A posteriori error estimates for non-linear parabolic equations

R. Verfürth

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

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

Date: December 2004

Summary: We consider space-time discretizations of non-linear parabolic equations. The temporal discretizations in particular cover the implicit Euler scheme and the mid-point rule. For linear equations they correspond to the well-known A-stable θ-schemes. The spatial discretizations consist of standard conforming finite element spaces that can vary from one time-level to the other. The spatial meshes may be locally refined, but must be isotropic. For these discretizations we derive a residual a posteriori error estimator which yields upper and lower bounds on the error. The ratio of upper and lower bounds does not depend on any mesh-size in space or time nor on any relation between both. In particular there is no restriction on the relative size of the temporal and spatial mesh-sizes.

Key words: a posteriori error estimates; non-linear parabolic equations; space-time finite elements

AMS Subject classification: 65N30, 65N15, 65J15

1. Introduction

We consider non-linear parabolic equations

$$\begin{aligned} \frac{\partial u}{\partial t} - \operatorname{div} \underline{a}(x,u,\nabla u) + b(x,u,\nabla u) &= 0 \text{ in } \Omega \times (0,T] \\ u &= 0 \text{ on } \Gamma \times (0,T] \\ u(\cdot,0) &= u_0 \text{ in } \Omega \end{aligned} \tag{1.1}$$

in a bounded space-time cylinder with a *convex two-dimensional* polygonal cross-section $\Omega \subset \mathbb{R}^2$ having a Lipschitz boundary Γ. The final time T is arbitrary, but kept fixed in what follows. The coefficients $\underline{a} : \Omega \times \mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}^2$ and $b : \Omega \times \mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}$ must be continuously differentiable with Lipschitz-continuous derivatives. They have to satisfy suitable growth conditions so that problem (1.1) admits an appropriate variational formulation (cf. Sections 2 and 4 for details). Some sample problems satisfying these conditions are given in Section 4.

We consider space-time discretizations of problem (1.1). The temporal discretizations in particular cover the implicit Euler scheme and the mid-point rule. For linear problems they correspond to the well-known A-stable θ-schemes. The spatial discretizations consist of standard conforming finite element spaces that can vary from one time-level to the other. The spatial meshes may be locally refined, but must be isotropic.

For these discretizations we derive a residual a posteriori error estimator. It consists of two contributions: a spatial error estimator and a temporal one.
The spatial contribution is a standard residual a posteriori error estimator for the non-linear elliptic equation arising from the time-discretization of problem (1.1). It consists of an element residual on a space-time cylinder $K \times [t_{n-1}, t_n]$ and of an edge residual on the lateral boundary $\partial K \times [t_{n-1}, t_n]$.
The evaluation of the temporal error estimator requires at each time-level the solution of a discrete Poisson problem. This term can be interpreted as an edge residual on the bottom $K \times \{t_{n-1}\}$ measured in an appropriate dual norm. The additional work of solving a supplementary discrete problem at each time-level is the price we have to pay for making computable this dual norm. This extra work is comparable to the one required by the now popular estimators that are based on the solution of suitable discrete adjoint problems [3].
We prove that the error estimator yields upper and lower bounds for the error measured in a suitable $L^r(0,T;W_0^{1,\rho}(\Omega))$-norm (cf. Section 2 for a definition of these spaces and their norms). The ratio of the upper and lower bounds does not depend on any mesh-size in space or time nor on any relation between these parameters.

The present results should be compared to our old results in [10]:

(1) Here, we consider standard time-discretizations which in particular cover the implicit Euler scheme and the midpoint-rule. In [10], we used non-standard time-discretizations which could be interpreted as implicit Runge-Kutta schemes and which covered the Crank-Nicolson scheme as method of lowest order.
(2) Here, the ratio of upper and lower error bounds is independent of any mesh-size in space and time and of any relation between both parameters. In [10], the ratio of the upper and lower error bounds is proportional to $1 + h^2\tau^{-1} + h^{-2}\tau$ where h and τ denote the local mesh-sizes in space and time respectively.
(3) The present analysis and the one in [10] both depart from an abstract non-linear equation $F(u) = 0$ with a continuously differentiable mapping $F : X \to Y^*$ between appropriate function spaces (Y^* denoting the dual of Y). Here, X carries a stronger topology than Y (cf. Section 2), in [10] these rôles are reversed.
(4) In [10] we do not have to solve additional discrete problems in order to evaluate the error estimator.

The article is organized as follows. In Section 2 we introduce the relevant function spaces and their norms. Section 3 gives the finite element discretization. Departing from the abstract error estimate of [9, Proposition 2.1] we prove in Section 4 that the error is equivalent to a residual which is defined in a suitable dual space. This residual is split into two contributions: a spatial residual and a temporal one. In Section 5 we derive upper and lower bounds for the spatial residual. The temporal residual is treated in Section 6. Combining these results we obtain in Section 7 a preliminary a posteriori error estimator. It yields upper and lower bounds on the error which are independent of the mesh-sizes in the sense described above. This preliminary error estimator, however, is not suited for practical computations since it incorporates the dual norm of a suitable residual. The residual itself is easy to evaluate, but the dual norm is not directly accessible. To overcome this difficulty we present in Section 8 some $W^{1,q}$-stability results for the Laplacian both in analytic and discrete form. These results require that the cross-section Ω is two-dimensional and convex. Based on these results we present in Section 9 the error estimator in its final form. The computation of the dual norm here is replaced by the evaluation of the solution of an auxiliary discrete Poisson problem.

2. Function spaces

For any bounded open subset ω of Ω with Lipschitz boundary γ we denote by $W^{k,p}(\omega), k \in \mathbb{N}, 1 \le p \le \infty$, $L^p(\omega) = W^{0,p}(\omega)$, and $L^p(\gamma)$ the usual Sobolev and Lebesgue spaces equipped with the standard norms (cf. [1], [6, Vol. 3, Chap. IV]):

$$\begin{aligned} \|u\|_{k,p;\omega} &= \left\{ \sum_{|\alpha| \le k} \int_\omega |D^\alpha u(x)|^p dx \right\}^{1,p}, p < \infty, \\ \|u\|_{k,\infty;\omega} &= \max_{|\alpha| \le k} \operatorname*{ess.sup}_{x \in \omega} |D^\alpha u(x)| \end{aligned}$$

and

$$\begin{aligned} \|u\|_{p;\gamma} &= \left\{ \int_\gamma |u(x)|^p ds(x) \right\}^{1,p}, p < \infty, \\ \|u\|_{\infty;\gamma} &= \operatorname*{ess.sup}_{x \in \gamma} |u(x)|. \end{aligned}$$

Here, $\alpha \in \mathbb{N}^2$ is a multi-index, $|\alpha| = \alpha_1 + \alpha_2$, and ds denotes the length element of the curve γ.

Set

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

and denote its dual space by

$$W^{-1,p'}(\Omega) = W_0^{1,p}(\Omega)^* \quad \text{for } 1 < p < \infty.$$

Here and in the sequel, p' denotes the dual exponent of p defined by $\frac{1}{p} + \frac{1}{p'} = 1$. The duality pairing between $W_0^{1,p}(\Omega)$ and $W^{-1,p'}(\Omega)$ will always be denoted by $\langle .,. \rangle$ where the relevant Lebesgue exponent p will be apparent from the context.

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

$$\begin{aligned} \|u\|_{L^p(a,b;V)} &= \left\{ \int_a^b \|u(\cdot,t)\|_V^p dt \right\}^{1/p}, p < \infty, \\ \|u\|_{L^\infty(a,b;V)} &= \operatorname*{ess.sup}_{t\in(a,b)} \|u(\cdot,t)\|_V \end{aligned}$$

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

$$W^p(a,b;V,W) = \left\{ u \in L^p(a,b;V) : \frac{\partial u}{\partial t} \in L^p(a,b;W) \right\}$$

equipped with the norm

$$\begin{aligned} \|u\|_{W^p(a,b;V,W)} &= \left\{ \int_a^b \|u(\cdot,t)\|_V^p dt + \int_a^b \|\frac{\partial u}{\partial t}(\cdot,t)\|_W^p dt \right\}^{1/p}, p < \infty \\ \|u\|_{W^\infty(a,b;V,W)} &= \operatorname*{ess.sup}_{t\in(a,b)} \max \left\{ \|u(\cdot,t)\|_V \ , \ \|\frac{\partial u}{\partial t}(\cdot,t)\|_W \right\}. \end{aligned}$$

Here the partial derivative $\frac{\partial u}{\partial t}$ must be interpreted in the distributional sense [6, loc.cit]. If $p > 1$, we know from [6, Vol. 5, Chap. XVIII, §1, Proposition 9] that for any $u \in W^p(a,b;V,W)$ the traces $u(\cdot,a)$ and $u(\cdot,b)$ are defined as elements of W.

A function u is called a weak solution of problem (1.1) if there are parameters $r, p, \rho, \pi \in (1,\infty)$ such that $u \in W^r(0,T;W_0^{1,\rho}(\Omega), W^{-1,\pi}(\Omega))$,

$$u(\cdot,0) = u_0 \quad \text{in } W^{-1,\pi}(\Omega) \tag{2.1}$$

and

$$\int_0^T \langle \frac{\partial u}{\partial t}(\cdot,t), v(\cdot,t)\rangle dt + \int_0^T \int_\Omega \{\underline{a}(x,u,\nabla u)\nabla v + b(x,u,\nabla u)v\}\, dxdt = 0 \qquad \forall v \in L^{p'}(0,T;W_0^{1,\pi'}(\Omega)) \tag{2.2}$$

(cf. [2]). Note that $W_0^{1,\rho}(\Omega)$ is continuously embedded in $W^{-1,\pi}(\Omega)$ for all ρ and π since $\Omega \subset \mathbb{R}^2$.

3. Finite element discretization

For the discretization we choose an integer $N \geq 1$ and intermediate times $0 = t_0 < t_1 < \ldots < t_N = T$ and set $\tau_n = t_n - t_{n-1}$, $1 \leq n \leq N$. With each intermediate time t_n, $0 \leq n \leq N$, we associate a partition $\mathcal{T}_{h,n}$ of Ω and a corresponding finite element space $X_{h,n}$. These have to satisfy the following conditions:

(1) *Affine equivalence:* every element $K \in \mathcal{T}_{h,n}$ is either a triangle or a parallelogram.

(2) *Admissibility:* any two elements are either disjoint or share a vertex or a complete edge.

(3) *Shape regularity:* for any element K the ratio of its diameter h_K to the diameter ρ_K of the largest inscribed ball is bounded uniformly with respect to all partitions $\mathcal{T}_{h,n}$ and to N.

(4) *Transition condition:* for $1 \leq n \leq N$ there is an affinely equivalent, admissible, and shape-regular partition $\widetilde{\mathcal{T}}_{h,n}$ such that it is a refinement of both $\mathcal{T}_{h,n}$ and $\mathcal{T}_{h,n-1}$ and such that

$$\sup_{1 \leq n \leq N} \sup_{K \in \widetilde{\mathcal{T}}_{h,n}} \sup_{K' \in \mathcal{T}_{h,n}; K \subset K'} \frac{h_{K'}}{h_K} < \infty.$$

(5) Each $X_{h,n}$ consists of continuous functions which vanish on Γ and which are piecewise polynomials, the degrees being bounded uniformly with respect to all partitions $\mathcal{T}_{h,n}$ and to N.

(6) Each $X_{h,n}$ contains the space of continuous, piecewise linear finite elements corresponding to $\mathcal{T}_{h,n}$.

Triangular and quadrilateral elements may be mixed. Condition (2) excludes hanging nodes. Condition (3) is a standard one and allows for locally refined meshes. However, it excludes anisotropic elements. Condition (4) is due to the simultaneous presence of finite element functions defined on different grids. Usually the partition $\mathcal{T}_{h,n}$ is obtained from $\mathcal{T}_{h,n-1}$ by a combination of refinement and of coarsening. In this case Condition (4) only restricts the coarsening. It must not be too abrupt nor too strong.

We choose a parameter $\theta \in [\frac{1}{2}, 1]$ and keep it fixed in what follows. Then the space-time discretization of problem (1.1) consists in finding $u_h^n \in X_{h,n}, 0 \leq n \leq N$, such that

$$u_h^0 = \pi_0 u_0 \tag{3.1}$$

and, for $n = 1, \cdots, N$, and all $v_h \in X_{h,n}$

$$\int_\Omega \frac{u_h^n - u_h^{n-1}}{\tau_n} v_h dx + \int_\Omega \left\{ \underline{a}(x, u_h^{n\theta}, \nabla u_h^{n\theta}) \nabla v_h + b(x, u_h^{n\theta}, \nabla u_h^{n\theta}) v_h \right\} dx = 0 \tag{3.2}$$

where

$$u_h^{n\theta} = \theta u_h^n + (1 - \theta) u_h^{n-1} \tag{3.3}$$

and π_0 denotes the L^2-projection onto $X_{h,0}$.

With every solution $(u_h^n)_{0\le n\le N}$ of problems (3.1) and (3.2) we associate two functions $u_{h\tau}$ and $\widetilde{u}_{h\tau}$. The function $u_{h\tau}$ is *piecewise affine* on the intervals $[t_{n-1}, t_n]$, $1 \le n \le N$, and equals u_h^n at time t_n. The function $\widetilde{u}_{h\tau}$ is *piecewise constant* on the intervals $(t_{n-1}, t_n]$, $1 \le n \le N$, and equals $u_h^{n\theta}$ on $(t_{n-1}, t_n]$. Since the function $t \to u_{h\tau}(.,t)$ is continuous and piecewise affine with values in $W_0^{1,\rho}(\Omega)$, it is differentiable in the distributional sense [6, Vol. 5, Chap. XVIII, §1] and its weak derivative satisfies

$$\frac{\partial u_{h\tau}}{\partial t} = \frac{u_h^n - u_h^{n-1}}{\tau_n} \quad \text{on } (t_{n-1}, t_n). \tag{3.4}$$

4. The equivalence of error and residual

As usual for non-linear problems, our a posteriori error estimates are based on the abstract error estimate of [9, Proposition 2.1]. For completeness we shortly recall this result. Given two Banach spaces X and Y with norms $\|\cdot\|_X$ and $\|\cdot\|_Y$ we denote by $\mathcal{L}(X,Y)$ the space of continuous linear mappings of X into Y and equip it with its standard norm

$$\|A\|_{\mathcal{L}(X,Y)} = \sup_{x\in X\setminus\{0\}} \frac{\|Ax\|_Y}{\|x\|_X}.$$

$ISOM(X,Y)$ denotes the space of all invertible $A \in \mathcal{L}(X,Y)$ with $A^{-1} \in \mathcal{L}(Y,X)$. Given a continuously differentiable map $F : X \to Y^*$ of X into the dual Y^* of Y we look for solutions of the non-linear equation

$$F(u) = 0. \tag{4.1}$$

4.1 Lemma. *[9, Proposition 2.1]. Let $u \in X$ be a solution of problem (4.1). Assume that u is regular, i.e. $DF(u) \in ISOM(X,Y^*)$ and that DF is locally Lipschitz continuous at u, i.e. there are constants $\gamma > 0$ and $R_0 > 0$ such that*

$$\|DF(v) - DF(w)\|_{\mathcal{L}(X,Y^*)} \le \gamma \|v - w\|_X$$

holds for all $v, w \in X$ with $\|v-u\|_X \le R_0$ and $\|w-u\|_X \le R_0$. Set

$$R = \min\left\{R_0 \ , \ \gamma^{-1}\|DF(u)^{-1}\|^{-1}_{\mathcal{L}(Y^*,X)} \ , \ 2\gamma^{-1}\|DF(u)\|_{\mathcal{L}(X,Y^*)}\right\}.$$

Then the following error estimate holds for all $v \in X$ with $\|v-u\| \le R$:

$$\begin{aligned}
&\frac{1}{2}\|DF(u)\|^{-1}_{\mathcal{L}(X,Y^*)}\|F(v)\|_{Y^*} \\
&\le \|v-u\|_X \\
&\le 2\|DF(u)^{-1}\|_{\mathcal{L}(Y^*,X)}\|F(v)\|_{Y^*}.
\end{aligned}$$

For abbreviation we define a function $G : L^r(0,T;W_0^{1,\rho}(\Omega)) \to L^p(0,T;W^{-1,\pi}(\Omega))$ by

$$\langle G(u), v\rangle = \int_\Omega \{\underline{a}(x,u,\nabla u)\nabla v + b(x,u,\nabla u)v\}\,dt \quad \text{a.e. in } (0,T). \tag{4.2}$$

Then the weak formulation (2.1), (2.2) of problem (1.1) fits into the framework of Lemma 4.1 with

$$\begin{aligned} X &= W^r(0,T;W_0^{1,\rho}(\Omega),W^{-1,\pi}(\Omega)), \\ Y &= W_0^{1,\pi'}(\Omega) \times L^{p'}(0,T;W_0^{1,\pi'}(\Omega)), \\ \langle F(u),(v_1,v_2)\rangle &= \begin{pmatrix} \langle u(\cdot,0)-u_0,v_1\rangle \\ \int_0^T \{\langle \frac{\partial u}{\partial t}, v_2\rangle + \langle G(u),v_2\rangle\}\,dt \end{pmatrix}. \end{aligned} \tag{4.3}$$

We therefore obtain:

4.2 Lemma. *Let u and $u_{h\tau}$ be solutions of problems (2.1), (2.2) and (3.1), (3.2). Assume that u and the function F of equation (4.3) satisfy the conditions of Lemma 4.1 and that*

$$\|u-u_{u\tau}\|_{W^r(0,T;W_0^{1,\rho}(\Omega),W^{-1,\pi}(\Omega))} \le R.$$

Then there are two constants c_* and c^* such that

$$\begin{aligned} &c_*\left\{\|u_0-u_h^0\|_{-1,\pi} + \|\frac{\partial u_{h\tau}}{\partial t} + G(u_{h\tau})\|_{L^p(0,T;W^{-1,\pi}(\Omega))}\right\} \\ &\le \|u-u_{h\tau}\|_{W^r(0,T;W_0^{1,\rho}(\Omega),W^{-1,\pi}(\Omega))} \\ &\le c^*\left\{\|u_0-u_h^0\|_{-1,\pi} + \|\frac{\partial u_{h\tau}}{\partial t} + G(u_{h\tau})\|_{L^p(0,T;W^{-1,\pi}(\Omega))}\right\}. \end{aligned}$$

4.3 Remark. Assume that G is locally Lipschitz continuous at the solution u of problems (2.1), (2.2), i.e., there are two constants $\gamma > 0$ and $R_0 > 0$ such that

$$\|DG(v)-DG(w)\|_{\mathcal{L}(L^r(0,T;W_0^{1,\rho}(\Omega)),L^p(0,T;W^{-1,\pi}(\Omega)))} \le \gamma\|v-w\|_{L^r(0,T;W_0^{1,\rho}(\Omega))}$$

holds for all $v,w \in L^r(0,T;W_0^{1,\rho}(\Omega))$ with $\|u-v\|_{L^r(0,T;W_0^{1,\rho}(\Omega))} \le R_0$ and $\|u-w\|_{L^r(0,T;W_0^{1,\rho}(\Omega))} \le R_0$. Then the function F of equation (4.3) satisfies the Lipschitz condition of Lemma 4.1 with the same constants γ and R_0.

Some examples of problems following into the present category are given by:

(1) The *heat equation with convection and non-linear diffusion coefficient:*

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

(2) The *non-stationary equation of prescribed mean curvature:*

$$\begin{aligned}
&\underline{a}(x,u,\nabla u) = [1+|\nabla u|^2]^{-1/2}\nabla u,\\
&b(x,u,\nabla u) = -f \in L^2(\Omega),\\
&\rho > 2,\ \pi = \frac{\rho}{2},\ r \geq 2\rho,\ p = \frac{r}{2}.
\end{aligned}$$

(3) The *non-stationary α-Laplacian:*

$$\begin{aligned}
&\underline{a}(x,u,\nabla u) = |\nabla u|^{\alpha-2}\nabla u \quad ,\alpha \geq 2,\\
&b(x,u,\nabla u) = -f \in L^{\alpha'}(\Omega),\\
&\rho = \alpha,\ \pi = \alpha',\ r > 6,\ p = \frac{r}{3}.
\end{aligned}$$

(4) The *non-stationary subsonic flow of an irrational, ideal, compressible gas:*

$$\begin{aligned}
&\underline{a}(x,u,\nabla u) = \left[1 - \frac{\gamma-1}{2}|\nabla u|^2\right]^{1/(\gamma-1)} \nabla u \quad ,\gamma > 1,\\
&b(x,u,\nabla u) = -f \in L^{\pi}(\Omega),\\
&\rho = \frac{2\gamma}{\gamma-1},\ \pi = \frac{2\gamma}{\gamma+1},\ r > 6,\ p = \frac{r}{3}.
\end{aligned}$$

Lemma 4.2 states that the error $u - u_{h\tau}$ and the residual $\frac{\partial u_{h\tau}}{\partial t} + G(u_{h\tau})$ are equivalent. In the following sections we will derive computable upper and lower bounds on the residual. To this end we split it into a spatial and a temporal contribution and set

$$\langle R_h(u_{h\tau}), v\rangle = \langle \frac{\partial u_{h\tau}}{\partial t} + G(\widetilde{u}_{h\tau}), v\rangle \tag{4.4}$$

and

$$\langle R_\tau(u_{h\tau}), v\rangle = \langle G(\widetilde{u}_{h\tau}) - G(u_{h\tau}), v\rangle. \tag{4.5}$$

Obviously we have

$$\frac{\partial u_{h\tau}}{\partial t} + G(u_{h\tau}) = R_h(u_{h\tau}) + R_\tau(u_{h\tau}). \tag{4.6}$$

Since $\frac{\partial u_{h\tau}}{\partial t} = \frac{u_h^n - u_h^{n-1}}{\tau_n}$ (cf. (3.4)) and $\widetilde{u}_{h\tau} = u_h^{n\theta}$ on each interval $(t_{n-1}, t_n]$, problem (3.2) is equivalent to

$$\langle R_h(u_{h\tau}), v_h\rangle = 0 \quad \forall v_h \in X_{hn}, 1 \leq n \leq N. \tag{4.7}$$

5. Estimation of the spatial residual

For the estimation of the spatial residual $R_h(u_{h\tau})$ we need some additional notations. We denote by $\widetilde{\mathcal{E}}_{h,n}, 1 \le n \le N$, the set of all edges of $\widetilde{\mathcal{T}}_{h,n}$. With each edge $E \in \widetilde{\mathcal{E}}_{h,n}$ we associate a unit vector $\underline{n}_E$ orthogonal to E such that it points to the outward of Ω if E is part of the boundary. For every edge E that is not contained in the boundary we denote by $[.]_E$ the jump across E in direction $\underline{n}_E$. The quantity $[.]_E$ of course depends on the orientation of $\underline{n}_E$, but quantities of the form $[\underline{n}_E \cdot .]_E$ are independent thereof. With each edge E that is not contained in the boundary, we associate the set ω_E which is the union of the two elements that share E. If E is a boundary-edge, ω_E simply is the unique element that has E as an edge.

For every n between 1 and N we denote by $\mathcal{N}_{h,n}$ the set of all element vertices in $\mathcal{T}_{h,n}$ that do not lie on the boundary. With every vertex $x \in \mathcal{N}_{h,n}$ we associate the nodal bases function λ_x which is uniquely defined by the properties

$$\lambda_{x|K} \in R_1(K)\ \forall K \in \mathcal{T}_{h,n},\ \lambda_x(y) = 0\ \forall y \in \mathcal{N}_{h,n} \backslash \{x\},\ \lambda_x(x) = 1.$$

Here, as usual, $R_k(K)$ denotes the set of all polynomials of total degree k, if K is a triangle, and of maximal degree k, if K is a quadrilateral. The support of a nodal bases function λ_x is denoted by ω_x and consists of all elements in $\mathcal{T}_{h,n}$ that share the vertex x. Denote by π_x the $L^2(\omega_x)$-projection onto R_1 defined by

$$\int_{\omega_x} \pi_x v w = \int_{\omega_x} v w \quad \forall\, w \in R_1.$$

Then the interpolation operator $I_{h,n}$ of Clément [5] corresponding to $\mathcal{T}_{h,n}$ is defined by

$$I_{h,n} v = \sum_{x \in \mathcal{N}_{h,n}} \lambda_x (\pi_x v)(x). \tag{5.1}$$

Due to condition (6) of Section 3 $I_{h,n}$ maps $L^1(\Omega)$ into a subspace of $X_{h,n}$.

5.1 Lemma. *[9, Lemma 3.1] The following error estimates hold for all $1 \le p \le \infty$, all $v \in W_0^{1,p}(\Omega)$, all $1 \le n \le N$, all $K \in \mathcal{T}_{h,n}$, and all $E \in \mathcal{E}_{h,n}$:*

$$\|v - I_{h,n} v\|_{0,p;K} \le c_1 h_K \|v\|_{1,p;\widetilde{\omega}_k},$$

$$\|v - I_{h,n} v\|_{p;E} \le c_2 h_E^{1-\frac{1}{p}} \|v\|_{1,p;\widetilde{\omega}_E}.$$

Here $\widetilde{\omega}_K$ and $\widetilde{\omega}_E$ consists of all elements in $\mathcal{T}_{h,n}$ that share at least a vertex with K or E, respectively. The constants c_1 and c_2 only depend on the ratios h_K/ρ_K.

For every element $K \in \widetilde{\mathcal{T}}_{h,n}$ and every edge $E \in \widetilde{\mathcal{E}}_{h,n}$ we denote by $\mathcal{N}_K$ and $\mathcal{N}_E$ the set of its vertices and set

$$\psi_K = \gamma_K \prod_{x \in \mathcal{N}_K} \lambda_x,$$

$$\psi_E = \gamma_E \prod_{x \in \mathcal{N}_E} \lambda_x.$$

The constants γ_k and γ_E are chosen such that ψ_K and ψ_E equal 1 at the barycentre of K and E, respectively. The support of ψ_K is contained in K and $\|\psi_K\|_{0,\infty;K} = 1$. Similarly, the support of ψ_E is contained in ω_E and $\|\psi_E\|_{0,\infty;\omega_E} = \|\psi_E\|_{\infty;E} = 1$.

5.2 Lemma. *[9, Lemma 3.3] The following estimates hold for all $1 \le p \le \infty$, all $k \in \mathbb{N}$, all $1 \le n \le N$, all $K \in \widetilde{\mathcal{T}}_{h,n}$, all $v \in R_k(K)$, all $E \in \widetilde{\mathcal{E}}_{h,n}$, and all $\sigma \in R_k(E)$:*

$$\begin{aligned}
\sup_{w \in R_k(K)} \frac{\int_K v\psi_K w}{\|w\|_{0,p';K}} &\ge c_3 \|v\|_{0,p;K}, \\
\|\psi_K v\|_{1,p;K} &\le c_4 h_K^{-1} \|v\|_{0,p;K}, \\
\sup_{\eta \in R_k(E)} \frac{\int_E \sigma\psi_E \eta}{\|\eta\|_{p';E}} &\ge c_5 \|\sigma\|_{p;E}, \\
\|\psi_E \sigma_E\|_{1,p;\omega_E} &\le c_6 h_E^{-1+\frac{1}{p}} \|\sigma\|_{p;E}, \\
\|\psi_E \sigma\|_{0,p;\omega_E} &\le c_7 h_E^{1/p} \|\sigma\|_{p;E}.
\end{aligned}$$

Here, a polynomial σ defined on an edge is continued in the canonical way to a polynomial defined on $\mathbb{R}^2$. The constants $c_3, \cdots, c_7$ only depend on the polynomial degree k and on the ratios h_K/ρ_K.

We choose an integer ℓ and denote for every n between 1 and N by $\underline{a}_{h,n}(x, u_h^{n\theta}, \nabla u_h^{n\theta})$ and $b_{h,n}(x, u_h^{n\theta}, \nabla u_h^{n\theta})$ the L^2-projections of $\underline{a}(x, u_h^{n\theta}, \nabla u_h^{n\theta})$ and $b(x, u_h^{n\theta}, \nabla u_h^{n\theta})$ onto discontinuous vector-fields respectively functions which are piecewise polynomials of degree ℓ on the elements of $\widetilde{\mathcal{T}}_{h,n}$. With this notation we define element residuals $R_K, K \in \widetilde{\mathcal{T}}_{h,n}, 1 \le n \le N$, by

$$R_K = \frac{u_h^n - u_h^{n-1}}{\tau_n} - \operatorname{div} \underline{a}_{h,n}(x, u_h^{n\theta}, \nabla u_h^{n\theta}) + b_{h,n}(x, u_h^{n\theta}, \nabla u_h^{n\theta}), \tag{5.2}$$

edge residuals $R_E, E \in \widetilde{\mathcal{E}}_{h,n}, 1 \le n \le N$, by

$$R_E = \begin{cases} [\underline{n}_E \cdot \underline{a}_{h,n}(x, u_h^{n\theta}, \nabla u_h^{n\theta})]_E & \text{if } E \not\subset \Gamma, \\ 0 & \text{if } E \subset \Gamma, \end{cases} \tag{5.3}$$

elementwise data errors $D_K, K \in \widetilde{\mathcal{T}}_{h,n}, 1 \le n \le N$, by

$$\begin{aligned} D_K =& \underline{a}(x, u_h^{n\theta}, \nabla u_h^{n\theta}) - \underline{a}_{h,n}(x, u_h^{n\theta}, \nabla u_h^{n\theta}) \\ &+ b(x, u_h^{n\theta}, \nabla u_h^{n\theta}) - b_{h,n}(x, u_h^{n\theta}, \nabla u_h^{n\theta}), \end{aligned} \tag{5.4}$$

and edgewise data errors $D_E, E \in \widetilde{\mathcal{E}}_{h,n}, 1 \le n \le N$, by

$$D_E = \begin{cases} [\underline{n}_E \cdot (\underline{a}(x, u_h^{n\theta}, \nabla u_h^{n\theta}) - \underline{a}_{h,n}(x, u_h^{n\theta}, \nabla u_h^{n\theta}))]_E & \text{if } E \not\subset \Gamma, \\ 0 & \text{if } E \subset \Gamma. \end{cases} \tag{5.5}$$

The choice of the parameter ℓ is influenced by the polynomial degree of the finite element spaces $X_{h,n}$ and by the smoothness of the coefficients $\underline{a}, b$. The simplest choice of course is $\ell = 0$.

With these preparations we are now ready to bound the spatial residual.

5.3 Lemma. *For every n between 1 and N define a spatial error indicator η_h^n by*

$$\eta_h^n = \left\{ \sum_{K \in \widetilde{\mathcal{T}}_{h,n}} h_K^\pi \|R_K\|_{0,\pi;K}^\pi + \sum_{E \in \widetilde{\mathcal{E}}_{h,n}} h_E \|R_E\|_{\pi;E}^\pi \right\}^{1/\pi} \tag{5.6}$$

and a spatial data error indicator Θ_h^n by

$$\Theta_h^n = \left\{ \sum_{K \in \widetilde{\mathcal{T}}_{h,n}} h_K^\pi \|D_K\|_{0,\pi;K}^\pi + \sum_{E \in \widetilde{\mathcal{E}}_{h,n}} h_E \|D_E\|_{\pi;E}^\pi \right\}^{1/\pi} . \tag{5.7}$$

Then there are functions $w_n \in W_0^{1,\pi'}(\Omega), 1 \le n \le N$, and constants $c^\dagger$ and $c_\dagger$ such that on each interval $(t_{n-1}, t_n], 1 \le n \le N$, the following estimates hold:

$$\|R_h(u_{h\tau})\|_{-1,\pi} \le c^\dagger \{\eta_h^n + \Theta_h^n\} \tag{5.8}$$

and

$$\begin{aligned} (\eta_h^n)^\pi \le& \langle R_h(u_{h\tau}), w_n \rangle + \Theta_h^n \|w_n\|_{1,\pi'}, \\ \|w_n\|_{1,\pi'} \le& c_\dagger (\eta_h^n)^{\pi-1}. \end{aligned} \tag{5.9}$$

The constants $c_\dagger$ and $c^\dagger$ depend on the ratios h_K/ρ_K. The constant $c^\dagger$ in addition depends on the ratios $h_{K'}/h_K$ in condition (4) of Section 3. The constant $c_\dagger$ in addition depends on the polynomial degree ℓ.

Proof. Choose an integer n between 1 and N and keep it fixed in what follows. Integration by parts on the elements of $\widetilde{\mathcal{T}}_{h,n}$ yields the following L^2-representation of the residual

$$\begin{aligned} \langle R_h(u_{h\tau}), v \rangle = & \sum_{K \in \widetilde{\mathcal{T}}_{h,n}} \int_K R_K v + \sum_{E \in \widetilde{\mathcal{E}}_{h,n}} \int_E R_E v \\ & + \sum_{K \in \widetilde{\mathcal{T}}_{h,n}} \int_K D_K v + \sum_{E \in \widetilde{\mathcal{E}}_{h,n}} \int_E D_E v. \end{aligned} \tag{5.10}$$

Lemma 5.1 and Hölder's inequality therefore imply for all $v \in W_0^{1,\pi'}(\Omega)$

$$\langle R_h(u_{h\tau}), v - I_{h,n} v \rangle \le c \|v\|_{1,\pi} \{\eta_h^n + \Theta_h^n\}. \tag{5.11}$$

The constant c only depends on the constants c_1 and c_2 of Lemma 5.1 and the ratios h_K/ρ_K.

Since $I_{h,n}$ maps $L^1(\Omega)$ into a subspace of $X_{h,n}$, equations (4.7) and (5.11) prove the upper bound (5.8).

From Lemma 5.2 we conclude that for every element $K \in \widetilde{\mathcal{T}}_{h,n}$ and every edge $E \in \widetilde{\mathcal{E}}_{h,n}$ there are polynomials v_K and σ_E such that

$$\begin{aligned}
\int_K R_K \psi_K v_K &= \|R_K\|_{0,\pi;K}^{\pi}, \\
\|v_K\|_{0,\pi';K} &\leq c_3^{-1} \|R_K\|_{0,\pi;K}^{\pi-1}, \\
\int_E R_E \psi_E \sigma_E &= \|R_E\|_{\pi;E}^{\pi}, \\
\|\sigma_E\|_{\pi';E} &\leq c_5^{-1} \|R_E\|_{\pi;E}^{\pi-1}.
\end{aligned}$$

Set

$$w_n = \gamma_1 \sum_{K \in \widetilde{\mathcal{T}}_{h,n}} h_K^{\pi} \psi_K v_K + \gamma_2 \sum_{E \in \widetilde{\mathcal{E}}_{h,n}} h_E \psi_E \sigma_E.$$

The constants γ_1 and γ_2 are arbitrary at present and will be determined below. The subsequent arguments are based on the following observations:

- the supports of the ψ_K are mutually disjoint,
- the support of a ψ_K intersects the support of at most four different ψ_E's,
- the support of a ψ_E intersects the support of at most two ψ_K's,
- the support of a ψ_E intersects the support of almost two other ψ_E's.

Since $(\pi - 1)\pi' = \pi$, Lemma 5.2 therefore yields

$$\begin{aligned}
\|w_n\|_{1,\pi'}^{\pi'} &= \sum_{K \in \widetilde{\mathcal{T}}_{h,n}} \|w_n\|_{1,\pi';K}^{\pi'} \\
&\leq 5^{\pi'-1} \gamma_1^{\pi'} \sum_{K \in \widetilde{\mathcal{T}}_{h,n}} h_K^{\pi\pi'} \|\psi_K v_K\|_{1,\pi';K}^{\pi'} \\
&\quad + 5^{\pi'-1} \gamma_2^{\pi'} \sum_{K \in \widetilde{\mathcal{T}}_{h,n}} \sum_{E \subset \partial K} h_E^{\pi'} \|\psi_E \sigma_E\|_{1,\pi';K}^{\pi'} \\
&\leq 5^{\pi'-1} \gamma_1^{\pi'} \sum_{K \in \widetilde{\mathcal{T}}_{h,n}} c_3^{-\pi'} c_4^{\pi'} h_K^{(\pi-1)\pi'} \|R_K\|_{0,\pi;K}^{\pi'(\pi-1)} \\
&\quad + 5^{\pi'-1} \gamma_2^{\pi'} \sum_{K \in \widetilde{\mathcal{T}}_{h,n}} \sum_{E \subset \partial K} c_5^{-\pi'} c_6^{\pi'} h_E^{\pi'+1-\pi'} \|R_E\|_{\pi;E}^{\pi'(\pi-1)} \\
&\leq 5^{\pi'} \max\{\gamma_1, \gamma_2\}^{\pi'} \max\{c_3^{-1} c_4, c_5^{-1} c_6\}^{\pi'} (\eta_h^n)^{\pi} \\
&= 5^{\pi'} \max\{\gamma_1, \gamma_2\}^{\pi'} \max\{c_3^{-1} c_4, c_5^{-1} c_6\}^{\pi'} (\eta_h^n)^{(\pi-1)\pi'}.
\end{aligned}$$

This proves that

$$\|w_n\|_{1,\pi'} \leq 5 \max\{\gamma_1, \gamma_2\} \max\{c_3^{-1} c_4, c_5^{-1} c_6\} (\eta_h^n)^{\pi-1}. \tag{5.12}$$

Since $h_E \leq h_K$ for all edges E of any element K, Lemma 5.2 also implies that

$$\begin{aligned}
&\sum_{K\in\widetilde{\mathcal{T}}_{h,n}} \int_K R_K w_n + \sum_{E\in\widetilde{\mathcal{E}}_{h,n}} \int_E R_E w_n \\
&= \gamma_1 \sum_{K\in\widetilde{\mathcal{T}}_{h,n}} h_K^\pi \int_K R_K \psi_K v_K \\
&\quad + \gamma_2 \sum_{E\in\widetilde{\mathcal{E}}_{h,n}} h_E \int_E R_E \psi_E \sigma_E \\
&\quad + \gamma_2 \sum_{K\in\widetilde{\mathcal{T}}_{h,n}} \sum_{E\subset\partial K} h_E \int_K R_K \psi_E \sigma_E \\
&\geq \gamma_1 \sum_{K\in\widetilde{\mathcal{T}}_{h,n}} h_K^\pi \|R_K\|_{0,\pi;K}^\pi \\
&\quad + \gamma_2 \sum_{E\in\widetilde{\mathcal{E}}_{h,n}} h_E \|R_E\|_{\pi;E}^\pi \\
&\quad - \gamma_2 \sum_{K\in\widetilde{\mathcal{T}}_{h,n}} \sum_{E\subset\partial K} c_5^{-1} c_7 h_K h_E^{\frac{1}{\pi'}} \|R_K\|_{0,\pi;K} \|R_E\|_{\pi;E}^{\pi-1}.
\end{aligned}$$

Using Young's inequality $ab \leq \frac{1}{\pi}\, a^\pi + \frac{1}{\pi'}\, b^{\pi'}$ with $a = \varepsilon^{-\frac{1}{\pi'}} c_5^{-1} c_7 h_K \|R_K\|_{0,\pi;K}$ and $b = \varepsilon^{\frac{1}{\pi'}} h_E^{\frac{1}{\pi'}} \|R_E\|_{\pi;E}^{\pi-1}$ and arbitrary $\varepsilon > 0$ and taking into account that $\pi'(\pi-1) = \pi$, this gives

$$\begin{aligned}
&\sum_{K\in\widetilde{\mathcal{T}}_{h,n}} \int_K R_K w_n + \sum_{E\in\widetilde{\mathcal{E}}_{h,n}} \int_E R_E w_n \\
&\geq \gamma_1 \sum_{K\in\widetilde{\mathcal{T}}_{h,n}} h_K^\pi \|R_K\|_{0,\pi;K}^\pi \\
&\quad + \gamma_2 \sum_{E\in\widetilde{\mathcal{E}}_{h,n}} h_E \|R_E\|_{\pi;E}^\pi \\
&\quad - \gamma_2 \sum_{K\in\widetilde{\mathcal{T}}_{h,n}} \sum_{E\in\partial K} \left\{ \frac{1}{\pi} \varepsilon^{-\frac{\pi}{\pi'}} c_5^{-\pi} c_7^{-\pi} h_K^\pi \|R_K\|_{0,\pi;K}^\pi + \frac{\varepsilon}{\pi'} h_E \|R_E\|_{\pi;E}^\pi \right\} \\
&\geq (\gamma_1 - 4\frac{\gamma_2}{\pi} \varepsilon^{-\frac{\pi}{\pi'}} c_5^{-\pi} c_7^\pi) \sum_{K\in\widetilde{\mathcal{T}}_{h,n}} h_K^\pi \|R_K\|_{0,\pi;K}^\pi \\
&\quad + \gamma_2 (1 - 2\frac{\varepsilon}{\pi'}) \sum_{E\in\widetilde{\mathcal{E}}_{h,n}} h_E \|R_E\|_{\pi;E}^\pi.
\end{aligned}$$

This proves that

$$\begin{aligned}&\sum_{K\in\widetilde{\mathcal{T}}_{h,n}}\int_K R_K w_n+\sum_{E\in\widetilde{\mathcal{E}}_{h,n}}\int_E R_E w_n\\&\geq \min\left\{\gamma_1-\gamma_2\frac{4}{\pi}\varepsilon^{1-\pi}c_5^{-\pi}c_7^{\pi},\ \gamma_2(1-\frac{2\varepsilon}{\pi'})\right\}(\eta_h^n)^\pi.\end{aligned}\tag{5.13}$$

From Lemma 5.2 we also obtain

$$\begin{aligned}&\sum_{K\in\widetilde{\mathcal{T}}_{h,n}}\int_K D_K w_n+\sum_{E\in\widetilde{\mathcal{E}}_{h,n}}\int_E D_E w_n\\&=\gamma_1\sum_{K\in\widetilde{\mathcal{T}}_{h,n}}h_K^\pi\int_K D_K\psi_K v_K+\gamma_2\sum_{E\in\widetilde{\mathcal{E}}_{h,n}}h_E\int_E D_E\psi_E\sigma_E\\&\quad+\gamma_2\sum_{K\in\widetilde{\mathcal{T}}_{h,n}}\sum_{E\subset\partial K}h_E\int_K D_K\psi_E\sigma_E\\&\leq\gamma_1\sum_{K\in\widetilde{\mathcal{T}}_{h,n}}c_3^{-1}h_K^\pi\|D_K\|_{0,\pi;K}\|R_K\|_{0,\pi;K}^{\pi-1}\\&\quad+\gamma_2\sum_{E\in\widetilde{\mathcal{E}}_{h,n}}c_5^{-1}h_E\|D_E\|_{\pi;E}\|R_E\|_{\pi;E}^{\pi-1}\\&\quad+\gamma_2\sum_{K\in\widetilde{\mathcal{T}}_{h,n}}c_5^{-1}c_7h_Kh_E^{\frac{1}{\pi'}}\|D_K\|_{0,\pi;K}\|R_E\|_{\pi;E}^{\pi-1}.\end{aligned}$$

Applying Hölder's inequality and using once more the relation $\pi'(\pi-1)=\pi$, this proves

$$\begin{aligned}&\sum_{K\in\widetilde{\mathcal{T}}_{h,n}}\int_K D_K w_n+\sum_{E\in\widetilde{\mathcal{E}}_{h,n}}\int_E D_E w_n\\&\leq\ \max\{\gamma_1,\gamma_2\}\max\{c_3^{-1},c_5^{-1},c_5^{-1}c_7\}5\Theta_h^n(\eta_h^n)^{\pi-1}.\end{aligned}\tag{5.14}$$

Now we choose (recall that $1<\pi,\pi'<\infty$)

$$\varepsilon=\frac{1}{2}\ ,\ \gamma_2=\frac{\pi'}{\pi'-1}\ ,\ \gamma_1=1+4\cdot 2^{\pi-1}c_5^{-\pi}c_7^{\pi}.$$

This gives

$$\min\left\{\gamma_1-\gamma_2\frac{4}{\pi}\varepsilon^{1-\pi}c_5^{-\pi}c_7^{\pi},\gamma_2(1-\frac{2\varepsilon}{\pi'})\right\}=1.\tag{5.15}$$

Equations (5.10), (5.12), (5.13), (5.14), and (5.15) prove estimate (5.9). □

6. Estimation of the temporal residual

Recall that the function $u_{h\tau}$ is piecewise affine on the intervals $[t_{n-1}, t_n]$ and equals u_h^n at time t_n and that the function $\widetilde{u}_{h\tau}$ is piecewise constant on the intervals $(t_{n-1}, t_n]$ and equals $u_h^{n\theta} = \theta u_h^n + (1-\theta)u_h^{n-1}$ on $(t_{n-1}, t_n]$. The following lemma provides us with sharp upper and lower bounds for the temporal residual.

6.1 Lemma. *Define the residual $r_\tau^n \in W^{-1,\pi}(\Omega)$, $1 \le n \le N$, by*

$$\begin{aligned}\langle r_\tau^n, v\rangle = \int_\Omega \Big\{ & \nabla v \cdot \underline{a}_p(x, u_h^{n\theta}, \nabla_h^{n\theta}) \cdot \nabla(u_h^n - u_h^{n-1}) \\ & + \nabla v \cdot \underline{a}_u(x, u_h^{n\theta}, \nabla u_h^{n\theta})(u_h^n - u_h^{n-1}) \\ & + v b_p(x, u_h^{n\theta}, \nabla u_h^{n\theta}) \cdot \nabla(u_h^n - u_h^{n-1}) \\ & + v b_u(x, u_h^{n\theta}, \nabla u_h^{n\theta})(u_h^n - u_h^{n-1}) \Big\} dx \\ & \forall v \in W_0^{1,\pi'}(\Omega),\end{aligned} \tag{6.1}$$

where the indices u and p denote the derivatives of the corresponding function with respect to the second respectively third variable. Assume that the function G defined in equation (4.2) satisfies the Lipschitz condition of Remark 4.3 and that $\|u - u_{h\tau}\|_{L^r(0,T;W_0^{1,\rho}(\Omega))} \le R_0$ and $\|u - \widetilde{u}_{h\tau}\|_{L^r(0,T;W_0^{1,\rho}(\Omega))} \le R_0$. Then the following upper and lower bounds for the temporal residual are valid:

$$\begin{aligned}&\|R_\tau(u_{h\tau})\|_{L^p(0,T;W^{-1,\pi}(\Omega))} \\ \le & \left\{ \sum_{n=1}^N \tau_n \|r_\tau^n\|_{-1,\pi}^p \right\}^{1,p} + \frac{\gamma}{2} \left\{ \sum_{n=1}^N \tau_n \|u_h^n - u_h^{n-1}\|_{1,\rho}^r \right\}^{2/r}\end{aligned} \tag{6.2}$$

and

$$\begin{aligned}&\left\{ \sum_{n=1}^N \tau_n \|r_\tau^n\|_{-1,\tau}^p \right\}^{1/p} \\ \le & \ 2\{p+1\}^{1/p} \Bigg\{ \|R_\tau(u_{h\tau})\|_{L^p(0,T;W^{-1,\pi}(\Omega))} \\ & \qquad + \frac{\gamma}{2} \left\{ \sum_{n=1}^N \tau_n \|u_h^n - u_h^{n-1}\|_{1,\rho}^r \right\}^{2/r} \Bigg\}.\end{aligned} \tag{6.3}$$

Proof. From the definition (4.5) of the temporal residual we immediately obtain

$$\begin{aligned}R_\tau(u_{h\tau}) &= G(\widetilde{u}_{h\tau}) - G(u_{h\tau}) \\ &= \int_0^1 DG(u_{h\tau} + s(\widetilde{u}_{h\tau} - u_{h\tau})(\widetilde{u}_{h\tau} - u_{h\tau}) ds \\ &= DG(\widetilde{u}_{h\tau})(\widetilde{u}_{h\tau} - u_{h\tau}) \\ &\quad + \int_0^1 [DG(u_{h\tau} + s(\widetilde{u}_{h\tau} - u_{h\tau})) - DG(\widetilde{u}_{h\tau})](\widetilde{u}_{h\tau} - u_{h\tau}) ds \\ &= R_\tau^{(1)}(u_{h\tau}) + R_\tau^{(2)}(u_{h\tau}).\end{aligned} \tag{6.4}$$

From the definition of the functions $u_{h\tau}$ and $\widetilde{u}_{h\tau}$ we conclude that

$$\widetilde{u}_{h\tau} - u_{h\tau} = (\theta - \frac{t - t_{n-1}}{\tau_n})(u_h^n - u_h^{n-1}) \quad \text{on } (t_{n-1}, t_n]. \tag{6.5}$$

A straight forward calculation yields for all $q \in (1, \infty)$ and all n between 1 and N

$$\int_{t_{n-1}}^{t_n} |\theta - \frac{t - t_{n-1}}{\tau_n}|^q dt = \tau_n \int_0^1 |\theta - z|^q dz = \tau_n \frac{1}{q+1}\{\theta^{q+1} + (1-\theta)^{q+1}\}. \tag{6.6}$$

Equations (6.5) and (6.6) in particular imply

$$\|u_{h\tau} - \widetilde{u}_{h\tau}\|_{L^r(0,T;W_0^{1,\rho}(\Omega))} \leq \left\{ \sum_{n=1}^N \tau_n \|u_h^n - u_h^{n-1}\|_{1,\rho}^r \right\}^{1/r}.$$

This estimate and the Lipschitz continuity of DG yield an upper bound for the term $R_\tau^{(2)}$ in equation (6.4):

$$\begin{aligned} \|R_\tau^{(2)}(u_{h\tau})\|_{L^p(0,T;W^{-1,\pi}(\Omega))} &\leq \frac{1}{2}\gamma \|u_{h\tau} - \widetilde{u}_{h\tau}\|^2_{L^r(0,T;W_0^{1,\rho}(\Omega))} \\ &\leq \frac{1}{2}\gamma \left\{ \sum_{n=1}^N \tau_n \|u_h^n - u_h^{n-1}\|_{1,\rho}^r \right\}^{2/r}. \end{aligned} \tag{6.7}$$

Since this is the second term on the right-hand sides of estimates (6.2) and (6.3), it remains to prove that $\|R_\tau^{(1)}(u_{h\tau})\|_{L^p(0,T;W^{-1,\pi}(\Omega))}$ is bounded from above and from below by the corresponding multiples of $\{\sum_{n=1}^N \tau_n \|r_\tau^n\|_{-1,\pi}^p\}^{1/p}$. Equations (4.2), (6.1), and (6.5) yield

$$R_\tau^{(1)}(u_{h\tau}) = (\theta - \frac{t - t_{n-1}}{\tau_n}) r_\tau^n \quad \text{on } (t_{n-1}, t_n]. \tag{6.8}$$

Combining this with equation (6.6) and observing that

$$2^{-q} \leq \theta^{q+1} + (1-\theta)^{q+1} \leq 1$$

for all $q \in (1, \infty)$ and all $\theta \in [\frac{1}{2}, 1]$, we obtain the estimate

$$\begin{aligned} &\left\{ \frac{1}{p+1} 2^{-p} \right\}^{1/p} \left\{ \sum_{n=1}^N \tau_n \|r_\tau^n\|_{-1,\pi}^p \right\}^{1/p} \\ &\leq \|R_\tau^{(1)}(u_{h\tau})\|_{L^p(0,T;W^{-1,\pi}(\Omega))} \\ &\leq \left\{ \sum_{n=1}^N \tau_n \|r_\tau^n\|_{-1,\pi}^p \right\}^{1/p}. \end{aligned}$$

This completes the proof of estimates (6.2) and (6.3). ∎

7. A preliminary a posteriori error estimate

The following lemma gives a posteriori error bounds which are reliable and efficient in the sense described in the Introduction. However, they are not suited for practical computations since they incorporate Sobolev norms of a negative order. In Section 9 we will bound these terms by computable quantities.

7.1 Lemma. *Assume that the functions* $F, G, u, u_{h\tau}$*, and* $\widetilde{u}_{h\tau}$ *satisfy the conditions of Lemma 4.2, Remark 4.3, and Lemma 6.1. Then the error between the solution* u *of problems (2.1), (2.2) and the solution* $u_{h\tau}$ *of problems (3.1), (3.2) is bounded from above by*

$$
\begin{aligned}
&\|u - u_{h\tau}\|_{W^r(0,T;W_0^{1,\rho}(\Omega),W^{-1,\pi}(\Omega))} \\
&\quad \le c^\sharp \Bigg\{ \|u_0 - \pi_0 u_0\|_{-1,\pi} + \left\{ \sum_{n=1}^N \tau_n [(\eta_h^n)^p + \|r_\tau^n\|_{-1,\pi}^p] \right\}^{1/p} \\
&\qquad + \left\{ \sum_{n=1}^N \tau_n (\Theta_h^n)^p \right\}^{1/p} + \left\{ \sum_{n=1}^N \tau_n \|u_h^n - u_h^{n-1}\|_{1,\rho}^r \right\}^{2/r} \Bigg\}
\end{aligned}
\tag{7.1}
$$

and from below by

$$
\begin{aligned}
&\left\{ \sum_{n=1}^N \tau_n \ [(\eta_h^n)^p + \|r_\tau^n\|_{-1,\pi}^p] \right\}^{1/p} \\
&\quad \le c_\sharp \Bigg\{ \|u - u_{h\tau}\|_{W^r(0,T;W_0^{1,\rho}(\Omega),W^{-1,\pi}(\Omega))} \\
&\qquad + \left\{ \sum_{n=1}^N \tau_n (\Theta_h^n)^p \right\}^{1/p} + \left\{ \sum_{n=1}^N \tau_n \|u_h^n - u_h^{n-1}\|_{1,\rho}^r \right\}^{2/r} \Bigg\}.
\end{aligned}
\tag{7.2}
$$

The quantities η_h^n*,* Θ_h^n *and* r_τ^n *are defined in equations (5.6), (5.7) and (6.1) respectively.*

Proof. The upper bound (7.1) immediately follows from the decomposition (4.6) of the residual and Lemmas 4.2, 5.3 and 6.1.

In view of Lemma 4.2, the lower bound (7.2) is established once we have proved that

$$
\begin{aligned}
&\left\{ \sum_{n=1}^N \tau_n \ [(\eta_h^n)^p + \|r_\tau^n\|_{-1,\pi}^p] \right\}^{1/p} \\
&\quad \le c_\sharp \Bigg\{ \|\frac{\partial u_{h\tau}}{\partial t} + G(u_{h\tau})\|_{L^p(0,T;W^{-1,\pi}(\Omega))} \\
&\qquad + \left\{ \sum_{n=1}^N \tau_n (\Theta_h^n)^p \right\}^{1/p} + \left\{ \sum_{n=1}^N \tau_n \|u_h^n - u_h^{n-1}\|_{1,\rho}^r \right\}^{2/r} \Bigg\}.
\end{aligned}
\tag{7.3}
$$

We start with the r_τ^n-term on the left-hand side of estimate (7.3). From the decomposition (4.6) and Lemmas 5.3 and 6.1 we obtain

$$\begin{aligned}
&\left\{\sum_{n=1}^{N} \tau_n \|r_\tau^n\|_{-1,\pi}^p\right\}^{1/p} \\
&\le 2\{p+1\}^{1/p}\left\{\|R_\tau(u_{h\tau})\|_{L^p(0,T;W^{-1,\pi}(\Omega))}\right. \\
&\qquad\left. + \frac{\gamma}{2}\left\{\sum_{n=1}^{N} \tau_n \|u_h^n - u_h^{n-1}\|_{1,\rho}^r\right\}^{2/r}\right\} \\
&\le 2\{p+1\}^{1/p}\left\{\|\frac{\partial u_{h\tau}}{\partial t} + G(u_{h\tau})\|_{L^p(0,T;W^{-1,\pi}(\Omega))}\right. \\
&\qquad + c^\dagger\left\{\sum_{n=1}^{N} \tau_n [(\eta_h^n)^p + (\Theta_h^n)^p]\right\}^{1/\rho} \\
&\qquad\left. + \frac{\gamma}{2}\left\{\sum_{n=1}^{N} \tau_n \|u_h^n - u_h^{n-1}\|_{1,\rho}^r\right\}^{2/r}\right\}.
\end{aligned} \tag{7.4}$$

To bound the η_h^n-term on the left-hand side of estimate (7.3) we set

$$w = \sum_{n=1}^{N} (\alpha+1)(\frac{t-t_{n-1}}{\tau_n})^\alpha (\eta_h^n)^{p-\pi} w_n \chi_{(t_{n-1},t_n]}(t),$$

where the functions w_n are as in Lemma 5.3 and $\chi_{(t_{n-1},t_n]}$ denotes the characteristic function of the interval $(t_{n-1}, t_n]$. The parameter $\alpha \ge 0$ is arbitrary at present and will be fixed later. Since $0 \le (\frac{t-t_{n-1}}{\tau_n})^\alpha \le 1$ on $[t_{n-1}, t_n]$ and since $(p-1)p' = p$, we obtain from the second line of estimate (5.9) that

$$\begin{aligned}
&\|w\|_{L^{p'}(0,T;W_0^{1,\pi'}(\Omega))} \\
&\le c_\dagger(\alpha+1)\left\{\sum_{n=1}^{N} \tau_n (\eta_h^n)^{(\pi-1)p'} (\eta_h^n)^{(p-\pi)p'}\right\}^{1/p'} \\
&= c_\dagger(\alpha+1)\left\{\sum_{n=1}^{N} \tau_n (\eta_h^n)^p\right\}^{\frac{p-1}{p}}.
\end{aligned} \tag{7.5}$$

Since

$$\int_{t_{n-1}}^{t_n} (\alpha+1)(\frac{t-t_{n-1}}{\tau_n})^\alpha dt = \tau_n,$$

the first line of estimate (5.9) implies that

$$\begin{aligned}
&\sum_{n=1}^{N} \tau_n (\eta_h^n)^p \\
\leq\ & \int_0^T \langle R_h(u_{h\tau}), w\rangle + \sum_{n=1}^{N} \tau_n \Theta_h^n c_\dagger (\eta_h^n)^{\pi-1} (\eta_h^n)^{p-\pi} \\
\leq\ & \int_0^T \langle R_h(u_{h\tau}), w\rangle + c_\dagger \left\{ \sum_{n=1}^{N} \tau_n (\Theta_h^n)^p \right\}^{1/p} \left\{ \sum_{n=1}^{N} \tau_n (\eta_h^n)^p \right\}^{\frac{p-1}{p}} .
\end{aligned} \tag{7.6}$$

Equations (4.6) and (6.4) yield

$$\begin{aligned}
\int_0^T \langle R_h(u_{h\tau}), w\rangle = & \int_0^\tau \langle \frac{\partial u_{h\tau}}{\partial t} + G(u_{h\tau}), w\rangle \\
& - \int_0^T \langle R_\tau^{(1)}(u_{h\tau}), w\rangle - \int_0^T \langle R_\tau^{(2)}(u_{h\tau}), w\rangle.
\end{aligned} \tag{7.7}$$

Estimate (7.5) gives

$$\begin{aligned}
&\int_0^T \langle \frac{\partial u_{h\tau}}{\partial t} + G(u_{h\tau}), w\rangle \\
\leq & \| \frac{\partial u_{h\tau}}{\partial t} + G(u_{h\tau}) \|_{L^p(0,T;W^{-1,\pi}(\Omega))} c_\dagger (\alpha+1) \left\{ \sum_{n=1}^{N} \tau_n (\eta_h^n)^p \right\}^{\frac{p-1}{p}} .
\end{aligned} \tag{7.8}$$

Estimates (6.7) and (7.5) imply

$$\begin{aligned}
&\int_0^T \langle R_\tau^{(2)}(u_{h\tau}), w\rangle \\
\leq & \frac{\gamma}{2} \left\{ \sum_{n=1}^{N} \tau_n \| u_h^n - u_h^{n-1} \|_{1,\rho}^r \right\}^{2/r} c_\dagger (\alpha+1) \left\{ \sum_{n=1}^{N} \tau_n (\eta_h^n)^p \right\}^{\frac{p-1}{p}} .
\end{aligned} \tag{7.9}$$

Equation (6.8) and the second line of estimate (5.9) finally yield

$$\begin{aligned}
&\int_0^T \langle R_\tau^{(1)}(u_{h\tau}), w\rangle \\
= & \sum_{n=1}^{N} \langle r_\tau^n, w_n\rangle (\eta_h^n)^{p-\pi} (\alpha+1) \int_{t_{n-1}}^{t_n} (\theta - \frac{t-t_{n-1}}{\tau_n})(\frac{t-t_{n-1}}{\tau_n})^\alpha dt \\
= & \sum_{n=1}^{N} \langle r_\tau^n, w_n\rangle (\eta_h^n)^{p-\pi} (\theta - \frac{\alpha+1}{\alpha+2}) \tau_n \\
\leq & c_\dagger |\theta - \frac{\alpha+1}{\alpha+2}| \sum_{n=1}^{N} \tau_n \| r_\tau^n \|_{-1,\pi} (\eta_h^n)^{p-1} \\
\leq & c_\dagger |\theta - \frac{\alpha+1}{\alpha+2}| \left\{ \sum_{n=1}^{N} \tau_n \| r_\tau^n \|_{-1,\pi}^p \right\}^{1/p} \left\{ \sum_{n=1}^{N} \tau_n (\eta_h^n)^p \right\}^{\frac{p-1}{p}} .
\end{aligned} \tag{7.10}$$

Equation (7.7) and estimates (7.6), (7.8), (7.9), and (7.10) give

$$
\begin{aligned}
&\left\{\sum_{n=1}^{N}\tau_n(\eta_h^n)^p\right\}^{1/p}\\
\le& c_\dagger(\alpha+1)\|\frac{\partial u_{h\tau}}{\partial t}+G(u_{h\tau})\|_{L^p(0,T;W^{-1,\pi}(\Omega))}\\
&+c_\dagger(\alpha+1)\frac{\gamma}{2}\left\{\sum_{n=1}^{N}\tau_n\|u_h^n-u_h^{n-1}\|_{1,\rho}^r\right\}^{2/r}\\
&+c_\dagger\left\{\sum_{n=1}^{N}\tau_n(\Theta_h^n)^p\right\}^{1/p}\\
&+c_\dagger|\theta-\frac{\alpha+1}{\alpha+2}|\left\{\sum_{n=1}^{N}\tau_n\|r_\tau^n\|_{-1,\pi}^p\right\}^{1/p}.
\end{aligned}
$$

Inserting estimate (7.4) in this inequality we finally arrive at

$$
\begin{aligned}
&\left\{\sum_{n=1}^{N}\tau_n(\eta_h^n)^p\right\}^{1/p}\\
\le& c_\dagger[(\alpha+1)+|\theta-\frac{\alpha+1}{\alpha+2}|2\{p+1\}^{1/p}]\|\frac{\partial u_{h\tau}}{\partial t}+G(u_{h\tau})\|_{L^p(0,T;W^{-1,\pi}(\Omega))}\\
&+c_\dagger[1+|\theta-\frac{\alpha+1}{\alpha+2}|c^\dagger 2\{p+1\}^{1/p}]\left\{\sum_{n=1}^{N}\tau_n(\Theta_h^n)^p\right\}^{1/p}\\
&+\frac{\gamma}{2}c_\dagger[(\alpha+1)+|\theta-\frac{\alpha+1}{\alpha+2}|2\{p+1\}^{1/p}]\left\{\sum_{n=1}^{N}\tau_n\|u_h^n\cdot u_h^{n-1}\|_{1,\rho}^r\right\}^{2/r}\\
&+c_\dagger c^\dagger|\theta-\frac{\alpha+1}{\alpha+2}|2\{p+1\}^{1/p}\left\{\sum_{n=1}^{N}\tau_n(\eta_h^n)^p\right\}^{1/p}.
\end{aligned}
\tag{7.11}
$$

Now we choose the parameter α such that the η_h^n-term on the right-hand side of estimate (7.11) is balanced by the term on the left-hand side. For the mid-point rule, i.e. $\theta=\frac{1}{2}$, this is obvious: We simply choose $\alpha=0$ and the η_h^n-term of the right-hand side of (7.11) vanishes. In the case $\frac{1}{2}<\theta\le 1$ we choose

$$\alpha=\frac{2K(2\theta-1)}{2K(1-\theta)+1}\ \text{with}\ \ K=c_\dagger 2\{p+1\}^{1/p}.$$

This choice ensures that

$$c_\dagger c^\dagger|\theta-\frac{\alpha+1}{\alpha+2}|2\{p+1\}^{1/p}\le\frac{1}{2}\ \text{and}\ \ \alpha\le 2K.$$

Hence, estimate (7.11) takes the form

$$\begin{aligned}&\left\{\sum_{n=1}^{N}\tau_n(\eta_h^n)^p\right\}^{1/p}\\ \le c&\left\{\|\frac{\partial u_{h\tau}}{\partial t}+G(u_{h\tau})\|_{L^p(0,T;W^{-1,\pi}(\Omega))}\right.\\ &\left.+\left\{\sum_{n=1}^{N}\tau_n(\Theta_h^n)^p\right\}^{1/p}+\left\{\sum_{n=1}\tau_n\|u_h^n-u_h^{n-1}\|_{1,\rho}^r\right\}^{2/r}\right\}^{1/p}\end{aligned} \tag{7.12}$$

with a constant c that only depends on p, $c_\dagger$ and $c^\dagger$. Estimates (7.4) and (7.12) prove inequality (7.3) and thus complete the proof of the lower bound (7.2). □

8. $W^{1,q}$-stability results for the Laplacian

The results of the next section are based on the $W^{1,q}$-stability of the Laplacian both in its analytical and discrete form. We start with the analytical case. The following result is well-known for domains with smooth C^1-boundary (cf. [8]). For polygonal domains, however, we are not aware of a proof. Recall that for $q \in [1,\infty]$ the dual Lebesgue exponent is denoted by $q' \in [1,\infty]$ and is defined by $\frac{1}{q}+\frac{1}{q'}=1$.

8.1 Lemma. *For every convex, bounded, polygonal domain $\Omega \subset \mathbb{R}^2$ and every $q \in [1,\infty]$ there is a constant $\alpha_q > 0$ such that*

$$\inf_{v\in W_0^{1,q}(\Omega)}\sup_{w\in W_0^{1,q'}(\Omega)}\frac{\int_\Omega \nabla v\cdot\nabla w}{\|\nabla v\|_{0,q}\|\nabla w\|_{0,q'}} \ge \alpha_q. \tag{8.1}$$

The constant α_q only depends on q and on the maximum interior angle at the vertices of Ω.

Proof. Inequality (8.1) is proved in [8] for domains $\Omega \subset \mathbb{R}^n, n \ge 2$, with smooth C^1-boundary. The proof is based on the following three auxiliary results:

1. Inequality (8.1) holds for $\mathbb{R}^n$.
2. Inequality (8.1) holds for the half-space $H_+ = \{x \in \mathbb{R}^n : x_1 > 0\}$.
3. Inequality (8.1) holds for domains $H_\omega = \{x \in \mathbb{R}^n : x_1 > \omega(x_2,\dots,x_n)\}$ with functions $\omega \in C^1(\mathbb{R}^{n-1})$ satisfying $\omega(0)=0$ and $\|\nabla\omega\|_{L^\infty(\mathbb{R}^{n-1})} \ll 1$.

The third result is the only point where the smoothness of the boundary comes into play. The smoothness condition on ω can be relaxed to the condition that ω should be Lipschitz continuous and that its Lipschitz constant is sufficiently small. This, however, does not help us since it would require that the interior angles at the

vertices of Ω should be sufficiently close to π. Instead we must prove that inequality (8.1) holds for domains $H_c = \{x \in \mathbb{R}^2 : |x_2| \le cx_1\}$ with $c > 0$.
To verify this, choose a parameter $c > 0$ and keep it fixed. Then introduce polar coordinates to transform H_c to the strip $\{(r, \varphi) : r > 0, |\varphi| \le \alpha\}$ where $\alpha = \arctan c$. Next apply the scaling $r \to \frac{2\alpha}{\pi} r$, $\varphi \to \frac{2\alpha}{\pi}\varphi$ to transform to the strip $\{(s, \psi) : s > 0, |\varphi| \le \frac{\pi}{2}\}$. Then transform back to cartesian coordinates. The combination of these transformations transforms H_c to the half space H_+. Now, we already know that inequality (8.1) holds for H_+. Hence it also holds in polar-coordinates. Since the left-hand side of (8.1) is invariant under scalings, inequality (8.1) holds in polar coordinates on the strip $\{(r, \varphi) : r > 0, |\varphi| \le \alpha\}$ and thus on H_c.
Once we know that we may replace H_ω by H_c, the rest of the proof of the lemma proceeds as in [8]. □

8.2 Remark. When Ω is not convex but has a re-entrant corner with angle $\omega > \pi$, Lemma 8.1 can at best hold for Lebesgue exponents $q \in [1, \frac{2\omega}{\omega - \pi})$. This is due to the fact that the singular solution $r^{\frac{\pi}{\omega}} \sin(\frac{\pi}{\omega}\varphi)$ of the Laplacian is in $W^{1,q}(\Omega)$ only for this realm of Lebesgue exponents.
In the proof of Lemma 8.1 the convexity is reflected by the fact that the transformation from polar to cartesian coordinates is globally invertible in the vicinity of convex corners. For non-convex corners it is only locally invertible.

Now we come to the discrete case.

8.3 Lemma. *Consider a convex, bounded, polygonal domain $\Omega \subset \mathbb{R}^2$ and an arbitrary affine equivalent, admissible and shape regular partition $\mathcal{T}$ of Ω. Denote by*

$$S^1(\mathcal{T}) = \{v \in C(\Omega) : v|_K \in R_1(K)\ \forall K \in \mathcal{T}, v = 0 \text{ on } \mathcal{T}\} \tag{8.2}$$

the space of continuous piecewise linear finite element functions corresponding to $\mathcal{T}$. Then for every $q \in [1, \infty]$ there is a constant $\beta_q > 0$ such that

$$\inf_{v_{\mathcal{T}} \in S^1(\mathcal{T})} \sup_{w_{\mathcal{T}} \in S^1(\mathcal{T})} \frac{\int_\Omega \nabla v_{\mathcal{T}} \nabla w_{\mathcal{T}}}{\|\nabla v_{\mathcal{T}}\|_{0,q} \|\nabla w_{\mathcal{T}}\|_{0,q'}} \ge \beta_q. \tag{8.3}$$

The constant β_q only depends on q, on the maximum interior angle at the vertices of Ω, and on the shape parameter $\sup_{K \in \mathcal{T}} h_K/\rho_K$ of $\mathcal{T}$.
Proof. We denote by $R_{\mathcal{T}} : W_0^{1,1}(\Omega) \to S^1(\mathcal{T})$ the Ritz projection which is defined by

$$\int_\Omega \nabla(R_{\mathcal{T}} v) \nabla w_{\mathcal{T}} = \int_\Omega \nabla v \nabla w_{\mathcal{T}} \quad \forall v \in W_0^{1,1}(\Omega), w_{\mathcal{T}} \in S^1(\mathcal{T}).$$

Consider first the case $q \in [1, 2]$. Then we have $q' \ge 2$. From [7] and [4, Chap. 7] we know that $R_{\mathcal{T}}$ is stable in the $W^{1,q'}$-norm, i.e., there is a constant $c_{q'} > 0$ such that

$$\|\nabla(R_{\mathcal{T}} w)\|_{0,q'} \le c_{q'} \|\nabla w\|_{0,q'} \quad \forall w \in W_0^{1,q'}(\Omega).$$

The constant $c_{q'}$ only depends on q', on the maximum interior angle at a vertex of Ω, and on the shape parameter of $\mathcal{T}$.
Consider an arbitrary function $v_\mathcal{T} \in S^1(\mathcal{T})$ and a number $\delta \in (0,1)$. From Lemma 8.1 we know that there is a function $w_\delta \in W_0^{1,q'}(\Omega)$ with $\|\nabla w_\delta\|_{0,q'} = 1$ and

$$\int_\Omega \nabla v_\mathcal{T} \nabla w_\delta = \delta \alpha_q \|\nabla v_\mathcal{T}\|_{0,q}.$$

Together with the stability of the Ritz projection this implies

$$\begin{aligned}\sup_{w_\mathcal{T} \in S^1(\mathcal{T})} \frac{\int_\Omega \nabla v_\mathcal{T} \nabla w_\mathcal{T}}{\|\nabla v_\mathcal{T}\|_{0,q} \|\nabla w_\mathcal{T}\|_{0,q'}} \geq & \frac{\int_\Omega \nabla v_\mathcal{T} \nabla (R_\mathcal{T} w_\delta)}{\|\nabla v_\mathcal{T}\|_{0,q} \|\nabla (R_\mathcal{T} w_\delta)\|_{0,q'}} \\ \geq & \frac{1}{c_{q'}} \frac{\int_\Omega \nabla v_\mathcal{T} \nabla (R_\mathcal{T} w_\delta)}{\|\nabla v_\mathcal{T}\|_{0,q} \|\nabla w_\delta\|_{0,q'}} \\ = & \frac{1}{c_{q'}} \frac{\int_\Omega \nabla v_\mathcal{T} \nabla w_\delta}{\|\nabla v_\mathcal{T}\|_{0,q} \|\nabla w_\delta\|_{0,q'}} \\ \geq & \frac{\delta \alpha_q}{c_{q'}}.\end{aligned}$$

Since δ and $v_\mathcal{T}$ were arbitrary this proves inequality (8.3) with $\beta_q = \frac{\alpha_q}{c_{q'}}$.
One easily checks that (8.3) implies the stability of $R_\mathcal{T}$ in the $W^{1,q}$-norm with $c_q = \frac{1}{\beta_q} = \frac{c_{q'}}{\alpha_q}$.
Now consider the case $q > 2$. This implies $1 < q' < 2$. Since we already have established the stability of $R_\mathcal{T}$ in the $W^{1,q'}$-norm with $c_{q'} = \frac{1}{\beta_{q'}} = \frac{c_q}{\alpha_{q'}}$, we can proceed as in the case $q \geq 2$ and obtain inequality (8.3) with $\beta_q = \frac{\alpha_q}{c_{q'}} = \frac{\alpha_q \alpha_{q'}}{c_q}$. □

9. The final a posteriori error estimate

In this section we make computable the error estimator of Lemma 7.1 by replacing the negative Sobolev norms of the $r_\mathcal{T}^n$-terms by computable quantities. This will be done with the help of suitable auxiliary discrete Poisson equations.
As in Section 5 we choose an integer ℓ and denote for every n between 1 and N by

$$\underline{a}_{p;h,n}(x, u_h^{n\theta}, \nabla u_h^{n\theta}), \underline{a}_{u;h,n}(x, u_h^{n\theta}, \nabla u_h^{n\theta}), b_{p;h,n}(x, u_h^{n\theta}, \nabla u_h^{h\theta}), b_{u;h,n}(x, u_h^{n\theta}, \nabla u_h^{n\theta})$$

the L^2-projections of

$$\underline{a}_p(x, u_h^{n\theta}, \nabla u_h^{n\theta}), \underline{a}_u(x, u_h^{n\theta}, \nabla u_h^{n\theta}), b_p(x, u_h^{n\theta}, \nabla u_h^{n\theta}), b_u(x, u_h^{n\theta}, \nabla u_h^{n\theta})$$

on discontinuous tensors, vector-fields and functions respectively which are piecewise polynomials of degree ℓ on the elements of $\widetilde{\mathcal{T}}_{h,n}$. For abbreviation we set for every element $K \in \widetilde{\mathcal{T}}_{h,n}$, every edge $E \in \widetilde{\mathcal{E}}_{h,n}$ and every n between 1 and N

$$\begin{aligned}
\widetilde{R}_K =& - \operatorname{div}[\underline{a}_{p;h,n}(x, u_h^{n\theta}, \nabla u_h^{n\theta}) \cdot \nabla(u_h^n - u_h^{n-1})] \\
& - \operatorname{div}[\underline{a}_{u;h,n}(x, u_h^{n\theta}, \nabla u_h^{n\theta})(u_h^n - u_h^{n-1})] \\
& + b_{p;h,n}(x, u_h^{n\theta}, \nabla u_h^{n\theta}) \cdot \nabla(u_h^n - u_h^{n-1}) \\
& + b_{u;h,n}(x, u_h^{n\theta}, \nabla u_h^{n\theta})(u_h^n - u_h^{n-1}), \\
\widetilde{R}_E =& \underline{a}_{p;h,n}(x, u_h^{n\theta}, \nabla u_h^{n\theta}) \cdot \nabla(u_h^n - u_h^{n-1}) \\
& + \underline{a}_{u;h,n}(x, u_h^{n\theta}, \nabla u_h^{n\theta})(u_h^n - u_h^{n-1}), \\
\widetilde{D}_K =& \operatorname{div}[(\underline{a}_p(x, u_h^{n\theta}, \nabla u_h^{n\theta}) - \underline{a}_{p;h,n}(x, u_h^{n\theta}, \nabla u_h^{n\theta})) \cdot \nabla(u_h^n - u_h^{n-1})] \\
& - \operatorname{div}[(\underline{a}_u(x, u_h^{n\theta}, \nabla u_h^{n\theta}) - \underline{a}_{u;h,n}(x, u_h^{n\theta}, \nabla u_h^{n\theta}))(u_h^n - u_h^{n-1})] \\
& + (b_p(x, u_h^{n\theta}, \nabla u_h^{n\theta}) - b_{p;h,n}(x, u_h^{n\theta}, \nabla u_h^{n\theta})) \cdot \nabla(u_h^n - u_h^{n-1}) \\
& + (b_u(x, u_h^{n\theta}, \nabla u_h^{n\theta}) - b_{u;h,n}(x, u_h^{n\theta}, \nabla u_h^{n\theta}))(u_h^n - u_h^{n-1}), \\
\widetilde{D}_E =& (\underline{a}_p(x, u_h^{n\theta}, \nabla u_h^{n\theta}) - \underline{a}_{p;h,n}(x, u_h^{n\theta}, \nabla u_h^{n\theta})) \cdot \nabla(u_h^n - u_h^{n-1}) \\
& + (\underline{a}_u(x, u_h^{n\theta}, \nabla u_h^{n\theta}) - \underline{a}_{u;h,n}(x, u_h^{n\theta}, \nabla u_h^{n\theta}))(u_h^n - u_h^{n-1}).
\end{aligned}$$

Of course, the right-hand sides of the above equations must always be interpreted as the restriction of corresponding functions to the relevant element or edge.
Recall the definitions (6.1) of the residuals r_τ^n and (8.2) of the spaces $S^1(\mathcal{T})$.

9.1 Lemma. *For every integer n between 1 and N denote by $\widetilde{u}_h^n \in S^1(\widetilde{\mathcal{T}}_{h,n})$ the unique solution of the discrete Poisson equation*

$$\int_\Omega \nabla \widetilde{u}_h^n \nabla v_h = \langle r_\tau^n, v_h \rangle \quad \forall v_h \in S^1(\widetilde{\mathcal{T}}_{h,n}). \tag{9.1}$$

Define the error indicator $\widetilde{\eta}_h^n$ by

$$\widetilde{\eta}_h^n = \left\{ \sum_{K \in \widetilde{\mathcal{T}}_{h,n}} h_K^\pi \| \widetilde{R}_K + \Delta \widetilde{u}_h^n \|_{0,\pi;K}^\pi + \sum_{E \in \widetilde{\mathcal{E}}_{h,n}} h_E \| [\underline{n}_E \cdot (\nabla \widetilde{u}_h^n - \widetilde{R}_E)]_E \|_{\pi;E}^\pi \right\}^{1/\pi} \tag{9.2}$$

and the data error $\widetilde{\Theta}_h^n$ by

$$\widetilde{\Theta}_h^n = \left\{ \sum_{K \in \widetilde{\mathcal{T}}_{h,n}} h_K^\pi \| \widetilde{D}_K \|_{0,\pi;K}^\pi + \sum_{E \in \widetilde{\mathcal{E}}_{h,n}} h_E \| \widetilde{D}_E \|_{\pi;E}^\pi \right\}^{1/\pi}. \tag{9.3}$$

Then there are two constants $\widetilde{c}_\dagger$ and $\widetilde{c}^\dagger$, which only depend on the polynomial degree ℓ and on the ratios h_K/ρ_K, such that

$$\begin{aligned} \|r_\tau^n\|_{-1,\pi} &\le \widetilde{c}^\dagger\{\widetilde{\eta}_h^n + \|\nabla\widetilde{u}_h^n\|_{0,\pi} + \widetilde{\Theta}_h^n\}, \\ \widetilde{\eta}_h^n + \|\nabla\widetilde{u}_h^n\|_{0,\pi} &\le \widetilde{c}_\dagger\{\|r_\tau^n\|_{-1,\pi} + \widetilde{\Theta}_h^n\}. \end{aligned} \tag{9.4}$$

Proof. We choose an integer n between 1 and N and keep it fixed.
Lemma 8.1 implies that the Poisson equation

$$\int_\Omega \nabla\widetilde{U}^n\nabla v = \langle r_\tau^n, v\rangle \quad \forall v \in W_0^{1,\pi'}(\Omega)$$

admits a unique solution $\widetilde{U}^n \in W_0^{1,\pi}(\Omega)$ and that

$$\|\nabla\widetilde{U}^n\|_{0,\pi} \le \frac{1}{\alpha_\pi}\|r_\tau^n\|_{-1,\pi}.$$

The definition of the negative Sobolev norms on the other hand yields

$$\|r_\tau^n\|_{-1,\pi} \le \|\nabla\widetilde{U}^n\|_{0,\pi}.$$

Lemma 8.3 similarly gives

$$\|\nabla\widetilde{u}_h^n\|_{0,\pi} \le \frac{1}{\beta_\pi}\|r_\tau^n\|_{-1,\pi}.$$

The triangle inequality therefore implies

$$\begin{aligned} &\frac{1}{3}\min\{\alpha_\pi,\beta_\pi\}\{\|\nabla\widetilde{u}_h^n\|_{0,\pi} + \|\nabla(\widetilde{U}^n - \widetilde{u}_h^n)\|_{0,\pi}\} \\ &\le \|r_\tau^n\|_{-1,\pi} \\ &\le \|\nabla\widetilde{u}_h^n\|_{0,\pi} + \|\nabla(\widetilde{U}^n - \widetilde{u}_h^n)\|_{0,\pi}. \end{aligned} \tag{9.5}$$

Using standard arguments (cf. e.g. [9]) we infer from Lemmas 5.1 and 5.2 that

$$\begin{aligned} \|\nabla(\widetilde{U}^n - \widetilde{u}_h^n)\|_{0,\pi} &\le c\{\widetilde{\eta}_h^n + \widetilde{\Theta}_h^n\}, \\ \widetilde{\eta}_h^n &\le C\{\|\nabla(\widetilde{U}^n - \widetilde{u}_h^n)\|_{0,\pi} + \widetilde{\Theta}_h^n\} \end{aligned} \tag{9.6}$$

with constants c and C which only depend on the polynomial degree ℓ and the ratios h_K/ρ_K. Combining estimates (9.5) and (9.6) we arrive at the desired estimate (9.4) of $\|r_\tau^n\|_{-1,\pi}$. □

A standard duality argument for the L^2-projection onto finite element spaces yields

$$\|u_0 - \pi_0 u_0\|_{-1,\pi} \le c\left\{\sum_{K\in\mathcal{T}_{h,0}} h_K^\pi \|u_0 - \pi_0 u_0\|_{0,\pi;K}^\pi\right\}^{1/\pi}.$$

Combining this estimate and Lemmas 7.1 and 9.1 proves our final result:

9.2 Theorem. *If the conditions of Lemma 7.1 are satisfied the error between the solution u of problems (2.1), (2.2) and the solution $u_{h\mathcal{T}}$ of problems (3.1), (3.2) is bounded from above by*

$$\begin{aligned}
&\|u - u_{h\mathcal{T}}\|_{W^r(0,T;W_0^{1,\rho}(\Omega),W^{-1,\pi}(\Omega))} \\
&\le \tilde{c}^\sharp \Bigg\{ \left\{ \sum_{n=1}^{N} \tau_n [(\eta_h^n)^p + (\tilde{\eta}_h^n)^p + \|\nabla \tilde{u}_h^n\|_{0,\pi}^p] \right\}^{1/p} \\
&\quad + \left\{ \sum_{n=1}^{N} \tau_n [(\Theta_h^n)^p + (\tilde{\Theta}_h^n)^p] \right\}^{1/p} \\
&\quad + \left\{ \sum_{K \in \mathcal{T}_{n,0}} h_K^\pi \|u_0 - \pi_0 u_0\|_{0,\pi;K}^\pi \right\}^{1/\pi} \\
&\quad + \left\{ \sum_{n=1}^{N} \tau_n \|u_h^n - u_h^{n-1}\|_{1,\rho}^r \right\}^{2/r} \Bigg\}
\end{aligned} \tag{9.7}$$

and from below by

$$\begin{aligned}
&\left\{ \sum_{n=1}^{N} \tau_n [(\eta_h^n)^p + (\tilde{\eta}_h^n)^p + \|\nabla \tilde{u}_h^n\|_{0,\pi}^p] \right\}^{1/p} \\
&\le \tilde{c}_\sharp \Bigg\{ \|u - u_{h\mathcal{T}}\|_{W^r(0,T;W_0^{1,\rho}(\Omega))} \\
&\quad + \left\{ \sum_{n=1}^{N} \tau_n [(\Theta_h^n)^p + (\tilde{\Theta}_h^n)^p] \right\}^{1/p} \\
&\quad + \left\{ \sum_{n+1}^{N} \tau_n \|u_h^n - u_h^{n-1}\|_{1,\rho}^r \right\}^{2/r} \Bigg\}.
\end{aligned} \tag{9.8}$$

The quantities $\eta_h^n, \Theta_h^n, \tilde{\eta}_h^n, \tilde{\Theta}_h^n$, and $\tilde{u}_h^n$ are defined in equations (5.6), (5.7), (9.2), (9.3), and (9.1) respectively.

9.3 Remark. The left-hand side of estimate (9.8) is our error indicator. Its first term controls the error of the space-discretization and can be used for adapting the spatial mesh. The second and third terms on the left-hand side of (9.8) control the error of the time-discretization and can be used to adapt the temporal mesh. The last terms on the right-hand sides of estimates (9.7) and (9.8) are not present for linear differential equations. They control the linearization error that is implicit in the discretization. Since they are computable, they can be used to control this linearization error. These contributions are of high-order in the sense that up to a different L^p-norm, they are similar to the square of the error indicator. The second and third term on the right-hand side of estimate (9.7) and the second term on the right-hand side of estimate (9.8) are data errors. In contrast to linear problems they not only involve given data but the discrete solution as well.

9.4 Remark. Estimate (9.8) is based on the lower bound of Lemma 4.2. This in turn follows from the lower bound in the abstract error estimate of Lemma 4.1. The latter only involves the Fréchet derivative $DF(u)$. Applied to differential equations this corresponds to a linearized differential operator. Since such operators have a local effect, the lower bounds can be localized. Therefore, as for linear differential equations (cf. [12]), estimate (9.8) has an analogue that is local with respect to time.

10. References

[1] R. A. Adams: Sobolev Spaces. Academic Press, New York, 1975

[2] H. Amann: Linear and Quasilinear Parabolic Problems. Volume I: Abstract Linear Theory. Birkhäuser, Basel, 1995

[3] G. Bangerth, R. Rannacher: Adaptive Finite Element Methods for Differential Equations. Birkhäuser, Basel, 2003

[4] S. C. Brenner, L. R. Scott: The Mathematical Theory of Finite Element Methods. Springer, Berlin - Heidelberg - New York, 1994

[5] P. Clément: Approximation by finite element functions using local regularization. RAIRO Anal. Numér. 9, 77 – 84 (1975)

[6] R. Dautray and J.-L. Lions: Mathematical Analysis and Numerical Methods for Science and Technology. Springer, Berlin - Heidelberg - New York, 1992

[7] R. Rannacher, L. R. Scott: Some optimal error estimates for piecewise linear finite element approximations. Math. Comput. 38, 437 – 445 (1982)

[8] C. G. Simader, H. Sohr: The Dirichlet Problem for the Laplacian in Bounded and Unbounded Domains. Longman, London, 1996

[9] R. Verfürth: A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques. Teubner-Wiley, Stuttgart, 1996

[10] ———: A posteriori error estimates for nonlinear problems: $L^r(0,T;W^{1,\rho}(\Omega))$–error estimates for finite element discretizations of parabolic equations. Numer. Meth. for PDE 14, 487 – 518 (1998)

[11] ———: Error estimates for some quasi-interpolation operators. Modél. Math. et Anal. Numér. 33, 695 – 713 (1999)

[12] ———: A posteriori error estimates for finite element discretizations of the heat equation. Calcolo 40, 195 – 212 (2003)