time methods can be applied to nonlinear partial differential equations as well, see, e.g., [14] for a space-time DG approach for the Navier–Stokes equations, problems with moving interfaces such as in fluid–structure interaction problems seem to be more challenging. In particular, when the interface is unknown, even a coupled system of linear partial differential equations becomes nonlinear in the geometric description. It is obvious that there are much more challenging and interesting applications around, and we believe that space-time discretizations of partial differential equations can turn to be a method of choice.

References

- [1] W. Bangerth, M. Geiger, R. Rannacher: Adaptive Galerkin finite element methods for the wave equation. *Comput. Meth. Appl. Math.* 10 (2010) 3–48.
- [2] M. Behr: Simplex space-time meshes in finite element simulations. *Int. J. Numer. Meth. Fluids* 57 (2008) 1421–1434.
- [3] D. Boffi, F. Brezzi, M. Fortin: *Mixed Finite Element Methods and Applications*. Springer Series in Computational Mathematics, vol. 44, Springer, Berlin, 2013.
- [4] S. Brenner, R. L. Scott: *The Mathematical Theory of Finite Element Methods*. Springer, New York, 1994.
- [5] P. G. Ciarlet: *The Finite Element Method for Elliptic Problems*. North–Holland, 1978.
- [6] M. Costabel: Boundary integral operators for the heat equation. *Integr. Equ. Oper. Theory* 13 (1990) 498–552.

- [7] K. Eriksson, C. Johnson: Adaptive finite element methods for parabolic problems. I. A linear model problem. *SIAM J. Numer. Anal.* 28 (1991) 43–77.
- [8] M. J. Gander, M. Neumüller: Analysis of a new space-time parallel multigrid algorithm for parabolic problems. *arXiv:1411.0519*, 2014.
- [9] M. J. Gander, M. Neumüller: Analysis of a time multigrid algorithm for DG-discretizations in time. *arXiv:1409.5254v2*, 2014.
- [10] T. J. R. Hughes, G. M. Hulbert: Space–time finite element methods for elastodynamics: formulations and error estimates. *Comp. Meth. Appl. Mech. Engrg.* 66 (1988) 339–363.
- [11] J. Lang: Adaptive multilevel solution of nonlinear parabolic PDE systems. *Lecture Notes in Computational Science and Engineering*, vol. 16, Springer, Berlin, 2001.
- [12] C. Mollet: Stability of Petrov–Galerkin discretizations: Application to the space–time weak formulation for parabolic evolution problems. *Comput. Meth. Appl. Math.* 14 (2014) 231–255.
- [13] A. D. Mont: Adaptive unstructured space–time meshing for four–dimensional space–time discontinuous Galerkin finite element methods. PhD thesis, University of Illinois, 2011.
- [14] M. Neumüller: Space–time methods. Fast solvers and applications. *Monographic Series TU Graz, Computation in Engineering and Science*, vol. 20, TU Graz, 2013.
- [15] M. Neumüller, O. Steinbach: Refinement of flexible space–time finite element meshes and discontinuous Galerkin methods. *Comput. Visual. Sci.* 14 (2011) 189–205
- [16] M. Schmich, B. Vexler: Adaptivity with dynamic meshes for space–time finite element discretizations of parabolic equations. *SIAM J. Sci. Comput.* 30 (2008) 369–393.
- [17] C. Schwab, R. Stevenson: Space–time adaptive wavelet methods for parabolic evolution problems. *Math. Comput.* 78 (2009) 1293–1318.
- [18] O. Steinbach: *Numerical Approximation Methods for Elliptic Boundary Value Problems. Finite and Boundary Elements.* Springer, New York, 2008.
- [19] S. Thite: Adaptive spacetime meshing for discontinuous Galerkin methods. *Comput. Geom. Theory Appl.* 42 (2009) 20–44.
- [20] V. Thomée: *Galerkin Finite Element Methods for Parabolic Problems.* Springer, New York, 2006.
- [21] K. Urban, A. T. Patera: An improved error bound for reduced basis approximation of linear parabolic problems. *Math. Comput.* 83 (2014) 1599–1615.)

- [22] R. Verfürth: A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinement. John-Wiley & Sons, Chichester, 1996.