MATHEMATICS OF COMPUTATION
Volume 67, Number 224, October 1998, Pages 1335–1360
S 0025-5718(98)01011-4

A POSTERIORI ERROR ESTIMATES FOR NONLINEAR PROBLEMS. $L^r(0,T;L^\rho(\Omega))$-ERROR ESTIMATES FOR FINITE ELEMENT DISCRETIZATIONS OF PARABOLIC EQUATIONS

R. VERFÜRTH

ABSTRACT. Using the abstract framework of [9] we analyze a residual a posteriori error estimator for space-time finite element discretizations of quasilinear parabolic pdes. The estimator gives global upper and local lower bounds on the error of the numerical solution. The finite element discretizations in particular cover the so-called θ-scheme, which includes the implicit and explicit Euler methods and the Crank-Nicholson scheme.

1. INTRODUCTION

We analyze a residual a posteriori error estimator for space-time finite element discretizations of parabolic pdes. Each space-time element $K \times J$ contributes the weighted sum of three terms:

1. the residual of the computed numerical solution with respect to the strong form of the differential operator evaluated on $K \times J$,
2. the jump across $\partial K \times J$ of that trace operator which naturally connects the strong and the weak formulation of the differential equation, and
3. the jump of the numerical solution across $K \times \partial J$.

Here, K stands for an arbitrary element of the spatial mesh and J denotes an arbitrary interval of the time mesh. We could also extend our analysis to error estimators which are based on the solution of auxiliary local time-dependent problems. We do not follow this line here, in order not to overload the presentation.

In order to construct our a posteriori error estimator and to prove that it yields upper and lower bounds on the error, we use the techniques introduced in [9] and consider in Section 2 abstract nonlinear problems of the form

$$F(u) = 0 \tag{1.1}$$

and corresponding discretizations of the form

$$F_h(u_h^0) = 0. \tag{1.2}$$

Here, $F \in C^1(X, Y^*)$ and $F_h \in C(X_h, Y_h^*), X_h \subset X$ and $Y_h \subset Y$ are finite dimensional subspaces of the Banach spaces X and Y, and $*$ denotes the dual of a Banach space.

Received by the editor March 21, 1995 and, in revised form, May 3, 1996 and January 3, 1997.
1991 *Mathematics Subject Classification.* Primary 65N30, 65N15, 65J15, 76D05.
Key words and phrases. A posteriori error estimates; quasilinear parabolic pdes; space-time finite elements; θ-scheme.

If $u_0 \in X$ is a solution of problem (1.1) such that $DF(u_0)$ is an isomorphism of X onto Y^* and DF is Lipschitz continuous at u_0, we know from Proposition 2.1 in [9] that

$$\underline{c}\|F(u)\|_{Y^*} \leq \|u - u_0\|_X \leq \overline{c}\|F(u)\|_{Y^*} \tag{1.3}$$

holds for all u in a suitable neighbourhood of u_0. The constants $\underline{c}$ and $\overline{c}$ depend on $DF(u_0)$ and $DF(u_0)^{-1}$. They measure the sensitivity of the infinite dimensional problem (1.1) with respect to small perturbations. For a simple model problem we derive explicit bounds for $\underline{c}$ and $\overline{c}$ in Section 4.

When applying estimate (1.3) to an approximate solution $u_h \in X_h$ of problem (1.2) one must evaluate the residual $\|F(u_h)\|_{Y^*}$. This is as expensive as the solution of the original problem (1.1) since it amounts in the solution of an infinite dimensional maximization problem. In order to obtain error estimates which are better amenable to practical calculations, we approximate the left and right-hand sides of inequality (1.3) by $\|\widetilde{F}_h(u_h)\|_{\widetilde{Y}_h^*}$ and $\|(Id_Y - R_h)^*\widetilde{F}_h(u_h)\|_{Y^*}$, respectively. Here $\widetilde{F}_h(u_h)$ is obtained by locally projecting $F(u_h)$ onto suitable finite element spaces, $\widetilde{Y}_h$ consists of appropriate test functions having a local support, and R_h is a suitable quasi-interpolation operator.

For parabolic pdes, these general results lead to error estimates in an $L^r(0,T;\ W_0^{1,\rho}(\Omega))$-norm. The space Y then consists of functions in $L^{p'}(0,T;W_0^{1,\pi'}(\Omega))$ having their time derivative in $L^{p'}(0,T;W^{-1,\pi'}(\Omega))$. Due to the non-local nature of the $W^{-1,\pi'}(\Omega)$-norm we get into troubles when deriving lower bounds on the error. This problem is tackled in [10]. Here, we circumvent this difficulty by imposing a weaker $L^r(0,T;L^\rho(\Omega))$-norm on X and a stronger $L^{p'}(0,T;W^{2,\pi'}(\Omega) \cap W_0^{1,\pi'}(\Omega))$-norm on Y. The corresponding spaces will be denoted by X_- and Y_+, respectively. In particular, the functions in Y_+ now have time derivatives in $L^{p'}(0,T;L^{\pi'}(\Omega))$.

In Section 4 we apply the general results of the previous sections to scalar quasilinear parabolic pdes of 2nd order. Problem (1.1) then is a variational formulation which is weaker than the weak formulation and stronger than the very weak formulation of the pde. The discrete problem (1.2) is a Petrov-Galerkin discretization. The trial space X_h consists of functions which are discontinuous in time and piecewise polynomials of degree $k \geq 0$. The test space Y_h consists of functions which are continuous in time, piecewise polynomials of degree $k+1$, and which vanish at the final time T. This discretization corresponds to an implicit $(k+1)$-stage Runge-Kutta scheme. When applied to a linear problem its stability function is the $(k+1)$-st diagonal Padé approximation. For $k = 0$ we in particular obtain the Crank-Nicholson scheme. By slightly modifying the basis functions of Y_h we may also recover the popular θ-scheme for all $\theta \in [0,1]$. This in particular covers the explicit ($\theta = 0$) and implicit ($\theta = 1$) Euler schemes.

We obtain global upper and local lower bounds for the error measured in an $L^r(0,T;L^\rho(\Omega))$-norm. The upper and lower bounds differ by a factor $1 + \tau^{-1}h^2 + \tau h^{-2}$. Here, h and τ are the local mesh-sizes in space and time, respectively. This factor reflects the fact that the differential operator is of 2nd order with respect to the space variables but only of 1st order with respect to the time variable. The local lower bounds may be combined to a global lower bound of the same type. In Remark

4.5 we briefly comment on error estimates with respect to an $L^r(0,T;W^{1,\rho}(\Omega))$-norm and the corresponding difficulties. A more detailed analysis, also including the time-dependent incompressible Navier-Stokes equations, may be found in [10].

When applied to the corresponding particular examples, our error estimates are similar to those obtained in [7], [8]. However, only upper bounds on the error are established there. Moreover, the techniques and, most important, the discretizations considerably differ from ours. The discontinuous Galerkin method of [7], [8] is non-conforming with respect to both the weak and the very weak formulations of a parabolic pde. It corresponds to an implicit $(k+1)$-stage Runge-Kutta method having the $(k+1)$-st subdiagonal Padé approximation as stability function. In particular, the lowest order scheme ($k=0$) corresponds to the implicit Euler method, and the Crank-Nicholson scheme is not covered by this family of discretizations.

In what follows we will always adopt the following **convention**:

$$\begin{aligned} a \preceq b &\iff a \le cb, \\ a \simeq b &\iff a \preceq b \text{ and } b \preceq a. \end{aligned}$$

Here, the constant c must not depend on any mesh size.

2. Abstract error estimates

Let X, Y be two Banach spaces with norms $\|.\|_X$ and $\|.\|_Y$. For any element $u \in X$ and any real number $R > 0$ set $B_X(u,R) := \{v \in X : \|u-v\|_X < R\}$. We denote by $\mathcal{L}(X,Y)$ and $\mathrm{Isom}(X,Y) \subset \mathcal{L}(X,Y)$ the Banach space of continuous linear maps of X in Y equipped with the operator norm $\|.\|_{\mathcal{L}(X,Y)}$ and the open subset of linear homeomorphisms of X onto Y. By $Y^* := \mathcal{L}(Y,\mathbb{R})$ and $\langle .,. \rangle_Y$ we denote the dual space of Y and the corresponding duality pairing. Finally, $A^* \in \mathcal{L}(Y^*, X^*)$ denotes the adjoint of a given operator $A \in \mathcal{L}(X,Y)$.

Let $F \in C^1(X, Y^*)$ be a given continuously differentiable function. Given a solution $u_0 \in X$ of problem (1.1) and an arbitrary element $u \in X$ "close" to u_0, we may estimate the error $\|u - u_0\|_X$ by the residual $\|F(u)\|_{Y^*}$ (cf. Proposition 2.1 in [9]). For parabolic pde's we thus obtain control on the $L^r(0,T;W_0^{1,\rho}(\Omega))$-norm of the error. However, we are interested in controlling the $L^r(0,T;L^\rho(\Omega))$-norm of the error. In order to achieve this within the present abstract framework, we must enlarge the space X and reduce the space Y. We therefore consider three additional Banach spaces X_-, X_+, and Y_+ such that $X_+ \subset X \subset X_-$ and $Y_+ \subset Y$ with continuous and dense injections. Here, the $+/-$ sign indicates a space with a stronger/weaker norm. We assume that X_- is reflexive.

Proposition 2.1. *Let $u_0 \in X$ be a solution of problem (1.1). Assume that $u_0 \in X_+$, that $DF(u_0)^* \in \mathrm{Isom}(Y_+, X_-^*)$, and that there are two numbers $R_0 > 0$ and $\beta > 0$ such that*

$$\|[DF(u_0) - DF(u_0 + tw)]w\|_{Y_+^*} \le \beta t \|w\|_{X_+} \|w\|_{X_-} \quad \forall w \in B_{X_+}(0, R_0), t \in [0,1]. \tag{2.1}$$

Set

$$R := \min\{R_0, \beta^{-1} \|DF(u_0)^{*^{-1}}\|^{-1}_{\mathcal{L}(X_-^*, Y_+)}, 2\beta^{-1} \|DF(u_0)^*\|_{\mathcal{L}(Y_+, X_-^*)}\}.$$

Then the following error estimate holds for all $u \in B_{X_+}(u_0, R)$:

(2.2)

$$\frac{1}{2}\|DF(u_0)^*\|^{-1}_{\mathcal{L}(Y_+,X_-^*)}\|F(u)\|_{Y_+^*} \le \|u-u_0\|_{X_-} \le 2\|DF(u_0)^{*^{-1}}\|_{\mathcal{L}(X_-^*,Y_+)}\|F(u)\|_{Y_+^*}.$$

Proof. Let $u \in B_{X_+}(u_0, R)$. Consider an arbitrary element $w \in X_-^*$ and set $\varphi := DF(u_0)^{*^{-1}} w \in Y_+$. We then have

$$\begin{aligned}\langle u-u_0, w\rangle_{X_-} =& \langle DF(u_0)(u-u_0), \varphi\rangle_{Y_+}\\ =& \langle F(u), \varphi\rangle_{Y_+}\\ &+ \int_0^1 \langle [DF(u_0) - DF(u_0 + t(u-u_0))](u-u_0), \varphi\rangle_{Y_+} dt.\end{aligned}$$

Inequality (2.1) and the continuity of $DF(u_0)^{*^{-1}}$ imply that

$$\begin{aligned}&|\int_0^1 \langle [DF(u_0) - DF(u_0 + t(u-u_0))](u-u_0), \varphi\rangle_{Y_+} dt|\\ &\le \frac{1}{2}\beta\|u-u_0\|_{X_+}\|u-u_0\|_{X_-}\|\varphi\|_{Y_+}\\ &\le \frac{1}{2}\beta\|DF(u_0)^{*^{-1}}\|_{\mathcal{L}(X_-^*,Y_+)} R\|u-u_0\|_{X_-}\|w\|_{X_-^*}.\end{aligned}$$

Combined with the above representation of $\langle u-u_0, w\rangle_{X_-}$, this yields

$$\langle u-u_0, w\rangle_{X_-} \le \{\|DF(u_0)^{*^{-1}}\|_{\mathcal{L}(X_-^*,Y_+)}\|F(u)\|_{Y_+^*} + \frac{1}{2}\|u-u_0\|_{X_-}\}\|w\|_{X_-^*}.$$

Since X_- is reflexive and $w \in X_-^*$ was arbitrary, this implies the upper bound of estimate (2.2).

In the same way, we obtain

$$\begin{aligned}&\langle F(u), \varphi\rangle_{Y_+}\\ &= \langle u-u_0, w\rangle_{X_-} - \int_0^1 \langle [DF(u_0) - DF(u_0 + t(u-u_0))](u-u_0), \varphi\rangle_{Y_+} dt\\ &\le \|u-u_0\|_{X_-}\|w\|_{X_-^*} + \frac{1}{2}\beta\|u-u_0\|_{X_+}\|u-u_0\|_{X_-}\|\varphi\|_{Y_+}\\ &\le 2\|DF(u_0)^*\|_{\mathcal{L}(Y_+,X_-^*)}\|u-u_0\|_{X_-}\|\varphi\|_{Y_+}.\end{aligned}$$

Since $\varphi \in Y_+$ is arbitrary, this proves the lower bound of estimate (2.2). ☐

The condition $DF(u_0)^* \in \mathrm{Isom}(Y_+, X_-^*)$ of Proposition 2.1 is more restrictive than the assumption $DF(u_0) \in \mathrm{Isom}(X, Y^*)$ which is needed to bound $\|u-u_0\|_X$ by $\|F(u)\|_{Y^*}$ (cf. Proposition 2.1 in [9]). For pdes, it is equivalent to an additional regularity condition. For linear problems, i.e. when DF is constant, one may extend F by continuity to a continuously differentiable map of X_- to Y_+^*. Then the space X_+ is not needed. For nonlinear problems, however, this extension may be impossible, or the derivative of the extension may be no longer Lipschitz continuous. This is the place where the space X_+ comes into play.

The factor $\|DF(u_0)^*\|_{\mathcal{L}(Y_+,X_-^*)}$ in the lower bound of estimate (2.2) corresponds to a differential operator which is local and the norm of which can be estimated in terms of its coefficients. The factor $\|DF(u_0)^{*^{-1}}\|_{\mathcal{L}(X_-^*,Y_+)}$ in the upper bound of estimate (2.2) is much more severe. In some applications, in particular when X_- is a Hilbert space, $\|DF(u_0)^{*^{-1}}\|_{\mathcal{L}(X_-^*,Y_+)}$ can be replaced by $\|DF(u_0)^{*^{-1}} w\|_{Y_+}$, where $w \in X_-^*$ satisfies $\|w\|_{X_-^*} = 1$ and $\langle u - u_0, w\rangle_{X_-} = \|u - u_0\|_{X_-}$. The quantity $\|DF(u_0)^{*-1}w\|_{Y_+}$ may be estimated numerically by approximately solving a discrete analogue of the corresponding adjoint pde.

Let $X_h \subset X_+$ and $Y_h \subset Y$ be finite dimensional subspaces and $F_h \in C(X_h, Y_h^*)$ be an approximation of F. Given an approximate solution $u_h \in X_h$ of problem (1.2), Proposition 2.1 allows us to estimate the error $\|u_0 - u_h\|_{X_-}$ by the residual $\|F(u_h)\|_{Y_+^*}$. The evaluation of the latter, however, is a difficult task since it amounts in the solution of an infinite dimensional maximization problem. In order to obtain an approximation of the residual which is easier to compute, we introduce a restriction operator $R_h \in \mathcal{L}(Y, Y_h)$, a finite dimensional subspace $\widetilde{Y}_h \subset Y_+$, and an approximation $\widetilde{F}_h : X_h \to Y^*$ of F at u_h. In the context of pdes $\widetilde{F}_h$ is obtained by locally freezing the coefficients of the differential operator. We equip $\widetilde{Y}_h$ with the norm of Y_+.

Proposition 2.2. *The following estimates hold:*

(2.3)
$$\begin{aligned}\|F(u_h)\|_{Y_+^*} \le{}& \|(Id_{Y_+} - R_h)^*\widetilde{F}_h(u_h)\|_{Y_+^*} + \|(Id_{Y_+} - R_h)^*[F(u_h) - \widetilde{F}_h(u_h)]\|_{Y_+^*}\\ &+ \|R_h^*[F(u_h) - F_h(u_h)]\|_{Y_+^*} + \|R_h^*F_h(u_h)\|_{Y_+^*}\end{aligned}$$

and

$$\|\widetilde{F}_h(u_h)\|_{\widetilde{Y}_h^*} \le \|F(u_h)\|_{\widetilde{Y}_h^*} + \|F(u_h) - \widetilde{F}_h(u_h)\|_{\widetilde{Y}_h^*}. \tag{2.4}$$

If there is a constant c_0, which does not depend on h, such that

$$\|(Id_{Y_+} - R_h)^*\widetilde{F}_h(u_h)\|_{Y_+^*} \le c_0\|\widetilde{F}_h(u_h)\|_{\widetilde{Y}_h^*}, \tag{2.5}$$

then both $\|(Id_{Y_+} - R_h)^*\widetilde{F}_h(u_h)\|_{Y_+^*}$ and $\|\widetilde{F}_h(u_h)\|_{\widetilde{Y}_h^*}$ yield upper and lower bounds for the residual $\|F(u_h)\|_{Y_+^*}$.

Proof. Estimate (2.3) follows from the identity

$$\begin{aligned}\langle F(u_h), \varphi\rangle_{Y_+} ={}& \langle \widetilde{F}_h(u_h), \varphi - R_h\varphi\rangle_{Y_+} + \langle F(u_h) - \widetilde{F}_h(u_h), \varphi - R_h\varphi\rangle_{Y_+}\\ &+ \langle F(u_h) - F_h(u_h), R_h\varphi\rangle_{Y_+} + \langle F_h(u_h), R_h\varphi\rangle_{Y_+}\end{aligned}$$

which holds for all $\varphi \in Y_+$. Estimate (2.4) follows from the triangle inequality. The statement concerning upper and lower bounds for the residual is an obvious consequence of inequality (2.5). ◻

The second terms on the right-hand sides of estimates (2.3) and (2.4) measure the quality of the approximation $\widetilde{F}_h(u_h)$ to $F(u_h)$. Usually, they are higher order terms when compared with $\|(Id_{Y_+} - R_h)^*\widetilde{F}_h(u_h)\|_{Y_+^*}$ and $\|\widetilde{F}_h(u_h)\|_{\widetilde{Y}_h^*}$. The term $\|R_h^*[F(u_h) - F_h(u_h)]\|_{Y_+^*}$ is the consistency error of the discretization. It can be estimated a priori and vanishes in the applications of Section 4. The term $\|R_h^*F_h(u_h)\|_{Y_+^*}$ measures the residual of the algebraic equation (1.2) and can be evaluated by standard methods. Inequality (2.5) is a non-trivial condition since it claims that a supremum with respect to an infinite dimensional space may be bounded from above by a supremum with respect to a finite dimensional space.

Propositions 2.1 and 2.2 will be used in the following way. Any upper bound for $\|(Id_{Y_+} - R_h)^*\widetilde{F}_h(u_h)\|_{Y_+^*}$ yields a reliable error estimator. If in addition the error estimator yields a lower bound for $\|\widetilde{F}_h(u_h)\|_{\widetilde{Y}_h^*}$, it is also efficient. Here, we adopt the standard convention that an error estimator is called *reliable* (resp. *efficient*) if it yields an upper (resp. lower) bound for the error. Moreover, upper and lower bounds may contain multiplicative constants which do not depend on the discretization parameter h.

3. Auxiliary results

Function spaces. Let Ω be a bounded, connected, open domain in $\mathbb{R}^n, n \geq 2$, with polyhedral boundary Γ. For any open subset ω of Ω with Lipschitz boundary γ, we denote by $W^{k,p}(\omega), k \in \mathbb{N}, 1 \leq p \leq \infty, L^p(\omega) := W^{0,p}(\omega)$, and $L^p(\gamma)$ the usual Sobolev and Lebesgue spaces equipped with the standard norms (cf. [1] and Vol. 3, Chap. IV in [6]). Let

$$W_0^{1,p}(\Omega) := \{u \in W^{1,p}(\Omega) : u = 0 \text{ on } \Gamma\}$$

and set for $1 < p < \infty$

$$W^{-1,p'}(\Omega) := W_0^{1,p}(\Omega)^*.$$

Here, p' denotes the dual exponent of p defined by $\frac{1}{p'} + \frac{1}{p} = 1$. In what follows, a prime will always denote the dual of a given Lebesgue exponent.

Let V and W be two Banach spaces such that $V \subset W$ with continuous and dense injection. Given two real numbers a and b with $a < b$, we denote by $L^p(a, b; V), 1 \leq p < \infty$, the space of measurable functions u defined on (a, b) with values in V such that the function $t \to \|u(., t)\|_V$ is in $L^p((a, b))$. $L^p(a, b; V)$ is a Banach space equipped with the norm

$$\|u\|_{L^p(a,b;V)} := \left\{ \int_a^b \|u(., t)\|_V^p dt \right\}^{1/p}$$

(cf. [6], Vol. 5, Chap. XVIII, §1). Slightly changing the notation of [6], we further consider the Banach space

$$W^p(a, b; V, W) := \{u \in L^p(a, b; V) : \partial_t u \in L^p(a, b; W)\}$$

equipped with the norm

$$\|u\|_{W^p(a,b;V,W)} := \left\{ \int_a^b \|u(., t)\|_V^p dt + \int_a^b \|\partial_t u(., t)\|_W^p dt \right\}^{1/p}.$$

Here, the partial derivative $\partial_t u$ must be interpreted in the distributional sense (cf. [6], loc. cit.). For all smooth functions $\varphi \in \mathcal{D}((a, b))$ it satisfies the identity

$$\int_a^b \partial_t u(., t)\varphi(t) dt = -\int_a^b u(., t)\varphi'(t) dt,$$

where the integrals are taken in W. From Proposition 9 in [6], loc. cit., it follows that the traces $u(., a)$ and $u(., b)$ are defined as elements of W, provided $p > 1$. We therefore set for $1 < p < \infty$

$$W_b^p(a, b; V, W) := \{u \in W^p(a, b; V, W) : u(., b) = 0\}.$$

Given an interval $I \subset \mathbb{R}$, a suitable subset $\omega \subset \Omega$, and numbers $1 \leq p, \pi < \infty$, we will use the following **abbreviation**:

$$L_\pi^p(\omega \times I) := L^p(I; L^\pi(\omega)), \quad W_\pi^p(\omega \times I) := W^p(I; W^{2,\pi}(\omega), L^\pi(\omega)).$$

Finite element partition. Denote by $T > 0$ an arbitrary but fixed time. Let $\mathcal{I}_\tau = \{[t_j, t_{j+1}] : 1 \le j \le N_\tau\}, \tau > 0$, be a family of partitions of $[0, T]$ with $0 = t_1 < t_2 < ... < t_{N_\tau+1} = T$. For $1 \le j \le N_\tau$ set

$$J_j := [t_j, t_{j+1}]\ ,\ \tau_j := t_{j+1} - t_j.$$

We assume that the family $\mathcal{I}_\tau$ is *shape regular*, i.e., the ratios τ_j/τ_{j+1} and τ_{j+1}/τ_j are bounded from above independently of j and τ.

With each $1 \le j \le N_\tau$, we associate a partition $\mathcal{T}_j$ of Ω into n-simplices. We denote by $\mathcal{E}_j$ the set of the interior faces of $\mathcal{T}_j$. For $K \in \mathcal{T}_j$ and $E \in \mathcal{E}_j$ let h_K, ρ_K, and h_E be the diameter of K, the diameter of the largest ball inscribed into K, and the diameter of E. We assume that the partitions $\mathcal{T}_j$ satisfy the following two conditions:

(1) *Admissibility:* Any two simplices of $\mathcal{T}_j$ either are disjoint or share a complete smooth submanifold of their boundaries.
(2) *Shape regularity:* The ratio h_K/ρ_K is bounded from above independently of $K \in \mathcal{T}_j, j$, and τ.

Condition (2) allows the use of locally refined meshes. It implies that the ratio h_K/h_E, for all $K \in \mathcal{T}_j$, all faces E of K, and all j, is bounded from above and from below by constants which do not depend on K, E, j, and τ.

Denote by

$$\mathcal{P}_\tau := \{K \times J_j : 1 \le j \le N_\tau, K \in \mathcal{T}_j\}$$

the partition of the space-time cylinder $\Omega \times [0, T]$ into prisms which is induced by $\mathcal{I}_\tau$ and the $\mathcal{T}_j$'s.

For any $E \in \mathcal{E}_j, 1 \le j \le N_\tau$, and any piecewise continuous function u, we denote by $[u]_E$ the jump of u across E in an arbitrary but fixed direction n_E orthogonal to E. Finally, we introduce the following neighbourhoods of elements and points:

$$\begin{aligned}
&U(t_j) := [t_{\max\{1,j-1\}}, t_{\min\{N_\tau+1,j+1\}}], \quad 1 \le j \le N_\tau,\\
&U(J_j) := U(t_j) \cup U(t_{j+1}) = [t_{\max\{1,j-1\}}, t_{\min\{N_\tau+1,j+2\}}], \quad 1 \le j \le N_\tau,\\
&U(K) := \bigcup_{\substack{K'\cap K\neq\emptyset\\ K'\in\mathcal{T}_j}} K', \quad K \in \mathcal{T}_j, 1 \le j \le N_\tau,\\
&U(E) := \bigcup_{\substack{E\cap K'\neq\emptyset\\ K'\in\mathcal{T}_j}} K', \quad E \in \mathcal{E}_j, 1 \le j \le N_\tau,\\
&U(P) := \bigcup_{\substack{P\cap P'\neq\emptyset\\ P'\in\mathcal{P}_\tau}} P', \quad P \in \mathcal{P}_\tau,\\
&\omega_K := \bigcup_{\substack{K'\cap K\in\mathcal{E}_j\\ K'\in\mathcal{T}_j}} K', \quad K \in \mathcal{T}_j, 1 \le j \le N_\tau,\\
&\omega_E := \bigcup_{\substack{E\subset\partial K'\\ K'\in\mathcal{T}_j}} K', \quad E \in \mathcal{E}_j, 1 \le j \le N_\tau.
\end{aligned} \tag{3.1}$$

Here, $K \cap K' \in \mathcal{E}_j$ means that K and K' share a complete $(n-1)$-dimensional smooth submanifold of their boundaries.

Finite element spaces. Denote by $\mathbb{P}_k, k \geq 0$, the space of polynomials (in x) of degree at most k. Given an admissible partition $\mathcal{T}_h$ of Ω, we define the finite element spaces (in x) as usual:

$$\begin{aligned}
&S_h^{k,-1} := \{u : \Omega \to \mathbb{R} : u|_K \in \mathbb{P}_k \quad \forall K \in \mathcal{T}_h\}, \quad k \geq 0, \\
&S_h^{k,0} := S_h^{k,-1} \cap C(\overline{\Omega}), \quad k \geq 1, \\
&S_{h,0}^{k,0} := \{u \in S_h^{k,0} : u = 0 \text{ on } \Gamma\}, \quad k \geq 1.
\end{aligned} \tag{3.2}$$

For any $k \geq 1$ and any $1 \leq p \leq \infty$ we obviously have $S_h^{k,0} \subset W^{1,p}(\Omega)$ and $S_{h,0}^{k,0} \subset W_0^{1,p}(\Omega)$.

Let $V_1, ..., V_{N_\tau}$ be finite element subspaces of $C(\overline{\Omega})$ associated with the partitions $\mathcal{T}_1, ..., \mathcal{T}_{N_\tau}$ introduced above. For $k \geq 0$, we define finite element spaces in space and time by

$$S_\tau^{k,-1}(V_{h(\tau)}) := \operatorname{span}\{\chi_{J_j}(t) t^\mu v_{j,\mu}(x) : 0 \leq \mu \leq k, 1 \leq j \leq N_\tau, v_{j,\mu} \in V_j\}. \tag{3.3}$$

Here, χ_{J_j} denotes the characteristic function of the interval J_j. The elements of $S_\tau^{k,-1}(V_{h(\tau)})$ are discontinuous at the intermediate points $t_2, ..., t_{N_\tau}$. But the left-sided limits $u(., t_j - 0) := \lim_{0<t\to 0} u(., t_j - t)$ exist for all $2 \leq j \leq N_\tau + 1$; and the right-sided limits $u(., t_j + 0) := \lim_{0<t\to 0} u(., t_j + t)$ exist for all $1 \leq j \leq N_\tau$. We have $S_\tau^{k,-1}(V_{h(\tau)}) \subset L_\pi^p(\Omega \times (0, T))$ for all $1 \leq p, \pi < \infty$. But due to the discontinuities at the intermediate points $t_2, ..., t_{N_\tau}$, the space $S_\tau^{k,-1}(V_{h(\tau)})$ is not contained in any of the spaces $W^p(0, T; W^{1,\pi}(\Omega), L^\pi(\Omega)), 1 \leq p, \pi < \infty$. In order to obtain conforming approximations of these spaces, we denote by $\lambda_j(t), 1 \leq j \leq N_\tau + 1$, the piecewise linear function corresponding to $\mathcal{I}_\tau$ which takes the value 1 at the point t_j and which vanishes at all other points $t_i, 1 \leq i \leq N_\tau + 1, i \neq j$. Set $b_j(t) := 4\lambda_j(t)\lambda_{j+1}(t), 1 \leq j \leq N_\tau$, and for $\theta \in [0, 1]$ define

$$\begin{aligned}
&\lambda_j^{(\theta)}(t) := \lambda_j(t) + \frac{3}{2}(\theta - \frac{1}{2})[b_j(t) - b_{j-1}(t)] \,, 2 \leq j \leq N_\tau, \\
&\lambda_1^{(\theta)}(t) := \lambda_1(t) + \frac{3}{2}(\theta - \frac{1}{2}) b_1(t), \\
&\lambda_{N_\tau+1}^{(\theta)}(t) := \lambda_{N_\tau+1}(t) - \frac{3}{2}(\theta - \frac{1}{2}) b_{N_\tau}(t).
\end{aligned} \tag{3.4}$$

Obviously, the functions $\lambda_j^{(1/2)}$ and λ_j coincide. For $k \geq 2$ we then set

$$\begin{aligned}
S_\tau^{\theta;1,0}(V_{h(\tau)}) &:= \operatorname{span}\{\lambda_j^{(\theta)}(t) v_j(x) : 1 \leq j \leq N_\tau, v_j \in V_j\}, \\
S_\tau^{\theta;k,0}(V_{h(\tau)}) &:= S_\tau^{\theta;1,0}(V_{h(\tau)}) \oplus \operatorname{span}\{\chi_{J_j}(t) t^\mu b_j(t) w_{j,\mu}(x) : \\
&\qquad\qquad 0 \leq \mu \leq k-2, 1 \leq j \leq N_\tau, w_{j,\mu} \in V_j\}.
\end{aligned} \tag{3.5}$$

For all $k \geq 1$ and $1 \leq p, \pi < \infty$ we have $S_\tau^{\theta;k,0}(V_{h(\tau)}) \subset W_T^p(0, T; W^{1,\pi}(\Omega), L^\pi(\Omega))$. Moreover, on the open subintervals $(t_j, t_{j+1}), 1 \leq j \leq N_\tau$, the distributional derivative $\partial_t u$ of $u \in S_\tau^{\theta;k,0}(V_{h(\tau)})$ coincides with the classical partial derivative. For abbreviation, we introduce the space

$$\theta_\tau := S_\tau^{\theta;1,0}(V_{h(\tau)}) \text{ with } V_j = S_{j,0}^{1,0} \,, 1 \leq j \leq N_\tau. \tag{3.6}$$

It will play a fundamental role in deriving reliable error estimates. Figure 1 shows the functions $\lambda_j^{(\theta)}$ for some values of θ. The following lemma collects some properties of these functions. We omit its straightforward proof (cf. Lemma 3.1 in [10]).

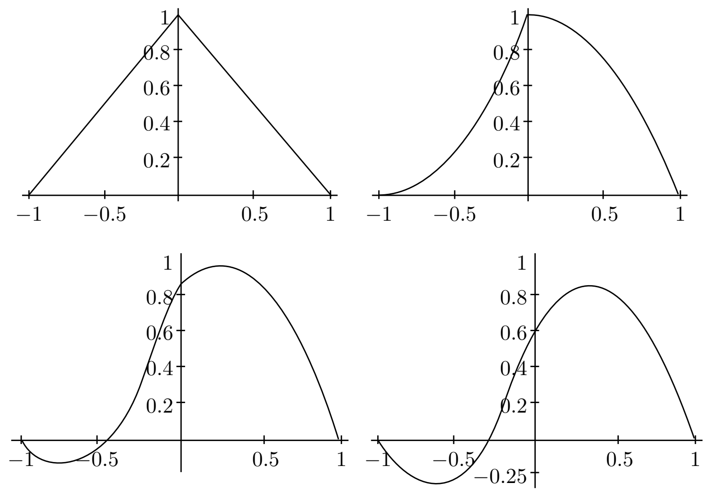

FIGURE 1. $\lambda_j^{(1/2)}$ top left, $\lambda_j^{(2/3)}$ top right, $\lambda_j^{(5/6)}$ bottom left, $\lambda_j^{(1)}$ bottom right

Lemma 3.1. *The functions $\lambda_j^{(\theta)}, 1 \le j \le N_\tau + 1, \theta \in [0,1]$, defined in (3.4) have the following properties:*

$$\begin{aligned}
&\int_{t_j}^{t_{j+1}} \lambda_j^{(\theta)}(t)dt = \theta\tau_j, \quad \forall 1 \le j \le N_\tau, \theta \in [0,1],\\
&\int_{t_{j-1}}^{t_j} \lambda_j^{(\theta)}(t)dt = (1-\theta)\tau_{j-1}, \quad \forall 2 \le j \le N_\tau + 1, \theta \in [0,1],\\
&\lambda_j^{(\theta)}(t) + \lambda_{j+1}^{(\theta)}(t) = 1 \text{ on } J_j, \quad \forall 1 \le j \le N_\tau, \theta \in [0,1],\\
&\lambda_j^{(\theta)}(t) = \lambda_{j+1}^{(1-\theta)}(t_j + t_{j+1} - t) \text{ on } J_j, \quad \forall 1 \le j \le N_\tau, \theta \in [0,1],\\
&\lambda_j^{(\theta)}(t) \ge 0 \qquad \forall 1 \le j \le N_\tau + 1, \theta \in [\frac{1}{3}, \frac{2}{3}], t \in [0,T],\\
&|\lambda_j^{(\theta)}(t)| \le \frac{4}{3} \qquad \forall 1 \le j \le N_\tau + 1, \theta \in [0,1], t \in [0,T].
\end{aligned}$$

Interpolation in space. Let $\mathcal{T}_h$ be an admissible and shape regular partition of Ω into n-simplices. We denote by $I_h : L^1(\Omega) \to S_{h,0}^{1,0}$ the quasi-interpolation operator of Clément (cf. [5] and Exercise 3.2.3 in [4]).

Lemma 3.2. *The operator I_h satisfies the following error estimates for all $K \in \mathcal{T}_h, E \in \mathcal{E}_h$, and $1 \le p < \infty$:*

$$\begin{aligned}
\|u - I_h u\|_{W^{k,p}(K)} &\preceq h_K^{l-k} \|u\|_{W^{l,p}(U(K))} \quad \forall 0 \le k \le l \le 2, u \in W^{l,p}(U(K)),\\
\|u - I_h u\|_{L^p(E)} &\preceq h_E^{l-\frac{1}{p}} \|u\|_{W^{l,p}(U(E))} \quad \forall 1 \le l \le 2, u \in W^{l,p}(U(E)),\\
\|I_h u\|_{L^p(K)} &\preceq h_K^2 \|u\|_{W^{2,p}(U(K))} + \|u\|_{L^p(K)} \quad \forall u \in W^{2,p}(U(K)).
\end{aligned}$$

Proof. The first two inequalities follow from [5] and Exercise 3.2.3 in [4]. The third estimate follows from the first one and the triangle inequality. $\square$

Interpolation in time. Let V and W be Banach spaces as above. For $1 \leq j \leq N_\tau + 1$ we denote by $\pi_j : W^1(0,T;V,W) \to V$ the $L^2(U(t_j))$-projection, i.e.

$$\pi_j u = \frac{1}{\tau_j + \tau_{j-1}} \int_{t_{j-1}}^{t_{j+1}} u(.,t)dt. \tag{3.7}$$

Here, we formally set

$$\tau_0 := \tau_{N_\tau+1} := 0\ ,\ t_0 := 0\ ,\ t_{N_\tau+2} := T.$$

Note that the integral in (3.7) is taken in V.

Lemma 3.3. *The operator π_j satisfies the following estimates for all $1 \leq j \leq N_\tau + 1, u \in W^p(0,T;V,W)$, and $1 \leq p < \infty$:*

$$\begin{aligned}
\|\pi_j u\|_{L^p(U(t_j);Z)} &\leq \|u\|_{L^p(U(t_j);Z)}, \quad Z = V, W,\\
\|u - \pi_j u\|_{L^p(U(t_j);W)} &\leq 2(\tau_j + \tau_{j-1})\|\partial_t u\|_{L^p(U(t_j);W)},\\
\|u(.,t_j) - \pi_j u\|_W &\leq (\tau_j + \tau_{j-1})^{\frac{1}{p'}}\|\partial_t u\|_{L^p(U(t_j);W)}.
\end{aligned}$$

Proof. From (3.7) and Hölder's inequality we obtain

$$\|\pi_j u\|_{L^p(U(t_j);Z)} = (\tau_j + \tau_{j-1})^{\frac{1}{p}-1}\| \int_{t_{j-1}}^{t_{j+1}} u(.,t)dt\|_Z \leq \|u\|_{L^p(U(t_j);Z)}.$$

From Hölder's inequality we get, for all $t \in U(t_j)$,

$$\|u(.,t) - u(.,t_j)\|_W = \| \int_{t_j}^{t} \partial_t u(.,\sigma)d\sigma\|_W \leq |t - t_j|^{\frac{1}{p'}}\|\partial_t u\|_{L^p(U(t_j);W)}.$$

Taking the p-th power and integrating from t_{j-1} to t_{j+1}, this yields

$$\|u - u(.,t_j)\|_{L^p(U(t_j);W)} \leq (\tau_j + \tau_{j-1})\|\partial_t u\|_{L^p(U(t_j);W)}. \tag{3.8}$$

Combining estimate (3.8) with the first assertion of the lemma, we obtain

$$\begin{aligned}
&\|u - \pi_j u\|_{L^p(U(t_j);W)}\\
&\quad\leq \|u - u(.,t_j)\|_{L^p(U(t_j);W)} + \|u(.,t_j) - \pi_j u\|_{L^p(U(t_j);W)}\\
&\quad= \|u - u(.,t_j)\|_{L^p(U(t_j);W)} + \|\pi_j[u(.,t_j) - u]\|_{L^p(U(t_j);W)}\\
&\quad\leq 2(\tau_j + \tau_{j-1})\|\partial_t u\|_{L^p(U(t_j);W)}.
\end{aligned}$$

Estimate (3.8), the definition of π_j and Hölder's inequality finally yield

$$\begin{aligned}
\|u(.,t_j) - \pi_j u\|_W &= \|\pi_j[u(.,t_j) - u]\|_W\\
&\leq (\tau_j + \tau_{j-1})^{-1}\int_{t_{j-1}}^{t_{j+1}} \|u(.,t_j) - u(.,t)\|_W dt\\
&\leq (\tau_j + \tau_{j-1})^{-1+\frac{1}{p'}}\|u - u(.,t_j)\|_{L^p(U(t_j);W)}\\
&\leq (\tau_j + \tau_{j-1})^{\frac{1}{p'}}\|\partial_t u\|_{L^p(U(t_j);W)}.
\end{aligned}$$

This proves the last assertion of the lemma. ∎

Interpolation in space and time. Now, we combine the operators I_j, corresponding to $\mathcal{T}_j$, and π_j, and define the operator $I_\tau : W^p(0,T;W_0^{1,\pi}(\Omega),L^\pi(\Omega)) \to \theta_\tau, 1 \le p,\pi < \infty$, by

$$I_\tau u := \sum_{j=1}^{N_\tau} \lambda_j^{(\theta)}(t)\pi_j I_j u. \tag{3.9}$$

Note that π_j and I_j commute for all $1 \le j \le N_\tau$.

Lemma 3.4. *The operator I_τ satisfies the following error estimates for all $Q = K \times J_j$, $1 \le j \le N_\tau, K \in \mathcal{T}_j, E \in \mathcal{E}_j, 1 \le p,\pi < \infty$, $\theta \in [0,1]$, and $u \in W_\pi^p(U(Q))$:*

$$\|u - I_\tau u\|_{L_\pi^p(Q)} \preceq (h_K^2 + \tau_j)\|u\|_{W_\pi^p(U(Q))},$$
$$\|u - I_\tau u\|_{L_\pi^p(E\times J_j)} \preceq h_K^{-\frac{1}{\pi}}(h_K^2 + \tau_j)\|u\|_{W_\pi^p(U(Q))},$$
$$\|u(.,t_j) - I_\tau u(.,t_j)\|_{L^\pi(K)} \preceq \tau_j^{-1/p}(h_K^2 + \tau_j)\|u\|_{W_\pi^p(U(Q))}.$$

Proof. The definition of I_τ and Lemma 3.1 imply that the splitting

$$u - I_\tau u = \lambda_j^{(\theta)}\{u - I_j\pi_j u\} + \lambda_{j+1}^{(\theta)}\{u - I_{j+1}\pi_{j+1}u\}$$

holds on J_j. Invoking Lemma 3.1 once more, we obtain

$$\|u - I_\tau u\|_{L_\pi^p(Q)} \le \frac{4}{3}\{\|u - I_j\pi_j u\|_{L_\pi^p(Q)} + \|u - I_{j+1}\pi_{j+1}u\|_{L_\pi^p(Q)}\}.$$

Let $k = j$ or $k = j+1$. Lemmas 3.2 and 3.3 and the shape regularity of $\mathcal{I}_\tau$ yield

$$\begin{aligned}
&\|u - I_k\pi_k u\|_{L_\pi^p(Q)} \\
&\quad \le \|u - I_k u\|_{L_\pi^p(Q)} + \|I_k[u - \pi_k u]\|_{L_\pi^p(Q)} \\
&\quad \preceq h_K^2\|u\|_{L^p(J_j,W^{2,\pi}(U(K)))} + \|u - \pi_k u\|_{L_\pi^p(Q)} \\
&\quad \preceq (h_K^2 + \tau_j)\|u\|_{W_\pi^p(U(Q))}.
\end{aligned}$$

This proves the first estimate of the lemma. The second one is established in the same way.

In order to prove the third estimate, we now write

$$\begin{aligned}
u(.,t_j) - I_\tau u(.,t_j) &= u(.,t_j) - \pi_j I_j u(.,t_j) \\
&= u(.,t_j) - \pi_j u(.,t_j) + \pi_j[u(.,t_j) - I_j u(.,t_j)].
\end{aligned}$$

Lemmas 3.2 and 3.3 and the definition of π_j then yield

$$\begin{aligned}
&\|u(.,t_j) - I_\tau u(.,t_j)\|_{L^\pi(K)} \\
&\quad \le \|u(.,t_j) - \pi_j u(.,t_j)\|_{L^\pi(K)} + \|\pi_j[u(.,t_j) - I_j u(.,t_j)]\|_{L^\pi(K)} \\
&\quad \le \tau_j^{\frac{1}{p'}}\|\partial_t u\|_{L_\pi^p(U(Q))} + \tau_j^{\frac{1}{p'}-1}\|u - I_j u\|_{L_\pi^p(U(Q))} \\
&\quad \preceq \tau_j^{-1/p}(h_K^2 + \tau_j)\|u\|_{W_\pi^p(U(Q))}.
\end{aligned}$$

This completes the proof. $\square$

Local cut-off functions. Denote by $\hat{K} := \{\hat{x} \in \mathbb{R}^n : \sum_{i=1}^n \hat{x}_i \le 1, \hat{x}_j \ge 0, 1 \le j \le n\}$ the reference n-simplex and set $\hat{E} := \hat{K} \cap \{\hat{x} \in \mathbb{R}^n : \hat{x}_n = 0\}$. Let

$$\varphi_{\hat{K}}(\hat{x}) := (n+1)^{n+1}[1 - \sum_{i=1}^n \hat{x}_i] \prod_{j=1}^n \hat{x}_j, \quad \varphi_{\hat{E}}(\hat{x}) := n^n[1 - \sum_{i=1}^n \hat{x}_i] \prod_{j=1}^{n-1} \hat{x}_j.$$

Given an arbitrary n-simplex K and a face E of K, we denote by $F_K : \hat{K} \to K, \hat{x} \to x := F_K(\hat{x}) = b_K + B_K\hat{x}$ an invertible affine mapping such that $\hat{K}$ is mapped onto K and $\hat{E}$ is mapped onto E. Let $F_E : \hat{E} \to E$ be the tranformation induced by F_K and denote by β_K its Gram determinant. One easily checks that $\beta_K = \det(B_K'^t B_K')^{1/2}$. Here B_K' denotes the matrix which is obtained by discarding the last column of B_K.

Let $\hat{J} := [0,1]$ be the unit interval, and set

$$\varphi_{\hat{J}}(t) := 4t(1-t).$$

Given an arbitrary interval $J = [a,b], a < b$, we denote by $F_J : \hat{J} \to J$ the invertible affine mapping which maps $\hat{J}$ onto J and 0 onto a. Set

$$\hat{Q} := \hat{K} \times \hat{J}\ ,\ \partial\hat{Q}_L := \hat{E} \times \hat{J}\ ,\ \partial\hat{Q}_B := \hat{K} \times \{0\}$$

and define the transformations $F_Q : \hat{Q} \to K \times J, F_{\partial Q_L} : \partial\hat{Q}_L \to E \times J$, and $F_{\partial Q_B} : \partial\hat{Q}_B \to K \times \{0\}$ by

$$F_Q := (F_K, F_J)\ ,\ F_{\partial Q_L} := (F_E, F_J)\ ,\ F_{\partial Q_B} := (F_K, a).$$

Finally, we denote by $V_{\hat{Q}} \subset L^\infty(\hat{Q}), V_{\partial\hat{Q}_L} \subset L^\infty(\partial\hat{Q}_L)$, and $V_{\partial\hat{Q}_B} \subset L^\infty(\partial\hat{Q}_B)$ three arbitrary finite dimensional spaces which are kept fixed in what follows. In applications, these spaces will be subspaces of appropriate spaces of polynomials. We set

$$\begin{aligned} V_Q &:= \{\hat{u} \circ F_Q^{-1} : \hat{u} \in V_{\hat{Q}}\}, \\ V_{\partial Q_L} &:= \{\hat{\sigma} \circ F_{\partial Q_L}^{-1} : \hat{\sigma} \in V_{\partial\hat{Q}_L}\}, \\ V_{\partial Q_B} &:= \{\hat{v} \circ F_{\partial Q_B}^{-1} : \hat{v} \in V_{\partial\hat{Q}_B}\}. \end{aligned}$$

Let $\mathcal{T}_h$ be an admissible and shape regular partition of Ω into n-simplices. Given an arbitrary simplex $K \in \mathcal{T}_h$, denote by $\lambda_{K1}, ..., \lambda_{K(n+1)}$ its barycentric coordinates, and set

$$\psi_K := \begin{cases} \{(n+1)^{n+1} \prod_{i=1}^{n+1} \lambda_{Ki}\}^2 & \text{on } K, \\ 0 & \text{on } \Omega \backslash K. \end{cases} \tag{3.10}$$

The function ψ_K obviously has the following properties:

$$\begin{aligned} &\psi_K = \{\varphi_{\hat{K}} \circ F_K^{-1}\}^2 \quad \text{on } K, \\ &0 \le \psi_K(x) \le 1 \quad \forall x \in \Omega, \\ &\max_{x \in K} \psi_K(x) = 1, \\ &\psi_K = 0 \text{ on } \partial K, \\ &\nabla\psi_K = 0 \text{ on } \partial K, \\ &\psi_K \in C^1(\Omega). \end{aligned} \tag{3.11}$$

Given $E \in \mathcal{E}_h$, denote by K_1 and K_2 the two simplices adjacent to E, and enumerate their vertices so that the vertices of E are numbered first. Set

$$\psi_E := \begin{cases} \{n^{2n} \prod_{i=1}^{n} (\lambda_{K_1 i} \lambda_{K_2 i})\}^2 & \text{on } \omega_E, \\ 0 & \text{on } \Omega \backslash \omega_E. \end{cases} \tag{3.12}$$

One easily checks that the function ψ_E has the following properties:

$$\begin{aligned} &\psi_E = \{\varphi_{\hat{E}} \circ F_E^{-1}\}^4 \quad \text{on } E, \\ &0 \le \psi_E(x) \le c(n, \max_{K \in \mathcal{T}_h} \frac{h_K}{\rho_K}) \quad \forall x \in \Omega, \\ &\max_{x \in E} \psi_E(x) = 1, \\ &\psi_E = 0 \text{ on } \partial\omega_E, \\ &\nabla\psi_E = \text{ on } \partial\omega_E, \\ &\psi_E \in C^1(\Omega). \end{aligned} \tag{3.13}$$

In particular, the restriction of ψ_E to E depends only on the vertices of E.

For $1 \le j \le N_\tau$, we set

$$\psi_j(t) := \begin{cases} \frac{4}{\tau_j^2}(t - t_j)(t_{j+1} - t) & \text{if } t \in J_j, \\ 0 & \text{if } t \notin J_j. \end{cases} \tag{3.14}$$

Obviously, we have

$$\begin{aligned} &\psi_j = \psi_{\hat{J}} \circ F_{J_j}^{-1}, \\ &0 \le \psi_j(t) \le 1 \quad \forall t \in \mathbb{R}, \\ &\max_{t \in J_j} \psi_j(t) = 1, \\ &\psi_j = 0 \text{ on } \partial J_j, \\ &\psi_j \in C(\mathbb{R}). \end{aligned} \tag{3.15}$$

Let $1 \le j \le N_\tau, K \in \mathcal{T}_j$, and $E \in \mathcal{E}_j$ be arbitrary. We then use the following **abbreviations**:

$$Q := K \times J_j \ , \ \partial Q_B := K \times \{t_j\} \ , \ \partial Q_L := E \times J_j. \tag{3.16}$$

We define a continuation operator $P_j : L^\infty(\partial Q_B) \to L^\infty(K \times U(t_j))$ by

$$P_j u(x, t) := u(x, t_j) \quad \forall (x, t) \in K \times U(t_j). \tag{3.17}$$

Next, we want to define a continuation operator $P_E : L^\infty(\partial Q_L) \to L^\infty(\omega_E \times J_j)$. To this end, we denote by $x_E = (x_{E1}, ..., x_{En})$ a Euclidean coordinate system such that E is contained in the set $\{x_{En} = 0\}$. We then set $x'_E := (x_{E1}, ..., x_{E(n-1)})$ and define

$$P_E\sigma(x_E, t) := \begin{cases} \psi_E(x'_E, 0)\sigma(x'_E, 0, t & \text{if } (x'_E, 0) \in E, \\ 0 & \text{if } (x'_E, 0) \notin E. \end{cases} \tag{3.18}$$

Note that, without any restriction on the partition $\mathcal{T}_j$, the factor ψ_E in the definition of P_E ensures that $\nabla(P_E\sigma) \in C(\omega_E \times J_j)$ if $\nabla_{x'_E}\sigma \in C(E \times J_j)$. The factor ψ_E may be dropped if $\mathcal{T}_j$ has the following property: For any $E \in \mathcal{E}_j$, the orthogonal projections of all vertices of ω_E onto the plane $\{x_{E_n} = 0\}$ lie inside E. In two dimensions this means that the triangulation $\mathcal{T}_j$ is weakly acute.

Lemma 3.5. *Let $1 \le j \le N_\tau, K \in \mathcal{T}_j$, and $E \in \mathcal{E}_j$ be arbitrary, and recall the abbreviations (3.16). Assume that the functions in $V_{\hat{Q}}, V_{\partial\hat{Q}_L}$, and $V_{\partial\hat{Q}_B}$ are continuously differentiable with respect to the time variable and twice continuously differentiable with respect to the space variable. Then the following estimates hold for all $u \in V_Q, \sigma \in V_{\partial Q_L}, v \in V_{\partial Q_B}$, and $1 < p, \pi < \infty$:*

$$
\begin{aligned}
\|u\|_{L^p_\pi(Q)} &\simeq \sup_{w \in V_Q} \frac{\int_Q u\psi_j\psi_K v}{\|v\|_{L^{p'}_{\pi'}(Q)}}\,,\\
h_K^{-2}\|\psi_j\psi_K u\|_{L^p_\pi(Q)} &\simeq \|\nabla^2(\psi_j\psi_K u)\|_{L^p_\pi(Q)}\,,\\
\tau_j^{-1}\|\psi_j\psi_K u\|_{L^p_\pi(Q)} &\simeq \|\partial_t(\psi_j\psi_K u)\|_{L^p_\pi(Q)}\,,\\
\|\sigma\|_{L^p_\pi(\partial Q_L)} &\simeq \sup_{\chi \in V_{\partial Q_L}} \frac{\int_{\partial Q_L} \sigma\psi_j\psi_E P_E\chi}{\|\chi\|_{L^{p'}_{\pi'}(\partial Q_L)}}\,,\\
h_E^{-2}\|\psi_j\psi_E P_E\sigma\|_{L^p_\pi(\omega_E\times J_j)} &\simeq \|\nabla^2(\psi_j\psi_E P_E\sigma)\|_{L^p_\pi(\omega_E\times J_j)}\,,\\
\tau_j^{-1}\|\psi_j\psi_E P_E\sigma\|_{L^p_\pi(\omega_E\times J_j)} &\simeq \|\partial_t(\psi_j\psi_E P_E\sigma)\|_{L^p_\pi(\omega_E\times J_j)}\,,\\
\|\psi_j\psi_E P_E\sigma\|_{L^p_\pi(\omega_E\times J_j)} &\preceq h_E^{1/\pi}\|\sigma\|_{L^p_\pi(\partial Q_L)}\,,\\
\|v\|_{L^\pi(K)} &\simeq \sup_{w \in V_{\partial Q_B}} \frac{\int_K v\psi_K w}{\|w\|_{L^{\pi'}(K)}},\\
h_K^{-2}\|\lambda_j\psi_K P_j v\|_{L^p_\pi(U(Q))} &\simeq \|\nabla^2(\lambda_j\psi_K P_j v)\|_{L^p_\pi(U(Q))}\,,\\
\tau_j^{-1}\|\lambda_j\psi_K P_j v\|_{L^p_\pi(U(Q))} &\simeq \|\partial_t(\lambda_j\psi_K P_j v)\|_{L^p_\pi(U(Q))}\,,\\
\|\lambda_j\psi_K P_j v\|_{L^p_\pi(U(Q))} &\preceq \tau_j^{1/p}\|v\|_{L^\pi(K)}\,.
\end{aligned}
$$

Proof. We prove the estimates concerning u. The upper bound in the first estimate follows from Hölder's inequality and $0 \le \psi_K \le 1, 0 \le \psi_j \le 1$. In order to prove the lower bound of the first estimate, one easily checks that the mapping

$$
\hat{u} \to \sup_{\hat{v} \in V_{\hat{Q}}} \frac{\int_{\hat{Q}} \hat{u}\psi_{\hat{j}}\psi_{\hat{K}}\hat{v}}{\|\hat{v}\|_{L^{p'}_{\pi'}(\hat{Q})}}
$$

defines a norm on $V_{\hat{Q}}$. Since $\dim V_{\hat{Q}} < \infty$, there is a constant $\hat{c}$ such that

$$
\hat{c}\|\hat{u}\|_{L^p_\pi(\hat{Q})} \le \sup_{\hat{v} \in V_{\hat{Q}}} \frac{\int_{\hat{Q}} \hat{u}\psi_{\hat{j}}\psi_{\hat{K}}\hat{v}}{\|\hat{v}\|_{L^{p'}_{\pi'}(\hat{Q})}} \quad \forall \hat{u} \in V_{\hat{Q}}.
$$

The definition of V_Q therefore yields

$$
\begin{aligned}
\sup_{v \in V_Q} \frac{\int_Q u\psi_j\psi_K v}{\|v\|_{L^{p'}_{\pi'}(Q)}} &= \tau_j^{1-\frac{1}{p'}}|\det B_K|^{1-\frac{1}{\pi'}} \sup_{\hat{v} \in V_{\hat{Q}}} \frac{\int_{\hat{Q}} u \circ F_Q\psi_{\hat{j}}\psi_{\hat{K}}\hat{v}}{\|\hat{v}\|_{L^{p'}_{\pi'}(\hat{Q})}}\\
&\ge \hat{c}\tau_j^{\frac{1}{p}}|\det B_K|^{\frac{1}{\pi}}\|u \circ F_Q\|_{L^p_\pi(\hat{Q})}\\
&= \hat{c}\|u\|_{L^p_\pi(Q)}.
\end{aligned}
$$

In order to establish the second and third estimates, we observe that the mappings $\hat{u} \to \|\nabla^2(\psi_{\hat{j}}\psi_{\hat{K}}\hat{u})\|_{L^p_\pi(\hat{Q})}$ and $\hat{u} \to \|\partial_t(\psi_{\hat{j}}\psi_{\hat{K}}\hat{u})\|_{L^p_\pi(\hat{Q})}$ define norms on the finite dimensional space $V_{\hat{Q}}$ which are equivalent to the standard norm $\|.\|_{L^p_\pi(\hat{Q})}$. The desired estimates now follow in the usual way by transforming to $\hat{Q}$, using the equivalence of norms there, and transforming back to Q.

The estimates concerning σ and v are established in the same way, taking into account that $P_E\sigma$ and $P_j v$ are constant along lines perpendicular to E and with respect to time, respectively. For a more detailed proof see Lemma 3.8 in [10]. $\square$

4. Quasilinear parabolic equations of 2nd order

Variational setting. As a model problem we consider the parabolic boundary value problem

$$\begin{aligned} \partial_t u - \nabla \cdot \underline{a}(x,u,\nabla u) &= b(x,u,\nabla u) && \text{in } \Omega \times (0,T), \\ u &= 0 && \text{on } \Gamma \times (0,T), \\ u(.,0) &= u_0 && \text{in } \Omega. \end{aligned} \tag{4.1}$$

Here, $b \in C^1(\Omega \times \mathbb{R} \times \mathbb{R}^n, \mathbb{R})$ and $\underline{a} \in C^1(\Omega \times \mathbb{R} \times \mathbb{R}^n, \mathbb{R}^n)$ are such that the matrix $A(x,y,z) := (\frac{1}{2}(\partial_{z_j} a_i(x,y,z) + \partial_{z_i} a_j(x,y,z)))_{1 \le i,j \le n}$ is positive definite for all $x \in \Omega, y \in \mathbb{R}, z \in \mathbb{R}^n$. Moreover, $T > 0$ denotes an arbitrary final time which is kept fixed.

Under suitable growth conditions on $\underline{a}, b$, and their derivatives there are real numbers $1 < p, r, \pi, \rho < \infty$ such that problem (4.1) fits into the abstract framework of Section 2 with

$$\begin{aligned} X &:= L^r(0,T; W_0^{1,\rho}(\Omega)), \\ Y &:= W_T^{p'}(0,T; W_0^{1,\pi'}(\Omega), W^{-1,\pi'}(\Omega)), \\ \langle F(u), \varphi \rangle &:= -\langle u_0, \varphi(.,0) \rangle_{W^{1,\pi}} \\ &\quad - \int_0^T \langle u(.,t), \partial_t \varphi(.,t) \rangle_{W^{1,\pi}} dt \\ &\quad + \int_0^T \int_\Omega \{\underline{a}(x,u,\nabla u)\nabla\varphi - b(x,u,\nabla u)\varphi\} dx dt. \end{aligned} \tag{4.2}$$

We recall that a prime denotes the dual of a Lebesgue exponent, i.e. $\frac{1}{p} + \frac{1}{p'} = 1$, etc. In order to ensure that F is well defined, we always assume that

$$\rho \ge \pi. \tag{4.3}$$

Within the framework of Section 2, we further set

$$\begin{aligned} Y_+ &:= W_T^{p'}(0,T; W^{2,\pi'}(\Omega) \cap W_0^{1,\pi'}(\Omega), L^{\pi'}(\Omega)), \\ X_- &:= L^r(0,T; L^\rho(\Omega)), \\ X_+ &:= L^r(0,T; W_0^{1,\sigma}(\Omega)), \end{aligned} \tag{4.4}$$

where $\rho \le \sigma < \infty$.

In order to better understand the flavour of problem (1.1) and definition (4.2), we recall the notions of weak and very weak solutions of problem (4.1) (cf. [2], Sections 11 and 13). A function $u \in W^r(0,T; W_0^{1,\rho}(\Omega), W^{-1,\pi}(\Omega))$ is called a *weak solution* of problem (4.1) if

$$u(.,0) = u_0 \text{ in } W^{-1,\pi}(\Omega)$$

and

$$\int_0^T \langle \partial_t u(.,t), \varphi(.,t) \rangle_{W^{1,\pi'}} + \int_0^T \int_\Omega \{\underline{a}(x,u,\nabla u)\nabla\varphi - b(x,u,\nabla u)\varphi\} dx dt = 0$$

for all $\varphi \in L^{p'}(0,T;W^{1,\pi'}(\Omega))$. It is a *very weak solution* if $u \in X_-$ and

$$- \langle u_0, \varphi(.,0)\rangle_{L^\rho} - \int_0^T \langle u(.,t), \partial_t \varphi(.,t)\rangle_{L^\rho} dt$$
$$+ \int_0^T \{\langle \underline{a}(.,u,\nabla u), \nabla\varphi\rangle_{W^{1,\pi'}} - \langle b(x,u,\nabla u), \varphi\rangle_{W^{1,\pi'}}\} dt = 0$$

for all $\varphi \in Y_+$. Obviously, every solution of problem (1.1) is a very weak solution of problem (4.1). Conversely, every very weak solution of problem (4.1) is a solution of problem (1.1) if it is contained in X. Using integration by parts with respect to the time variable, one sees that every weak solution of problem (4.1) is a solution of problem (1.1), and that the converse is true if the solution of problem (1.1) is contained in $W^r(0,T;W_0^{1,\rho}(\Omega),W^{-1,\pi}(\Omega))$. In this sense, a solution of problem (1.1) is weaker than a weak solution and stronger than a very weak solution of problem (4.1).

One easily checks that $DF(u)^* \in \mathrm{Isom}(Y_+, X_-^*)$ if and only if the adjoint linearized problem

$$\begin{aligned}
-\partial_t w - \nabla\cdot(A(x,u,\nabla u)\nabla w) + \partial_y \underline{a}(x,u,\nabla u)\cdot\nabla w& \\
+\nabla\cdot(\nabla_z b(x,u,\nabla u)w) - \partial_y b(x,u,\nabla u)w &= g \text{ in } \Omega\times(0,T),\\
w &= 0 \text{ on } \Gamma\times(0,T),\\
w(.,T) &= 0 \text{ in } \Omega
\end{aligned}$$

admits, for each $g \in X_-^*$, a unique solution $w \in Y_+$ such that $\|w\|_{Y_+} \preceq \|g\|_{X_-^*}$.

Examples. We consider two particular examples:

(1) *A heat equation with nonlinear source term*:

$$\begin{aligned}
&\underline{a}(x,u,\nabla u) = \nabla u,\\
&b(x,u,\nabla u) = f(u),\\
&f \in C^1(\mathbb{R},\mathbb{R}),\\
&|f'(s)| \le \gamma \qquad \forall s \in \mathbb{R},\\
&p = r = \pi = \rho = 2.
\end{aligned}$$

(2) *A nonlinear convection-diffusion equation*:

$$\begin{aligned}
&\underline{a}(x,u,\nabla u) := k(u)\nabla u,\\
&b(x,u,\nabla u) := f - \underline{c}(x,u)\cdot\nabla u,\\
&f \in L^\infty(\Omega), \underline{c} \in C^1(\Omega\times\mathbb{R},\mathbb{R}^n), k \in C^2(\mathbb{R},\mathbb{R}),\\
&k(s) \ge \alpha > 0, |k^{(l)}(s)| \le \gamma \quad \forall s \in \mathbb{R}, l \in \{0,1,2\},\\
&|\partial_s \underline{c}(x,s)| \le \gamma \quad \forall x \in \Omega, s \in \mathbb{R},\\
&\rho = \pi \in (n,4), p > \frac{n}{n-1}, r \ge 2p.
\end{aligned}$$

If in example (1) the constant γ is sufficiently small, we may use an energy estimate and a perturbation argument to get explicit bounds on $\|DF(u)^{*^{-1}}\|_{\mathcal{L}(X_-^*,Y_+)}$ in terms of the norm of the inverse Laplacian with homogeneous Dirichlet boundary conditions. More precisely, denote by $c_\Delta := \|(-\Delta)^{-1}\|_{\mathcal{L}(L^2(\Omega),W_0^{1,2}(\Omega)\cap W^{2,2}(\Omega))}$ the norm of the inverse Laplacian with homogeneous Dirichlet boundary conditions. This quantity only depends on the geometry of Ω. Inserting $\underline{a}$ and b given above into the adjoint linearized problem, we immediadely see that $DF(u)^* = L - N$, where L is the operator associated with the time-reversed heat equation

$$
\begin{aligned}
-\partial_t v - \Delta v &= g \quad \text{in } \Omega \times (0,T), \\
v &= 0 \quad \text{on } \Gamma \times (0,T), \\
v(.,T) &= 0 \quad \text{in } \Omega,
\end{aligned}
\tag{4.5}
$$

and where the operator N is given by

$$
\langle Nv, \varphi \rangle_{Y_+} = \int_0^T \int_\Omega f'(u) v \varphi. \tag{4.6}
$$

From equations (4.5) and (4.6) we deduce that

$$
\|L\|_{\mathcal{L}(Y_+, X_-^*)} \le 2, \quad \|N\|_{\mathcal{L}(Y_+, X_-^*)} \le \gamma,
$$

and, hence,

$$
\|DF(u)^*\|_{\mathcal{L}(Y_+, X_-^*)} \le 2 + \gamma.
$$

Multiplying the first equation of (4.5) by $-\partial_t v$, integrating over $\Omega \times (0,T)$, and integrating by parts with respect to the space variable, we conclude that

$$
\|\partial_t v\|_{L_2^2(\Omega \times (0,T))} \le \|g\|_{L_2^2(\Omega \times (0,T))}. \tag{4.7}
$$

Writing the first equation of (4.5) in the form

$$
-\Delta v = g + \partial_t v
$$

and using the estimate (4.7), we obtain on the other hand that

$$
\begin{aligned}
\|v\|_{L^2(0,T;W^{2,2}(\Omega))} &\le c_\Delta \sqrt{2} \{ \|g\|_{L_2^2(\Omega \times (0,T))} + \|\partial_t v\|_{L_2^2(\Omega \times (0,T))} \} \\
&\le c_\Delta \sqrt{8} \|g\|_{L_2^2(\Omega \times (0,T))}.
\end{aligned}
\tag{4.8}
$$

Estimates (4.7) and (4.8) yield

$$
\|L^{-1}\|_{\mathcal{L}(X_-^*, Y_+)} \le 1 + \sqrt{8} c_\Delta.
$$

Assume that $\gamma(1 + \sqrt{8} c_\Delta) < 1$. A standard perturbation argument then gives

$$
\begin{aligned}
\|DF(u)^{*^{-1}}\|_{\mathcal{L}(X_-^*, Y_+)} &\le \|L^{-1}\|_{\mathcal{L}(X_-^*, Y_+)} \big[1 - \|L^{-1}\|_{\mathcal{L}(X_-^*, Y_+)} \|N\|_{\mathcal{L}(Y_+, X_-^*)} \big]^{-1} \\
&\le \frac{1 + \sqrt{8} c_\Delta}{1 - \gamma(1 + \sqrt{8} c_\Delta)}.
\end{aligned}
$$

Finite element discretization. For the discretization of problem (4.1) we proceed as in Section 3. We choose a family $\mathcal{I}_\tau$ of shape regular partitions of the interval $[0,T]$. With each time $t_j, 1 \le j \le N_\tau$, we associate an admissible and shape regular partition $\mathcal{T}_j$ of Ω into n-simplices and a finite element space $V_j \subset W_0^{1,\sigma}(\Omega)$ corresponding to $\mathcal{T}_j$ and consisting of affine equivalent finite elements in the sense of [4]. We choose an integer k and a parameter $\theta \in [0,1]$, and set

$$
\begin{aligned}
X_h &:= S_\tau^{k,-1}(V_{h(\tau)}), \\
Y_h &:= S_\tau^{\theta;k+1,0}(V_{h(\tau)}), \\
\langle F_h(u_h), \varphi_h \rangle_{Y_h} &:= \langle F(u_h), \varphi_h \rangle_Y \quad \forall u_h \in X_h, \varphi_h \in Y_h.
\end{aligned}
\tag{4.9}
$$

For simplicity, in (4.9) we use the parameter h for the mesh sizes both in space and in time. We recall that the spaces on the right-hand side of (4.9) are defined in (3.3) and (3.5) and that $X_h \subset X_+$ and $Y_h \subset Y$. Hence, the discretization

(4.9) is conforming. It is also consistent, i.e $\|R_h^*[F(u_h) - F_h(u_h)]\|_{Y_+^*} = 0$. It is a Petrov-Galerkin discretization since the test and trial spaces are different: The trial functions are discontinuous in time, piecewise polynomials of degree k; the test functions are continuous in time, piecewise polynomials of degree $k+1$.

Relation to Runge-Kutta schemes. In order to better understand the flavour of problem (1.2) and definition (4.9) we rewrite F_h. Recalling that the functions in Y_h are continuous at the intermediate times $t_2, ..., t_{N_\tau}$ and vanish at the final time T, and using integration by parts on each time interval, for $u_h \in X_h$ and $\varphi_h \in Y_h$ we get

$$\begin{aligned}
&\langle F(u_h), \varphi_h\rangle_Y \\
&\quad = \int_\Omega [u_h(x, 0+0) - u_0(x)]\varphi_h(x, 0)dx \\
&\qquad + \sum_{j=2}^{N_\tau} \int_\Omega [u_h(x, t_j + 0) - u_h(x, t_j - 0)]\varphi_h(x, t_j)dx \\
&\qquad + \sum_{j=1}^{N_\tau} \int_{t_j}^{t_{j+1}} \int_\Omega [\partial_t u_h \varphi_h + \underline{a}(x, u_h, \nabla u_h)\nabla\varphi_h - b(x, u_h, \nabla u_h)\varphi_h]dxdt.
\end{aligned}$$

Using the convention that

$$u_h(., 0-0) := u_0, \tag{4.10}$$

we may write this in the compact form

$$\begin{aligned}
&\langle F(u_h), \varphi_h\rangle_Y \\
&\quad = \sum_{j=1}^{N_\tau} \Big\{ \int_\Omega [u_h(x, t_j + 0) - u_h(x, t_j - 0)]\varphi_h(x, t_j)dx \\
&\qquad\qquad + \int_{t_j}^{t_{j+1}} \int_\Omega [\partial_t u_h \varphi_h + \underline{a}(x, u_h, \nabla u_h)\nabla\varphi_h - b(x, u_h, \nabla u_h)\varphi_h]dxdt \Big\}.
\end{aligned} \tag{4.11}$$

We first consider the case $k = 0$ and set

$$u_h^j := u_h(., t_j + 0) = u_h(., t_{j+1} - 0) \quad \forall 1 \le j \le N_\tau.$$

Observing that u_h is piecewise constant on the time intervals, inserting $\varphi_h = \lambda_j^{(\theta)} v_j$, $2 \le j \le N_\tau, v_j \in V_j$, as a test function in (4.11) and using Lemma 3.1, we obtain

$$\begin{aligned}
&\langle F_h(u_h), \varphi_h\rangle_{Y_h} \\
&\quad = \int_\Omega \{u_h^j - u_h^{j-1}\} v_j dx \\
&\qquad + \theta\tau_j \int_\Omega \{\underline{a}(x, u_h^j, \nabla u_h^j)\nabla v_j - b(x, u_h^j, \nabla u_h^j) v_j\}dx \\
&\qquad + (1-\theta)\tau_{j-1} \int_\Omega \{\underline{a}(x, u_h^{j-1}, \nabla u_h^{j-1})\nabla v_j - b(x, u_h^{j-1}, \nabla u_h^{j-1}) v_j\}dx.
\end{aligned}$$

Inserting $\varphi_h = \lambda_1^{(\theta)} v_1, v_1 \in V_1$, as a test function in (4.11), we similarly get

$$\langle F_h(u_h), \varphi_h \rangle_{Y_h} = \int_\Omega \{u_h^1 - u_0\} v_1 dx + \theta\tau_1 \int_\Omega \{\underline{a}(x, u_h^1, \nabla u_h^1)\nabla v_1 - b(x, u_h^1, \nabla u_h^1) v_1\} dx.$$

Hence, in the case $k = 0$, problem (1.2) yields the popular θ-scheme. In particular, the parameters $\theta = 0, \theta = 1$, and $\theta = \frac{1}{2}$ correspond to the explicit Euler scheme, the implicit Euler scheme, and the trapezoidal rule (Crank-Nicholson scheme). Thus the time discretization is of first order unless $\theta = \frac{1}{2}$; in this case it is of second order. Moreover, the time discretization is A-stable if $\theta \geq \frac{1}{2}$.

Next we consider the case $k \geq 1$. Denote by $\hat{p}_l, 0 \leq l \leq k-1$, a set of orthonormal polynomials of degree l with respect to the weight function $4t(1-t)$ on $[0, 1]$. Let $\hat{q}_l(t) := \int_0^t \hat{p}_l(s) ds, 0 \leq l \leq k-1$, and $\hat{q}_{-1}(t) = 1$. For $1 \leq j \leq N_\tau$, set

$$p_{l,j} := \hat{p}_l \circ F_{J_j}^{-1}\ ,\ q_{m,j} := \hat{q}_m \circ F_{J_j}^{-1}, \quad 0 \leq l \leq k-1, -1 \leq m \leq k-1.$$

Then every $u_h \in X_h$ and every $\varphi_h \in Y_h$ have unique representations of the form

$$u_h = \sum_{j=1}^{N_\tau} \chi_{J_j}(t) \Big\{ \sum_{\mu=-1}^{k-1} q_{\mu,j}(t) v_{\mu,j}(x) \Big\},$$

$$\varphi_h = \sum_{j=1}^{N_\tau} \lambda_j^{(\theta)}(t) w_j(x) + \sum_{j=1}^{N_\tau} \psi_j(t) \chi_{J_j}(t) \Big\{ \sum_{\mu=0}^{k-1} p_{\mu,j}(t) w_{\mu,j}(x) \Big\}$$

with $v_{\mu,j}, w_{\mu,j}, w_j \in V_j, 1 \leq j \leq N_\tau$. Consider a fixed $j \in \{2, ..., N_\tau\}$ and insert $\varphi_h = \psi_j(t)\chi_{J_j}(t) p_{\mu,j}(t) w_{\mu,j}(x), 0 \leq \mu \leq k-1$, as a test function in (4.11). We then get

$$\begin{aligned}
&\langle F_h(u_h), \varphi_h \rangle_{Y_h} \\
&\quad = \int_{t_j}^{t_{j+1}} \int_\Omega \{\partial_t u_h \varphi_h + \underline{a}(x, u_h, \nabla u_h)\nabla \varphi_h - b(x, u_h, \nabla u_h)\varphi_h\} dx dt \\
&\quad = \int_\Omega v_{\mu,j}(x) w_{\mu,j}(x) dx \\
&\qquad + \int_{t_j}^{t_{j+1}} \{ \int_\Omega [\underline{a}(x, u_h, \nabla u_h)\nabla w_{\mu,j} - b(x, u_h, \nabla u_h) w_{\mu,j}] dx\} \psi_j(t) p_{\mu,j}(t) dt.
\end{aligned}$$

Inserting $\varphi_h = \lambda_j^{(\theta)}(t) w_j(x)$ as a test function in (4.11), we obtain on the other hand

$$\begin{aligned}
&\langle F_h(u_h), \varphi_h \rangle_{Y_h} \\
&\quad = \int_\Omega \{v_{-1,j}(x) - u_h(x, t_j - 0)\} w_j(x) dx \\
&\qquad + \int_{t_{j-1}}^{t_{j+1}} \{ \int_\Omega [\partial_t u_h w_j + \underline{a}(x, u_h, \nabla u_h)\nabla w_j - b(x, u_h, \nabla u_h) w_j] dx\} \lambda_j^{(\theta)}(t) dt.
\end{aligned}$$

With the obvious modifications these expressions also hold for $j = 1$. Hence, the coefficients $v_{0,j}, ..., v_{k-1,j}$ are functions of the coefficient $v_{-1,j}$, and the latter is determined by the values of u_h on the previous time interval. Thus problem (1.2) amounts in a $(k+1)$-stage implicit Runge-Kutta scheme. A lengthy but straightforward calculation shows that, for linear problems, $k \in \{1,2\}$, and $\theta = \frac{1}{2}$, this scheme corresponds to the $(k+1)$-st diagonal Padé approximation. In particular, the time discretization then is of order $2k+2$ and A-stable.

Remark 4.1. When we write problem (1.2) in the form (4.11) it strongly resembles the discontinuous Galerkin method (cf. e.g. [7], [8]). In the discontinuous Galerkin method, however, the test and trial spaces are identical, and both consist of discontinuous in time, piecewise polynomials of degree k. In particular, the case $k = 0$ corresponds to the implicit Euler scheme. Due to the discontinuities at the intermediate times $t_2, ..., t_{N_\tau - 1}$ the discontinuous Galerkin method is non-conforming with respect to both the standard weak formulation of problem (4.1) and the formulation (4.2). This complicates its analysis within the framework of Secction 2. This difficulty is overcome in [10]. A different analysis of the discontinuous Galerkin method is given in [7], [8].

Definition of $R_h, \widetilde{F}_h$, and $\widetilde{Y}_h$. In order to put the discretization in the framework of Section 2, we assume that Y_h contains the space θ_τ defined in (3.6). This is equivalent to assuming that the space discretization at least consists of linear elements, i.e. $V_j \supset S^{1,0}_{j,0}, 1 \le j \le N_\tau$. As restriction operator R_h we use the operator I_τ defined in (3.9). For the construction of $\widetilde{F}_h$ and $\widetilde{Y}_h$ we define integers μ, ν and approximations $\underline{a}_h$ of $\underline{a}$ and b_h of b as follows:

$$
\begin{aligned}
\underline{a}_h(x, u_h, \nabla u_h) &:= \begin{cases} \underline{a}(x, u_h, \nabla u_h), & \text{if } \underline{a}(x, v_h, \nabla v_h) \\ & \in S^{\mu,-1}_\tau(S^{\mu,-1}_{h(\tau)})\, \forall v_h \in X_h, \\ \displaystyle\sum_{Q \in \mathcal{P}_\tau} \pi_{1,Q}\underline{a}(x, u_h, \nabla u_h), \mu := 1, & \text{otherwise}, \end{cases} \\
b_h(x, u_h, \nabla u_h) &:= \begin{cases} b(x, u_h, \nabla u_h), & \text{if } b(x, v_h, \nabla v_h) \\ & \in S^{\nu,-1}_\tau(S^{\nu,-1}_{h(\tau)})\, \forall v_h \in X_h, \\ \displaystyle\sum_{Q \in \mathcal{P}_\tau} \pi_{0,Q} b(x, u_h, \nabla u_h), \nu := 0, & \text{otherwise}. \end{cases}
\end{aligned} \tag{4.12}
$$

Here, $u_h \in X_h$ is arbitrary and $\pi_{0,Q}$ and $\pi_{1,Q}$ denote the $L^2(Q)$-projections onto the spaces of polynomials of degree at most 0 and 1 in the variables x and t, respectively. Now, $\widetilde{F}_h$ is defined in the same way as F with $\underline{a}$ and b replaced by $\underline{a}_h$ and b_h, respectively, and

$$
\begin{aligned}
\widetilde{Y}_h := \operatorname{span}\{&\psi_j\psi_K v, \psi_j\psi_E P_E\sigma, \psi_j\psi_K P_j w : 1 \le j \le N_\tau, K \in \mathcal{T}_j, E \in \mathcal{E}_j, \\
& v \in \widetilde{\mathbb{P}}_{m|K\times J_j}, \sigma \in \widetilde{\mathbb{P}}_{m|E\times J_j}, w \in \widetilde{\mathbb{P}}_{m|K\times\{t_j\}}\}.
\end{aligned} \tag{4.13}
$$

Here, $m := \max\{\mu - 1, \nu\}$ and $\widetilde{\mathbb{P}}_m$ denotes the space of polynomials of degree at most m in the variables x and t.

The estimators. Given $Q = K \times J_j \in \mathcal{P}_\tau$, we recall the abbreviation (3.16) and set

$$
\begin{aligned}
\varepsilon_{Q,\pi} &:= (h_K^2 + \tau_j)\|\nabla \cdot [\underline{a}(.,u_h,\nabla u_h) - \underline{a}_h(.,u_h,\nabla u_h)] \\
&\qquad\qquad + [b(.,u_h,\nabla u_h) - b_h(.,u_h,\nabla u_h)]\|_{L^p_\pi(Q)} \\
&\quad + h_K^{\frac{1}{\pi}-1}(h_K^2 + \tau_j)\|\underline{n}_E \cdot [\underline{a}(.,u_h,\nabla u_h) - \underline{a}_h(.,u_h,\nabla u_h)]_E\|_{L^p_\pi(\partial Q_L)}, \\
\eta_{Q,\pi} &:= (h_K^2 + \tau_j)\|\partial_t u_h - \nabla \cdot \underline{a}_h(.,u_h,\nabla u_h) - b_h(.,u_h,\nabla u_h)\|_{L^p_\pi(Q)} \\
&\quad + h_K^{\frac{1}{\pi}-1}(h_K^2 + \tau_j)\|\underline{n}_E \cdot [\underline{a}_h(.,u_h,\nabla u_h)]_E\|_{L^p_\pi(\partial Q_L)} \\
&\quad + \tau_j^{\frac{1}{p}-1}(h_K^2 + \tau_j)\|u_h(.,t_j+0) - u_h(.,t_j-0)\|_{L^\pi(K)}.
\end{aligned}
\tag{4.14}
$$

The quantity $\varepsilon_{Q,\pi}$ obviously measures the quality of the approximation of $\underline{a}$ and b by $\underline{a}_h$ and b_h respectively, and can be estimated explicitly. Below, we will show that it yields upper bounds on the second terms on the right-hand sides of estimates (2.3) and (2.4). Note that in our second example

$$
\varepsilon_{Q,\pi} \le h_K^2 \|f - \pi_{0,Q} f\|_{L^p_\pi(Q)} + h_K^3 \|\nabla u_h\|_{L^p_\pi(Q)},
$$

if $k = 0$ and $V_j = S^{1,0}_{j,0}, 1 \le j \le N_\tau$.

Estimation of $\|(Id_{Y_+} - R_h)^* \widetilde{F}_h(u_h)\|_{Y_+^*}$ **and** $\|(Id_{Y_+} - R_h)^*[F(u_h) - \widetilde{F}_h(u_h)]\|_{Y_+^*}$**.** Next, we will derive upper bounds for the first and second terms on the right-hand side of inequality (2.3). Recalling equation (4.11) and using, for the space variables, integration by parts elementwise, we obtain for all $\varphi \in Y$

$$
\begin{aligned}
&\langle F(u_h), \varphi\rangle_Y \\
&\quad = \sum_{j=1}^{N_\tau} \Big\{ \sum_{K \in \mathcal{T}_j} \{ \int_K [u_h(x,t_j+0) - u_h(x,t_j-0)]\varphi(x,t_j)dx \\
&\qquad\qquad + \int_{t_j}^{t_{j+1}} \int_K [\partial_t u_h - \nabla \cdot \underline{a}(x,u_h,\nabla u_h) - b(x,u_h,\nabla u_h)]\varphi dx dt\} \\
&\qquad + \sum_{E \in \mathcal{E}_j} \int_{t_j}^{t_{j+1}} \int_E \underline{n}_E \cdot [\underline{a}(x,u_h,\nabla u_h)]_E \varphi ds dt \Big\}
\end{aligned}
\tag{4.15}
$$

and

$$
\begin{aligned}
&\langle \widetilde{F}_h(u_h), \varphi\rangle_Y \\
&\quad = \sum_{j=1}^{N_\tau} \Big\{ \sum_{K \in \mathcal{T}_j} \{ \int_K [u_h(x,t_j+0) - u_h(x,t_j-0)]\varphi(x,t_j)dx \\
&\qquad\qquad + \int_{t_j}^{t_{j+1}} \int_K [\partial_t u_h - \nabla \cdot \underline{a}_h(x,u_h,\nabla u_h) - b_h(x,u_h,\nabla u_h)]\varphi dx dt\} \\
&\qquad + \sum_{E \in \mathcal{E}_j} \int_{t_j}^{t_{j+1}} \int_E \underline{n}_E \cdot [\underline{a}_h(x,u_h,\nabla u_h)]_E \varphi ds dt \Big\}.
\end{aligned}
\tag{4.16}
$$

Let $\varphi \in Y_+ = W^{p'}_{\pi'}(\Omega \times (0,T))$ be arbitrary. From Lemma 3.4 we obtain for $Q = K \times I_j \in P_\tau$, $1 \leq j \leq N_\tau$, $K \in \mathcal{T}_j$, and $E \subset \partial K \backslash \Gamma$ the estimates

$$\begin{aligned}
&\int_K [u_h(x, t_j + 0) - u_h(x, t_j - 0)][\varphi(x, t_j) - I_\tau \varphi(x, t_j)]dx \\
&\quad \leq \| u_h(., t_j + 0) - u_h(., t_j - 0) \|_{L^\pi(K)} \| \varphi - I_\tau \varphi \|_{L^{p'}_{\pi'}(\partial Q_B)} \\
&\quad \preceq \tau_j^{\frac{1}{p} - 1} (h_K^2 + \tau_j) \| u_h(., t_j + 0) - u_h(., t_j - 0) \|_{L^\pi(K)} \| \varphi \|_{W^{p'}_{\pi'}(K \times U(t_j))}, \\
&\int_{t_j}^{t_{j+1}} \int_K [\partial_t u_h - \nabla \cdot \underline{a}_h(x, u_h, \nabla u_h) - b_h(x, u_h, \nabla u_h)][\varphi - I_\tau \varphi] dx dt \\
&\quad \leq \| \partial_t u_h - \nabla \cdot \underline{a}_h(., u_h, \nabla u_h) - b_h(., u_h, \nabla u_h) \|_{L^p_\pi(Q)} \| \varphi - I_\tau \varphi \|_{L^{p'}_{\pi'}(Q)} \\
&\quad \preceq (h_K^2 + \tau_j) \| \partial_t u_h - \nabla \cdot \underline{a}_h(., u_h, \nabla u_h) - b_h(., u_h, \nabla u_h) \|_{L^p_\pi(Q)} \| \varphi \|_{W^{p'}_{\pi'}(U(Q))}, \\
&\int_{t_j}^{t_{j+1}} \int_E \underline{n}_E \cdot [\underline{a}_h(x, u_h, \nabla u_h)]_E [\varphi - I_\tau \varphi] ds dt \\
&\quad \leq \| \underline{n}_E \cdot [\underline{a}_h(., u_h, \nabla u_h)]_E \|_{L^p_\pi(\partial Q_L)} \| \varphi - I_\tau \varphi \|_{L^{p'}_{\pi'}(\partial Q_L)} \\
&\quad \preceq h_E^{\frac{1}{\pi} - 1} (h_K^2 + \tau_j) \| \underline{n}_E \cdot [\underline{a}_h(., u_h, \nabla u_h)]_E \|_{L^p_\pi(\partial Q_L)} \| \varphi \|_{W^{p'}_{\pi'}(U(\partial Q_L))}.
\end{aligned}$$

Inserting these estimates in (4.16), using Hölder's inequality for finite sums, and recalling the definition (4.14), we conclude that

$$\begin{aligned}
&\langle \widetilde{F}_h(u_h), \varphi - I_\tau \varphi \rangle \\
&\preceq \Big\{ \sum_{j=1}^{N_\tau} \{ \sum_{K \in \mathcal{T}_j} \eta_{Q,\pi}^\pi \}^{p/\pi} \Big\}^{1/p} \Big\{ \sum_{j=1}^{N_\tau} \{ \sum_{K \in \mathcal{T}_j} \| \varphi \|^{\pi'}_{W^{p'}_{\pi'}(U(K \times J_j))} \}^{p'/\pi'} \Big\}^{1/p'}.
\end{aligned}$$

Assume that $p' \leq \pi'$ or, equivalently, $p \geq \pi$. Then Jensen's inequality implies that

$$\Big\{ \sum_{j=1}^{N_\tau} \{ \sum_{K \in \mathcal{T}_j} \| \varphi \|^{\pi'}_{W^{p'}_{\pi'}(U(K \times J_j))} \}^{p'/\pi'} \Big\}^{1/p'} \preceq \| \varphi \|_{W^{p'}_{\pi'}(\Omega \times (0,T))}.$$

Using the abbreviation

$$\eta := \Big\{ \sum_{j=1}^{N_\tau} \{ \sum_{K \in \mathcal{T}_j} \eta_{Q,\pi}^\pi \}^{p/\pi} \Big\}^{1/p}, \tag{4.17}$$

we have thus shown that

$$\|(Id_{Y_+} - R_h)^* \widetilde{F}_h(u_h)\|_{Y_+^*} \preceq \eta. \tag{4.18}$$

Replacing $\underline{a}_h$ and b_h by $\underline{a}_h - a$ and $b_h - b$, respectively, we conclude with the same arguments that

$$\|(Id_{Y_+} - R_h)^* [\widetilde{F}_h(u_h) - F(u_h)]\|_{Y_+^*} \preceq \varepsilon, \tag{4.19}$$

where

$$\varepsilon := \Big\{ \sum_{j=1}^{N_\tau} \{ \sum_{K \in \mathcal{T}_j} \varepsilon_{Q,\pi}^{\pi} \}^{p/\pi} \Big\}^{1/p}. \tag{4.20}$$

Estimation of $\|\widetilde{F}_h(u_h)\|_{\widetilde{Y}_h^*}$ and $\|\widetilde{F}_h(u_h) - F(u_h)\|_{\widetilde{Y}_h^*}$. Now, we will bound the terms in inequality (2.4). Given a subset Q of $\Omega \times (0,T)$, we set, for abbreviation, $\widetilde{Y}_{h|Q} := \{\varphi \in \widetilde{Y}_h : \mathrm{supp}\varphi \subset Q\}$. In order to bound the second term on the right-hand side of estimate (2.4), we conclude from the shape regularity of the partitions and a standard scaling argument that the estimate

$$\|\varphi\|_{L^{p'}_{\pi'}(K' \times J')} + h_K^{\frac{1}{\pi'}} \|\varphi\|_{L^{p'}_{\pi'}(\partial K' \times J')} \preceq h_K^2 \|\varphi\|_{L^{p'}(J', W^{2,\pi'}(K'))}$$

holds for all $\varphi \in \widetilde{Y}_{h|U(Q)}, Q = K \times J_j \in \mathcal{P}_\tau, J' \subset U(J_j), K' \subset U(K)$. Combining this with equations (4.14) – (4.16), and using Hölder's inequality, we obtain the following estimate for all $Q \in \mathcal{P}_\tau$:

$$\|F(u_h) - \widetilde{F}_h(u_h)\|_{\widetilde{Y}^*_{h|U(Q)}} \preceq \sum_{Q' \subset U(Q)} \varepsilon_{Q',\pi}, \tag{4.21}$$

where $\widetilde{Y}_h$ is equipped with the norm of Y_+.

In order to derive lower bounds for the left-hand side of estimate (2.4), consider an arbitrary $Q = K \times J_j \in \mathcal{P}_\tau$. From Lemma 3.5 with $V_{\hat{Q}} = \widetilde{\mathbb{P}}_m$ and equation (4.14) we then obtain

(4.22)
$$\begin{aligned}
&\|\partial_t u_h - \nabla \cdot \underline{a}_h(., u_h, \nabla u_h) - b_h(., u_h, \nabla u_h)\|_{L^p_\pi(Q)} \\
&\quad \preceq \sup_{v \in \widetilde{\mathbb{P}}_m} \|v\|^{-1}_{L^{p'}_{\pi'}(Q)} \int_Q [\partial_t u_h - \nabla \cdot \underline{a}_h(., u_h, \nabla u_h) - b_h(., u_h, \nabla u_h)] \psi_j \psi_K v \\
&\quad = \sup_{v \in \widetilde{\mathbb{P}}_m} \|v\|^{-1}_{L^{p'}_{\pi'}(Q)} \langle \widetilde{F}_h(u_h), \psi_j \psi_K v \rangle_{Y_+} \\
&\quad \le \|\widetilde{F}_h(u_h)\|_{\widetilde{Y}^*_{h|Q}} \sup_{v \in \widetilde{\mathbb{P}}_m} \|v\|^{-1}_{L^p_{\pi'}(Q)} \|\psi_j \psi_K v\|_{Y_+} \\
&\quad \preceq \{h_K^{-2} + \tau_j^{-1}\} \|\widetilde{F}_h(u_h)\|_{\widetilde{Y}^*_{h|Q}}.
\end{aligned}$$

Estimate (4.22), equation (4.16) and Lemma 3.5 with $V_{\partial\hat{Q}_L} = \widetilde{\mathbb{P}}_{m|\hat{E}\times\hat{J}}$ yield for all $E \subset \partial K\backslash\Gamma$

(4.23)

$$\begin{aligned}
&\|\underline{n}_E \cdot [\underline{a}_h(.,u_h,\nabla u_h)]_E\|_{L^p_\pi(\partial Q_L)} \\
&\quad\preceq \sup_{\sigma\in\widetilde{\mathbb{P}}_{m|\partial Q_L}} \|\sigma\|^{-1}_{L^{p'}_{\pi'}(\partial Q_L)} \int_{\partial Q_L} \underline{n}_E \cdot [\underline{a}_h(.,u_h,\nabla u_h)]_E \psi_j\psi_E P_E\sigma \\
&\quad\leq \sup_{\sigma\in\widetilde{\mathbb{P}}_{m|\partial Q_L}} \|\sigma\|^{-1}_{L^{p'}_{\pi'}(\partial Q_L)} \Big\{\langle \widetilde{F}_h(u_h), \psi_j\psi_E P_E\sigma\rangle_{Y_+} \\
&\qquad - \sum_{K'\subset\omega_E} \int_{J_j}\int_{K'} [\partial_t u_h - \nabla\cdot\underline{a}_h(.,u_h,\nabla u_h) - b_h(.,u_h,\nabla u_h)]\psi_j\psi_E P_E\sigma dxdt\Big\} \\
&\quad\leq \sup_{\sigma\in\widetilde{\mathbb{P}}_{m|\partial Q_L}} \|\sigma\|^{-1}_{L^{p'}_{\pi'}(\partial Q_L)} \Big\{\|\widetilde{F}_h(u_h)\|_{\widetilde{Y}_{h|\omega_E\times J_j}} \|\psi_j\psi_E P_E\sigma\|_{Y_+} \\
&\qquad + \sum_{K'\subset\omega_E} \|\partial_t u_h - \nabla\cdot\underline{a}_h(.,u_h,\nabla u_h) - b_h(.,u_h,\nabla u_h)\|_{L^p_\pi(K'\times J_j)} \\
&\qquad\qquad \cdot \|\psi_j\psi_E P_E\sigma\|_{L^{p'}_{\pi'}(K'\times J_j)}\Big\} \\
&\quad\preceq h_E^{\frac{1}{\pi'}}\{h_E^{-2} + \tau_j^{-1}\}\|\widetilde{F}_h(u_h)\|_{\widetilde{Y}^*_{h|\omega_E\times J_j}}.
\end{aligned}$$

From estimate (4.22), equation (4.16) and Lemma 3.5 with $V_{\partial\hat{Q}_B} = \widetilde{\mathbb{P}}_{m|\hat{K}\times\{0\}}$ we further conclude that

(4.24)

$$\begin{aligned}
&\|u_h(.,t_j+0) - u_h(.,t_j-0)\|_{L^\pi(K)} \\
&\quad\preceq \sup_{w\in\widetilde{\mathbb{P}}_{m|K\times\{t_j\}}} \|w\|^{-1}_{L^{\pi'}(K)} \int_K [u_h(.,t_j+0) - u_h(.,t_j-0)]\lambda_j\psi_K P_j w dx \\
&\quad= \sup_{w\in\widetilde{\mathbb{P}}_{m|K\times\{t_j\}}} \|w\|^{-1}_{L^{\pi'}(K)} \Big\{< \widetilde{F}_h(u_h), \lambda_j\psi_K P_j w >_{Y_+} \\
&\qquad - \sum_{J\subset U(t_j)} \int_J\int_K [\partial_t u_h - \nabla\cdot\underline{a}_h(.,u_h,\nabla u_h) - b_h(.,u_h,\nabla u_h)]\lambda_j\psi_K P_j w dxdt\Big\} \\
&\quad\leq \sup_{w\in\widetilde{\mathbb{P}}_{m|K\times\{t_j\}}} \|w\|^{-1}_{L^{\pi'}(K)} \Big\{\|\widetilde{F}_h(u_h)\|_{\widetilde{Y}^*_{h|K\times U(t_j)}} \|\lambda_j\psi_K P_j w\|_{Y_+} \\
&\qquad + \sum_{J\subset U(t_j)} \|\partial_t u_h - \nabla\cdot\underline{a}_h(.,u_h,\nabla u_h) - b_h(.,u_h,\nabla u_h)\|_{L^p_\pi(K\times J)} \\
&\qquad\qquad \cdot \|\lambda_j\psi_K P_j w\|_{L^{p'}_{\pi'}(K\times J)}\Big\} \\
&\quad\preceq \tau_j^{\frac{1}{p'}}\{h_K^{-2} + \tau_j^{-1}\}\|\widetilde{F}_h(u_h)\|_{\widetilde{Y}^*_{h|K\times U(t_j)}}.
\end{aligned}$$

Estimates (4.22) – (4.24) and the definition (4.14) yield

$$\eta_{Q,\pi} \preceq \{1 + \tau_j h_K^{-2} + \tau_j^{-1}h_K^2\}\|\widetilde{F}_h(u_h)\|_{\widetilde{Y}^*_{h|U(Q)}}, \tag{4.25}$$

where $\widetilde{Y}_h$ is endowed with the norm of Y_+.

A posteriori error estimates. Combining estimates (4.18), (4.19), (4.21), and (4.25) with Propositions 2.1, 2.2 and recalling that $\|R_h^*[F(u_h) - F_h(u_h)]\|_{Y_+^*} = 0$, we obtain the following result.

Proposition 4.2. *Let u be a regular solution of problem (4.1) in the sense of Proposition 2.1 and definition (4.2), and let u_h be an approximation of u in the sense of Proposition 2.2 and definition (4.5). Suppose that $p \geq \pi$. Then the following a posteriori error estimates hold:*

$$\|u - u_h\|_{L_\rho^r(\Omega\times(0,T))} \preceq \{\eta + \varepsilon + \|R_h^* F_h(u_h)\|_{Y_+^*}\}$$

and

$$\eta_{Q,\pi} \preceq \{1 + \tau_j^{-1} h_K^2 + \tau_j h_K^{-2}\}\{\|u - u_h\|_{L_\rho^r(U(Q))} + \sum_{Q'\subset U(Q)} \varepsilon_{Q',\pi}\}$$
$$\forall Q = K \times J_j \in \mathcal{P}_\tau.$$

The quantities $\varepsilon_{Q,\pi}$, $\eta_{Q,\pi}$, η, and ε are given by equations (4.14), (4.17), and (4.20).

Remark 4.3. The local lower bounds for $\|u - u_h\|_{L_\rho^r(\Omega\times(0,T))}$ can be combined in the standard way to the global lower bound

$$\eta \preceq \max_{1\leq j\leq N_\tau} \max_{K\in\mathcal{T}_j} \{1 + \tau_j^{-1} h_K^2 + \tau_j h_K^{-2}\} \cdot \{\|u - u_h\|_{L_\rho^r(\Omega\times(0,T))} + \varepsilon\}.$$

The factor $1 + \tau_j^{-1} h_K^2 + \tau_j h_K^{-2}$ in this estimate and the second one of Proposition 4.2 reflects the fact that the differential operator is of 2nd order with respect to the space variables but only of 1st order with respect to the time variable.

Remark 4.4. If $p < \pi$ one may still obtain upper bounds on the error. Since, in this case, Jensen's inequality cannot be used in estimating $\langle \widetilde{F}_h(u_h), \varphi - I_\tau\varphi\rangle_{Y_+}$ and $\langle \widetilde{F}_h(u_h) - F(u_h), \varphi - I_\tau\varphi\rangle_{Y_+}$, one must now proceed as follows:

1. Bound the space-integrals by using Hölder's inequality and Lemma 3.2.
2. On each time-level add all contributions to that level and apply Hölder's inequality for finite sums.
3. Bound the remaining time-integrals by using Hölder's inequality and Lemma 3.3.
4. Add all time-levels and use Hölder's inequality for finite sums.

Remark 4.5. One can establish similar estimates for the $L^r(0,T;W_0^{1,\rho}(\Omega))$-norm of the error (cf. Proposition 4.1 in [10]). To this end one must replace Y_+, $\eta_{Q,\pi}$, $\varepsilon_{Q,\pi}$, and $1 + \tau_j^{-1} h_K^2 + \tau_j h_K^{-2}$ by $Y, h_K^{-1}\eta_{Q,\pi}, h_K^{-1}\varepsilon_{Q,\pi}$ and $\sigma_{n,\pi}(h_K) + \tau_j^{-1} h_K^2 \sigma_{n,\pi}(h_K) + \tau_j h_K^{-2}$, resp., where

$$\sigma_{n,\pi}(h) = \begin{cases} 1 & \text{if } \pi < n, \\ |\ln h| & \text{if } \pi = n, \\ h^{n/\pi - 1} & \text{if } \pi > n. \end{cases}$$

Moreover, the local lower bounds may be combined to global lower bounds at the expense of an additional factor $\max\{h_K^{-1} : K \in \mathcal{T}_j, 1 \leq j \leq N_\tau\}$. This factor and the term $\sigma_{n,\pi}(h_K)$ are due to the non-local nature of the $W^{-1,\pi'}(\Omega)$-norm, which allows only for weaker Poincaré and inverse inequalities (cf. Lemma 3.5 and Remarks 3.6, 3.7 in [10]).

References

[1] R. A. Adams: Sobolev spaces. Academic Press, New York, 1975 MR **56:**9247

[2] H. Amann: Nonhomogeneous linear and quasilinear elliptic and parabolic boundary value problems. Function Spaces, Differential Operators and Nonlinear Analysis (Friedrichroda, 1992), Teubner, Stuttgart, 1993, pp. 9–126. MR **94m:**35153

[3] J. Baranger and H. El Amri: Estimateurs a posteriori d'erreur pour le calcul adaptif d'écoulements quasi-Newtoniens. RAIRO Modél Math. Anal. Numer. 25, 31 – 48 (1991) MR **91m:**76070

[4] P. G. Ciarlet: The finite element method for elliptic problems. North Holland, Amsterdam, 1978 MR **58:**25001

[5] P. Clément: Approximation by finite element functions using local regularization. RAIRO Anal. Numér. 9 R-2, 77 – 84 (1975) MR **53:**4569

[6] R. Dautray and J.-L. Lions: Mathematical analysis and numerical methods for science and technology. Vol. 5, Springer, Berlin - Heidelberg - New York, 1992 MR **92k:**00006

[7] K. Eriksson and C. Johnson: Adaptive finite element methods for parabolic problems IV. Nonlinear problems. SIAM J. Numer Anal. **32** (1995), 1729–1749. MR **96i:**65081

[8] ———, Adaptive finite element methods for parabolic problems V. Long-time integration. SIAM J. Numer. Anal. **32** (1995), 1750–1763 MR **96i:**65082

[9] R. Verfürth: A posteriori error estimates for nonlinear problems. Finite element discretizations of elliptic equations. Math. Comput. 62 (206), 445 – 475 (1994) MR **94j:**65136

[10] ———, A posteriori error estimates for nonlinear problems. Finite element discretizations of parabolic equations. Report 180, Ruhr-Universität Bochum (1995) http://www.num1.ruhr-uni-bochum.de/rv/preprints.html

Fakultät für Mathematik, Ruhr-Universität Bochum, D-44780 Bochum, Germany

E-mail address: `rv@silly.num1.ruhr-uni-bochum.de`