Electronic Transactions on Numerical Analysis.
Volume 65, pp. 387–413, 2026.

ISSN 1068–9613.
DOI: 10.1553/etna_vol65s387

ETNA
Kent State University and
Johann Radon Institute (RICAM)

ADAPTIVE LEAST-SQUARES SPACE-TIME FINITE ELEMENT METHODS*

CHRISTIAN KÖTHE†, RICHARD LÖSCHER‡, AND OLAF STEINBACH‡

Abstract. We consider the numerical solution of an abstract operator equation $Bu = f$ by using a least-squares approach. We assume that $B : X \to Y^*$ is an isomorphism and that $A : Y \to Y^*$ implies a norm in Y, where X and Y are Hilbert spaces. The minimizer of the least-squares functional $\frac{1}{2} \|Bu - f\|^2_{A^{-1}}$, i.e., the solution of the operator equation, is then characterized by the gradient equation $Su = B^* A^{-1} f$ with an elliptic and self-adjoint operator $S := B^* A^{-1} B : X \to X^*$. When introducing the adjoint $p = A^{-1}(f - Bu)$, we end up with a saddle point formulation to be solved numerically by using a mixed finite element method. Based on a discrete inf-sup stability condition, we derive related a priori error estimates. While the adjoint p is zero by construction, its approximation p_h serves as an a posteriori error indicator to drive an adaptive scheme when discretized appropriately. While this approach can be applied to rather general equations, here we consider second-order linear partial differential equations, including the Poisson equation, the heat equation, and the wave equation, in order to demonstrate its potential, which allows one to use almost arbitrary space-time finite element meshes for the adaptive solution of time-dependent partial differential equations.

Key words. least-squares methods, space-time finite element methods, a posteriori error indicator, adaptivity, Poisson equation, heat equation, wave equation

AMS subject classifications. 65M60, 65M12, 65M50, 65N30, 65N12, 65N50

1. Introduction.

The use of Galerkin finite element methods for the numerical solution of partial differential equations is well established; see, e.g., the textbooks [13, 48] and many others. However, for time-dependent problems it is a common procedure to first discretize the spatial part using a finite element method and then apply either a time-stepping method, e.g., an explicit or implicit Euler scheme (see, e.g., the monograph [54]), or discontinuous Galerkin methods in time; see, e.g., [42] and the references given therein. Recently, interest in discretizing space and time at once has been rising, resulting in so-called space-time discretization methods; see, e.g., [34]. Although the discretization in space and time leads to larger systems of algebraic equations to be solved, these methods bring the advantage of having full control of the discretization in space and time simultaneously, allowing for space-time adaptivity. Moreover, space-time methods offer more flexibility in the construction of efficient solvers than time-stepping methods since preconditioning and parallelization in the space-time domain are applicable, and mandatory; see, e.g., [28, 37] for space-time solvers in the case of parabolic equations. The derivation of space-time formulations usually results in Petrov-Galerkin schemes (see, e.g., [2, 47, 49, 55] in the case of the heat equation), where for the numerical treatment it is crucial to establish a related discrete inf-sup stability condition. This becomes even more involved in the case of the wave equation, where a CFL condition is required; see, e.g., [51]. Possible approaches to overcome such a restriction are the use of discontinuous Galerkin methods, e.g., [22, 42, 43], or introducing a suitable transformation operator such as the (modified) Hilbert transformation [40, 44, 51]. Another approach is to replace the direct variational formulation by a least-squares/minimal residual equation. This has been studied in the context of first-order least-squares systems in, e.g., [25, 26, 27, 45], and in the context of minimal residual Petrov–Galerkin discretizations in [2, 19, 53], among

*Received December 20, 2024. Accepted June 2, 2026. Published online on August 18, 2026. Recommended by Martin Gander.
†Johann Radon Institute for Computational and Applied Mathematics, Altenberger Strasse 69, 4040 Linz, Austria (christian.koethe@ricam.oeaw.ac.at).
‡Corresponding author: Olaf Steinbach. Institut für Angewandte Mathematik, TU Graz, Steyrergasse 30, 8010 Graz, Austria ({loescher, o.steinbach}@math.tugraz.at).

others. For the latter approach it is well known that the Galerkin discretization results in a mixed system, where the second variable is the Riesz lift of the residual of the primal variable. This is also the point of view that we will take.

From the well-established theory of least-squares methods, with a tremendous overview by Bochev and Gunzberger [7] (see also [5] and [4, 6]), the least-squares formulations of second-order partial differential equations come with the advantage of offering an error estimator for free, but they also double the degrees of freedom. To apply the theory, it is of main interest to consider so-called practical methods that allow one to measure the residual in a localizable norm, i.e., in a non-negative Sobolev norm allowing at least L^2-regularity. Trimmed to this interest, Führer and Karkulik introduced a first-order system least-squares method (FOSLS) for parabolic problems [27], showing the applicability and power of least-squares methods in the space-time setting. However, the first-order reformulation assumes higher regularity for the source term in order to show convergence of the method. This is unnatural to a certain extent since source terms of partial differential equations are usually considered as functionals acting on the test space and thus belong to Sobolev spaces of negative order. At this point we want to mention that recently in [25, 26], FOSLSs with a source term in a negative-order Sobolev space were analyzed by replacing the load by its finite element approximation.

In this paper, we will formulate and analyze a least-squares method for the solution of abstract operator equations that overcomes this problem, i.e., we will be able to formulate the method even when the source is of minimal regularity. For X and Y being Sobolev spaces of non-negative order, we consider the problem of finding $u \in X$ such that $Bu = f \in Y^*$, where $B : X \to Y^*$ is an isomorphism. Then the problem is equivalent to minimizing the residual $\frac{1}{2}\|Bv - f\|^2_{Y^*}$ over all $v \in X$. As Y^* is the dual space of Y, the residual is measured in a Sobolev norm with negative index. The main idea is now based on the lifting of the negative-order Sobolev norm using a bijective operator $A : Y \to Y^*$ in order to keep the method practical. Then the problem is equivalent to minimizing the functional $\frac{1}{2}\|Bv - f\|^2_{A^{-1}}$, for which the minimizer $u \in X$ is characterized as the unique solution of the gradient equation $B^*A^{-1}(Bu - f) = 0$. Introducing the auxiliary variable $p = A^{-1}(f - Bu) \in Y$, we end up with a saddle point formulation. This saddle point problem has a similar form as those obtained from optimal control with energy regularization; see [35, 36, 39]. The discretization follows by standard means, using conforming finite element spaces of lowest order. To ensure uniqueness, we need a discrete inf-sup stability condition. Though, $p \equiv 0$ is valid on the continuous level, for the discrete Lagrange multiplier in general it holds that $p_h \neq 0$. Thus, it can be used to define an a posteriori error estimator that is shown to be efficient and reliable using a so-called saturation assumption [16, 21]. The idea of lifting the norm was already considered in [9, 10] for elliptic partial differential equations, or more recently in [41] for a larger class of problems, and in [17] by adding a stabilization term when considering ill-posed problems. In [2] this concept was already applied in the context of space-time methods to parabolic equations, with a discretization on a tensor-product mesh. The novelty of our method is the application to space-time problems on completely unstructured meshes and the practicability even when using negative Sobolev norms as well as the proof of an efficient and reliable error estimator in this setting. Moreover, the use of completely unstructured space-time meshes allows for full space-time adaptivity and, later on, for a parallel solution of the resulting linear systems of algebraic equations.

This paper is organized as follows. In Section 2 we consider the least-squares approach for the solution of an operator equation $Bu = f$ in an abstract sense. We derive the saddle point formulation and show ellipticity of the related operator $S = B^*A^{-1}B$. Based on a discrete inf-sup condition, we show unique solvability of the discrete system and derive related a priori error estimates. Using a so-called saturation assumption, we are able to prove that

ETNA
Kent State University and
Johann Radon Institute (RICAM)

the use of the dual p_h as error estimator is efficient and reliable. In Section 3 we apply the approach to the Poisson equation, showing that it is also related to the well-known $h - \frac{h}{2}$ estimator [15, 24]. Section 4 deals with the application of the approach to the heat equation, while the wave equation is considered in Section 5. In all cases we provide numerical examples to illustrate our theoretical findings. Finally, in Section 6 we give a conclusion, where we mention possible extensions and further work that needs to be done.

2. Abstract setting. Let $X \subset H_X \subset X^*$ and $Y \subset H_Y \subset Y^*$ be Gelfand triples of Hilbert spaces, where X^* and Y^* are the duals of X and Y with respect to H_X and H_Y, respectively, and with the duality pairings $\langle g, v\rangle_{H_X}$, for $g \in X^*$, $v \in X$, and $\langle f, q\rangle_{H_Y}$, for $f \in Y^*$ and $q \in Y$. Let $A : Y \to Y^*$ be a bounded linear operator that is assumed to be self-adjoint and elliptic in Y. Therefore, A implies a norm in Y, i.e., $\|q\|_Y := \sqrt{\langle Aq, q\rangle_{H_Y}}$, for $q \in Y$. For $f \in Y^*$, the norm is given by duality,

$$\|f\|_{Y^*} := \sup_{0\neq q\in Y} \frac{\langle f, q\rangle_{H_Y}}{\|q\|_Y}. \tag{2.1}$$

LEMMA 2.1. *The dual norm* (2.1) *allows the representations*

$$\|f\|^2_{Y^*} = \|p_f\|^2_Y = \langle Ap_f, p_f\rangle_{H_Y} = \langle f, p_f\rangle_{H_Y} = \langle A^{-1}f, f\rangle_{H_Y}, \tag{2.2}$$

where $p_f \in Y$ is the unique solution of the variational formulation

$$\langle Ap_f, q\rangle_{H_Y} = \langle f, q\rangle_{H_Y} \qquad \text{for all } q \in Y. \tag{2.3}$$

Proof. Since $A : Y \to Y^*$ is bounded and elliptic, $p_f = A^{-1}f$ is well defined as the unique solution of the variational formulation (2.3), implying $\|p_f\|_Y > 0$ when $\|f\|_{Y^*} > 0$. From the definition (2.1), we have

$$\|f\|_{Y^*} = \sup_{0\neq q\in Y} \frac{\langle f, q\rangle_{H_Y}}{\|q\|_Y} \geq \frac{\langle f, p_f\rangle_{H_Y}}{\|p_f\|_Y}, \qquad \text{i.e.,} \quad \langle f, p_f\rangle_{H_Y} \leq \|f\|_{Y^*}\|p_f\|_Y.$$

Hence, we obtain

$$\|p_f\|^2_Y = \langle Ap_f, p_f\rangle_{H_Y} = \langle f, p_f\rangle_{H_Y} \leq \|f\|_{Y^*}\|p_f\|_Y, \qquad \text{i.e.,} \quad \|p_f\|_Y \leq \|f\|_{Y^*}.$$

On the other hand, (2.1) implies, when using (2.3) and the Cauchy–Schwarz inequality $\langle Ap_f, q\rangle_{H_Y} \leq \|p_f\|_Y\|q\|_Y$,

$$\|f\|_{Y^*} = \sup_{0\neq q\in Y} \frac{\langle f, q\rangle_{H_Y}}{\|q\|_Y} = \sup_{0\neq q\in Y} \frac{\langle Ap_f, q\rangle_{H_Y}}{\|q\|_Y} \leq \|p_f\|_Y,$$

and hence, $\|f\|_{Y^*} = \|p_f\|_Y$ follows. With this we finally obtain

$$\|f\|^2_{Y^*} = \|p_f\|^2_Y = \langle Ap_f, p_f\rangle_{H_Y} = \langle f, p_f\rangle_{H_Y} = \langle f, A^{-1}f\rangle_{H_Y}. \qquad \square$$

Let $B : X \to Y^*$ be a bounded linear operator that satisfies an inf-sup condition, i.e., there exist positive constants c_1^B and c_2^B such that

$$\|Bv\|_{Y^*} \leq c_2^B\,\|v\|_X, \qquad \sup_{0\neq q\in Y} \frac{\langle Bv, q\rangle_{H_Y}}{\|q\|_Y} \geq c_1^B\,\|v\|_X \qquad \text{for all } v \in X. \tag{2.4}$$

In addition, we assume that B is surjective. Then, $B : X \to Y^*$ is an isomorphism. Due to the assumptions made, we conclude unique solvability of the operator equation of finding $u \in X$ such that $Bu = f$ in Y^* is satisfied. The solution of the operator equation $Bu = f$ in Y^* is equivalent to the minimization of a quadratic functional for $v \in X$,

$$\begin{aligned}\mathcal{J}(v) &= \frac{1}{2}\,\|Bv - f\|^2_{Y^*} \;=\; \frac{1}{2}\,\langle A^{-1}(Bv - f), Bv - f\rangle_{H_Y} \\ &= \frac{1}{2}\,\langle B^* A^{-1} Bv, v\rangle_{H_X} - \langle B^* A^{-1} f, v\rangle_{H_X} + \frac{1}{2}\,\langle A^{-1} f, f\rangle_{H_Y},\end{aligned}$$

whose minimizer $u \in X$ is given as solution of the gradient equation

$$B^* A^{-1}(Bu - f) = 0\,,$$

i.e., we have to solve the operator equation

$$Su := B^* A^{-1} Bu = B^* A^{-1} f\,. \tag{2.5}$$

Note that $B^* : Y \to X^*$ is the adjoint of $B : X \to Y^*$, i.e.,

$$\langle B^* q, v\rangle_{H_X} := \langle q, Bv\rangle_{H_Y} \qquad \text{for all } v \in X,\; q \in Y,$$

satisfying

$$\|B^* q\|_{X^*} = \sup_{0 \neq v \in X} \frac{\langle B^* q, v\rangle_{H_X}}{\|v\|_X} = \sup_{0 \neq v \in X} \frac{\langle q, Bv\rangle_{H_Y}}{\|v\|_X} \leq \sup_{0 \neq v \in X} \frac{\|q\|_Y \|Bv\|_{Y^*}}{\|v\|_X} \leq c_2^B\,\|q\|_Y,$$

for all $q \in Y$.

LEMMA 2.2. *The operator $S := B^* A^{-1} B : X \to X^*$ is bounded and elliptic, i.e.,*

$$\|Su\|_{X^*} \leq c_2^S\,\|u\|_X, \qquad \langle Su, u\rangle_{H_X} \geq c_1^S\,\|u\|^2_X \qquad \textit{for all } u \in X,$$

where $c_2^S = [c_2^B]^2$, $c_1^S = [c_1^B]^2$.

Proof. For $u \in X$ we first have, using (2.2) for $f = Bu$,

$$\begin{aligned}\|Su\|_{X^*} &= \sup_{0 \neq v \in X} \frac{\langle Su, v\rangle_{H_X}}{\|v\|_X} = \sup_{0 \neq v \in X} \frac{\langle A^{-1} Bu, Bv\rangle_{H_Y}}{\|v\|_X} \\ &\leq \sup_{0 \neq v \in X} \frac{\|A^{-1} Bu\|_Y \|Bv\|_{Y^*}}{\|v\|_X} = \sup_{0 \neq v \in X} \frac{\|Bu\|_{Y^*} \|Bv\|_{Y^*}}{\|v\|_X} \leq [c_2^B]^2\,\|u\|_X\,.\end{aligned}$$

In addition, we define $p_u = A^{-1} Bu \in Y$ to obtain

$$\|p_u\|^2_Y = \langle Ap_u, p_u\rangle_{H_Y} = \langle B^* A^{-1} Bu, u\rangle_{H_X} = \langle Su, u\rangle_{H_X}\,.$$

From the inf-sup stability condition in (2.4), we then conclude

$$c_1^B\,\|u\|_X \leq \sup_{0 \neq q \in Y} \frac{\langle Bu, q\rangle_{H_Y}}{\|q\|_Y} = \sup_{0 \neq q \in Y} \frac{\langle Ap_u, q\rangle_{H_Y}}{\|q\|_Y} \leq \|p_u\|_Y,$$

i.e.,

$$[c_1^B]^2\,\|u\|^2_X \leq \|p_u\|^2_Y \;=\; \langle Su, u\rangle_{H_X}. \qquad \square$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

In fact, $S \coloneqq B^* A^{-1} B : X \to X^*$ defines an equivalent norm in X,

$$\|v\|_S \coloneqq \sqrt{\langle Sv, v\rangle_{H_X}} = \sqrt{\langle A^{-1}Bv, Bv\rangle_{H_Y}} = \|Bv\|_{Y^*}\,,$$

satisfying

$$c_1^B\,\|v\|_X \le \|v\|_S \le c_2^B\,\|v\|_X \qquad \text{for all } v \in X. \tag{2.6}$$

The variational formulation of the operator equation (2.5) is to find $u \in X$ such that

$$\langle Su, v\rangle_{H_X} = \langle B^* A^{-1} f, v\rangle_{H_X} \qquad \text{for all } v \in X, \tag{2.7}$$

which is uniquely solvable for all $f \in Y^*$.

Instead of the operator equation (2.5) and when using $p \coloneqq A^{-1}(f - Bu)$, we now consider the equivalent coupled operator equation

$$Ap + Bu = f, \qquad B^* p = 0\,,$$

where by construction we have $p = 0$. The related variational formulation is to find a pair $(u, p) \in X \times Y$ such that

$$\langle Ap, q\rangle_{H_Y} + \langle Bu, q\rangle_{H_Y} = \langle f, q\rangle_{H_Y}\,, \qquad \langle p, Bv\rangle_{H_Y} = 0 \tag{2.8}$$

is satisfied for all $(v, q) \in X \times Y$.

For the Galerkin discretization of (2.8) we introduce conforming finite element spaces $X_H = \text{span}\{\varphi_k\}_{k=1}^{M_X} \subset X$ and $Y_h = \text{span}\{\psi_i\}_{i=1}^{M_Y} \subset Y$ which are defined with respect to some admissible but mutually different decompositions of the computational domain into shape-regular simplicial finite elements of mesh size H and h, respectively; see, e.g., [13, 48]. Then the Galerkin variational formulation of (2.8) is to find $(u_H, p_h) \in X_H \times Y_h$ such that

$$\langle Ap_h, q_h\rangle_{H_Y} + \langle Bu_H, q_h\rangle_{H_Y} = \langle f, q_h\rangle_{H_Y}\,, \qquad \langle p_h, Bv_H\rangle_{H_Y} = 0 \tag{2.9}$$

is satisfied for all $(v_H, q_h) \in X_H \times Y_h$. This is equivalent to the linear system of algebraic equations

$$\begin{bmatrix} A_h & B_h \\ B_h^\top & \end{bmatrix} \begin{bmatrix} \underline{p} \\ \underline{u} \end{bmatrix} = \begin{bmatrix} \underline{f} \\ \underline{0} \end{bmatrix},$$

where, for $i, j = 1, \ldots, M_Y$ and $k = 1, \ldots, M_X$,

$$A_h[j,i] = \langle A\psi_i, \psi_j\rangle_{H_Y}\,, \qquad B_h[j,k] = \langle B\varphi_k, \psi_j\rangle_{H_Y}\,, \qquad f_j = \langle f, \psi_j\rangle_{H_Y}\,.$$

Since A_h is symmetric and positive definite, and hence invertible, we obtain the Schur complement system

$$B_h^\top A_h^{-1} B_h \underline{u} = B_h^\top A_h^{-1} \underline{f}\,. \tag{2.10}$$

To ensure unique solvability of (2.10), we assume the discrete inf-sup stability condition

$$c_S\,\|v_H\|_X \le \sup_{0 \ne q_h \in Y_h} \frac{\langle Bv_H, q_h\rangle_{H_Y}}{\|q_h\|_Y} \qquad \text{for all } v_H \in X_H. \tag{2.11}$$

In this case, the symmetric Schur complement matrix $S_h \coloneqq B_h^\top A_h^{-1} B_h$ is positive definite and hence invertible.

LEMMA 2.3. *Assume the discrete inf-sup stability condition (2.11) to be satisfied. Then the Schur complement matrix $S_h = B_h^\top A_h^{-1} B_h$ is positive definite, i.e., the inequalities for norm equivalence hold,*

$$[c_S]^2\,\|v_H\|_X^2 \;\le\; (S_h \underline{v}, \underline{v}) \;\le\; [c_2^B]^2\,\|v_H\|_X^2, \tag{2.12}$$

for all $\underline{v} \in \mathbb{R}^{M_X} \Leftrightarrow v_H \in X_H$.

Proof. For $\underline{v} \in \mathbb{R}^{M_X} \Leftrightarrow v_H \in X_H$, we at first have

$$\begin{aligned}(S_h\underline{v},\underline{v}) &= (A_h^{-1}B_h\underline{v}, B_h\underline{v}) = (\underline{p}_v, B_h\underline{v}) = \langle Bv_H, p_{v,h}\rangle_{H_Y}\\ &= \langle Ap_{v,h}, p_{v,h}\rangle_{H_Y} = \|p_{v,h}\|_Y^2,\end{aligned}$$

where $\underline{p}_v = A_h^{-1}B_h\underline{v} \Leftrightarrow p_{v,h} \in Y_h$ solves

$$\langle Ap_{v,h}, q_h\rangle_{H_Y} = \langle Bv_H, q_h\rangle_{H_Y} \qquad \text{for all } q_h \in Y_h.$$

On the other hand, the discrete inf-sup stability condition (2.11) implies

$$c_S\,\|v_H\|_X \le \sup_{0\neq q_h\in Y_h} \frac{\langle Bv_H, q_h\rangle_{H_Y}}{\|q_h\|_Y} = \sup_{0\neq q_h\in Y_h} \frac{\langle Ap_{v,h}, q_h\rangle_{H_Y}}{\|q_h\|_Y} \le \|p_{v,h}\|_Y,$$

and hence,

$$c_S^2\,\|v_H\|_X^2 \;\le\; \|p_{v,h}\|_Y^2 \;=\; (S_h\underline{v},\underline{v})$$

follows. Note that $p_{v,h} \in Y_h$ is the Galerkin projection of $p_v = A^{-1}Bv_H$, satisfying $\|p_{v,h}\|_Y \le \|p_v\|_Y$. This also implies

$$(S_h\underline{v},\underline{v}) = \|p_{v,h}\|_Y^2 \le \|p_v\|_Y^2 = \langle A^{-1}Bv_H, Bv_H\rangle_{H_Y} = \|v_H\|_S^2 \le [c_2^B]^2\,\|v_H\|_X^2. \qquad \square$$

For the solution $\underline{u} \in \mathbb{R}^{M_X} \Leftrightarrow u_H \in X_H$ of (2.10) we now obtain, since A_h is symmetric and positive definite,

$$(S_h\underline{u},\underline{u}) = (B_h^\top A_h^{-1}B_h\underline{u},\underline{u}) = (B_h^\top A_h^{-1}\underline{f},\underline{u}) = (A_h^{-1}\underline{f}, B_h\underline{u}) \le \|\underline{f}\|_{A_h^{-1}}\|B_h\underline{u}\|_{A_h^{-1}}.$$

When using

$$\|B_h\underline{u}\|_{A_h^{-1}}^2 = (A_h^{-1}B_h\underline{u}, B_h\underline{u}) = (S_h\underline{u},\underline{u}),$$

we further conclude

$$\begin{aligned}(S_h\underline{u},\underline{u}) &\le \|\underline{f}\|_{A_h^{-1}}^2 = (A_h^{-1}\underline{f},\underline{f}) = (A_h\underline{p}_f,\underline{p}_f)\\ &= \langle Ap_{f,h}, p_{f,h}\rangle_{H_Y} = \|p_{f,h}\|_Y^2 \le \|p_f\|_Y^2 = \|f\|_{Y^*}^2,\end{aligned}$$

where $\underline{p}_f = A_h^{-1}\underline{f} \in \mathbb{R}^{M_Y} \Leftrightarrow p_{f,h} \in Y_h$ is the Galerkin approximation of $p_f = A^{-1}f$. With (2.12) we finally obtain

$$[c_S]^2\,\|u_H\|_X^2 \le (S_h\underline{u},\underline{u}) \le \|f\|_{Y^*}^2.$$

For arbitrary $u \in X$ and $f = Bu$, the solution of (2.9) defines the Galerkin projection $u_H = G_H u$, satisfying

$$\|G_H u\|_X = \|u_H\|_X \le \frac{1}{c_S}\,\|Bu\|_{Y^*} \le \frac{c_2^B}{c_S}\,\|u\|_X\,, \qquad \text{i.e.,} \quad \|G_H\|_{X\to X} \le \frac{c_2^B}{c_S}\,. \tag{2.13}$$

Note that $G_H : X \to X_H$ is indeed a projection. For $\underline{w} \in \mathbb{R}^{M_X} \Leftrightarrow w_H \in X_H$ we have

ETNA
Kent State University and
Johann Radon Institute (RICAM)

$f = Bw_H$, implying $\underline{f} = B_h\underline{w}$. In this case, (2.10) is

$$S_h\underline{u} = B_h^\top A_h^{-1} B_h \underline{u} = B_h^\top A_h^{-1}\underline{f} = B_h^\top A_h^{-1} B_h\underline{w} = S_h\underline{w},$$

and since S_h is invertible, $\underline{u} = \underline{w}$, i.e., $w_H = u_H = G_H w_H$ follows.

Now, as in [2, Theorem 3.1], we can prove the following best-approximation result. Although the basic idea of proof is as in [2], our arguments are mainly of algebraic nature.

LEMMA 2.4. *Let $u \in X$ be the unique solution of (2.7), and let $\underline{u} \in \mathbb{R}^{M_X} \Leftrightarrow u_H \in X_H$ be the unique solution of (2.10), respectively. Then there holds*

$$\|u - u_H\|_X \le \frac{c_2^B}{c_S} \inf_{v_H \in X_H} \|u - v_H\|_X. \tag{2.14}$$

Proof. The assertion follows from the projection property $v_H = G_H v_H$ for all $v_H \in X_H$, the estimate (2.13), and $\|I - G_H\|_{X\to X} = \|G_H\|_{X\to X}$ (see [56, Lemma 5]):

$$\|u - u_H\|_X = \|(I - G_H)(u - v_H)\|_X \le \|I - G_H\|_{X\to X}\|u - v_H\|_X \le \frac{c_2^B}{c_S}\|u - v_H\|_X,$$

for all $v_H \in X_H$. $\square$

While in the continuous case we have $p \coloneqq A^{-1}(f - Bu) \equiv 0$, this is in general not the case for $p_h \in Y_h$. If the Galerkin matrix B_h is invertible, then so is $B_h^\top$, and hence, $p_h \equiv 0$ follows in this particular case. But in general, the discrete inf-sup stability condition (2.11) involves spaces such that $\dim(X_H) \neq \dim(Y_h)$. Thus, B_h is not a square matrix and hence not invertible at all. But then we can use $\|p_h\|_Y$ to define an a posteriori error indicator for $\|u - u_H\|_X$. For the next two results, see also [14, Theorem 2.1].

LEMMA 2.5. *Let $(u_H, p_h) \in X_H \times Y_h$ be the unique solution of (2.9). Then,*

$$\|p_h\|_Y \le \|u - u_H\|_S \le c_2^B \|u - u_H\|_X\,. \tag{2.15}$$

Proof. When subtracting the Galerkin formulation (2.9) from (2.8), for $q = q_h \in Y_h \subset Y$ and $v = v_H \in X_H \subset X$, this gives the Galerkin orthogonalities

$$\langle A(p - p_h), q_h\rangle_{H_Y} + \langle B(u - u_H), q_h\rangle_{H_Y} = 0, \qquad \langle p - p_h, Bv_H\rangle_{H_Y} = 0,$$

for all $(v_H, q_h) \in X_H \times Y_h$. In particular for $p \equiv 0$ this gives

$$\langle Ap_h, q_h\rangle_{H_Y} = \langle B(u - u_H), q_h\rangle_{H_Y}, \quad \langle Bv_H, p_h\rangle_{H_Y} = 0 \quad \text{for all } (v_H, q_h) \in X_H \times Y_h.$$

Hence, when choosing $q_h = p_h \in Y_h$, we further conclude

$$\|p_h\|_Y^2 = \langle Ap_h, p_h\rangle_{H_Y} = \langle B(u - u_H), p_h\rangle_{H_Y} \le \|B(u - u_H)\|_{Y^*}\|p_h\|_Y,$$

i.e.,

$$\|p_h\|_Y \le \|B(u - u_H)\|_{Y^*} = \|u - u_H\|_S,$$

and the assertion finally follows from (2.6). $\square$

It remains to verify the reliability of the error estimator $\|p_h\|_Y$. In addition to (2.9) we therefore consider the variational formulation to find $(w_H, r_h) \in X_H \times Y_h$ such that

$$\langle Ar_h, q_h\rangle_{H_Y} - \langle Bw_H, q_h\rangle_{H_Y} = 0, \qquad \langle r_h, Bv_H\rangle_{H_Y} = \langle r, Bv_H\rangle_{H_Y} \tag{2.16}$$

is satisfied for all $(v_H, q_h) \in X_H \times Y_h$ when $r \in Y$ is given. This is equivalent to the linear system $S_h \underline{w} = \underline{g}$, $g_\ell = \langle r, B\varphi_\ell \rangle_{H_Y}$, for $\ell = 1, \dots, M_X$, and hence uniquely solvable due to Lemma 2.3. In particular, for $q_h = r_h$ and $v_H = w_H$, (2.16) gives

$$\|r_h\|_Y^2 = \langle A r_h, r_h \rangle_{H_Y} = \langle B w_H, r_h \rangle_{H_Y} = \langle B w_H, r \rangle_{H_Y} \le c_2^B \, \|w_H\|_X \|r\|_Y \, .$$

On the other hand, the discrete inf-sup condition (2.11) implies

$$c_S \, \|w_H\|_X \le \sup_{0 \ne q_h \in Y_h} \frac{\langle B w_H, q_h \rangle_{H_Y}}{\|q_h\|_Y} = \sup_{0 \ne q_h \in Y_h} \frac{\langle A r_h, q_h \rangle_{H_Y}}{\|q_h\|_Y} \le \|r_h\|_Y ,$$

i.e.,

$$\|r_h\|_Y \le \frac{c_2^B}{c_S} \, \|r\|_Y \qquad \text{for all } r \in Y$$

follows. The solution $r_h = \Pi_h r$ of (2.16) therefore defines the Fortin operator $\Pi_h : Y \to Y_h$ satisfying (see also [8])

$$\langle r - \Pi_h r, B v_H \rangle_{H_Y} = 0 \quad \text{for all } v_H \in X_H, \quad \|\Pi_h r\|_Y \le \frac{c_2^B}{c_S} \, \|r\|_Y \quad \text{for all } r \in Y. \tag{2.17}$$

LEMMA 2.6. *Let* $(u_H, p_h) \in X_H \times Y_h$ *be the unique solution of the Galerkin variational formulation* (2.9)*, and let* $u = B^{-1} f$*. Then the error estimator* $\|p_h\|_Y$ *is reliable modulo a data-oscillation term*

$$\operatorname{osc}(f) := \sup_{0 \ne q \in Y} \frac{\langle f, (I - \Pi_h) q \rangle_{H_Y}}{\|q\|_Y}, \tag{2.18}$$

satisfying

$$\|u - u_H\|_X \le \frac{1}{c_S} \frac{c_2^B}{c_1^B} \, \|p_h\|_Y + \frac{1}{c_1^B} \operatorname{osc}(f) \, . \tag{2.19}$$

Proof. Using the inf-sup condition in (2.4) and the properties (2.17) of the Fortin operator Π_h gives

$$\begin{aligned}
c_1^B \, \|u - u_H\|_X &\le \sup_{0 \ne q \in Y} \frac{\langle B(u - u_H), q \rangle_{H_Y}}{\|q\|_Y} \\
&= \sup_{0 \ne q \in Y} \frac{\langle B(u - u_H), q - \Pi_h q \rangle_{H_Y} + \langle B(u - u_H), \Pi_h q \rangle_{H_Y}}{\|q\|_Y} \\
&= \sup_{0 \ne q \in Y} \frac{\langle Bu, q - \Pi_h q \rangle_{H_Y} + \langle A p_h, \Pi_h q \rangle_{H_Y}}{\|q\|_Y} \\
&= \sup_{0 \ne q \in Y} \frac{\langle f, q - \Pi_h q \rangle_{H_Y} + \langle A p_h, \Pi_h q \rangle_{H_Y}}{\|q\|_Y} \le \operatorname{osc}(f) + \frac{c_2^B}{c_S} \, \|p_h\|_Y ,
\end{aligned}$$

from which (2.19) follows. ◻

Using (2.17), we obtain

$$\begin{aligned}
\operatorname{osc}(f) = \sup_{0 \ne q \in Y} \frac{\langle f, (I - \Pi_h) q \rangle_{H_Y}}{\|q\|_Y} &= \sup_{0 \ne q \in Y} \frac{\langle Bu, (I - \Pi_h) q \rangle_{H_Y}}{\|q\|_Y} \\
= \sup_{0 \ne q \in Y} \frac{\langle B(u - v_H), (I - \Pi_h) q \rangle_{H_Y}}{\|q\|_Y} &\le c_2^B \, \|I - \Pi_h\|_{Y \to Y} \|u - v_H\|_X ,
\end{aligned}$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

for all $v_H \in X_H$, and with $\|I - \Pi_h\|_{Y\to Y} = \|\Pi_h\|_{Y\to Y}$ (see again [56, Lemma 5]) we get

$$\operatorname{osc}(f) \leq \frac{[c_2^B]^2}{c_S} \inf_{v_H \in X_H} \|u - v_H\|_X,$$

i.e., the data-oscillation term is at least of the same order as the error $\|u - u_H\|_X$. However, for sufficiently smooth data we can expect that the data-oscillation term (2.18) is of higher order than $\|u - u_H\|_X$, in particular when using local polynomials of different orders to define the ansatz spaces Y_h and X_H, respectively; see also [14]. In this case, the error estimator $\|p_h\|_Y$ is asymptotically reliable. However, there is probably no need to drive an adaptive scheme when the solution is sufficiently regular.

For small data-oscillations it is shown in [21] that a so-called saturation assumption is valid. In such a case, we can give an alternative proof of the reliability estimate. For this we introduce an ansatz space $X_{\overline{H}} \subset X$ such that $X_H \subset X_{\overline{H}}$ is satisfied. As in (2.11) we assume the discrete inf-sup stability condition

$$\bar{c}_S \, \|v_{\overline{H}}\|_X \leq \sup_{0 \neq q_h \in Y_h} \frac{\langle B v_{\overline{H}}, q_h\rangle_{H_Y}}{\|q_h\|_Y} \qquad \text{for all } v_{\overline{H}} \in X_{\overline{H}}. \tag{2.20}$$

Due to $X_H \subset X_{\overline{H}}$, we have that (2.11) is a direct consequence of (2.20). Using (2.20) we can determine $(u_{\overline{H}}, \overline{p}_h) \in X_{\overline{H}} \times Y_h$ as the unique solution of the variational formulation such that

$$\langle A\overline{p}_h, q_h\rangle_{H_Y} + \langle B u_{\overline{H}}, q_h\rangle_{H_Y} = \langle f, q_h\rangle_{H_Y}, \qquad \langle \overline{p}_h, B v_{\overline{H}}\rangle_{H_Y} = 0 \tag{2.21}$$

is satisfied for all $(v_{\overline{H}}, q_h) \in X_{\overline{H}} \times Y_h$.

LEMMA 2.7. *Let $(u_H, p_h) \in X_H \times Y_h$ and $(u_{\overline{H}}, \overline{p}_h) \in X_{\overline{H}} \times Y_h$ be the unique solutions of the Galerkin variational formulations* (2.9) *and* (2.21)*, respectively. Assume the saturation assumption*

$$\|u - u_{\overline{H}}\|_X \leq \eta \, \|u - u_H\|_X \qquad \textit{for some } \eta \in (0,1). \tag{2.22}$$

Then the error estimator $\|p_h\|_Y$ is reliable, satisfying

$$\|u - u_H\|_X \leq \frac{1}{1-\eta} \, \frac{c_2^B}{\bar{c}_S^2} \, \|p_h\|_Y \, . \tag{2.23}$$

Proof. Subtracting the Galerkin variational formulation (2.21) from (2.9) gives the Galerkin orthogonality

$$\langle B(u_{\overline{H}} - u_H), q_h\rangle_{H_Y} = \langle A(p_h - \overline{p}_h), q_h\rangle_{H_Y} \qquad \text{for all } q_h \in Y_h.$$

From the discrete inf-sup stability condition (2.20), we conclude, recall that $u_H - u_{\overline{H}} \in X_{\overline{H}}$,

$$\begin{aligned}
\bar{c}_S \, \|u_H - u_{\overline{H}}\|_X &\leq \sup_{0 \neq q_h \in Y_h} \frac{\langle B(u_{\overline{H}} - u_H), q_h\rangle_{H_Y}}{\|q_h\|_Y} \\
&= \sup_{0 \neq q_h \in Y_h} \frac{\langle A(p_h - \overline{p}_h), q_h\rangle_{H_Y}}{\|q_h\|_Y} \leq \|p_h - \overline{p}_h\|_Y \, .
\end{aligned}$$

On the other hand, using the above Galerkin orthogonality and the second equation in (2.21) for $v_{\overline{H}} = u_{\overline{H}} - u_H$ gives

$$\begin{aligned}
\|p_h - \overline{p}_h\|_Y^2 &= \langle A(p_h - \overline{p}_h), p_h - \overline{p}_h\rangle_{H_Y} = \langle B(u_{\overline{H}} - u_H), p_h - \overline{p}_h\rangle_{H_Y} \\
&= \langle B(u_{\overline{H}} - u_H), p_h\rangle_{H_Y} \leq c_2^B \, \|u_{\overline{H}} - u_H\|_X \|p_h\|_Y \, .
\end{aligned}$$

Hence, we obtain

$$\|u_{\overline{H}} - u_H\|_X^2 \leq \frac{1}{\overline{c}_S^2}\, \|p_h - \overline{p}_h\|_Y^2 \leq \frac{c_2^B}{\overline{c}_S^2}\, \|u_{\overline{H}} - u_H\|_X \|p_h\|_Y ,$$

i.e.,

$$\|u_{\overline{H}} - u_H\|_X \leq \frac{c_2^B}{\overline{c}_S^2}\, \|p_h\|_Y \,.$$

Using the triangle inequality and the saturation assumption (2.22), we finally have

$$\|u - u_H\|_X \leq \|u - u_{\overline{H}}\|_X + \|u_{\overline{H}} - u_H\|_X \leq \eta\, \|u - u_H\|_X + \frac{c_2^B}{\overline{c}_S^2}\, \|p_h\|_Y ,$$

from which the assertion follows. □

It remains to define, for a given ansatz space $X_{\overline{H}}$, the test space Y_h such that the discrete inf-sup condition (2.20), and therefore (2.11), is satisfied. This can be achieved by assuming a sufficiently rich test space Y_h, as stated in the following abstract approach. However, as we will see later, this is not always required since one may establish (2.20) in a different way.

THEOREM 2.8. *For a given finite element space $X_{\overline{H}} \subset X$, let $Y_h \subset Y$ such that*

$$\sup_{0 \neq v_{\overline{H}} \in X_{\overline{H}}} \inf_{q_h \in Y_h} \frac{\|p_{v_{\overline{H}}} - q_h\|_Y}{\|p_{v_{\overline{H}}}\|_Y} \leq \delta \tag{2.24}$$

is satisfied for some $\delta \in (0,1)$. Then there holds the discrete inf-sup stability condition (2.20), i.e.,

$$c_1^B \sqrt{(1-\delta^2)}\, \|v_{\overline{H}}\|_X \leq \sup_{0 \neq q_h \in Y_h} \frac{\langle B v_{\overline{H}}, q_h \rangle_{H_Y}}{\|q_h\|_Y} \qquad \textit{for all } v_{\overline{H}} \in X_{\overline{H}}. \tag{2.25}$$

Proof. For an arbitrary but fixed $v_{\overline{H}} \in X_{\overline{H}}$, we define $p_{v_{\overline{H}}} = A^{-1} B v_{\overline{H}} \in Y$, and as in the proof of Lemma 2.2, we conclude

$$\|p_{v_{\overline{H}}}\|_Y^2 = \|v_{\overline{H}}\|_S^2 = \langle B v_{\overline{H}}, p_{v_{\overline{H}}} \rangle_{H_Y} \,.$$

In addition, we define $p_{v_{\overline{H}} h} \in Y_h$ as the unique solution of the Galerkin variational formulation

$$\langle A p_{v_{\overline{H}} h}, q_h \rangle_{H_Y} = \langle B v_{\overline{H}}, q_h \rangle_{H_Y} = \langle A p_{v_{\overline{H}}}, q_h \rangle_{H_Y} \qquad \text{for all } q_h \in Y_h,$$

satisfying the bound

$$\|p_{v_{\overline{H}} h}\|_Y \leq \|p_{v_{\overline{H}}}\|_Y ,$$

the orthogonal splitting

$$\|p_{v_{\overline{H}}}\|_Y^2 = \|p_{v_{\overline{H}} h}\|_Y^2 + \|p_{v_{\overline{H}}} - p_{v_{\overline{H}} h}\|_Y^2 ,$$

and Cea's lemma,

$$\|p_{v_{\overline{H}}} - p_{v_{\overline{H}} h}\|_Y \leq \inf_{q_h \in Y_h} \|p_{v_{\overline{H}}} - q_h\|_Y \,.$$

From (2.24) we then obtain

$$\|p_{v_{\overline{H}}} - p_{v_{\overline{H}} h}\|_Y \leq \delta\, \|p_{v_{\overline{H}}}\|_Y \,.$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

Hence, we can write

$$\sup_{0\neq q_h\in Y_h}\frac{\langle Bv_{\overline{H}},q_h\rangle_{H_Y}}{\|q_h\|_Y}=\sup_{0\neq q_h\in Y_h}\frac{\langle Ap_{v_{\overline{H}}h},q_h\rangle_{H_Y}}{\|q_h\|_Y}=\frac{\langle Ap_{v_{\overline{H}}h},p_{v_{\overline{H}}h}\rangle_{H_Y}}{\|p_{v_{\overline{H}}h}\|_Y}=\|p_{v_{\overline{H}}h}\|_Y,$$

as well as

$$\|p_{v_{\overline{H}}h}\|_Y^2\;=\;\|p_{v_{\overline{H}}}\|_Y^2-\|p_{v_{\overline{H}}}-p_{v_{\overline{H}}h}\|_Y^2\;\geq\;(1-\delta^2)\,\|p_{v_{\overline{H}}}\|_Y^2\;=\;(1-\delta^2)\,\|v_{\overline{H}}\|_S^2,$$

implying the inf-sup stability condition (2.25) when using (2.6). $\square$

In some applications, e.g., when considering space-time finite element methods for the heat equation as in [49], the discrete inf-sup condition can be established when using a (discretization-dependent) norm for the ansatz space X. In particular, let $\|\cdot\|_{X,h}:X\to\mathbb{R}$ define a norm on X_H satisfying $\|v\|_{X,h}\leq\|v\|_X$ for all $v\in X$, and assume that

$$\widetilde{c}_S\,\|v_H\|_{X,h}\leq\sup_{0\neq q_h\in Y_h}\frac{\langle Bv_H,q_h\rangle_{H_Y}}{\|q_h\|_Y}\qquad\text{for all } v_H\in X_H. \tag{2.26}$$

As in the proof of Lemma 2.3, replacing the discrete inf-sup stability condition (2.11) by (2.26), we then conclude

$$(S_h\underline{v},\underline{v})\geq[\widetilde{c}_S]^2\,\|v_H\|_{X,h}^2\qquad\text{for all } \underline{v}\in\mathbb{R}^{M_X}\Leftrightarrow v_H\in X_H.$$

This ensures unique solvability of (2.10), and the error estimate (2.14) remains valid,

$$\|u-u_H\|_{X,h}\leq\frac{c_2^B}{\widetilde{c}_S}\inf_{v_H\in X_H}\|u-v_H\|_X. \tag{2.27}$$

In what follows we discuss several applications of this abstract setting. Although our main interest is in the discretization of time-dependent partial differential equations such as the heat and the wave equation, we first consider an elliptic problem in order to present the main ideas of this approach for several examples that are well known in the literature. Crucial is the choice of the finite element spaces $X_{\overline{H}}$ and Y_h such that the discrete inf-sup condition (2.20) is satisfied. As we will see and depending on the particular partial differential equation to be solved, we may even consider $X_{\overline{H}}=Y_h$, or we may consider stable pairs of finite element functions which are defined with respect to the same finite element mesh. In the most general case we may choose first a finite element space X_H, then a space $X_{\overline{H}}$ to ensure the saturation condition (2.22), and finally Y_h to satisfy the discrete inf-sup stability condition (2.20), e.g., we may define the ansatz space Y_h with respect to a sufficiently small mesh size $h<c_0\overline{H}$ in order to meet (2.24).

3. Elliptic Dirichlet boundary value problem. As a first example we consider the Dirichlet boundary value problem for the Poisson equation,

$$-\Delta u(x)=f(x)\quad\text{for } x\in\Omega,\qquad u(x)=0\quad\text{for } x\in\Gamma, \tag{3.1}$$

where $\Omega\subset\mathbb{R}^n$, $n=2,3$, is a bounded Lipschitz domain with boundary $\Gamma=\partial\Omega$. With respect to the abstract setting, we have $H_X=H_Y=L^2(\Omega)$ and

$$X=Y=H_0^1(\Omega),\qquad A=B=-\Delta:H_0^1(\Omega)\to H^{-1}(\Omega),\qquad \|v\|_X=\|\nabla v\|_{L^2(\Omega)}.$$

We obviously have (2.4) with $c_1^B=c_2^B=1$. In this case, the abstract variational formulation (2.8) reads to find $(u,p)\in H_0^1(\Omega)\times H_0^1(\Omega)$ such that

$$\int_\Omega\nabla p(x)\cdot\nabla q(x)\,\mathrm{d}x+\int_\Omega\nabla u(x)\cdot\nabla q(x)\,\mathrm{d}x=\langle f,q\rangle_\Omega,\qquad\int_\Omega\nabla v(x)\cdot\nabla p(x)\,\mathrm{d}x=0$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

is satisfied for all $(v,q) \in H^1_0(\Omega) \times H^1_0(\Omega)$. Let $X_h := S^1_h(\Omega) \cap H^1_0(\Omega) = \operatorname{span}\{\varphi_k\}_{k=1}^{M_X}$, where $S^1_h(\Omega) = \operatorname{span}\{\varphi_k\}_{k=1}^{\widetilde{M}_X}$ denotes the standard finite element ansatz space of piecewise linear and continuous basis functions that are defined with respect to some admissible locally quasi-uniform decomposition of Ω into shape-regular simplicial finite elements τ^h_ℓ of local mesh size h_ℓ. By definition we have, for $0 \neq u_h \in X_h$,

$$\|\nabla u_h\|_{L^2(\Omega)} = \frac{\langle \nabla u_h, \nabla u_h\rangle_{L^2(\Omega)}}{\|\nabla u_h\|_{L^2(\Omega)}} \leq \sup_{0\neq q_h \in X_h} \frac{\langle \nabla u_h, \nabla q_h\rangle_{L^2(\Omega)}}{\|\nabla q_h\|_{L^2(\Omega)}},$$

i.e., (2.20) for $Y_h = X_{\overline{H}} = X_h$ with $\bar{c}_S = 1$. In the case $X_H = X_{\overline{H}} = X_h$, we finally obtain, due to $p_h = 0$, the standard finite element formulation for the Poisson equation, i.e.,

$$\int_\Omega \nabla u_h(x) \cdot \nabla q_h(x)\, \mathrm{d}x = \langle f, q_h\rangle_\Omega \qquad \text{for all } q_h \in X_h.$$

In order to use $\|\nabla p_h\|_{L^2(\Omega)}$ as a reliable and efficient error estimator, we need to introduce a finite element space X_H that is defined with respect to a coarser mesh. In fact, one may first define the ansatz space X_H and afterwards Y_h by applying some additional refinements. Hence, we consider the mixed variational formulation to find $(u_H, p_h) \in X_H \times Y_h$ such that

$$\begin{aligned} \int_\Omega \nabla p_h(x) \cdot \nabla q_h(x)\, \mathrm{d}x + \int_\Omega \nabla u_H(x) \cdot \nabla q_h(x)\, \mathrm{d}x &= \langle f, q_h\rangle_\Omega, \\ \int_\Omega \nabla v_H(x) \cdot \nabla p_h(x)\, \mathrm{d}x &= 0 \end{aligned} \tag{3.2}$$

is satisfied for all $(v_H, q_h) \in X_H \times Y_h$. According to Lemma 2.4 we conclude the error estimate

$$\|\nabla(u - u_H)\|_{L^2(\Omega)} \leq \inf_{v_H \in X_H} \|\nabla(u - v_H)\|_{L^2(\Omega)}.$$

Hence, assuming $u \in H^1_0(\Omega) \cap H^s(\Omega)$ for some $s \in [1,2]$, we finally obtain the error estimate

$$\|\nabla(u - u_H)\|_{L^2(\Omega)} \leq c\, H^{s-1}\, |u|_{H^s(\Omega)}.$$

In what follows we use the error estimator

$$\eta_h^2 := \|\nabla p_h\|^2_{L^2(\Omega)} = \int_\Omega |\nabla p_h(x)|^2\, \mathrm{d}x = \sum_{\ell=1}^{N_h} \int_{\tau^h_\ell} |\nabla p_h(x)|^2\, \mathrm{d}x = \sum_{\ell=1}^{N_h} \eta^2_{h,\ell}$$

with the local error indicators

$$\eta^2_{h,\ell} = \int_{\tau^h_\ell} |\nabla p_h(x)|^2\, \mathrm{d}x\,.$$

In order to define local indicators for the error $\|\nabla(u - u_H)\|_{L^2(\Omega)}$, we sum up all local contributions

$$\eta^2_{H,j} = \sum_{\tau^h_\ell \subset \tau^H_j} \eta^2_{h,\ell}\,.$$

Using the estimates (2.15) and (2.19), we have

$$\|\nabla p_h\|_{L^2(\Omega)} \leq \|\nabla(u - u_H)\|_{L^2(\Omega)} \leq \|\nabla p_h\|_{L^2(\Omega)} + \operatorname{osc}(f).$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

If the saturation assumption (2.22) is fulfilled, then we have due to (2.23)

$$\|\nabla(u-u_H)\|_{L^2(\Omega)} \leq \frac{1}{1-\eta}\,\|\nabla p_h\|_{L^2(\Omega)},$$

where the saturation assumption now reads, for $u_h \in X_h = Y_h$,

$$\|\nabla(u-u_h)\|_{L^2(\Omega)} \leq \eta\,\|\nabla(u-u_H)\|_{L^2(\Omega)} \qquad \text{for some } \eta \in (0,1).$$

For the particular choice $H = 2h$, i.e., one additional refinement to define Y_h from X_H, this is obviously related to the $h - \frac{h}{2}$ error estimator; see, e.g., [15, 24]. Although this error estimator is well established in the literature, we just present one example solving the mixed Galerkin finite element scheme (3.2). Note that in contrast to the standard $h - \frac{h}{2}$ error estimator, we compute (p_h, u_H) at once solving (3.2).

In our numerical experiment we use for the adaptive refinement the Dörfler marking [20] strategy with the parameter $\theta = 0.5$. The marked elements are then refined using newest vertex bisection. For the numerical solution of the Galerkin variational formulation (3.2), we use the ansatz space $X_H = S^1_H(\Omega) \cap H^1_0(\Omega)$, which consists of piecewise linear and globally continuous basis functions, and the test space $Y_H = S^2_H(\Omega) \cap H^1_0(\Omega)$ consisting of piecewise quadratic, globally continuous basis functions. Instead of choosing the test space as $Y_h = S^1_h(\Omega) \cap H^1_0(\Omega)$ on a uniformly refined mesh with $h = H/2$, we replace the mesh refinement by a higher polynomial degree, i.e., a p-refinement strategy. Since numerical experiments [38, Sect. 5.3] have revealed no noteworthy differences between the two approaches, all examples are carried out for the p-refined setting. All computations are done using the finite element software Netgen/NGSolve [46]. The resulting linear systems are solved using the sparse direct solver package Umfpack [18]. As an example, we consider the L-shaped domain

$$\Omega = (-1,1)^2 \setminus ([0,1] \times [-1,0]) \subset \mathbb{R}^2,$$

and we consider the Dirichlet boundary value problem (3.1) with the exact solution, given in polar coordinates,

$$u(r,\varphi) = r^{2/3}\,\sin\frac{2}{3}\varphi, \qquad u \in H^s(\Omega),\;\; s < \frac{5}{3}. \tag{3.3}$$

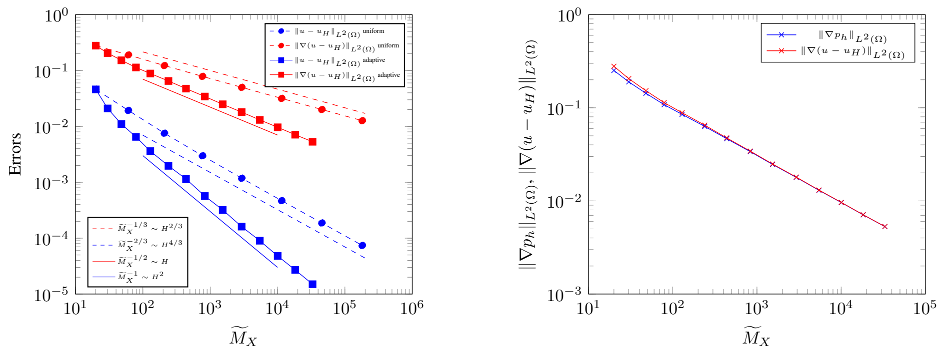
Due to the reduced regularity, we expect a reduced order of convergence of $\mathcal{O}(H^{2/3})$ for a uniform refinement when measuring the energy error $\|\nabla(u-u_H)\|_{L^2(\Omega)}$ and of $\mathcal{O}(H^{4/3})$ when considering the L^2-error $\|u-u_H\|_{L^2(\Omega)}$. When using the adaptive refinement strategy as described, we recover the optimal convergence of $\mathcal{O}(\widetilde{M}_X^{-1/2})$, which corresponds to $\mathcal{O}(H)$ in the energy norm, and of $\mathcal{O}(\widetilde{M}_X^{-1})$, which corresponds to $\mathcal{O}(H^2)$ in $L^2(\Omega)$, respectively. The related numerical results are given in Figure 3.1(a). In Figure 3.1(b) we present a comparison of the error $\|\nabla(u-u_H)\|_{L^2(\Omega)}$ and the error estimator $\eta_h = \|\nabla p_h\|_{L^2(\Omega)}$, which shows that the error estimator is effective.

REMARK 3.1. Instead of $X = Y = H^1_0(\Omega)$, $S = B^* A^{-1} B = -\Delta$, we may also consider the abstract setting with

$$\begin{aligned} &X = H_\Delta(\Omega) := \{v \in H^1_0(\Omega) : \Delta v \in L^2(\Omega)\}, && Y = L^2(\Omega), \\ &B = -\Delta : H_\Delta(\Omega) \to Y, && A = I : L^2(\Omega) \to L^2(\Omega). \end{aligned}$$

Then we have that $S = B^* A^{-1} B = \Delta^2$ is the Bi-Laplacian. For a conforming finite element discretization, in this case one may use tensor product meshes and quadratic B-splines to

ETNA
Kent State University and
Johann Radon Institute (RICAM)



(a) Errors $\|u - u_H\|_{L^2(\Omega)}$ and $\|\nabla(u - u_H)\|_{L^2(\Omega)}$ for adaptive and uniform refinement.

(b) Comparison of the error $\|\nabla(u - u_H)\|_{L^2(\Omega)}$ and the error estimator $\|\nabla p_h\|_{L^2(\Omega)}$.

FIG. 3.1. *Numerical results in the case of the singular function* (3.3).

define X_h and piecewise constant basis functions to define Y_h. For a related approach in the case of a distributed optimal control problem, see, e.g., [11, 12]. Alternatively, and applying integration by parts twice, one may exchange the roles of $X = L^2(\Omega)$ and $Y = H_\Delta(\Omega)$ in order to consider, e.g., less regular Dirichlet boundary conditions $u = g \in L^2(\Gamma)$; see, e.g., [3]. In addition, one can also apply the abstract theory as presented in this paper for the analysis of related boundary element methods [50].

4. Parabolic Dirichlet boundary value problem. As an example for a parabolic partial differential equation, we consider the Dirichlet problem for the heat equation,

$$\partial_t u(x,t) - \Delta_x u(x,t) = f(x,t) \quad \text{for } (x,t) \in Q := \Omega \times (0,T), \tag{4.1a}$$

$$u(x,t) = 0 \quad \text{for } (x,t) \in \Sigma := \Gamma \times (0,T), \tag{4.1b}$$

$$u(x,0) = 0 \quad \text{for } x \in \Omega, \tag{4.1c}$$

where $\Omega \subset \mathbb{R}^n$, $n = 1,2,3$, is a bounded Lipschitz domain with boundary $\Gamma = \partial\Omega$ and $T > 0$ is a given time horizon. In view of the abstract setting we have the Bochner spaces

$$X := L^2(0,T; H_0^1(\Omega)) \cap H^1_{0,}(0,T; H^{-1}(\Omega)),$$
$$Y := L^2(0,T; H_0^1(\Omega)), \qquad H_X = H_Y := L^2(Q),$$

where

$$H^1_{0,}(0,T;H^{-1}(\Omega)) := \{u \in L^2(Q) : \partial_t u \in L^2(0,T;H^{-1}(\Omega)), u(x,0) = 0, x \in \Omega\}.$$

The operators $A : Y \to Y^*$ and $B : X \to Y^*$ are defined in a variational sense satisfying

$$\langle Ap, q\rangle_Q := \langle \nabla_x p, \nabla_x q\rangle_{L^2(Q)}, \qquad \langle Bu, q\rangle_Q := \langle \partial_t u, q\rangle_Q + \langle \nabla_x u, \nabla_x q\rangle_{L^2(Q)},$$

for all $p, q \in Y$ and $u \in X$. The corresponding norms read

$$\|p\|_Y := \|\nabla_x p\|_{L^2(Q)}, \qquad \|u\|_X := \sqrt{\|\partial_t u\|^2_{Y^*} + \|u\|^2_Y}, \qquad \|\partial_t u\|_{Y^*} = \|w_u\|_Y,$$

where $w_u \in Y$ is the unique solution of the variational problem

$$\langle \nabla_x w_u, \nabla_x q\rangle_{L^2(Q)} = \langle \partial_t u, q\rangle_Q \qquad \text{for all } q \in Y. \tag{4.2}$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

The operator $A : Y \to Y^*$ is self-adjoint, bounded, and elliptic. Moreover, $B : X \to Y^*$ is bounded with $c_2^B = \sqrt{2}$, i.e.,

$$\|Bv\|_{Y^*} \leq \sqrt{2}\, \|v\|_X \qquad \text{for all } v \in X.$$

In order to prove an inf-sup stability condition, for $u \in X$ we first define $w_u \in Y$ as in (4.2). For $q := u + w_u \in Y$ we then have

$$\begin{aligned}
\langle Bu, q\rangle_Q &= \langle Bu, u + w_u\rangle_Q \\
&= \langle \partial_t u, u\rangle_Q + \langle \nabla_x u, \nabla_x u\rangle_{L^2(Q)} + \langle \partial_t u, w_u\rangle_Q + \langle \nabla_x u, \nabla_x w_u\rangle_{L^2(Q)} \\
&= \langle \nabla_x w_u, \nabla_x u\rangle_{L^2(Q)} + \langle \nabla_x u, \nabla_x u\rangle_{L^2(Q)} \\
&\qquad + \langle \nabla_x w_u, \nabla_x w_u\rangle_{L^2(Q)} + \langle \nabla_x u, \nabla_x w_u\rangle_{L^2(Q)} \\
&= \langle \nabla_x (u + w_u), \nabla_x (u + w_u)\rangle_{L^2(Q)} \;=\; \|q\|_Y^2 \,.
\end{aligned}$$

On the other hand,

$$\begin{aligned}
\|q\|_Y^2 &= \|\nabla_x (u + w_u)\|_{L^2(Q)}^2 \\
&= \langle \nabla_x u, \nabla_x u\rangle_{L^2(Q)} + \langle \nabla_x w_u, \nabla_x w_u\rangle_{L^2(Q)} + 2\,\langle \nabla_x w_u, \nabla_x u\rangle_{L^2(Q)} \\
&= \langle \nabla_x u, \nabla_x u\rangle_{L^2(Q)} + \langle \nabla_x w_u, \nabla_x w_u\rangle_{L^2(Q)} + 2\,\langle \partial_t u, u\rangle_Q \\
&\geq \|u\|_Y^2 + \|w_u\|_Y^2 \;=\; \|u\|_X^2
\end{aligned}$$

implies

$$\langle Bu, u + w_u\rangle_Q \geq \|u\|_X \|u + w_u\|_Y,$$

and therefore the inf-sup stability condition

$$\|u\|_X \leq \sup_{0 \neq q \in Y} \frac{\langle Bu, q\rangle_Q}{\|q\|_Y} \qquad \text{for all } u \in X$$

follows with $c_1^B = 1$. Note that this estimate is improved compared to that originally derived in [49], and it also follows from the "inf-sup identity"; see, e.g., [23, Theorem 2.1]. Moreover, $B : X \to Y^*$ is surjective; see, e.g., [35]. Thus, we can apply the abstract theory from Section 2. The variational formulation (2.8) now reads to find $(u, p) \in X \times Y$ such that

$$\begin{aligned}
&\int_Q \nabla_x p(x,t) \cdot \nabla_x q(x,t)\,\mathrm{d}x\,\mathrm{d}t \\
&\qquad + \int_Q \Big[\partial_t u(x,t)\, q(x,t) + \nabla_x u(x,t) \cdot \nabla_x q(x,t)\Big]\mathrm{d}x\,\mathrm{d}t = \langle f, q\rangle_Q, \\
&\qquad \int_Q \Big[\partial_t v(x,t)\, p(x,t) + \nabla_x v(x,t) \cdot \nabla_x p(x,t)\Big]\mathrm{d}x\,\mathrm{d}t = 0
\end{aligned} \tag{4.3}$$

is satisfied for all $(v, q) \in X \times Y$.

For the discretization of (4.3) we first consider an ansatz space $X_H = \text{span}\{\varphi_k\}_{k=1}^{M_X} \subset X$ of piecewise linear and continuous basis functions φ_k that are defined with respect to some admissible and locally quasi-uniform decomposition of the space-time domain Q into shape-regular simplicial finite elements q_ℓ^H of mesh size H_ℓ. In addition we define the ansatz space $Y_h = \text{span}\{\psi_i\}_{i=1}^{M_Y} \subset Y$ of piecewise linear and continuous basis functions ψ_i that are defined with respect to some refined decomposition of the space-time domain Q into finite elements q_i^h

ETNA
Kent State University and
Johann Radon Institute (RICAM)

of local mesh size h_i. From a practical point of view, we may use one additional refinement of the mesh that was used to define X_H, i.e., for $q_i^h \subset q_\ell^H$ we have $h_i = \frac{1}{2} H_\ell$. As an alternative, we may also use second-order form functions on q_ℓ^H to define Y_h. In both cases we have the inclusion $X_H \subset Y_h$, which enables us to prove a discrete inf-sup stability condition. For any $u \in X$ we define $w_{uh} \in Y_h$ as the unique solution of the Galerkin variational formulation

$$\langle \nabla_x w_{uh}, \nabla_x q_h \rangle_{L^2(Q)} = \langle \partial_t u, q_h \rangle_Q \qquad \text{for all } q_h \in Y_h.$$

With this we define the discrete norm

$$\|u\|_{X,h} := \sqrt{\|u\|_Y^2 + \|w_{uh}\|_Y^2} \le \|u\|_X,$$

and as in the continuous case—instead of $w_u \in Y$ we now use $w_{u_H h} \in Y_h$—we can prove the discrete inf-sup stability condition

$$\|u_H\|_{X,h} \le \sup_{0 \ne q_h \in Y_h} \frac{\langle B u_H, q_h \rangle_Q}{\|q_h\|_Y} \qquad \text{for all } u_H \in X_H.$$

Further, we see that the operator B is also bounded on the discrete level with respect to the discretization-dependent norm, i.e.,

$$\begin{aligned}
\langle Bu, q_h \rangle_Q &= \langle \partial_t u, q_h \rangle_Q + \langle \nabla_x u, \nabla_x q_h \rangle_{L^2(Q)} \\
&= \langle \nabla_x w_{uh}, \nabla_x q_h \rangle_{L^2(Q)} + \langle \nabla_x u, \nabla_x q_h \rangle_{L^2(Q)} \\
&\le \Big(\|\nabla_x w_{uh}\|_{L^2(Q)} + \|\nabla_x u\|_{L^2(Q)} \Big) \|\nabla_x q_h\|_{L^2(Q)} \le \sqrt{2}\, \|u\|_{X,h} \|q_h\|_Y,
\end{aligned}$$

for all $q_h \in Y_h$ and $u \in X$. Hence, we conclude unique solvability of the mixed space-time finite element variational formulation to find $(u_H, p_h) \in X_H \times Y_h$ such that

$$\begin{aligned}
&\int_Q \nabla_x p_h(x,t) \cdot \nabla_x q_h(x,t) \, \mathrm{d}x \, \mathrm{d}t \\
&\qquad + \int_Q \Big[\partial_t u_H(x,t)\, q_h(x,t) + \nabla_x u_H(x,t) \cdot \nabla_x q_h(x,t) \Big] \mathrm{d}x \, \mathrm{d}t = \langle f, q_h \rangle_Q, \\
&\qquad \int_Q \Big[\partial_t v_H(x,t)\, p_h(x,t) + \nabla_x v_H(x,t) \cdot \nabla_x p_h(x,t) \Big] \mathrm{d}x \, \mathrm{d}t = 0
\end{aligned} \tag{4.4}$$

is satisfied for all $(v_H, q_h) \in X_H \times Y_h$. The related error estimate now follows from the abstract estimate (2.27), i.e., with $c_2^B = \sqrt{2}$ and $\widetilde{c}_S = 1$ this gives

$$\|u - u_H\|_{X,h} \le \sqrt{2} \inf_{v_H \in X_H} \|u - v_H\|_X \, .$$

Hence, assuming $u \in H^s(Q)$ for some $s \in [1, 2]$, we conclude the error estimate (see, e.g., [49])

$$\|\nabla_x (u - u_H)\|_{L^2(Q)} \le \|u - u_H\|_{X,h} \le c\, H^{s-1}\, |u|_{H^s(Q)}.$$

For the error indicator there hold the estimates [32, Sect. 4.1.1]

$$\frac{1}{\sqrt{2}} \|\nabla_x p_h\|_{L^2(Q)} \;\le\; \|u - u_H\|_{X,h} \;\le\; \frac{\sqrt{2}}{1-\eta} \|\nabla_x p_h\|_{L^2(Q)}.$$

The efficiency estimate follows the lines of the proof of Lemma 2.5. Choosing $X_{\overline{H}} = Y_h \cap X$, the discrete inf-sup stability condition (2.20) holds with $\overline{c}_S = 1$ and with respect to the

ETNA
Kent State University and
Johann Radon Institute (RICAM)

mesh-dependent norm $\|\cdot\|_{X,h}$. If we assume the saturation assumption (2.22) with respect to the discrete norm $\|\cdot\|_{X,h}$, then we can prove reliability of the error estimator similar as in Lemma 2.6. Hence, and as in the case of the Poisson equation, we now use the global error estimator

$$\eta_h^2 = \|\nabla_x p_h\|^2_{L^2(Q)} = \int_Q |\nabla_x p_h(x,t)|^2 \,\mathrm{d}x\,\mathrm{d}t = \sum_{\ell=1}^{N_h} \int_{q_\ell^h} |\nabla_x p_h(x,t)|^2 \,\mathrm{d}x\,\mathrm{d}t = \sum_{\ell=1}^{N_h} \eta_{h,\ell}^2$$

with the local error indicators

$$\eta_{h,\ell}^2 = \int_{q_\ell^h} |\nabla_x p_h(x,t)|^2 \,\mathrm{d}x\,\mathrm{d}t, \qquad \eta_{H,j}^2 = \sum_{q_\ell^h \subset q_j^H} \eta_{h,\ell}^2,$$

to drive an adaptive refinement algorithm.

In the following we consider several numerical examples to verify our theoretical findings. We consider the Galerkin variational formulation (4.4) for the ansatz space $X_H = S_H^1(Q) \cap X$ of piecewise linear and continuous basis functions and the test space $Y_H = S_H^2(Q) \cap Y$ of piecewise second-order and continuous basis functions. The latter corresponds to the test space $Y_h = S_h^1(Q) \cap Y$ of piecewise linear continuous basis functions, which are defined with respect to a refined mesh with local mesh size $h = \frac{1}{2}H$. We use the global error estimator η_h to drive an adaptive algorithm using Dörfler marking with $\theta = 0.5$ and newest vertex bisection [20], and we compare these results with those obtained by a uniform refinement strategy. As before, all computations are done in Netgen/NGSolve [46], where we use the sparse direct solver package Umfpack [18] to solve the resulting linear system.

In the first example we use the one-dimensional spatial domain $\Omega = (0, 3)$ and the time horizon $T = 6$, i.e., we have the space-time domain $Q \coloneqq (0, 3) \times (0, 6) \subset \mathbb{R}^2$. As exact solution we consider the smooth function

$$u(x,t) \coloneqq \begin{cases} \frac{1}{2}(t-x-2)^3(x-t)^3 \sin\frac{\pi}{3}x & \text{for } x \le t \text{ and } t - x \le 2, \\ 0 & \text{otherwise,} \end{cases} \tag{4.5}$$

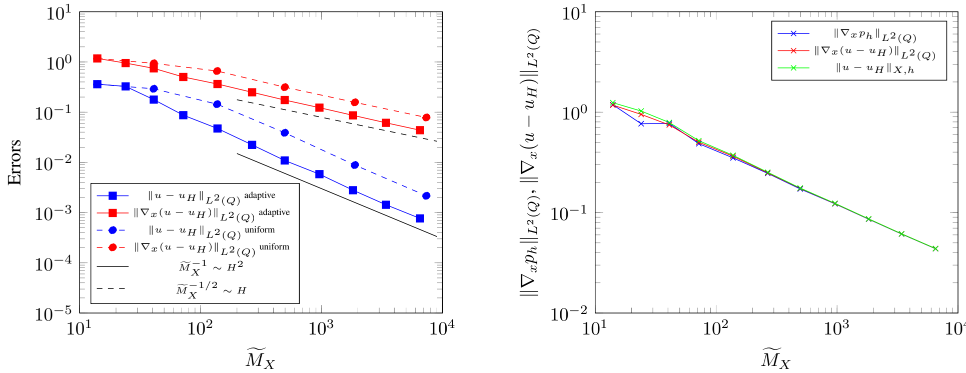
and we compute $f = \partial_t u - \Delta_x u$ accordingly. Since u is smooth, we expect optimal orders of convergence, i.e., $\mathcal{O}(H^2)$ when measuring the error in $L^2(Q)$ and $\mathcal{O}(H)$ in the energy norm $\|\nabla_x \cdot\|_{L^2(Q)}$. These rates are confirmed by the numerical results for both a uniform and an adaptive refinement strategy, as shown in Figure 4.1(a). In Figure 4.1(b) we present a comparison of the errors $\|\nabla_x(u - u_H)\|_{L^2(Q)}$ as well as $\|u - u_H\|_{X,h}$ with the error estimator $\|\nabla_x p_h\|_{L^2(Q)}$, where the curves are almost parallel. Moreover, in Table 4.1 we provide a comparison of the errors $\|\nabla_x(u - u_H)\|_{L^2(Q)}$ for both the uniform and adaptive refinement strategy. We observe that in the adaptive case, fewer degrees of freedom are needed to reach the same level of accuracy as in the case of a uniform refinement. Finally, in Figure 4.2 we present the related finite element meshes. Note that the solution is smooth but behaves similarly to a wave. This motivates the use of a mesh that is adaptive in space and time.

As a second example we consider the heat equation (4.1) in the space-time domain $Q = (0,1)^2 \subset \mathbb{R}^2$, i.e., $\Omega = (0,1)$ and $T = 1$ with homogeneous Dirichlet and initial conditions but with a discontinuous right-hand side

$$f(x,t) = \begin{cases} 1, & \text{for } (x,t) \in (0,1) \times \left(\frac{1}{10}, \frac{1}{2}\right) \text{ and } x - \frac{1}{10} \le t \le x - \frac{1}{20}, \\ 0, & \text{otherwise.} \end{cases} \tag{4.6}$$

Note that a similar example was also considered in [27]. Since the exact solution is unknown, in Figure 4.3 we only provide the error rates of $\|\nabla_x p_h\|_{L^2(Q)}$. As already observed in [27,

ETNA
Kent State University and
Johann Radon Institute (RICAM)



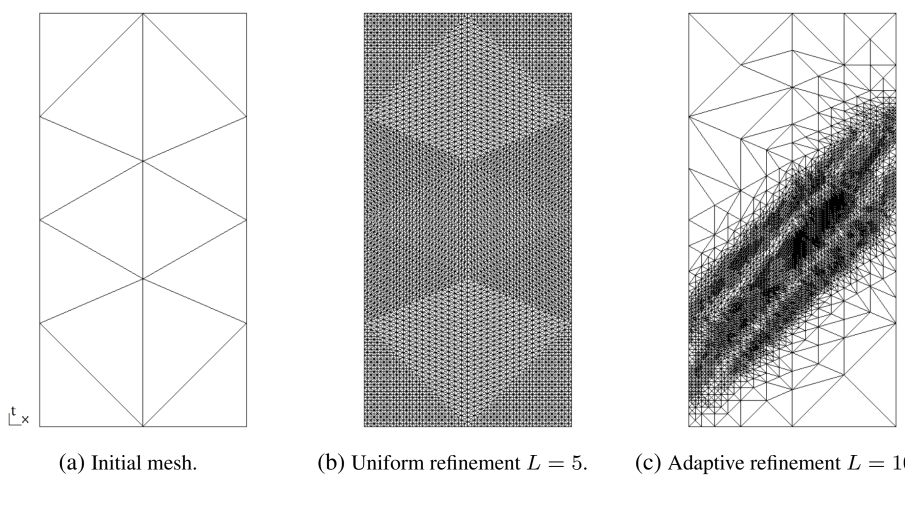
(a) Errors $\|\nabla_x(u-u_H)\|_{L^2(Q)}$ and $\|u-u_H\|_{L^2(Q)}$ for uniform and adaptive refinement strategies.

(b) History of $\|\nabla_x p_h\|_{L^2(Q)}$, $\|\nabla_x(u-u_H)\|_{L^2(Q)}$ for an adaptive refinement strategy.

FIG. 4.1. *Convergence results in the case of a smooth solution for the heat equation.*

TABLE 4.1
Comparison of the errors between uniform and adaptive refinement.

uniform			adaptive		
L	$\widetilde{M}_X$	$\|\nabla_x(u-u_H)\|_{L^2(Q)}$	L	$\widetilde{M}_X$	$\|\nabla_x(u-u_H)\|_{L^2(Q)}$
0	14	1.175×10^{0}	0	14	1.175×10^{0}
1	41	9.391×10^{-1}	2	41	7.463×10^{-1}
2	137	6.597×10^{-1}	4	138	3.637×10^{-1}
3	497	3.144×10^{-1}	6	496	1.745×10^{-1}
4	1889	1.568×10^{-1}	8	1825	8.636×10^{-2}
5	7361	7.840×10^{-2}	10	6524	4.377×10^{-2}



(a) Initial mesh. (b) Uniform refinement $L = 5$. (c) Adaptive refinement $L = 10$.

FIG. 4.2. *Space-time finite element meshes.*

ETNA
Kent State University and
Johann Radon Institute (RICAM)

Section 5.2.3], in the case of uniform refinement we obtain a reduced rate of $\widetilde{M}_X^{-\frac{1}{4}}$, which corresponds to $\mathcal{O}(H^{\frac{1}{2}})$, while in the case of an adaptive refinement we have the optimal rate of $\widetilde{M}_X^{-\frac{1}{2}}$, i.e., $\mathcal{O}(H)$. In Figure 4.4(a) we depict the numerical solution u_H obtained on refinement level $L = 9$ and in Figure 4.4(b) we provide the corresponding adaptive mesh.

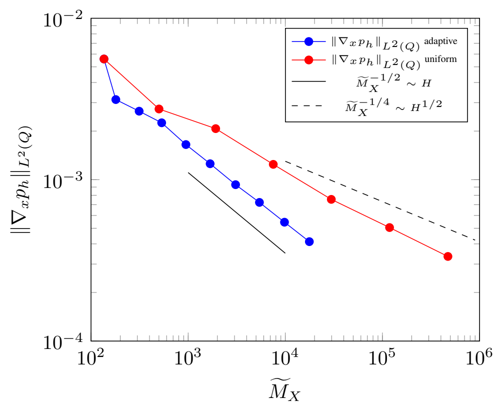
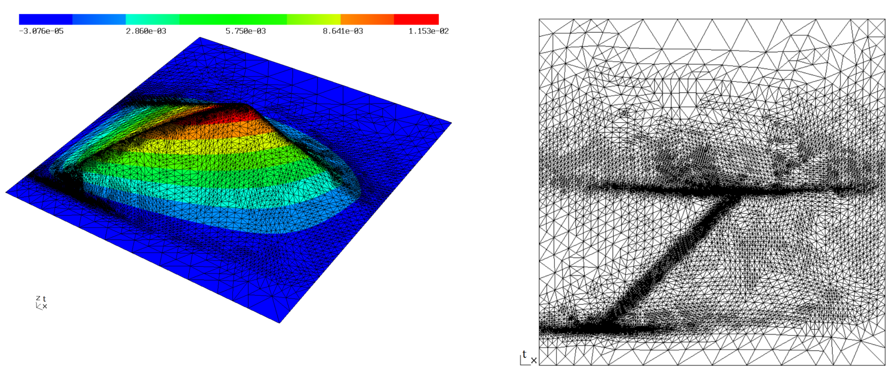


FIG. 4.3. *Error estimator* $\|\nabla_x p_h\|_{L^2(Q)}$ *in the case of a discontinuous right-hand side.*



(a) Solution u_H.

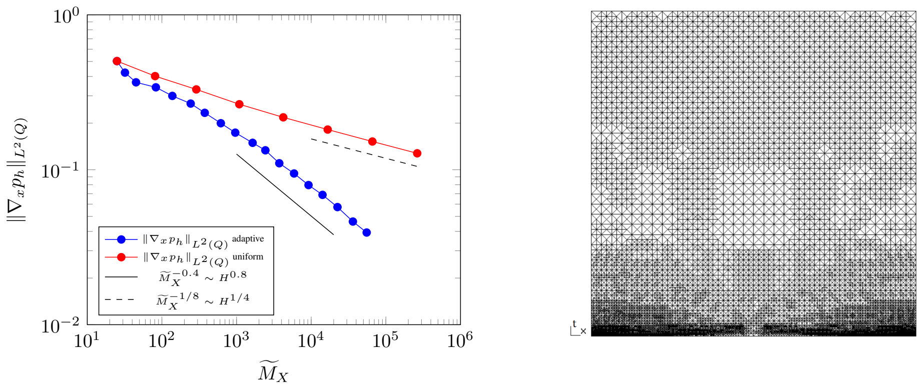
(b) Adaptive mesh, $\widetilde{M}_X = 17689$.

FIG. 4.4. *Singular solution of the heat equation for the discontinuous right-hand side* (4.6).

As a last example for the heat equation we consider again the space-time domain $Q = (0,1)^2$, with $f(x,t) = 2$ for $(x,t) \in Q$, and an inhomogeneous initial datum $u_0(x) = 1$, for $x \in (0,1)$. Obviously, we have $u_0 \in L^2(0,1)$ but $u_0 \notin H_0^1(\Omega)$, i.e., there is no compatibility with the homogeneous Dirichlet boundary conditions at $t = 0$. This results in a reduced order of convergence. The numerical implementation of the inhomogeneous initial datum is done via a homogenization approach similar to inhomogeneous Dirichlet boundary conditions. This is possible because in the space-time setting, the initial condition acts like a Dirichlet condition on the initial boundary. Since the exact solution is not known, in Figure 4.5(a) we provide the results for the error indicator $\|\nabla_x p_h\|_{L^2(Q)}$ while in Figure 4.5(b) we depict

ETNA
Kent State University and
Johann Radon Institute (RICAM)

the adaptive mesh obtained on refinement level $L = 17$. We observe a rate of $\widetilde{M}_X^{-1/8}$, i.e., $\mathcal{O}(H^{1/4})$ in the case of a uniform refinement, and of $\widetilde{M}_X^{-0.4}$, i.e., $\mathcal{O}(H^{0.8})$ for the adaptive refinement strategy. These rates are better than those observed in [27, Sect. 5.2.4] and similar as in [29, Sect. 4.2.1].



(a) Error estimator $\|\nabla_x p_h\|_{L^2(Q)}$.

(b) Adaptive mesh, $L = 17$, $\widetilde{M}_X = 55378$.

FIG. 4.5. *Numerical results in the case of incompatible initial conditions $u_0 \notin H_0^1(\Omega)$.*

REMARK 4.1. Similar to the case for the Poisson equation, we may also define

$$X := \Big\{ v \in L^2(0,T; L^2(\Omega)) \cap H^1_{0,}(0,T; H^{-1}(\Omega)) : \partial_t v - \Delta_x v \in L^2(Q) \Big\}, \quad Y = L^2(Q).$$

In this case we have $B = \partial_t - \Delta_x : X \to L^2(Q)$ and $A := I : L^2(Q) \to L^2(Q)$. For a conforming space-time finite element discretization, one can use tensor-product meshes with quadratic B-splines in the spatial directions and piecewise linear continuous basis functions in the temporal direction.

5. Hyperbolic Dirichlet boundary value problem.

In this section we apply the abstract theory to the Dirichlet boundary value problem for the wave equation,

$$\text{(5.1a)} \quad \Box u(x,t) := \partial_{tt} u(x,t) - \Delta_x u(x,t) = f(x,t) \quad \text{for } (x,t) \in Q := \Omega \times (0,T),$$

$$\text{(5.1b)} \quad u(x,t) = 0 \quad \text{for } (x,t) \in \Sigma := \Gamma \times (0,T),$$

$$\text{(5.1c)} \quad u(x,0) = \partial_t u(x,t)_{|t=0} = 0 \quad \text{for } x \in \Omega,$$

where, as in the parabolic case, $\Omega \subset \mathbb{R}^n$, $n = 1,2,3$, is a bounded Lipschitz domain with boundary $\Gamma := \partial\Omega$ and $T > 0$ is a given time horizon. Following [51], the space-time variational formulation of the wave equation (5.1) is to find $u \in H^{1,1}_{0;0,}(Q)$ such that

$$\text{(5.2)} \qquad -\langle \partial_t u, \partial_t q \rangle_{L^2(Q)} + \langle \nabla_x u, \nabla_x q \rangle_{L^2(Q)} = \langle f, q \rangle_Q$$

is satisfied for all $q \in H^{1,1}_{0;,0}(Q)$. Here we use the anisotropic Sobolev space

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

where $H^1_{0,}(0,T; L^2(\Omega))$ covers the zero initial condition $u(x,0) = 0$, for $x \in \Omega$, while $L^2(0,T; H^1_0(\Omega))$ includes the homogeneous Dirichlet boundary condition on Σ. Note that the

ETNA
Kent State University and
Johann Radon Institute (RICAM)

second initial condition $\partial_t u(x,t)_{|t=0} = 0$, for $x \in \Omega$, enters the variational formulation (5.2) in a natural way. A norm in $H^{1,1}_{0;0,}(Q)$ is given by the graph norm

$$\|u\|_{H^{1,1}_{0;0,}(Q)} := \sqrt{\|\partial_t u\|^2_{L^2(Q)} + \|\nabla_x u\|^2_{L^2(Q)}} = |u|_{H^1(Q)}.$$

Note that $H^{1,1}_{0;,0}(Q)$ is defined accordingly but with a zero terminal condition $q(x,T) = 0$ for $x \in \Omega$. Although the right-hand side of the variational formulation is well defined for $f \in [H^{1,1}_{0;,0}(Q)]^*$, in order to ensure a unique solution $u \in H^{1,1}_{0;0,}(Q)$, we have to assume $f \in L^2(Q)$; see e.g., [33, 51]. In fact, the solution operator mapping $f \in L^2(Q)$ to the solution $u \in H^{1,1}_{0;0,}(Q)$ of the variational formulation (5.2) does not define an isomorphism. Instead we have to enlarge the ansatz space in order to incorporate the second initial condition $\partial_t u(x,t)_{|t=0} = 0$ in an appropriate way. In what follows we consider a generalized variational formulation of the wave equation (see [52]): For the enlarged space-time domain $Q_- := \Omega \times (-T,T)$ and for $u \in L^2(Q)$, we define the zero extension

$$\widetilde{u}(x,t) := \begin{cases} u(x,t) & \text{for } (x,t) \in Q, \\ 0, & \text{otherwise.} \end{cases}$$

The application of the wave operator $\Box\widetilde{u}$ on Q_- is formulated as a distribution, i.e., for $\varphi \in C^\infty_0(Q_-)$ we define

$$\langle \Box\widetilde{u}, \varphi\rangle_{Q_-} := \int_{Q_-} \widetilde{u}(x,t)\,\Box\varphi(x,t)\,\mathrm{d}x\,\mathrm{d}t = \int_Q u(x,t)\,\Box\varphi(x,t)\,\mathrm{d}x\,\mathrm{d}t.$$

Now we are in a position to introduce the space

$$\mathcal{H}(Q) := \Big\{ u = \widetilde{u}_{|Q} : \widetilde{u} \in L^2(Q_-),\ \widetilde{u}_{|\Omega\times(-T,0)} = 0,\ \Box\widetilde{u} \in [H^1_0(Q_-)]^* \Big\},$$

with the graph norm

$$\|u\|_{\mathcal{H}(Q)} := \sqrt{\|u\|^2_{L^2(Q)} + \|\Box\widetilde{u}\|^2_{[H^1_0(Q_-)]^*}}\,.$$

The normed vector space $(\mathcal{H}(Q), \|\cdot\|_{\mathcal{H}(Q)})$ is a Banach space, and it holds true that (see [52, Lemma 3.5]) $H^{1,1}_{0;0,}(Q) \subset \mathcal{H}(Q)$, i.e.,

$$\|\Box\widetilde{u}\|_{[H^1_0(Q_-)]^*} \le \|u\|_{H^{1,1}_{0;0,}(Q)} \qquad \text{for all } u \in H^{1,1}_{0;0,}(Q).$$

Therefore, we can consider the space

$$\mathcal{H}_{0;0,}(Q) := \overline{H^{1,1}_{0;0,}(Q)}^{\|\cdot\|_{\mathcal{H}(Q)}} \subset \mathcal{H}(Q),$$

which will serve as ansatz space. For $u \in \mathcal{H}_{0;0,}(Q)$, an equivalent norm is given as (see [52, Lemma 3.6])

$$\|u\|_{\mathcal{H}_{0;0,}(Q)} = \|\Box\widetilde{u}\|_{[H^1_0(Q_-)]^*}.$$

For given $f \in [H^{1,1}_{0;,0}(Q)]^*$ we now consider the variational formulation to find $u \in \mathcal{H}_{0;0,}(Q)$ such that

$$b(u,q) := \langle Bu, q\rangle_Q = \langle \Box\widetilde{u}, \mathcal{E}q\rangle_{Q_-} = \langle f, q\rangle_Q \qquad \text{for all } q \in H^{1,1}_{0;,0}(Q), \tag{5.3}$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

where $\mathcal{E} : H^{1,1}_{0;,0}(Q) \to H^1_0(Q_-)$ is a suitable extension operator, e.g., by reflection in time. Note that we have (see [52, Lemma 3.5])

$$\langle \Box \widetilde{u}, \mathcal{E}q\rangle_{Q_-} = -\langle \partial_t u, \partial_t q\rangle_{L^2(Q)} + \langle \nabla_x u, \nabla_x q\rangle_{L^2(Q)}$$

for $u \in H^{1,1}_{0;0,}(Q) \subset \mathcal{H}_{0;0,}(Q)$, $q \in H^{1,1}_{0;,0}(Q)$. We conclude that the bilinear form within the variational formulation (5.3) is bounded for all $u \in \mathcal{H}_{0;0,}(Q)$ and $q \in H^{1,1}_{0;,0}(Q)$ and satisfies the inf-sup stability condition

$$\|u\|_{\mathcal{H}_{0;0,}(Q)} = \sup_{0\neq q\in H^{1,1}_{0;,0}(Q)} \frac{\langle \Box \widetilde{u}, \mathcal{E}q\rangle_{Q_-}}{\|q\|_{H^{1,1}_{0;,0}(Q)}} \qquad \text{for all } u \in \mathcal{H}_{0;0,}(Q).$$

In view of the abstract setting we therefore have

$$X := \mathcal{H}_{0;0,}(Q), \quad Y := H^{1,1}_{0;,0}(Q), \quad H_X = H_Y := L^2(Q), \quad B : X \to Y^*.$$

Finally, we define

$$\langle Ap, q\rangle_Q := \langle \partial_t p, \partial_t q\rangle_{L^2(Q)} + \langle \nabla_x p, \nabla_x q\rangle_{L^2(Q)} \qquad \text{for all } p, q \in Y = H^{1,1}_{0;,0}(Q),$$

which corresponds to the space-time Laplacian. Thus we can apply the abstract theory as given in Section 2. The variational formulation (2.8) now reads to find $(u, p) \in X \times Y$ such that

$$\langle \partial_t p, \partial_t q\rangle_{L^2(Q)} + \langle \nabla_x p, \nabla_x q\rangle_{L^2(Q)} + \langle \Box \widetilde{u}, \mathcal{E}q\rangle_{Q_-} = \langle f, q\rangle_Q, \qquad \langle \Box \widetilde{v}, \mathcal{E}p\rangle_{Q_-} = 0 \tag{5.4}$$

is satisfied for all $(v, q) \in X \times Y$.

Let $X_H \subset H^{1,1}_{0;0,}(Q) \subset \mathcal{H}_{0;0,}(Q) = X$ be the conforming finite element space of piecewise linear and continuous basis functions that are defined with respect to an admissible locally quasi-uniform decomposition of the space-time domain Q into shape-regular simplicial finite elements of mesh size H. Moreover, let $Y_h \subset H^{1,1}_{0;,0}(Q)$ be a second finite element space of piecewise linear and continuous basis functions that are defined with respect to a refined decomposition of the space-time domain into finite elements of mesh size h. The Galerkin discretization of (5.4) then reads to find $(u_H, p_h) \in X_H \times Y_h$ such that

$$\begin{aligned} &\langle \partial_t p_h, \partial_t q_h\rangle_{L^2(Q)} + \langle \nabla_x p_h, \nabla_x q_h\rangle_{L^2(Q)} \\ &\qquad - \langle \partial_t u_H, \partial_t q_h\rangle_{L^2(Q)} + \langle \nabla_x u_H, \nabla_x q_h\rangle_{L^2(Q)} = \langle f, q_h\rangle_Q \end{aligned} \tag{5.5}$$

and

$$-\langle \partial_t v_H, \partial_t p_h\rangle_{L^2(Q)} + \langle \nabla_x v_H, \nabla_x p_h\rangle_{L^2(Q)} = 0 \tag{5.6}$$

is satisfied for all $(v_H, q_h) \in X_H \times Y_h$. With $\|u\|_{\mathcal{H}_{0;0,}(Q)} \le \|u\|_{H^{1,1}_{0;0,}(Q)} = |u|_{H^1(Q)}$, for $u \in H^{1,1}_{0;0,}(Q)$, we conclude the following error estimate.

THEOREM 5.1. *Assume* (2.24) *such that the discrete inf-sup stability condition* (2.11) *is satisfied, and assume $u \in H^{1,1}_{0;0,}(Q) \cap H^s(Q)$ for some $s \in [1, 2]$. For the unique solution $(u_H, p_h) \in X_H \times Y_h$ of* (5.5) *and* (5.6)*, there holds the error estimate*

$$\begin{aligned} \|u - u_H\|_{\mathcal{H}_{0;0,}(Q)} &\le \frac{1}{c_S} \inf_{v_H\in X_H} \|u - v_H\|_{\mathcal{H}_{0;0,}(Q)} \\ &\le \frac{1}{c_S} \inf_{v_H\in X_H} |u - v_H|_{H^1(Q)} \le c\, H^{s-1} |u|_{H^s(Q)}. \end{aligned}$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

Proof. The estimate is a consequence of the best approximation result in Lemma 2.4 and standard approximation results of the discrete space X_H. ∎

If the finite element space Y_h is chosen appropriately as formulated in Theorem 2.8, then the discrete inf-sup stability condition (2.11) is satisfied and $\eta_h = |p_h|_{H^1(Q)}$ serves as an error indicator for $\|u - u_H\|_{\mathcal{H}_{0;0,}(Q)}$ satisfying (2.15) and (2.19) with $c_1^B = c_2^B = 1$, i.e.,

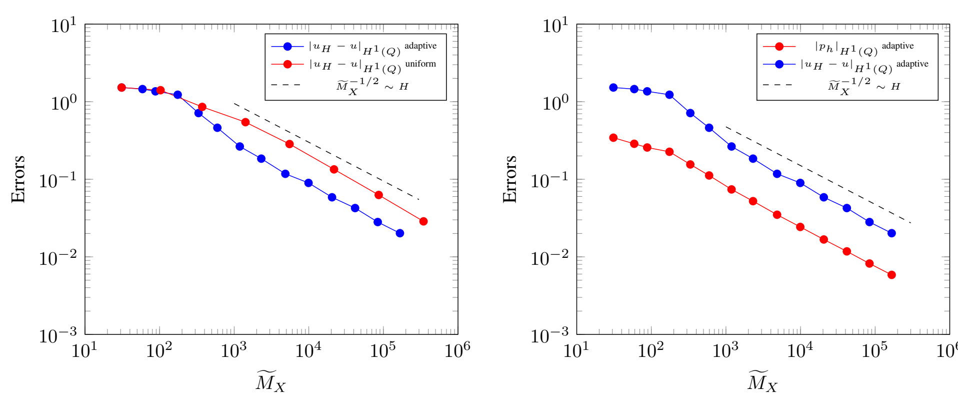
$$|p_h|_{H^1(Q)} \leq \|u - u_H\|_{\mathcal{H}_{0;0,}(Q)} \leq c\,|p_h|_{H^1(Q)} + \operatorname{osc}(f).$$

If in addition (2.20) and the saturation condition (2.22) are satisfied, then we get rid of the data-oscillation term in the reliability estimate, i.e., there holds

$$|p_h|_{H^1(Q)} \leq \|u - u_H\|_{\mathcal{H}_{0;0,}(Q)} \leq \frac{1}{1-\eta}\frac{1}{\bar{c}_S^2}\,|p_h|_{H^1(Q)}.$$

In our numerical examples we use the spaces $X_H = S_H^1(Q) \cap X$ and $Y_h = Y_H = S_H^2(Q) \cap Y$. The first space consists of piecewise linear continuous basis functions while the latter has piecewise quadratic continuous basis functions. For the marking strategy we use the Dörfler criterion [20] with parameter $\theta = 0.5$. As in the previous examples we refine all marked elements using newest vertex bisection and use the finite element software Netgen/NGSolve [46] together with the sparse direct solver package Umfpack [18] to solve the resulting linear systems.

As a first example we consider the smooth solution (4.5) as for the heat equation, and we compute $f = \partial_{tt} u - \Delta_x u$. In this case we expect to see optimal orders of convergence for the error in the energy norm. In Figure 5.1 we present the numerical results for both a uniform and an adaptive refinement as well as a comparison between the error $|u - u_H|_{H^1(Q)}$ and the estimator $|p_h|_{H^1(Q)}$. In all cases we observe a linear convergence, for both the error $|u - u_H|_{H^1(Q)}$ and the estimator $|p_h|_{H^1(Q)}$. Finally, in Figure 5.2 we present the initial mesh, and an adaptively refined mesh in the space-time domain.



(a) Comparison of the errors $|u - u_H|_{H^1(Q)}$ for adaptive and uniform refinement.

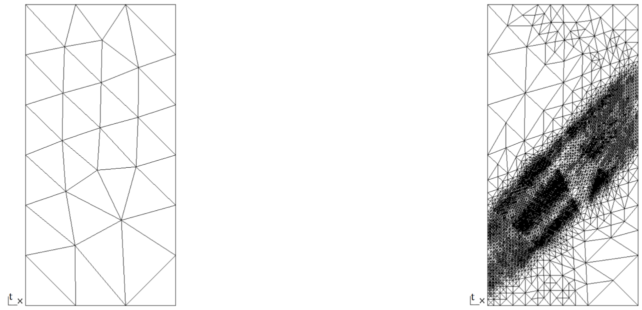
(b) Comparison between error and estimator in case of adaptive refinement.

FIG. 5.1. *Errors $|u - u_H|_{H^1(Q)}$ and estimators $|p_h|_{H^1(Q)}$ for adaptive and uniform refinements for the smooth solution* (4.5).

As a second example we consider the unit square $Q = (0,1)^2$ and choose the right-hand side f and the initial values $u(x,0) = u_0(x)$ and $\partial_t u(x,t)_{|t=0} = g(x)$ to be

$$f(x,t) = 2, \ \text{for } (x,t) \in Q, \qquad u_0(x) = 1, \ \text{for } x \in (0,1), \qquad g(x) = 0, \ \text{for } x \in (0,1).$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

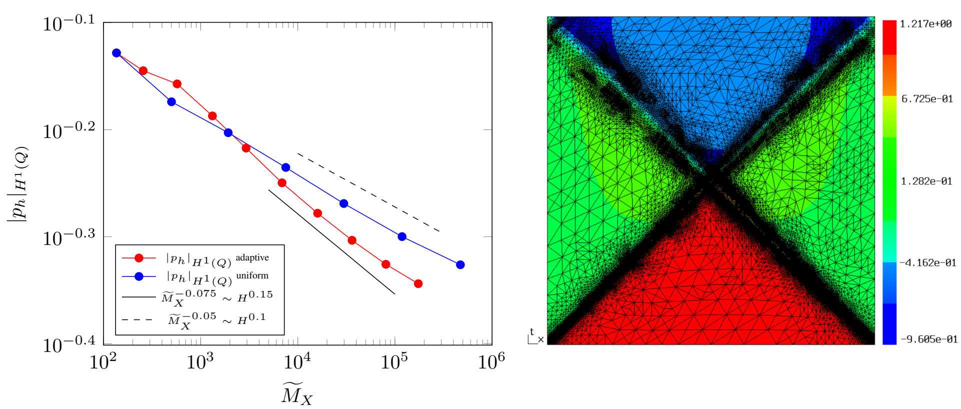


(a) Initial level $L = 0$, $\widetilde{M}_X = 31$.

(b) Refinement level $L = 9$, $\widetilde{M}_X = 9919$.

FIG. 5.2. *Meshes from the adaptive refinement for the solution* (4.5).

We have incorporated the inhomogeneous initial datum u_0 via a homogenization approach similar to that explained in the case of the heat equation. Note that the initial condition satisfies $u_0 \in L^2(0,1)$ but $u_0 \notin H^1_0(0,1)$, i.e., there is no compatibility with the homogeneous Dirichlet boundary conditions at $t = 0$. Therefore, we expect to see reduced orders of convergence. This is confirmed as shown in Figure 5.3(a), where we observe a rate of $\mathcal{O}(\widetilde{M}_X^{-0.05})$, i.e., $\mathcal{O}(H^{0.1})$ in the uniform refinement case, using $X_H = S^1_H(Q) \cap X$ and $Y_h = Y_H = S^2_H(Q) \cap Y$. Driving an adaptive refinement scheme with the Dörfler parameter $\theta = 0.9$, the rate could be improved to $\mathcal{O}(\widetilde{M}_X^{-0.075})$, i.e., $\mathcal{O}(H^{0.15})$. In Figure 5.3(b) we provide the numerical solution u_H plotted on an adaptive mesh, which is obtained on refinement level $L = 8$. We see strong refinements in the directions of $t = x$ and $t = 1 - x$.



(a) Error estimator.

(b) Numerical solution u_H.

FIG. 5.3. *Error estimator for adaptive and uniform refinement and the numerical solution* u_H *on an adaptive mesh with* $\widetilde{M}_X = 80654$ *in the case of a less regular solution violating the compatibility of initial and boundary conditions.*

REMARK 5.2. As for the Laplace equation we can also use a boundary element least-squares formulation in the case of the wave equation [31]. In particular in the one-dimensional case, one can prove the mesh condition $H > 2h$ in order to ensure stability even for adaptive meshes.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

6. Conclusions. In this paper we have formulated and analyzed least-squares methods for the numerical solution of abstract operator equations $Bu = f$. Applications involve elliptic, parabolic, and hyperbolic problems, with the Poisson equation, the heat equation, and the wave equation as examples. While we assume that $B : X \to Y^*$ is an isomorphism and B is given from the application, there is still some freedom in the choice of the underlying spaces X and Y. When considering the Laplace operator $B = -\Delta$, instead of $B : H_0^1(\Omega) \to H^{-1}(\Omega)$ we may also consider $Y = L^2(\Omega)$, implying $X = \{v \in H_0^1(\Omega) : \Delta v \in L^2(\Omega)\}$, or the other way around, i.e., an ultra-weak variational formulation using $X = L^2(\Omega)$ and $Y = \{v \in H_0^1(\Omega) : \Delta v \in L^2(\Omega)\}$. It is obvious that in all of these cases we have to use appropriate inf-sup stable finite element spaces X_H and Y_h. But these approaches can be applied to the heat and the wave equation as well; see, e.g., [30].

In addition to those simple model problems, we can apply this methodology to rather general problems, including flow problems, nonlinear equations, etc. Other applications involve the least-squares formulations of boundary integral equations [31, 50].

In this paper we have considered the stability and error analysis of adaptive space-time least-squares finite element methods. It is clear that for an efficient solution of the resulting huge linear systems of algebraic equations, we need to use appropriate preconditioned and parallel solution strategies. While the construction of these solvers was not within the scope of this paper, this will be done in future work. In particular, due to the structure of the linear system to be solved, we need to have preconditioners for A_h and for the discrete Schur complement $S_h = B_h^\top A_h^{-1} B_h$. Both matrices are symmetric and positive definite, independent of the particular choice of B. Possible solution strategies involve direct methods based on factorization as used in [37] or multigrid methods as considered in [28]. The use of efficient solution methods then also allows the numerical solution of partial differential equations in the four-dimensional space-time domain, as we already did for optimal control problems; see, e.g., [36].

Acknowledgement. This work has been supported by the Austrian Science Fund (FWF) under the Grant Collaborative Research Center TRR361/F90: CREATOR Computational Electric Machine Laboratory.

REFERENCES

[1] M. AINSWORTH AND J. T. ODEN, *A Posteriori Error Estimation in Finite Element Analysis*, Wiley, New York, 2000.
[2] R. ANDREEV, *Stability of sparse space-time finite element discretizations of linear parabolic evolution equations*, IMA J. Numer. Anal., 33 (2013), pp. 242–260.
[3] T. APEL, S. NICAISE, AND J. PFEFFERER, *Adapted numerical methods for the Poisson equation with L^2 boundary data in nonconvex domains*, SIAM J. Numer. Anal., 55 (2017), pp. 1937–1957.
[4] P. BOCHEV, L. DEMKOWICZ, J. GOPALAKRISHNAN, AND M. GUNZBURGER, *Minimum residual and least squares finite element methods [Editorial]*, Comput. Math. Appl., 68 (2014), p. 1479.
[5] P. BOCHEV AND M. GUNZBURGER, *Least-squares finite element methods*, in International Congress of Mathematicians. Vol. III, M. Sanz-Solé, J. Soria, J. L. Varona, and J. Verdera, eds., Eur. Math. Soc., Zürich, 2006, pp. 1137–1162.
[6] ———, *Least-squares methods for hyperbolic problems*, in Handbook of Numerical Methods for Hyperbolic Problems, R. Abgrall and C.-W. Shu, eds., Handb. Numer. Anal., 17, Elsevier/North-Holland, Amsterdam, 2016, pp. 289–317.
[7] ———, *Least-Squares Finite Element Methods*, Springer, New York, 2009.
[8] D. BOFFI, F. BREZZI, AND M. FORTIN, *Mixed Finite Element Methods and Applications*, Springer, Heidelberg, 2013.
[9] J. H. BRAMBLE, R. D. LAZAROV, AND J. E. PASCIAK, *A least-squares approach based on a discrete minus one inner product for first order systems*, Math. Comp., 66 (1997), pp. 935–955.
[10] ———, *Least-squares for second-order elliptic problems*, Comput. Methods Appl. Mech. Engrg., 152 (1998), pp. 195–210.

[11] S. C. BRENNER, *Finite element methods for elliptic distributed optimal control problems with pointwise state constraints (survey)*, in Advances in Mathematical Sciences, B. Acu, D. Danielli, M. Lewicka, A. Pati, R. V. Saraswathy, and M. Teboh-Ewungkem, eds., vol. 21 of Assoc. Women Math. Ser., Springer, Cham, 2020, pp. 3–16.
[12] S. C. BRENNER, J. GEDICKE, AND L.-Y. SUNG, *A symmetric interior penalty method for an elliptic distributed optimal control problem with pointwise state constraints*, Comput. Methods Appl. Math., 23 (2023), pp. 565–589.
[13] S. C. BRENNER AND L. R. SCOTT, *The Mathematical Theory of Finite Element Methods*, Springer, New York, 2008.
[14] C. CARSTENSEN, L. DEMKOWICZ, AND J. GOPALAKRISHNAN, *A posteriori error control for DPG methods*, SIAM J. Numer. Anal., 52 (2014), pp. 1335–1353.
[15] C. CARSTENSEN, M. FEISCHL, M. PAGE, AND D. PRAETORIUS, *Axioms of adaptivity*, Comput. Math. Appl., 67 (2014), pp. 1195–1253.
[16] C. CARSTENSEN, D. GALLISTL, AND J. GEDICKE, *Justification of the saturation assumption*, Numer. Math., 134 (2016), pp. 1–25.
[17] W. DAHMEN, H. MONSUUR, AND R. STEVENSON, *Least squares solvers for ill-posed PDEs that are conditionally stable*, ESAIM Math. Model. Numer. Anal., 57 (2023), pp. 2227–2255.
[18] T. A. DAVIS, *Algorithm 832: UMFPACK V4.3—an unsymmetric-pattern multifrontal method*, ACM Trans. Math. Software, 30 (2004), pp. 196–199.
[19] L. DEMKOWICZ AND J. GOPALAKRISHNAN, *The discontinuous Petrov–Galerkin method*, Acta Numer., 34 (2025), pp. 293–384.
[20] W. DÖRFLER, *A convergent adaptive algorithm for Poisson's equation*, SIAM J. Numer. Anal., 33 (1996), pp. 1106–1124.
[21] W. DÖRFLER AND R. H. NOCHETTO, *Small data oscillation implies the saturation assumption*, Numer. Math., 91 (2002), pp. 1–12.
[22] W. DÖRFLER, C. WIENERS, AND D. ZIEGLER, *Space-time discontinuous Galerkin methods for linear hyperbolic systems and the application to the forward problem in seismic imaging*, in Finite Volumes for Complex Applications IX—Methods, Theoretical Aspects, Examples—FVCA 9, Bergen, Norway, June 2020, R. Klöfkorn, E. Keilegavlen, F. A. Radu, and J. Fuhrmann, eds., vol. 323 of Springer Proc. Math. Stat., Springer, Cham, 2020, pp. 477–485.
[23] A. ERN, I. SMEARS, AND M. VOHRALÍK, *Guaranteed, locally space-time efficient, and polynomial-degree robust a posteriori error estimates for high-order discretizations of parabolic problems*, SIAM J. Numer. Anal., 55 (2017), pp. 2811–2834.
[24] S. FERRAZ-LEITE, C. ORTNER, AND D. PRAETORIUS, *Convergence of simple adaptive Galerkin schemes based on $h - h/2$ error estimators*, Numer. Math., 116 (2010), pp. 291–316.
[25] T. FÜHRER, *On a mixed FEM and a FOSLS with H^{-1} loads*, Comput. Methods Appl. Math., 24 (2024), pp. 355–370.
[26] T. FÜHRER, N. HEUER, AND M. KARKULIK, *MINRES for second-order PDEs with singular data*, SIAM J. Numer. Anal., 60 (2022), pp. 1111–1135.
[27] T. FÜHRER AND M. KARKULIK, *Space-time least-squares finite elements for parabolic equations*, Comput. Math. Appl., 92 (2021), pp. 27–36.
[28] M. J. GANDER AND M. NEUMÜLLER, *Analysis of a new space-time parallel multigrid algorithm for parabolic problems*, SIAM J. Sci. Comput., 38 (2016), pp. A2173–A2208.
[29] G. GANTNER AND R. STEVENSON, *Improved rates for a space-time FOSLS of parabolic PDEs*, Numer. Math., 156 (2024), pp. 133–157.
[30] J. HENNING, D. PALITTA, V. SIMONCINI, AND K. URBAN, *An ultraweak space-time variational formulation for the wave equation: analysis and efficient numerical solution*, ESAIM Math. Model. Numer. Anal., 56 (2022), pp. 1173–1198.
[31] D. HOONHOUT, R. LÖSCHER, O. STEINBACH, AND C. URZÚA-TORRES, *Stable least-squares space-time boundary element methods for the wave equation*, Adv. Comput. Math., 52 (2026), Paper No. 7, 30 pages.
[32] C. KÖTHE, *Adaptive Least-Squares Space-Time Finite Element Methods*, Monographic Series Computation in Engineering and Science, Band 48, TU Graz, Graz, 2025.
[33] O. A. LADYZHENSKAYA, *The Boundary Value Problems of Mathematical Physics*, Springer, 1985.
[34] U. LANGER AND O. STEINBACH, eds., *Space-time methods—applications to partial differential equations*, De Gruyter, Berlin, 2019.
[35] U. LANGER, O. STEINBACH, F. TRÖLTZSCH, AND H. YANG, *Unstructured space-time finite element methods for optimal control of parabolic equations*, SIAM J. Sci. Comput., 43 (2021), pp. A744–A771.
[36] U. LANGER, O. STEINBACH, AND H. YANG, *Robust space-time finite element methods for parabolic distributed optimal control problems with energy regularization*, Adv. Comput. Math., 50 (2024), Paper No. 24, 30 pages.
[37] U. LANGER AND M. ZANK, *Efficient direct space-time finite element solvers for parabolic initial-boundary value problems in anisotropic Sobolev spaces*, SIAM J. Sci. Comput., 43 (2021), pp. A2714–A2736.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

[38] R. LÖSCHER, *On Unified Frameworks for Space-Time Optimal Control and Least Squares Problems*, PhD thesis, TU Graz, Graz, 2023.
[39] R. LÖSCHER AND O. STEINBACH, *Space-time finite element methods for distributed optimal control of the wave equation*, SIAM J. Numer. Anal., 62 (2024), pp. 452–475.
[40] R. LÖSCHER, O. STEINBACH, AND M. ZANK, *Numerical results for an unconditionally stable space-time finite element method for the wave equation*, in Domain Decomposition Methods in Science and Engineering XXVI, S. C. Brenner, E. Chung, A. Klawonn, F. Kwok, J. Xu, and J. Zou, eds., vol. 145 of Lect. Notes Comput. Sci. Eng., Springer, Cham, 2022, pp. 625–632.
[41] H. MONSUUR, R. STEVENSON, AND J. STORN, *Minimal residual methods in negative or fractional Sobolev norms*, Math. Comp., 93 (2024), pp. 1027–1052.
[42] M. NEUMÜLLER, *Space-Time Methods. Fast Solvers and Applications*. Monographic Series Computation in Engineering and Science, Band 20, TU Graz, Graz, 2013.
[43] M. NEUMÜLLER AND O. STEINBACH, *Refinement of flexible space-time finite element meshes and discontinuous Galerkin methods*, Comput. Vis. Sci., 14 (2011), pp. 189–205.
[44] I. PERUGIA, C. SCHWAB, AND M. ZANK, *Exponential convergence of hp-time-stepping in space-time discretizations of parabolic PDEs*, ESAIM Math. Model. Numer. Anal., 57 (2023), pp. 29–67.
[45] A. SCHAFELNER AND P. S. VASSILEVSKI, *Numerical results for adaptive (negative norm) constrained first order system least squares formulations*, Comput. Math. Appl., 95 (2021), pp. 256–270.
[46] J. SCHÖBERL, *C++ implementation of finite elements in NGSolve*, Technical Report 30, Institute for Analysis and Scientific Computing, TU Wien, Wien, 2014.
[47] C. SCHWAB AND R. STEVENSON, *Space-time adaptive wavelet methods for parabolic evolution problems*, Math. Comp., 78 (2009), pp. 1293–1318.
[48] O. STEINBACH, *Numerical Approximation Methods for Elliptic Boundary Value Problems*, Springer, New York, 2008.
[49] ———, *Space-time finite element methods for parabolic problems*, Comput. Meth. Appl. Math., 15 (2015), pp. 551–566.
[50] ———, *An adaptive least squares boundary element method for elliptic boundary value problems*, Numer. Math., 158 (2026), pp. 653–669.
[51] O. STEINBACH AND M. ZANK, *Coercive space-time finite element methods for initial boundary value problems*, Electron. Trans. Numer. Anal., 52 (2020), pp. 154–194. https://etna.ricam.oeaw.ac.at/vol.52.2020/pp154-194.dir/pp154-194.pdf
[52] ———, *A generalized inf-sup stable variational formulation for the wave equation*, J. Math. Anal. Appl., 505 (2022), Paper No. 125457, 24 pages.
[53] R. STEVENSON AND J. WESTERDIEP, *Stability of Galerkin discretizations of a mixed space-time variational formulation of parabolic evolution equations*, IMA J. Numer. Anal., 41 (2021), pp. 28–47.
[54] V. THOMÉE, *Galerkin Finite Element Methods for Parabolic Problems*, Springer, Berlin, 2006.
[55] K. URBAN AND A. T. PATERA, *An improved error bound for reduced basis approximation of linear parabolic problems*, Math. Comp., 83 (2014), pp. 1599–1615.
[56] J. XU AND L. ZIKATANOV, *Some observations on Babuška and Brezzi theories*, Numer. Math., 94 (2003), pp. 195–202.