A POSTERIORI ERROR ANALYSIS IN THE MAXIMUM NORM FOR A PENALTY FINITE ELEMENT METHOD FOR THE TIME-DEPENDENT OBSTACLE PROBLEM

MATS BOMAN

Abstract. We consider finite element approximation of the parabolic obstacle problem. The analysis is based on a penalty formulation of the problem where the penalisation parameter is allowed to vary in space and time. We estimate the penalisation error in terms of the penalty parameter and the data of the equation. The penalised problem is discretised in space and time by means of a Discontinuous Galerkin method. We prove an a posteriori error estimate in the space-time maximum norm involving a residual and the stability property of a linearised adjoint problem.

1. Introduction

In this note we study numerical solution of the time dependent obstacle problem by means of a finite element method. Our work is based on a penalty formulation of the problem and concerns an a posteriori error estimate in the maximum norm. In our context the penalty method consists of the introduction of a penalised problem, a certain nonlinear partial differential equation involving a penalty parameter ϵ, whose solution converges to the solution of the time dependent obstacle problem as ϵ tends to zero. The penalty problem is approximated by means of a finite element method.

Using this approach Scholz [13], [14] proved optimal a priori error estimates in the energy norm, for the stationary and the time dependent obstacle problems. Optimal error estimates in the L_2 norm are not known.

Recently French, Larsson and Nochetto [5] proved an a posteriori error estimate in the maximum norm for the stationary obstacle problem.

In the present work we prove an a posteriori error estimate in the maximum norm for the time dependent obstacle problem. Our analysis allows the penalty parameter ϵ to vary in space and time, which might be useful in adaptive algorithms.

In our method there are two sources for the error. The first part, the penalisation error, comes from the use of the penalty problem. This part is estimated in maximum norm, in terms of the penalty parameter ϵ and data, using an a priori estimate. The second part, the discretisation error, comes from the finite element discretisation of the penalty problem. We use a Discontinuous Galerkin method, see [2] and [3], to discretise the penalty problem in space and time. As in [2] and [3] our a posteriori error estimate of the discretisation error is proved by means of a duality argument, involving a continuous linear adjoint problem. The analysis

1991 *Mathematics Subject Classification.* 65N60, 35K85.

Key words and phrases. parabolic obstacle problem, a posteriori error estimate, residual, pointwise, maximum norm, adaptive, finite element, penalty method, duality argument, maximal regularity, discontinuous Galerkin method.

relies on the regularity of the solution G of the adjoint problem. In our case the regularity of G depends on the choice of L_p-norm. We can only show the regularity estimate we need in the L_1-norm. Therefore we use an L_1 - L_∞ duality argument with G acting as a regularised Green function, a technique introduced in [11] and [5]. In order to prove the regularity estimate for G in the L_1-norm we use a maximal regularity estimate for linear parabolic problems, see [10].

The outline of the paper is as follows:

In Section 2 we state the parabolic obstacle problem and the associated penalised problem.

In Section 3 we introduce time and space discretisations, and formulate the Discontinuous Galerkin method for the penalised problem.

In Section 4 we state and prove our estimate of the penalisation error in the maximum norm in space and time. The error is estimated in terms of the penalty parameter ϵ and data.

In Section 5 we state our maximum norm estimate of the discretisation error. Our estimate allows ϵ to vary in space and time.

In Sections 6-11 we prove the estimate of the discretisation error. In Section 6 we develop an error representation formula expressing part of the error in terms of a residual and the solution of the linearised adjoint problem.

In Section 7 we estimate the residual in terms of computable quantities.

In Section 8 we prove a lemma concerning the regularity of the solution G of the adjoint problem in the case when the initial data is a regularised δ-function. Here G acts as an approximate Green function.

In Section 9 we put things together and conclude the proof of the estimate of the discretisation error.

In Section 10 we prove an a priori estimate in the maximum norm of the gradient of the solution of the penalty problem. This result is used in Section 9.

In Section 11 we extend some of the results in [3] concerning estimates of the L_2-projection with respect to weighted L_p-norms. These results are used in Section 7.

We conclude this section by introducing some notation. Let Ω be a bounded domain in $\mathbf{R}^d$, $d = 1, 2, 3$. We use the standard Lebesgue spaces $L_p(\omega)$ for $\omega \subset \Omega$, with the convention that $L_p = L_p(\Omega)$, and the corresponding Sobolev spaces $W_p^k(\omega)$, $W_p^k = W_p^k(\Omega)$, $H^k = W_2^k$ and $H_0^1 = \{u \in H^1 : u|_{\partial\Omega} = 0\}$. Moreover $(u,v)_\omega = \int_\omega uv\, dx$, $(u,v) = (u,v)_\Omega$. Let X be a Banach space with norm $\|\cdot\|_X$ and let I be an interval. We define $\|u\|_{L_p(I;X)} = (\int_I \|u(t)\|_X^p\, dt)^{1/p}$ for $1 \le p < \infty$ and $\|u\|_{L_\infty(I;X)} = \operatorname{ess\,sup}_I \|u(t)\|_X$ for $p = \infty$. We introduce the notation $D^j v(x) = \sqrt{\sum_{|\alpha|=j} |D^\alpha v(x)|^2}$, so that the W_p^m seminorm may be conveniently written $\|D^m v\|_p$.

2. The time dependent obstacle problem

We assume that Ω is a bounded, convex, polygonal domain in $\mathbf{R}^d$, $d = 1, 2, 3$. The time dependent obstacle problem can be formulated as the following differential inclusion: Find $u = u(x,t)$ such that

$$\begin{aligned} u_t - \Delta u + \omega(u-\psi) &\ni f, && \text{in } \Omega \times [0,T], \\ u &= 0, && \text{in } \partial\Omega \times [0,T], \\ u &= v_0, && \text{in } \Omega \times \{0\}, \end{aligned} \tag{2.1}$$

where the obstacle $\psi = \psi(x,t)$ and $f = f(x,t)$ and $v_0 = v_0(x)$ are given functions with $v_0(x) \geq \psi(x,0)$ for $x \in \Omega$. ω is the maximal monotone graph defined by

$$\omega(s) = \begin{cases} \{0\}, & s > 0, \\ (-\infty, 0], & s = 0, \\ \emptyset, & s < 0. \end{cases}$$

For sufficiently smooth data there exists a unique solution of the parabolic variational inequality (2.1) such that

$$u \in L_p(0,T; W_p^2(\Omega)), \quad u_t \in L_p(0,T; L_p(\Omega)),$$

for any $2 \leq p < \infty$, see Theorem 16.11.2 in [6], see also [7]. As a basis for our finite element analysis we use the following penalty formulation of (2.1): Find $u_\epsilon = u_\epsilon(x,t)$ such that

$$\begin{aligned} u_{\epsilon,t} - \Delta u_\epsilon + \epsilon^{-1}\beta(u_\epsilon) &= f, && \text{in } \Omega \times [0,T], \\ u_\epsilon &= 0, && \text{in } \partial\Omega \times [0,T], \\ u_\epsilon &= v_0, && \text{in } \Omega \times \{0\}, \end{aligned} \tag{2.2}$$

where $\epsilon = \epsilon(x,t)$ is a positive function and

$$\beta(w(x,t)) = (w(x,t) - \psi(x,t))^- = \min(w(x,t) - \psi(x,t), 0).$$

We note that $\beta(v_0) = 0$. For sufficiently smooth data there exists a unique solution of (2.2) such that

$$u_\epsilon \in L_p(0,T; W_p^2(\Omega)), \quad u_{\epsilon,t} \in L_p(0,T; L_p(\Omega)),$$

for any $2 \leq p < \infty$, see the proof of Theorem 16.10.1 in [6], see also [7].

Remark. Existence, uniqueness and regularity for (2.1) and (2.2) is stated and proved in [6] for a smooth domain and with initial value equal to zero. The proof is based on the corresponding result, due to V. A. Solonnikov, for the linear parabolic problem. The results in [6], for (2.1) and (2.2), can be extended to the case of a nonsmooth but convex domain with smooth initial value, by using the same proof as in [6] but based on another regularity result for the linear problem. Such a result can be achieved by combining two facts: Let Ω be a nonsmooth but convex domain. Then there is a maximal regularity result for the inhomogeneous linear parabolic problem with zero initial value, see example 3.2.B in [10]. Further, the Laplace operator Δ with homogeneous Dirichlet boundary condition is the generator of an analytic semigroup on $L_p(\Omega)$ for $1 < p < \infty$. See also Lemma 8.1 below.

3. The discretisation method

In this section we formulate a discretisation of (2.2) in space and time using the Discontinuous Galerkin method, see [2] and [3]. For the discretisation in space let $\mathcal{F} = \{\mathcal{T}\}$ be a family of triangulations, where a triangulation $\mathcal{T} = \{K\}$ is a partition of Ω into open simplices K which are face to face. Let $h_K = \operatorname{diam}(\overline{K})$

and let ρ_K denote the radius of the largest closed ball contained in $\overline{K}$. We assume that $\mathcal{F}$ is nondegenerate, i.e., we assume that there is a constant c_0 such that for all triangulations $\mathcal{T} \in \mathcal{F}$ we have

$$\max_{K\in\mathcal{T}} \frac{h_K}{\rho_K} \leq c_0. \tag{3.1}$$

To each triangulation $\mathcal{T} \in \mathcal{F}$ we associate a positive, piecewise constant function $h(x)$, defined on $\overline{\Omega}$ by

$$h|_K = h_K, \quad \forall K \in \mathcal{T}.$$

We also need a measure of the "regularity" of a triangulation. Therefore we introduce the quantity $\delta = \delta(\mathcal{T})$ as follows. Let $\mathcal{T} \in \mathcal{F}$ be a triangulation and K be simplex in $\mathcal{T}$. We define the set $S_K = \{K' \in \mathcal{T} : \overline{K'} \cap \overline{K} \neq \emptyset\}$ and

$$\delta = \max_{K\in\mathcal{T}} \max_{K'\in S_K} |1 - h_{K'}^2/h_K^2|. \tag{3.2}$$

To each triangulation $\mathcal{T} \in \mathcal{F}$ we have an associated function space $S = S(\mathcal{T})$, consisting of all continuous functions on $\overline{\Omega}$ which are polynomials of degree ≤ 1 on each $K \in T$ and vanish on $\partial\Omega$.

We now consider the discretisation in time. Let $0 = t_0 < \cdots < t_N = T$ be a partition of $[0,T]$ into subintervals $I_n = (t_{n-1}, t_n)$ of lengths $k_n = t_n - t_{n-1}$, and associate with each such time interval a triangulation $\mathcal{T}_n \in \mathcal{F}$ with corresponding function $h_n = h(\mathcal{T}_n)$ and function space $S_n = S(\mathcal{T}_n)$. We define the following function spaces:

$$V_n = \{\varphi : \overline{\Omega} \times I_n \to \mathbf{R} : v \text{ is constant in time and } v(\cdot,t) \in S_n\},$$

$$V = \{\varphi : \varphi|_{\overline{\Omega}\times I_n} \in V_n,\ n = 1, ..., N\}.$$

We discretise (2.2) as follows: Find $U_\epsilon \in V$ such that for $n = 1, 2, ..., N$,

$$\begin{aligned}\int_{I_n} \{(U_{\epsilon,t}, w) + (\nabla U_\epsilon, \nabla w) + (\epsilon^{-1}\beta(U_\epsilon), w)\}\, dt + ([U_\epsilon]_{n-1}, w_{n-1}^+)\\ = \int_{I_n} (f, w)\, dt, \quad \forall w \in V_n,\end{aligned} \tag{3.3}$$

where $U_{\epsilon 0}^- = v_0$ and

$$[\varphi]_n = \varphi_n^+ - \varphi_n^-, \quad \varphi_n^\pm = \lim_{s\to 0^+} \varphi(t_n \pm s).$$

4. The penalisation error

The following theorem estimates the penalty error $u - u_\epsilon$ in the maximum norm. Note that ϵ may depend on x and t and that the estimate is localised to the contact set. The proof is adapted from [5], where the stationary case is studied.

Theorem 4.1. *Let u and u_ϵ be the solutions of* (2.1) *and* (2.2), *respectively. Then*

$$\|u - u_\epsilon\|_{L_\infty(0,T;L_\infty(\Omega))} \leq \|\epsilon(f + \Delta\psi - \psi_t)\chi_{\hat{\Omega}}\|_{L_\infty(0,T;L_\infty(\Omega))}, \tag{4.1}$$

where $\hat{\Omega} = \hat{\Omega}(t) = \{x \in \Omega : u(x,t) - \psi(x,t) = 0,\ u_\epsilon(x,t) - \psi(x,t) \leq 0\}$ is the "contact set", and $\chi_{\hat{\Omega}(t)}$ is the characteristic function of $\hat{\Omega}(t)$.

Proof. In this proof it is convenient to use the notation $\beta_\epsilon(\cdot - \psi) = \beta(\cdot)/\epsilon$. We define

$$\begin{aligned} \Omega^-(t) &= \{x \in \Omega : u(x,t) - \psi(x,t) = 0\}, \\ \Omega^+(t) &= \{x \in \Omega : u(x,t) - \psi(x,t) > 0\}, \\ \Omega_\epsilon^-(t) &= \{x \in \Omega : u_\epsilon(x,t) - \psi(x,t) \le 0\}, \\ \Omega_\epsilon^+(t) &= \{x \in \Omega : u_\epsilon(x,t) - \psi(x,t) > 0\}, \end{aligned} \tag{4.2}$$

so that $\hat{\Omega}(t) = \Omega^-(t) \cap \Omega_\epsilon^-(t)$. Let $v = u - u_\epsilon$. Let $\chi_{\hat{\Omega}(t)}$ be the characteristic function of $\hat{\Omega}(t)$. We will show that

$$\|v(t)\|_{L_q(\Omega)} \le (|\Omega| T)^{1/q} \|\epsilon^{1/p}(f + \Delta\psi - \psi_t)\chi_{\hat{\Omega}}\|_{L_\infty(0,T;L_\infty(\Omega))}, \quad \forall t \in [0,T], \tag{4.3}$$

for all even integers $q \ge 2$, where $1/p + 1/q = 1$. Letting $q \to \infty$ we obtain (4.1).

In order to prove (4.3) we define

$$\begin{aligned} B(x,t) &:= -u_t(x,t) + \Delta u(x,t) + f(x,t) \in \omega(u(x,t) - \psi(x,t)), \\ B_\epsilon(x,t) &:= -u_{\epsilon,t}(x,t) + \Delta u_\epsilon(x,t) + f(x,t) = \beta_{\epsilon(x,t)}(u_\epsilon(x,t) - \psi(x,t)), \end{aligned} \tag{4.4}$$

so that, for any even integer $q \ge 2$,

$$(v_t, v^{q-1}) + (\nabla v, \nabla v^{q-1}) = (B_\epsilon - B, v^{q-1}),$$

where

$$(\nabla v, \nabla v^{q-1}) = (q-1)\|v^{-1+q/2}\nabla v\|_{L_2}^2 = \frac{4}{qp}\|\nabla(v^{q/2})\|_{L_2}^2.$$

Hence

$$\frac{1}{q}\frac{d}{dt}\|v\|_{L_q}^q + \frac{4}{qp}\|\nabla(v^{q/2})\|_{L_2}^2 = (B_\epsilon - B, v^{q-1}). \tag{4.5}$$

We now turn to the right hand side of (4.5). We first show that

$$(B_\epsilon - B, v^{q-1}) \le (B_\epsilon - B, v^{q-1})_{\hat{\Omega}(t)}. \tag{4.6}$$

This follows from (4.4) and the monotonicity of the graphs ω and $\beta_{\epsilon(x,t)}$. More precisely, if $x \in \Omega^+(t)$, then $B(x,t) = 0 = \beta_{\epsilon(x,t)}(u(x,t) - \psi(x,t))$, so that

$$(B_\epsilon - B)v^{q-1} = -(\beta_\epsilon(u_\epsilon - \psi) - \beta_\epsilon(u - \psi))(u_\epsilon - u)v^{q-2} \le 0 \quad \text{in } \Omega^+(t).$$

Here we used the monotonicity $(\beta_\epsilon(\xi) - \beta_\epsilon(\eta))(\xi - \eta) \ge 0$, and the assumption that q is an even integer. Similarly, if $x \in \Omega_\epsilon^+(t)$, then $B_\epsilon(x,t) = 0 \in \{0\} = \omega(u_\epsilon(x,t) - \psi(x,t))$, so that

$$(B_\epsilon - B)v^{q-1} \in -(\omega(u_\epsilon - \psi) - \omega(u - \psi))(u_\epsilon - u)v^{q-2} \subset \mathbf{R}^- \quad \text{in } \Omega_\epsilon^+(t),$$

since $(\omega(\xi) - \omega(\eta))(\xi - \eta) \subset \mathbf{R}^+$. Therefore, $(B_\epsilon - B, v^{q-1})_{\Omega^+(t)\cup\Omega_\epsilon^+(t)} \le 0$, and (4.6) follows.

It now remains to bound the right hand side of (4.6). In order to do so, we note that in $\Omega^-(t)$ we have $u = \psi$, so that $v = \psi - u_\epsilon$, and $\Delta u = \Delta\psi$ a.e., so that $B = f + \Delta\psi - \psi_t$ a.e. in $\Omega^-(t)$. Moreover, in $\Omega_\epsilon^-(t)$ we have $B_\epsilon = -\epsilon^{-1}(\psi - u_\epsilon)$ in view of (4.4). Summing up: in $\hat{\Omega}(t) = \Omega^-(t) \cap \Omega_\epsilon^-(t)$ we have $B = f + \Delta\psi - \psi_t$, $B_\epsilon = -\epsilon^{-1}v$, so that

$$\begin{aligned} (B_\epsilon - B, v^{q-1})_{\hat{\Omega}(t)} &= (-\epsilon^{-1}v - (f + \Delta\psi - \psi_t), v^{q-1})_{\hat{\Omega}(t)} \\ &= -\|\epsilon^{-1/q}v\|_{L_q(\hat{\Omega}(t))}^q - (f + \Delta\psi - \psi_t, v^{q-1})_{\hat{\Omega}(t)}. \end{aligned} \tag{4.7}$$

Using Hölder's and Young's inequalities

$$|(f,g)| \le \|f\|_{L_q}\|g\|_{L_p} \le \frac{1}{q}\|f\|_{L_q}^q + \frac{1}{p}\|g\|_{L_p}^p,$$

we get

$$\begin{aligned} |(f+\Delta\psi-\psi_t, v^{q-1})_{\hat{\Omega}(t)}| &= |(\epsilon^{1/p}(f+\Delta\psi-\psi_t), (\epsilon^{-1/q}v)^{q/p})_{\hat{\Omega}(t)}| \\ &\le \frac{1}{q}\|\epsilon^{1/p}(f+\Delta\psi-\psi_t)\|_{L_q(\hat{\Omega}(t))}^q \\ &\quad + (1-\frac{1}{q})\|\epsilon^{-1/q}v\|_{L_q(\hat{\Omega}(t))}^q. \end{aligned} \tag{4.8}$$

Combining (4.5), (4.6), (4.7) and (4.8), and multiplying by q gives

$$\frac{d}{dt}\|v\|_{L_q}^q + \frac{4}{p}\|\nabla(v^{q/2})\|_{L_2}^2 + \|\epsilon^{-1/q}v\|_{L_q(\hat{\Omega}(t))}^q \le \|\epsilon^{1/p}(f+\Delta\psi-\psi_t)\|_{L_q(\hat{\Omega}(t))}^q.$$

We integrate in time and use the fact that $v(0)=0$, to find

$$\begin{aligned} \|v(t)\|_{L_q(\Omega)}^q &\le \int_0^t \|\epsilon^{1/p}(f+\Delta\psi-\psi_t)\chi_{\hat{\Omega}}\|_{L_q(\Omega)}^q ds \\ &\le |\Omega|T\|\epsilon^{1/p}(f+\Delta\psi-\psi_t)\chi_{\hat{\Omega}}\|_{L_\infty(0,T;L_\infty(\Omega))}^q, \end{aligned}$$

which proves (4.3). □

5. The discretisation error

In the proof of our estimate of the discretisation error we use a crude a priori bound for $\|\nabla u_\epsilon\|_{L_\infty(\Omega)}$. In order to use this a priori estimate we need some additional assumptions on the penalty parameter $\epsilon = \epsilon(x,t)$. Let

$$\begin{aligned} h_{\min} &= \min_{1\le n\le N} \inf_{x\in\Omega} h_n(x), \\ h_{\max} &= \max_{1\le n\le N} \sup_{x\in\Omega} h_n(x). \end{aligned}$$

Let

$$\Omega_\epsilon^-(t) = \{x\in\Omega : u_\epsilon(x,t) - \psi(x,t) \le 0\},$$

and let

$$\begin{aligned} \epsilon_{\min} &= \inf_{t\in[0,T]} \inf_{x\in\Omega} \epsilon(x,t), \\ \epsilon_{\max} &= \sup_{t\in[0,T]} \sup_{x\in\Omega} \epsilon(x,t), \\ \epsilon_{t,\max} &= \sup_{t\in[0,T]} \sup_{x\in\Omega} |\epsilon_t(x,t)|. \end{aligned}$$

Let C be a positive constant and let $\zeta \ge 0$. We assume that

$$\begin{aligned} \epsilon_{t,\max} &\le C, \\ \frac{\epsilon_{\max}}{\epsilon_{\min}} &\le C h_{\min}^{-\zeta}, \\ \epsilon_{\min} &\ge h_{\min}^2. \end{aligned} \tag{5.1}$$

These assumptions are only used in Section 9 together with the a priori bound on $\|\nabla u_\epsilon\|_{L_\infty(\Omega)}$. It is likely that assumption (5.1) may be relaxed by using a better a priori estimate of u_ϵ, e.g., $\|u_\epsilon\|_{C^\alpha(\bar{(}\Omega))}$.

In order to state our estimate we introduce computable residuals. Let $\chi, \eta : I_n \times \Omega \to \mathbf{R}$ be arbitrary functions such that χ is constant in time on I_n and $\eta(t) \in V_n$ for $t \in I_n$. Let $[\partial_\nu U_\epsilon]$ denote the jump across ∂K in the outward normal derivative and let $*$ indicate that a term is not present if $V_{n-1} \subset V_n$. We define the following computable residuals:

$$
\begin{aligned}
R^t_{n,e} &= k_n^{-1}[U_\epsilon]_{n-1}, \\
R^t_{n,i} &= k_n^{-1}\int_{I_n} |\epsilon^{-1}\beta(U_\epsilon) - f - \chi|\,dt, \\
R^x_{n,e}|_K &= \|h_K^{-1}[\partial_\nu U^-_{\epsilon,n}]\|_{L_\infty(\partial K\setminus\partial\Omega)}, \quad K \in \mathcal{T}, \\
R^x_{n,i}|_{I_n} &= k_n^{-1}\int_{I_n} (k_n^{-1}[U_\epsilon]^*_{n-1} + \epsilon^{-1}\beta(U_\epsilon) - f - \eta)\,dt,
\end{aligned}
\tag{5.2}
$$

where index e and i refer to the edge part and the interior part of the residual and x and t refers to space and time discretisation. We now state our main theorem concerning the discretisation error. Recall that δ and ζ are defined by (3.2) and (5.1).

Theorem 5.1. *Let u_ϵ and U_ϵ be the solutions of* (2.2) *and* (3.3). *Let $\xi \geq \zeta$ be arbitrary. If δ is sufficiently small, then*

$$
\begin{aligned}
\max_{1\le n\le N} \|u_\epsilon(t_n) - U_{\epsilon n}^-\|_{L_\infty(\Omega)} \le C_1 h_{\min}^{\xi-\zeta}(1 + |\log h_{\min}|) \\
+ C_2 L_N \max_{1\le n\le N}\Big(k_n\|R^t_{n,e}\|_{L_\infty(\Omega)} + k_n\|R^t_{n,i}\|_{L_\infty(\Omega)} \\
+ \|h_n^2 R^x_{n,e}\|_{L_\infty(\Omega)} + \|h_n^2 R^x_{n,i}\|_{L_\infty(\Omega)}\Big),
\end{aligned}
$$

where

$$
L_N = \Big(1 + \log\frac{t_N}{k_N}\Big)\Big(\log\frac{t_N}{k_N} + \xi|\log h_{\min}|\Big)^2.
$$

Remark 5.1.1. The term $h_{\min}^{\xi-\zeta}(1+|\log h_{\min})|$ comes from the use of an approximate Green function in the proof of Theorem 5.1. We note that this term can be made arbitrary small by choosing ξ sufficiently large.

Remark 5.1.2. In our proof of Theorem 5.1 we use estimates of the L_2-projection in a weighted L_p-norm, in which δ is required to be small. It is possible to use other interpolants, for example, the Lagrange interpolant. In that case there is no condition on δ, but the residual $R^x_{n,i}$ is different: the star on $[U_\epsilon]_{n-1}$ and the term $\eta(t)$ have to be removed.

Remark 5.1.3. Let $e_N^- = (u_\epsilon - U_\epsilon)_N^-$ be the error at time t_N. The strategy of the proof of Theorem 5.1 is to derive an estimate of the form $\|e_N^-\|_{L_\infty} \le C\rho\|\nabla u_\epsilon(t_N)\|_{L_\infty} + C|(e_N^-, g)|$, where ρ can be chosen small and g is a regularised δ-function. $\|\nabla u_\epsilon\|_{L_\infty}$ is bounded by an a priori estimate. The term $|(e_N^-, g)|$ is estimated by means of an L_1-L_∞ duality argument involving a linearised adjoint problem, combined with regularity estimates of the solution of the adjoint problem.

6. The duality argument

In this section we derive an error representation formula expressing the quantity (e_N^-, g) in terms of a residual $r(U_\epsilon)$ and the solution $G(x,t)$ of a certain adjoint problem with data g.

In this section we use the notation $\beta_\epsilon(\cdot - \psi) = \beta(\cdot)/\epsilon$. In order to find the error representation formula we note that the equations (3.3) defining the finite element solution U_ϵ can be written in compact form as

$$\begin{aligned} &\sum_{n=1}^{N} \int_{I_n} \{(U_{\epsilon,t}, w) + (\nabla U_\epsilon, \nabla w) + (\beta_\epsilon(U_\epsilon - \psi), w)\}\, dt \\ &\quad + \sum_{n=1}^{N} ([U_\epsilon]_{n-1}, w_{n-1}^+) = \int_0^{t_N} (f, w)\, dt, \quad \forall w \in V, \end{aligned} \tag{6.1}$$

where $U_{\epsilon 0}^- = v_0$. Let

$$\begin{aligned} \mathcal{V} = \{w : w|_{I_n} &\text{ is smooth in time }, w(t) \in H_0^1(\Omega) \\ &\text{and } w_n^\pm \in H_0^1(\Omega) \text{ exist}\}. \end{aligned} \tag{6.2}$$

We note that the solution u_ϵ of (2.2) satisfies

$$\begin{aligned} &\sum_{n=1}^{N} \int_{I_n} \{(u_{\epsilon,t}, w) + (\nabla u_\epsilon, \nabla w) + (\beta_\epsilon(u_\epsilon - \psi), w)\}\, dt \\ &\quad + \sum_{n=1}^{N} ([u_\epsilon]_{n-1}, w_{n-1}^+) = \int_0^{t_N} (f, w)\, dt, \quad \forall w \in \mathcal{V}. \end{aligned} \tag{6.3}$$

We define the residual $r(U_\epsilon)$ of U_ϵ as a linear functional on $\mathcal{V}$:

$$\langle r(U_\epsilon), w\rangle = \sum_{n=1}^{N} \langle r(U_\epsilon), w\rangle_n, \quad \forall w \in \mathcal{V}, \tag{6.4}$$

where

$$\begin{aligned} \langle r(U_\epsilon), w\rangle_n = &\int_{I_n} \{(U_{\epsilon,t}, w) + (\nabla U_\epsilon, \nabla w) + (\beta_\epsilon(U_\epsilon - \psi) - f, w)\}\, dt \\ &+ ([U_\epsilon]_{n-1}, w_{n-1}^+). \end{aligned} \tag{6.5}$$

We note that by (6.1)

$$\langle r(U_\epsilon), w\rangle = 0, \quad \forall w \in V. \tag{6.6}$$

We define the bilinear form $E(\cdot, \cdot)$ by

$$\begin{aligned} E(\xi, \eta) = &\sum_{n=1}^{N} \int_{I_n} \{(\xi_t, \eta) + (\nabla \xi, \nabla \eta) + (b\xi, \eta)\}\, dt \\ &+ \sum_{n=1}^{N} ([\xi]_{n-1}, \eta_{n-1}^+), \quad \forall \xi, \eta \in \mathcal{V}, \end{aligned} \tag{6.7}$$

where $b = b(x,t)$ is defined by

$$b = \frac{\beta_\epsilon(u_\epsilon - \psi) - \beta_\epsilon(U_\epsilon - \psi)}{u_\epsilon - U_\epsilon}.$$

We note that

$$0 \le b(x,t) \le \frac{1}{\epsilon(x,t)} \le \frac{1}{\epsilon_{\min}}. \tag{6.8}$$

Let the error e be defined by $e = u_\epsilon - U_\epsilon$. Combining (6.3), (6.4), and (6.7) gives

$$E(e, w) = -\langle r(U_\epsilon), w\rangle, \quad \forall w \in \mathcal{V}. \tag{6.9}$$

Let us now consider the following adjoint problem: Given g, find $G = G(x,t)$ such that

$$\begin{aligned} -G_t - \Delta G + bG &= 0, \quad &&\text{in } \Omega \times [0, t_N], \\ G &= 0, &&\text{in } \partial\Omega \times [0, t_N], \\ G &= g, &&\text{in } \Omega \times \{t_N\}. \end{aligned} \tag{6.10}$$

We multiply (6.10) by $w \in \mathcal{V}$ and integrate to find

$$\begin{aligned} 0 &= \sum_{n=1}^{N} \int_{I_n} \{(w, -G_t - \Delta G + bG)\}\, dt \\ &= \sum_{n=1}^{N} \int_{I_n} \{(w_t, G) + (\nabla w, \nabla G) + (bw, G)\}\, dt \\ &\quad + \sum_{n=1}^{N} ([w]_{n-1}, G_{n-1}) + (w_0^-, G_0) - (w_N^-, G_N), \end{aligned}$$

where the second equality follows from a integration by parts in space and time and the continuity of G. We thus have, using also $G_N = g$,

$$(w_N^-, g) - (w_0^-, G_0) = E(w, G). \tag{6.11}$$

Since $U_{\epsilon,0}^- = v_0$, we have $e_0^- = 0$, so by choosing $w = e$ in (6.11) we get

$$(e_N^-, g) = E(e, G).$$

Using also (6.9) with $w = G$ we finally arrive at the following error representation formula

$$(e_N^-, g) = -\langle r(U_\epsilon), G\rangle. \tag{6.12}$$

7. An estimate of the residual

In this section we state and prove an estimate of the residual $\langle r(U_\epsilon), v\rangle$ in terms of the computable residuals defined in (5.2) and derivatives of v. Our result is based on the use of the approximation properties of the L_2-projection with respect to weighted L_1-norms, which we prove in Section 11. We note that it is here we need a "regularity" condition on the triangulation, measured by the quantity δ, defined in (3.2). The condition on δ is not needed if we use another interpolant, for example, the Lagrange interpolant. However, this would change the computable residuals as described in Remark 5.1.2.

Lemma 7.1. *Let $1 \le n \le N$ and let $r(U_\epsilon)$ be defined as in (6.5) and $R_{n,e}^t$, $R_{n,i}^t$, $R_{n,e}^x$, $R_{n,i}^x$ be defined as in (5.2). Let $v : \Omega \times I_n \to \mathbf{R}$ be a smooth function with $v(\cdot,t)|_{\partial\Omega} = 0$. For sufficiently small δ we have*

$$\begin{aligned} |\langle r(U_\epsilon), v\rangle_n| \le{} & \Big(k_n \|R_{n,e}^t\|_{L_\infty(\Omega)} + k_n \|R_{n,i}^t\|_{L_\infty(\Omega)}\Big) \\ & \times \min\Big(2\|v\|_{L_\infty(I_n;L_1(\Omega))}, \|v_t\|_{L_1(I_n;L_1(\Omega))}\Big) \\ & + C\Big(\|h_n^2 R_{n,e}^x\|_{L_\infty(\Omega)} + \|h_n^2 R_{n,i}^x\|_{L_\infty(\Omega)}\Big) \Big\|D^2 \int_{I_n} v\, dt\Big\|_{L_1(\Omega)}. \end{aligned}$$

Proof. In this proof we use the notation $\beta_\epsilon = \beta(U_\epsilon)/\epsilon$. Let P_n denote the orthogonal projection of $L_2(\Omega)$ onto S_n, i.e., if $\varphi \in L_2(\Omega)$, then $P_n\varphi \in S_n$ is defined by

$$(P_n\varphi, \chi) = (\varphi, \chi), \quad \forall \chi \in S_n.$$

Let J_n denote the L_2 projection in time onto functions, which are constant in time on I_n. Hence $J_n v = \frac{1}{k_n}\int_{I_n} v(t)\,dt$. By the orthogonality property (6.6) of $r(U_\epsilon)$ it follows that

$$\begin{aligned}\langle r(U_\epsilon), v\rangle_n &= \langle r(U_\epsilon), v - J_nP_nv\rangle_n \\ &= \langle r(U_\epsilon), v - J_nv\rangle_n + \langle r(U_\epsilon), J_nv - J_nP_nv\rangle_n.\end{aligned}$$

We first estimate $\langle r(U_\epsilon), v - J_nv\rangle_n$. Since U_ϵ is constant on I_n we have

$$\int_{I_n} (U_{\epsilon,t}, v - J_nv)\,dt = 0,$$

and

$$\int_{I_n} (\nabla U_\epsilon, \nabla(v - J_nv))\,dt = \Big(\nabla U_\epsilon, \nabla \int_{I_n} (v - J_nv)\,dt\Big) = 0.$$

By Hölder's inequality we get

$$|([U_\epsilon]_{n-1}, (v - J_nv)^+_{n-1})| \le \|[U_\epsilon]_{n-1}\|_{L_\infty(\Omega)}\|v^+_{n-1} - J_nv\|_{L_1(\Omega)}.$$

Let $\chi : I_n \times \Omega \to \mathbf{R}$ be any function such that $\chi|_{I_n}$ is constant in time. Then

$$\Big|\int_{I_n} (\beta_\epsilon - f, v - J_nv)\,dt\Big| \le \|\beta_\epsilon - f - \chi\|_{L_1(I_n;L_\infty)}\|v - J_nv\|_{L_\infty(I_n;L_1)}.$$

But

$$\begin{aligned}\|v^+_{n-1} - J_nv\|_{L_1} &\le \|v - J_nv\|_{L_\infty(I_n;L_1)} \\ &\le \min(2\|v\|_{L_\infty(I_n;L_1)}, \|v_t\|_{L_1(I_n;L_1)}),\end{aligned}$$

so that

$$\begin{aligned}|\langle r(U_\epsilon), v - J_nv\rangle| \le (k_n\|R^t_{n,e}\|_{L_\infty} &+ k_n\|R^t_{n,i}\|_{L_\infty}) \\ &\times \min(2\|v\|_{L_\infty(I_n;L_1)}, \|v_t\|_{L_1(I_n;L_1)}).\end{aligned}$$

In order to estimate $\langle r(U_\epsilon), J_nv - J_nP_nv\rangle_n$, we first note that

$$\int_{I_n} (U_{\epsilon,t}, J_nv - J_nP_nv)\,dt = 0.$$

Let $w = (I - P_n)\int_{I_n} v\,dt$. Then

$$\begin{aligned}\int_{I_n} (\nabla U_\epsilon, \nabla(J_nv - J_nP_nv)) &= (\nabla U_\epsilon, \nabla w) = \sum_{K\in\mathcal{T}} (\nabla U_\epsilon, \nabla w)_K \\ &= \sum_{K\in\mathcal{T}} \{-(\Delta U_\epsilon, w)_K + (\partial_\nu U_\epsilon, w)_{\partial K}\} \\ &= \sum_{K\in\mathcal{T}} (\partial_\nu U_\epsilon, w)_{\partial K} = \frac{1}{2}\sum_{K\in\mathcal{T}} ([\partial_\nu U_\epsilon], w)_{\partial K\setminus\partial\Omega},\end{aligned} \tag{7.1}$$

where $[\partial_\nu U_\epsilon]$ denotes the jump of the outward normal derivative and where we used that $\Delta U_\epsilon|_K = 0$, since U_ϵ is piecewise linear. By the trace inequality

$$\|w\|_{L_1(\partial K)} \le C\Big(h_K^{-1}\|w\|_{L_1(K)} + \|Dw\|_{L_1(K)}\Big),$$

and since $h_n|_K = h_K$, we have that the right hand side of (7.1) is less than or equal to

$$
\begin{aligned}
&C \max_{K\in\mathcal{T}} \|h_K[\partial_\nu U_\epsilon]\|_{L_\infty(\partial K\setminus\partial\Omega)} \sum_{K\in\mathcal{T}} \Big(h_K^{-2}\|w\|_{L_1(K)} + h_K^{-1}\|Dw\|_{L_1(K)} \Big) \\
&\quad \le C \max_{K\in\mathcal{T}} \|h_n^2 h_K^{-1}[\partial_\nu U_\epsilon]\|_{L_\infty(\partial K\setminus\partial\Omega)} \Big(\|h_n^{-2}w\|_{L_1(\Omega)} + \|h_n^{-1}Dw\|_{L_1(\Omega)} \Big).
\end{aligned}
$$

Finally, let $\eta(t) \in V_n$. We then have by Hölder's inequality that

$$
\begin{aligned}
([U_\epsilon]_{n-1}, &J_n(I-P_n)v) + \int_{I_n} (\beta_\epsilon - f, J_n(I-P_n)v)\,dt \\
&= \int_{I_n} (k_n^{-1}[U_\epsilon]^*_{n-1} + \beta_\epsilon - f - \eta, (I-P_n)J_n v)\,dt \\
&\le \Big\| h_n^2 k_n^{-1} \int_{I_n} (k_n^{-1}[U_\epsilon]^*_{n-1} + \beta_\epsilon - f - \eta)\,dt \Big\|_{L_\infty(\Omega)} \|h_n^{-2}w\|_{L_1(\Omega)}.
\end{aligned}
$$

By Lemma 11.5 we have for functions $\varphi \in H_0^1(\Omega) \cap W_1^2(\Omega)$, and for sufficiently small δ, that

$$
\begin{aligned}
\|h_n^{-2}(I-P_n)\varphi\|_{L_1(\Omega)} &\le C\|D^2\varphi\|_{L_1(\Omega)}, \\
\|h_n^{-1}D(I-P_n)\varphi\|_{L_1(\Omega)} &\le C\|D^2\varphi\|_{L_1(\Omega)}.
\end{aligned}
$$

This concludes the proof. □

8. The regularity of the solution of the adjoint problem.

In this section we prove regularity estimates for the solution G of (6.10) in the case when the data g is a regularised δ-function.

Our strategy is as follows: We first prove estimates in L_p, for the solution u of a linear, inhomogeneous, initial value problem using maximal regularity results from [10]. In our application it is important to keep track of the p-dependence as p tends to 1. The problem reads as follows: Let $1 < p \le 2$, let $f \in L_p(0,T;L_p)$ and let $g \in L_p(\Omega)$. Let u be the solution, in the $L_p(0,T;L_p)$ sense, of

$$
\begin{aligned}
u_t - \Delta u &= f, \quad \text{in } \Omega\times[0,T], \\
u &= 0, \quad \text{in } \partial\Omega\times[0,T], \\
u &= g, \quad \text{in } \Omega\times\{0\}.
\end{aligned}
\tag{8.1}
$$

We then apply these estimates to the adjoint problem (6.10) with $f = -bG$ and with the data g chosen to be a regularised δ-function. A key step in this part is to establish the estimate $\|bG\|_{L_p(0,t_N;L_p)} \le \epsilon_{\min}^{-1/p'}\|g\|_{L_p}$, where $p' = p/(p-1)$.

We now state the estimate of the solution u of (8.1).

Lemma 8.1. *Let $1 < p \le 2$, let $p' = p/(p-1)$ and let u be the solution of* (8.1). *Then, for $0 < \tau < T$,*

$$
\int_\tau^T \|u_t\|_{L_1}\,dt \le Cp'\Big(\log\frac{T}{\tau}\|g\|_{L_p} + T^{1/p'}\|f\|_{L_p(0,T;L_p)} \Big), \tag{8.2}
$$

$$
\int_\tau^T \|D^2u\|_{L_1}\,dt \le Cp'^2\Big(\log\frac{T}{\tau}\|g\|_{L_p} + T^{1/p'}\|f\|_{L_p(0,T;L_p)} \Big), \tag{8.3}
$$

$$
\Big\| D^2 \int_0^\tau u\,dt \Big\|_{L_1} \le Cp'\Big(\|g\|_{L_p} + p'\tau^{1/p'}\|f\|_{L_p(0,T;L_p)} \Big). \tag{8.4}
$$

Proof. We write $u = u_1 + u_2$, where

$$\begin{aligned} u_{1,t} - \Delta u_1 &= 0, \quad \text{in } \Omega \times [0,T], \\ u_1 &= 0, \quad \text{in } \partial\Omega \times [0,T], \\ u_1 &= g, \quad \text{in } \Omega \times \{0\}, \end{aligned} \tag{8.5}$$

and

$$\begin{aligned} u_{2,t} - \Delta u_2 &= f, \quad \text{in } \Omega \times [0,T], \\ u_2 &= 0, \quad \text{in } \partial\Omega \times [0,T], \\ u_2 &= 0, \quad \text{in } \Omega \times \{0\}. \end{aligned}$$

By using Hölder's inequality we get

$$\int_\tau^T \|u_t\|_{L_1}\,dt \le C\int_\tau^T \|u_t\|_{L_p}\,dt \le C\int_\tau^T (\|u_{1,t}\|_{L_p} + \|u_{2,t}\|_{L_p})\,dt.$$

By the analyticity of the semigroup generated by $-\Delta$ we have, for $1 < p \le 2$,

$$\|u_1(t)\|_{L_p} \le C\|g\|_{L_p}, \quad \|u_{1,t}(t)\|_{L_p} \le Cp't^{-1}\|g\|_{L_p}, \quad t > 0, \tag{8.6}$$

where $p' = p/(p-1)$. The p-dependence in (8.6) follows by keeping track of p in the proof of Theorem 7.3.6 in [12]. We thus obtain

$$\int_\tau^T \|u_{1,t}\|_{L_p}\,dt \le Cp'\int_\tau^T t^{-1}\|g\|_{L_p}\,dt \le Cp'\log\left(\frac{T}{\tau}\right)\|g\|_{L_p}. \tag{8.7}$$

In order to estimate $\int_\tau^T \|u_{2,t}\|_{L_p}\,dt$ we will use the following estimate, which follows from Theorem 3.1 in [10]:

$$\|u_2\|_{L_p(0,T;L_p)} + \|u_{2,t}\|_{L_p(0,T;L_p)} + \|\Delta u_2\|_{L_p(0,T;L_p)} \le Cp'\|f\|_{L_p(0,T;L_p)}. \tag{8.8}$$

The p-dependence in (8.8) follows from the use of the Marcinkiewicz interpolation theorem in its proof. We now use Hölder's inequality in time and (8.8), which give

$$\int_\tau^T \|u_{2,t}\|_{L_p}\,dt \le T^{1/p'}\|u_{2,t}\|_{L_p(0,T;L_p)} \le Cp'T^{1/p'}\|f\|_{L_p(0,T;L_p)}.$$

Together with (8.7) this proves (8.2). We recall the following elliptic regularity estimate: If Ω is a smooth domain or a convex domain, then

$$\|D^2 v\|_{L_p} \le Cp'\|\Delta v\|_{L_p}, \quad \forall v \in W_p^2 \cap H_0^1, \quad 1 < p \le 2. \tag{8.9}$$

The p-dependence in (8.9) is classical in the case of a smooth domain. In the case of a convex domain we argue as in [5]. Let $v = Tf$ be the solution of the Dirichlet problem $-\Delta v = f$ in Ω, $v = 0$ on $\partial\Omega$, and let D_{ij} be a partial derivative. It is well known [9] that the operator $D_{ij}T$ is bounded on L_2, i.e., it is strong type $(2,2)$; this is the case $p = 2$ of (8.9). Moreover, $D_{ij}T$ is weak type $(1,1)$; this is an unpublished result of Dahlberg, Verchota, and Wolff, a proof can be found in [8] and a generalisation in [4]. An application of the Marcinkiewicz interpolation theorem now yields (8.9).

By Hölder's inequality and (8.9) it follows that

$$\begin{aligned} \int_\tau^T \|D^2 u\|_{L_1}\,dt &\le C|\Omega|^{1/p'}p'\int_\tau^T \|\Delta u\|_{L_p}\,dt \\ &\le Cp'\int_\tau^T (\|\Delta u_1\|_{L_p} + \|\Delta u_2\|_{L_p})\,dt, \end{aligned}$$

where by (8.5) and (8.7)

$$\int_\tau^T \|\Delta u_1\|_{L_p}\,dt = \int_\tau^T \|u_{1,t}\|_{L_p}\,dt \le Cp' \log\Big(\frac{T}{\tau}\Big)\|g\|_{L_p},$$

and by (8.8)

$$\int_\tau^T \|\Delta u_2\|_{L_p}\,dt \le CT^{1/p'}\|\Delta u_2\|_{L_p(0,T;L_p)} \le Cp'T^{1/p'}\|f\|_{L_p(0,T;L_p)}.$$

This proves (8.3). Finally, by (8.9), (8.5), (8.6) and (8.8),

$$\begin{aligned}\Big\|D^2\int_0^\tau u\,dt\Big\|_{L_1} &\le Cp'\Big\|\Delta\int_0^\tau u\,dt\Big\|_{L_p}\\ &\le Cp'\Big(\Big\|\Delta\int_0^\tau u_1\,dt\Big\|_{L_p} + \Big\|\Delta\int_0^\tau u_2\,dt\Big\|_{L_p}\Big)\\ &\le Cp'\Big(\Big\|\int_0^\tau u_{1,t}\,dt\Big\|_{L_p} + \int_0^\tau \|\Delta u_2\|_{L_p}\,dt\Big)\\ &\le Cp'\big(2\|u_1\|_{L_\infty(0,\tau;L_p)} + \tau^{1/p'}\|\Delta u_2\|_{L_p(0,\tau;L_p)}\big)\\ &\le Cp'\big(2\|g\|_{L_p} + p'\tau^{1/p'}\|f\|_{L_p(0,T;L_p)}\big),\end{aligned}$$

which proves (8.4). □

We now return to the adjoint problem (6.10) and we have to be more precise concerning the data g. Let $x_0 \in \Omega$ be given and let g be such that

$$\int_{\mathbf{R}^d} g(x)\,dx = 1; \quad \operatorname{supp} g \subset \mathcal{B}(x_0;\rho); \quad 0 \le g(x) \le C\rho^{-d}. \tag{8.10}$$

Here $\mathcal{B}(x_0;\rho)$ denotes the closed ball with center at x_0 and with small radius ρ to be chosen.

We are now able to state and prove the main result of this section.

Lemma 8.2. *Let G be the solution of* (6.10) *with g as in* (8.10)*. For any $\alpha > 0$ there is a constant C such that if $\rho \le \epsilon_{\min}^\alpha$, then*

$$\begin{aligned}\|G\|_{L_\infty(0,t_N;L_1)} &\le 1,\\ \|G_t\|_{L_1(0,t_{N-1};L_1)} &\le C\Big(1+\log\frac{t_N}{k_N}\Big)\log\frac{t_N}{k_N\rho},\\ \|D^2G\|_{L_1(0,t_{N-1};L_1)} &\le C\Big(1+\log\frac{t_N}{k_N}\Big)\Big(\log\frac{t_N}{k_N\rho}\Big)^2,\\ \Big\|D^2\int_{t_{N-1}}^{t_N} G(t)\,dt\Big\|_{L_1} &\le C\Big(\log\frac{t_N}{k_N\rho}\Big)^2.\end{aligned}$$

Proof. It is convenient to use the change of variable $t \to t_N - t$. Note that this also reverses the mesh so that I_N becomes I_1 and k_N becomes k_1. In this variable the adjoint problem (6.10) takes the form:

$$\begin{aligned}G_t - \Delta G + bG = 0, &\quad \text{in } \Omega\times[0,t_N],\\ G = 0, &\quad \text{in } \partial\Omega\times[0,t_N],\\ G = g, &\quad \text{in } \Omega\times\{0\}.\end{aligned} \tag{8.11}$$

Let first g be arbitrary. We will prove that

$$\begin{aligned} &\|G\|_{L_\infty(0,t_N;L_1)} \le \|g\|_{L_1}, \\ &\|bG\|_{L_p(0,t_N;L_p)} \le \epsilon_{\min}^{-1/p'} \|g\|_{L_p}, \quad 1 \le p \le 2, \end{aligned} \tag{8.12}$$

where $p' = p/(p-1)$. From Lemma 8.1 with $f = -bG$ we then obtain,

$$\begin{aligned} \int_{k_1}^{t_N} \|G_t\|_{L_1}\,dt &\le Cp'\Big(\log\frac{t_N}{k_1} + t_N^{1/p'}\epsilon_{\min}^{-1/p'}\Big)\|g\|_{L_p} \\ &\le Cp'\Big(\log\frac{t_N}{k_1} + \Big(\frac{t_N}{k_1}\Big)^{1/p'}\Big)\epsilon_{\min}^{-1/p'}\|g\|_{L_p}, \quad 1 < p \le 2, \end{aligned} \tag{8.13}$$

since $k_{\max} \le C$ and $\epsilon_{\max} \le C$.

With g as in (8.10) we have from a direct calculation

$$\|g\|_{L_p} \le C\rho^{-d/p'},$$

and since $\rho \le \epsilon_{\min}^{\alpha}$ we conclude from (8.13) with $p' = \log\frac{t_N}{k_1\rho} \ge 2$, (note that $p' \ge 2$ if $\rho \le e^{-2}$), that

$$\begin{aligned} \int_{k_1}^{t_N} \|G_t\|_{L_1}\,dt &\le Cp'\Big(\log\frac{t_N}{k_1} + \Big(\frac{t_N}{k_1}\Big)^{1/p'}\Big)\rho^{-c/p'} \\ &\le C\Big(\Big(\frac{t_N}{k_1}\Big)^{-c/p'}\log\frac{t_N}{k_1} + \Big(\frac{t_N}{k_1}\Big)^{(1-c)/p'}\Big)p'\Big(\frac{t_N}{k_1\rho}\Big)^{c/p'} \\ &\le C\Big(\log\frac{t_N}{k_1} + 1\Big)\log\frac{t_N}{k_1\rho}, \end{aligned}$$

where $c = 1/\alpha + d$. Similarly we obtain

$$\int_{k_1}^{t_N} \|D^2G\|_{L_1}\,dt \le C\Big(\log\frac{t_N}{k_1} + 1\Big)\Big(\log\frac{t_N}{k_1\rho}\Big)^2,$$

and

$$\Big\|D^2\int_0^{k_1} G\,dt\Big\|_{L_1} \le C\Big(\log\frac{t_N}{k_1\rho}\Big)^2,$$

and the lemma follows by returning to the original time variable.

We now prove (8.12). Multiply (8.11) by $\frac{G}{\sqrt{G^2+\xi}}$, where $\xi > 0$. Integrate over Ω, and then integrate by parts. We find, by letting $\xi \to 0^+$,

$$(G_t, \operatorname{sign}(G)) + \|bG\|_{L_1} = \frac{d}{dt}\|G\|_{L_1} + \|bG\|_{L_1} \le 0.$$

After an integration in time we thus obtain, for $0 \le t \le t_N$,

$$\|G(t)\|_{L_1} + \int_0^t \|bG\|_{L_1}\,dt \le \|G(0)\|_{L_1} = \|g\|_{L_1}. \tag{8.14}$$

This proves the first inequality in (8.12). By a standard energy argument we get, for $0 \le t \le t_N$,

$$\|G(t)\|_{L_2}^2 + \int_0^t \|b^{\frac12}G\|_{L_2}^2\,dt \le \|G(0)\|_{L_2}^2 = \|g\|_{L_2}^2,$$

and we conclude by (6.8) that,

$$\|bG\|_{L_2(0,t_N;L_2)} \le \epsilon_{\min}^{-1/2}\|g\|_{L_2}. \tag{8.15}$$

The Riesz-Thorin theorem applied to the linear operator $g \mapsto bG$ gives in view of (8.14) and (8.15)

$$\begin{aligned}\sup_g \frac{\|bG\|_{L_p(0,t_N;L_p)}}{\|g\|_{L_p}} &\le \Big(\sup_g \frac{\|bG\|_{L_1(0,t_N;L_1)}}{\|g\|_{L_1}}\Big)^{1-2/p'} \Big(\sup_g \frac{\|bG\|_{L_2(0,t_N;L_2)}}{\|g\|_{L_2}}\Big)^{2/p'} \\ &\le \epsilon_{\min}^{-1/p'}.\end{aligned}$$

This proves the second inequality in (8.12). $\square$

9. Proof of Theorem 5.1.

In this section we conclude the proof of Theorem 5.1.

Proof. Let g be as in (8.10) and let $\pi : C(\overline{\Omega}) \to S_N$ be the Langrange interpolation operator. We extend u_ϵ and U_ϵ to be equal to zero outside of Ω. Let $e_N^- = u_\epsilon(t_N) - U_{\epsilon,N}^-$ and let $x_0 \in \Omega$ be such that $|e_N^-(x_0)| = \|e_N^-\|_{L_\infty}$. Let $\mathcal{B} = \mathcal{B}(x_0;\rho)$ be the closed ball with center at x_0 and radius ρ. By the mean value theorem and the continuity of e_N^- there is an $x_1 \in \Omega \cap \mathcal{B}$ such that $(e_N^-, g) = e_N^-(x_1)$. By the triangle inequality we have

$$\|e_N^-\|_{L_\infty} = |e_N^-(x_0)| \le |e_N^-(x_0) - e_N^-(x_1)| + |e_N^-(x_1)|. \tag{9.1}$$

If $x_2 \in \mathcal{B} \cap \Omega$ is such that $|\nabla \pi e_N^-(x_2)| = \|\nabla \pi e_N^-\|_{L_\infty(\mathcal{B}\cap\Omega)}$, then there is a triangle $K \in \mathcal{T}$ such that $x_2 \in \overline{K}$. Therefore

$$\begin{aligned}|e_N^-(x_0) - e_N^-(x_1)| &\le \rho \|\nabla e_N^-\|_{L_\infty(\mathcal{B}\cap\Omega)} \\ &\le \rho \|\nabla (e_N^- - \pi e_N^-)\|_{L_\infty(\mathcal{B}\cap\Omega)} + \rho \|\nabla \pi e_N^-\|_{L_\infty(\overline{K})} \\ &\le \rho \|\nabla u_\epsilon(t_N)\|_{L_\infty(\Omega)} + \rho \|\nabla \pi u_\epsilon(t_N)\|_{L_\infty(\Omega)} + \rho \|\nabla \pi e_N^-\|_{L_\infty(\overline{K})} \\ &\le C\rho \|\nabla u_\epsilon(t_N)\|_{L_\infty(\Omega)} + C\rho h_{\min}^{-1} \|e_N^-\|_{L_\infty(\Omega)},\end{aligned}$$

where we used the stability of π in L_∞ and W_∞^1, and an inverse estimate. For sufficiently small ρ, more precisely $\rho \le h_{\min}/(2C)$, we may subtract $C\rho h_{\min}^{-1}\|e_N^-\|_{L_\infty(\Omega)}$ from both sides of (9.1). Hence

$$\|e_N^-\|_{L_\infty} \le C\rho \|\nabla u_\epsilon(t_N)\|_{L_\infty} + C|(e_N^-, g)|. \tag{9.2}$$

By Lemma 10.1

$$\|\nabla u_\epsilon(t_N)\|_{L_\infty} \le C_1 \frac{\epsilon_{\max}}{\epsilon_{\min}} (\epsilon_{t,\max} + 1) + C_2 + C_3 |\log \epsilon_{\min}|,$$

where C_1, C_2, C_3 depend on data. Let us now chose ρ to be $\rho = h_{\min}^{\xi}$ for some $\xi \ge \zeta$. By condition (5.1) we thus have

$$\begin{aligned}\rho \|\nabla u_\epsilon(t_N)\|_{L_\infty(\Omega)} &\le C_1 h_{\min}^{\xi-\zeta} + h_{\min}^{\xi} C_2 + C_3 h_{\min}^{\xi} |\log h_{\min}| \\ &\le C h_{\min}^{\xi-\zeta}(1 + |\log h_{\min}|).\end{aligned} \tag{9.3}$$

We conclude by (9.2) and (9.3) that

$$\|e_N^-\|_{L_\infty} \le C h_{\min}^{\xi-\zeta}(1 + |\log h_{\min}|) + |(e_N^-, g)|.$$

Let G be the solution of (6.10) with g as in (8.10). By our choice of ρ and condition (5.1) we have that $\rho \le \epsilon_{\min}^{\alpha}$ for some $\alpha > 0$ so that Lemma 8.2 applies.

By combining (6.12), Lemma 7.1 and Lemma 8.2 we find

$$
\begin{aligned}
|(e_N^-, g)| &\le |\langle r(U_\epsilon), G\rangle| \le \sum_{1\le n\le N} |\langle r(U_\epsilon), G\rangle_n| \\
&\le C \max_{1\le n\le N} \Big(k_n \|R^t_{n,e}\|_{L_\infty} + k_n \|R^t_{n,i}\|_{L_\infty} \\
&\qquad\qquad + \|h_n^2 R^x_{n,e}\|_{L_\infty} + \|h_n^2 R^x_{n,i}\|_{L_\infty} \Big) \\
&\qquad \times \max \Big(\|G\|_{L_\infty(I_N, L_1)}, \|G_t\|_{L_1(0,t_{N-1};L_1)}, \\
&\qquad\qquad \|D^2 \int_{I_N} G\, dt\|_{L_1}, \|D^2 G\|_{L_1(0,t_{N-1},L_1)} \Big) \\
&\le C L_N \max_{1\le n\le N} \Big(k_n \|R^t_{n,e}\|_{L_\infty} + k_n \|R^t_{n,i}\|_{L_\infty} \\
&\qquad\qquad + \|h_n^2 R^x_{n,e}\|_{L_\infty} + \|h_n^2 R^x_{n,i}\|_{L_\infty} \Big),
\end{aligned}
$$

where

$$
L_N = \Big(1 + \log \frac{t_N}{k_N}\Big)\Big(\log \frac{t_N}{k_N} + \xi |\log h_{\min}|\Big)^2.
$$

□

10. A pointwise a priori estimate of ∇u_ϵ.

Let u_ϵ be the solution of (2.2). In the proof of Theorem 5.1 we need an estimate of $\|\nabla u_\epsilon\|_{L_\infty(0,T;\Omega)}$. The proof is an adaptation of the proof of Theorem 16.11.1 in [6]. Recall the definition of Ω_ϵ^- in (4.2).

Lemma 10.1. *Let u_ϵ be the solution of* (2.2). *There is a constant C such that $t \in [0,T]$:*

$$
\begin{aligned}
\|\nabla u_\epsilon(t)\|_{L_\infty(\Omega)} \le C\Big(&\frac{\epsilon_{\max}}{\epsilon_{\min}} (\epsilon_{t,\max} + 1) \|f - \psi_t + \Delta\psi\|_{L_\infty(0,T;\Omega_\epsilon^-(t))} \\
&+ \|\psi_t\|_{L_\infty(0,T;\Omega_\epsilon^-(t))} + \|f\|_{L_\infty(0,T;\Omega)} \\
&+ |\log \epsilon_{\min}|\, \|f_t\|_{L_\infty(0,T;\Omega)} + \|f(0) + \Delta v_0\|_{L_\infty(0,T;\Omega)} \Big).
\end{aligned}
\tag{10.1}
$$

Remark 10.1.1 The constant C in Lemma 10.1 depends on Ω, and the constant in the elliptic regularity estimate (8.9).

Proof. We write $\beta = \beta(u_\epsilon)$. By combining Sobolev's inequality, elliptic regularity (8.9), Hölder's inequality, and (2.2) we find

$$
\|\nabla u_\epsilon(t)\|_{L_\infty} \le C_q \|\Delta u_\epsilon(t)\|_{L_q} \le C_{q,p} \Big(\|u_{\epsilon,t}\|_{L_p} + \|f\|_{L_p} + \|\epsilon^{-1}\beta\|_{L_p} \Big),
\tag{10.2}
$$

where $p \ge q > 3$. We will use this inequality with q fixed close to 3 and p close to $2|\log \epsilon_{\min}|$. By using q and p we avoid a p dependence in the first inequality. Moreover, $C_{p,q} = C_q |\Omega|^{1/q - 1/p}$, which is bounded independently of p.

We first estimate $\|\epsilon^{-1}\beta\|_{L_p}$. We add $-\psi_t + \Delta\psi$ to both sides of (2.2) and multiply by β^{2k-1}, where k is a positive integer. We also integrate in space followed by an integration by parts using the fact that $\beta = u_\epsilon - \psi \in H_0^1$ inside Ω_ϵ^-, and $\beta = 0$ outside Ω_ϵ^-. We get

$$
\frac{1}{2k} \frac{d}{dt} \|\beta^{2k}\|_{L_1} + \frac{2k-1}{k^2} \|\nabla \beta^k\|_{L_2}^2 + \|\epsilon^{-1}\beta^{2k}\|_{L_1} = (\tilde{f}, \beta^{2k-1}),
\tag{10.3}
$$

where $\tilde f = f - \psi_t + \Delta\psi$. Using Hölder's inequality and the inequality $|ab| \le \frac{1}{p}|a|^p + \frac{1}{q}|b|^q$ for $1/p + 1/q = 1$ we find

$$
\begin{aligned}
|(\tilde f, \beta^{2k-1})| &= |(\epsilon^{(2k-1)/2k}\tilde f, \epsilon^{-(2k-1)/2k}\beta^{2k-1})| \\
&\le \frac{1}{2k}\|\epsilon^{(2k-1)/2k}\tilde f\|^{2k}_{L_{2k}(\Omega_\epsilon^-)} \\
&\quad + \frac{2k-1}{2k}\|\epsilon^{-(2k-1)/2k}\beta^{2k-1})\|^{2k/(2k-1)}_{L_{2k/(2k-1)}} \\
&= \frac{1}{2k}\|\epsilon^{2k-1}\tilde f^{2k}\|_{L_1(\Omega_\epsilon^-)} + \Big(1 - \frac{1}{2k}\Big)\|\epsilon^{-1}\beta^{2k}\|_{L_1}.
\end{aligned}
\tag{10.4}
$$

We subtract the last term from both sides of (10.3) and multiply by $2k$ which gives

$$
\begin{aligned}
\frac{d}{dt}\|\beta^{2k}\|_{L_1} + \frac{2(2k-1)}{k}\|\nabla\beta^k\|^2_{L_2} + \|\epsilon^{-1}\beta^{2k}\|_{L_1} \\
\le \|\epsilon^{2k-1}\tilde f^{2k}\|_{L_1(\Omega_\epsilon^-)}.
\end{aligned}
\tag{10.5}
$$

We note that $\epsilon_{\max}^{-1}\|\beta^{2k}\|_{L_1} \le \|\epsilon^{-1}\beta^{2k}\|_{L_1}$. Thus it follows from the differential inequality (10.5) that

$$
\begin{aligned}
\|\beta^{2k}\|_{L_\infty(0,T;L_1)} &\le \int_0^T e^{-(T-t)/\epsilon_{\max}}\|\epsilon^{(2k-1)}\tilde f^{2k}\|_{L_1(\Omega_\epsilon^-)}\,dt \\
&\le C\epsilon_{\max}^{2k}\|\tilde f\|^{2k}_{L_\infty(0,T;\Omega_\epsilon^-)},
\end{aligned}
$$

where we also used $\beta(u_\epsilon(0)) = 0$. We thus have, recalling the definition of $\tilde f$,

$$
\|\epsilon^{-1}\beta\|_{L_{2k}} \le C(\epsilon_{\max}/\epsilon_{\min})\|f - \psi_t + \Delta\psi\|_{L_\infty(0,T;\Omega_\epsilon^-)}.
\tag{10.6}
$$

Now we turn to the estimate of $\|u_{\epsilon,t}\|_{L_{2k}}$. The argument is formal, since it uses the assumption $u_{\epsilon,tt} \in L_2$. However, the argument can be justified by using difference quotients in time instead of time derivatives, see proof of Theorem 16.11.1 in [6]. We differentiate (2.2) with respect to time, which gives

$$
u_{\epsilon,tt} - \Delta u_{\epsilon,t} + \frac{-\epsilon_t}{\epsilon^2}\beta + \frac{1}{\epsilon}\beta_t = f_t.
\tag{10.7}
$$

Multiply (10.7) by $u_{\epsilon,t}^{2k-1}$ and integrate in space. Integration by parts combined with the fact $u_{\epsilon,t}|_{\partial\Omega} = 0$ gives

$$
\begin{aligned}
\frac{1}{2k}\frac{d}{dt}\|u_{\epsilon,t}^{2k}\|_{L_1} + \frac{2k-1}{k^2}\|\nabla u_{\epsilon,t}^k\|^2_{L_2} + \|\epsilon^{-1}u_{\epsilon,t}^{2k}\|_{L_1(\Omega_\epsilon^-)} \\
= \Big(\frac{1}{\epsilon}\Big(\frac{\epsilon_t}{\epsilon}\beta + \psi_t\Big), u_{\epsilon,t}^{2k-1}\Big)_{\Omega_\epsilon^-} + (f_t, u_{\epsilon,t}^{2k-1}).
\end{aligned}
\tag{10.8}
$$

As in (10.4) we have

$$
\begin{aligned}
\Big|\Big(\epsilon^{-1}\Big(\frac{\epsilon_t}{\epsilon}\beta + \psi_t\Big), u_{\epsilon,t}^{2k-1}\Big)_{\Omega_\epsilon^-}\Big| &\le \frac{1}{2k}\Big\|\epsilon^{-1}\Big(\frac{\epsilon_t}{\epsilon}\beta + \psi_t\Big)^{2k}\Big\|_{L_1(\Omega_\epsilon^-)} \\
&\quad + \Big(1 - \frac{1}{2k}\Big)\|\epsilon^{-1}u_{\epsilon,t}^{2k}\|_{L_1(\Omega_\epsilon^-)}.
\end{aligned}
\tag{10.9}
$$

By the same argument we also have

$$
|(f_t, u_{\epsilon,t}^{2k-1})| \le \Big(\frac{k}{\lambda_1}\Big)^{2k-1}\frac{1}{2k}\|f_t\|^{2k}_{L_{2k}} + \frac{\lambda_1}{k}\frac{2k-1}{2k}\|u_{\epsilon,t}^{2k}\|_{L_1}.
\tag{10.10}
$$

By combining (10.8), (10.9), (10.10) and the inequality $\|\nabla u_{\epsilon,t}^k\|_{L_2}^2 \geq \lambda_1 \|u_{\epsilon,t}^k\|_{L_2}^2$ we find, after a multiplication by $2k$,

$$\frac{d}{dt}\|u_{\epsilon,t}^{2k}\|_{L_1} + \lambda_1 \frac{2k-1}{k}\|u_{\epsilon,t}^{2k}\|_{L_1} \leq \Big\|\epsilon^{-1/2k}\Big(\frac{\epsilon_t}{\epsilon}\beta + \psi_t\Big)\Big\|_{L_{2k}(\Omega_\epsilon^-)}^{2k} + \Big(\frac{k}{\lambda_1}\Big)^{2k-1}\|f_t\|_{L_{2k}}^{2k}.$$

Hence, with $c = \lambda_1(2k-1)/k$,

(10.11)

$$\begin{aligned}\|u_{\epsilon,t}^{2k}(T)\|_{L_1} &\leq \int_0^T e^{-c(T-t)}\Big(\Big\|\epsilon^{-1/2k}\Big(\frac{\epsilon_t}{\epsilon}\beta + \psi_t\Big)\Big\|_{L_{2k}(\Omega_\epsilon^-)}^{2k} + \Big(\frac{k}{\lambda_1}\Big)^{2k-1}\|f_t\|_{L_{2k}}^{2k}\Big)\,dt \\ &\quad + \|u_{\epsilon,t}^{2k}(0)\|_{L_1} \\ &\leq C\Big(\frac{1}{\epsilon_{\min}}\Big(\Big\|\frac{\epsilon_t}{\epsilon}\beta\Big\|_{L_\infty(0,T;\Omega_\epsilon^-)}^{2k} + \|\psi_t\|_{L_\infty(0,T;\Omega_\epsilon^-)}^{2k}\Big) \\ &\quad + \Big(\frac{k}{\lambda_1}\Big)^{2k-1}\|f_t\|_{L_\infty(0,T;\Omega)}^{2k}\Big) + \|f(0) + \Delta v_0\|_{L_{2k}}^{2k}.\end{aligned}$$

In order to keep track of the k-dependence in the constant C in (10.11) we note that $C \leq k|\Omega|/(\lambda_1(2k-1)) \leq |\Omega|/\lambda_1$. Using (10.6) we get

$$\begin{aligned}\|u_{\epsilon,t}(T)\|_{L_{2k}} &\leq C\Big(\epsilon_{\min}^{-1/2k}\Big(\frac{\epsilon_{\max}\epsilon_{t,\max}}{\epsilon_{\min}}\|f - \psi_t + \Delta\psi\|_{L_\infty(0,T;\Omega_\epsilon^-)} \\ &\quad + \|\psi_t\|_{L_\infty(0,T;\Omega_\epsilon^-)}\Big) + \Big(\frac{k}{\lambda_1}\Big)^{(2k-1)/2k}\|f_t\|_{L_\infty(0,T;\Omega)}\Big) \\ &\quad + \|f(0) + \Delta v_0\|_{L_{2k}}.\end{aligned} \tag{10.12}$$

Let k be the smallest integer greater than $|\log(\epsilon_{\min})|$ and let $p = 2k$ in (10.12). For this particular choice of p we get

$$\begin{aligned}\|u_{\epsilon,t}(T)\|_{L_p} &\leq C\Big(\frac{\epsilon_{\max}\epsilon_{t,\max}}{\epsilon_{\min}}\|f - \psi_t + \Delta\psi\|_{L_\infty(0,T;\Omega_\epsilon^-)} + \|\psi_t\|_{L_\infty(0,T;\Omega_\epsilon^-)} \\ &\quad + |\log(\epsilon_{\min})|\,\|f_t\|_{L_\infty(0,T;\Omega)} + \|f(0) + \Delta v_0\|_{L_\infty(0,T;\Omega)}\Big).\end{aligned} \tag{10.13}$$

By combining (10.2), (10.6), (10.13), and Hölder's inequality the lemma follows. ☐

11. Estimates for the L_2-projection in a weighted L_p-norm.

In this section we prove the approximation result Lemma 11.5 , for the L_2-projection, which we use in the proof of Lemma 7.1. Our analysis is an extension of the corresponding result in Section 7 of [3], and the main ideas are the same. For a different approach see [1].

The proof of Lemma 11.5 and the corresponding result in [3] are based on stability estimates for the L_2-projection with respect to a weighted L_p-norm, which we now describe in more detail.

Let Ω be a polygonal domain, not necessarily convex, in $\mathbf{R}^d, d = 1, 2, 3, \ldots$. We use the same notation and assumptions on the family of triangulations $\mathcal{F}$ as in Section 3. In particular, c_0 is as defined in (3.1). Let a triangulation $\mathcal{T} \in \mathcal{F}$ be given, let $\delta = \delta(\mathcal{T})$ be as in (3.2) and let $S = S(\mathcal{T})$ be the function space defined in Section 3. Further, $h(x)$ is defined by $h|_K = h_K$ for all $K \in \mathcal{T}$. We recall that the L_2-projection $P: L_2(\Omega) \to S$ is defined by

$$(Pf, \chi) = (f, \chi), \quad \forall \chi \in S, \quad \forall f \in L_2(\Omega). \tag{11.1}$$

In [3] it is proved that

$$\|\varphi P u\|_{L_2} \le C\|\varphi u\|_{L_2},$$

under a certain assumption on the weight $\varphi \in C^1(\Omega)$. This inequality is then applied with $\varphi = \tilde{h}^{-2}$, where $\tilde{h} \in C^1(\Omega)$ is comparable with h. In order to satisfy the assumption about φ it is required that $\nabla \tilde{h}$ is sufficiently small. We extend this in two ways. First of all we replace the L_2-norm by the L_p-norm and prove

$$\|\varphi P u\|_{L_p} \le C\|\varphi u\|_{L_p}, \quad 1 \le p \le \infty.$$

This result is proved in [3] for the case $\varphi = 1$. Secondly, it turns out that it is sufficient to assume that φ is only piecewise C^1, i.e., $\varphi \in \mathcal{W}$, where

$$\mathcal{W} = \{\varphi \in W^1_\infty(\Omega) \cap C(\overline{\Omega}) : \quad \varphi|_K \in C^1(K), \quad \forall K \in \mathcal{T}\}.$$

This makes it possible to replace the function $\tilde{h}$ by the piecewise linear function $h_{\mathcal{T}}$, defined below, which allows us to express the condition on the triangulation as the requirement that the computable quantity δ is sufficiently small.

Given a positive number ν and a triangulation $\mathcal{T}$, we define a set $F^\nu_{\mathcal{T}}$ of weight functions associated with $\mathcal{T}$. A function $\varphi : \Omega \to \mathbf{R}$ belongs to $F^\nu_{\mathcal{T}}$, if $\varphi \in \mathcal{W}$, $\varphi > 0$ on $\overline{\Omega}$, and

$$|\nabla\varphi(x)| \le \nu h_K^{-1}\varphi(x), \quad \forall x \in K, \quad \forall K \in \mathcal{T}. \tag{11.2}$$

We now construct the function $h_{\mathcal{T}}$ for a given triangulation $\mathcal{T} \in \mathcal{F}$. Let n_i be a node in the triangulation $\mathcal{T}$, define the set $\mathcal{M}(n_i) = \{K \in \mathcal{T} : n_i \in \overline{K}\}$ and let $M(n_i)$ be the number of simplices in $\mathcal{M}(n_i)$. We define the function $h_{\mathcal{T}} \in C(\Omega)$ by: $h_{\mathcal{T}}|_K$ is linear for all $K \in \mathcal{T}$ and

$$h_{\mathcal{T}}(n_i) = \frac{1}{M(n_i)} \sum_{K' \in \mathcal{M}(n_i)} h_{K'}, \quad \forall n_i. \tag{11.3}$$

By definition $h_{\mathcal{T}} \in \mathcal{W}$. In the following lemma we state some properties of $h_{\mathcal{T}}$.

Lemma 11.1. *Assume that $\mathcal{T} \in \mathcal{F}$ with $\delta = \delta(\mathcal{T}) < 1$. Then the function $h_{\mathcal{T}}$ defined by* (11.3) *has the following properties:*

$$\begin{aligned} &h_K\sqrt{1-\delta} \le h_{\mathcal{T}}(x) \le h_K\sqrt{1+\delta}, \quad \forall x \in K, \quad \forall K \in \mathcal{T}, \\ &|\nabla h_{\mathcal{T}}(x)| \le 2c_0\delta, \quad \forall x \in K, \quad \forall K \in \mathcal{T}. \end{aligned}$$

Proof. Let $K \in \mathcal{T}$. It follows from $|1 - h^2_{K'}/h^2_K| \le \delta$, see (3.2), and $\delta < 1$ that

$$h_K\sqrt{1-\delta} \le h_{K'} \le h_K\sqrt{1+\delta}, \quad \forall K' \in S_K.$$

Since $h_{\mathcal{T}}$ is linear on K its maxima and minima of are attained at the nodes of K and hence by (11.3)

$$\begin{aligned} h_{\mathcal{T}}|_K(x) &\le \max_{n_i \in \partial K} \frac{1}{M(n_i)} \sum_{K' \in \mathcal{M}(n_i)} h_{K'} \le h_K\sqrt{1+\delta}, \\ h_{\mathcal{T}}|_K(x) &\ge \min_{n_i \in \partial K} \frac{1}{M(n_i)} \sum_{K' \in \mathcal{M}(n_i)} h_{K'} \ge h_K\sqrt{1-\delta}. \end{aligned} \tag{11.4}$$

Since $h_\mathcal{T}|_K$ is linear we also have, by (3.1) and (11.4),

$$|\nabla h_\mathcal{T}|_K(x)| \le \frac{\sup_{x\in K} h_\mathcal{T} - \inf_{x\in K} h_\mathcal{T}}{\rho_K} \le \frac{h_K}{\rho_K}\frac{2\delta}{\sqrt{1+\delta}+\sqrt{1-\delta}} \le 2c_0\delta.$$

□

We next note that Lemma 7.1 in [3], originally stated for weight functions $\varphi \in C^1(\Omega)$, is also valid for weight functions in $\mathcal{W}$. The proof holds without change, but for completeness we include it here. We recall that under assumption (3.1) there are two constants C_i and $C_{inv,1}$ such that, if ψ_i denotes the piecewise linear Lagrange node interpolant of ψ, then for all $K \in \mathcal{T}$ and all $\mathcal{T} \in \mathcal{F}$

$$\begin{aligned} \|\psi - \psi_i\|_{L_p(K)} &\le C_i h_K^2 \|D^2\psi\|_{L_p(K)}, \quad \forall \psi \in H^2(K), \quad 1 \le p \le \infty, \\ \|\nabla\chi\|_{L_p(K)} &\le C_{inv,1} h_K^{-1}\|\chi\|_{L_p(K)}, \quad \forall \chi \in S, \quad 1 \le p \le \infty. \end{aligned} \tag{11.5}$$

In the following proofs we use the notation $\overline{f}_K = \sup_{x\in K} f(x)$, $\underline{f}_K = \inf_{x\in K} f(x)$.

Lemma 11.2. *Assume that the family $\mathcal{F}$ satisfies (3.1) for some constant c_0. Then there exist positive constants $\nu_\mathcal{F}$ and C such that for any triangulation $\mathcal{T} \in \mathcal{F}$ and any $\varphi \in F_\mathcal{T}^{\nu_\mathcal{F}}$ we have*

$$\|\varphi Pu\|_{L_2(\Omega)} \le C\|\varphi u\|_{L_2(\Omega)}, \quad \forall u \in L_2(\Omega).$$

Proof. First we recall that

$$\|\psi - \psi_i\|_{L_\infty(K)} \le h_K\|\nabla\psi\|_{L_\infty(K)}, \quad \forall \psi \in \mathcal{W}, \quad \forall K \in \mathcal{T}. \tag{11.6}$$

Since φ is continuous, we may define the interpolant $(\varphi^2 Pu)_i$ and therefore, by (11.1), we get

$$\begin{aligned} \|\varphi Pu\|_{L_2(\Omega)}^2 &= (Pu, \varphi^2 Pu) = (Pu, \varphi^2 Pu - (\varphi^2 Pu)_i) + (u, (\varphi^2 Pu)_i) \\ &\le \|\varphi Pu\|_{L_2(\Omega)}\|\varphi^{-1}(\varphi^2 Pu - (\varphi^2 Pu)_i)\|_{L_2(\Omega)} \\ &\quad + \|\varphi u\|_{L_2(\Omega)}\|\varphi^{-1}(\varphi^2 Pu)_i\|_{L_2(\Omega)}. \end{aligned} \tag{11.7}$$

Since $(\varphi^2 Pu)_i = ((\varphi^2)_i Pu)_i$, we have here

$$\begin{aligned} \|\varphi^{-1}(\varphi^2 Pu - (\varphi^2 Pu)_i)\|_{L_2(K)} &\le \|\varphi^{-1}(\varphi^2 - (\varphi^2)_i)Pu\|_{L_2(K)} \\ &\quad + \|\varphi^{-1}[(\varphi^2)_i Pu - ((\varphi^2)_i Pu)_i]\|_{L_2(K)} = I_K + II_K. \end{aligned}$$

In order to estimate I_K and II_K we note that

$$\overline{\varphi}_K \le \underline{\varphi}_K + h_K\|\nabla\varphi\|_{L_\infty(K)},$$

where, by assumption (11.2),

$$\|\nabla\varphi\|_{L_\infty(K)} \le \nu {h_K}^{-1}\overline{\varphi}_K, \tag{11.8}$$

so that

$$\overline{\varphi}_K \le (1-\nu)^{-1}\underline{\varphi}_K, \tag{11.9}$$

if $\nu < 1$. Further by (11.6), (11.8), and (11.9) we get

$$\begin{aligned} I_K &\le \underline{\varphi}_K^{-1} h_K\|\nabla\varphi^2\|_{L_\infty(K)}\|Pu\|_{L_2(K)} \le 2h_K\frac{\overline{\varphi}_K}{\underline{\varphi}_K}\|\nabla\varphi\|_{L_\infty(K)}\|Pu\|_{L_2(K)} \\ &\le 2\nu\Big(\frac{\overline{\varphi}_K}{\underline{\varphi}_K}\Big)^2\|\varphi Pu\|_{L_2(K)} \le 2\nu(1-\nu)^{-2}\|\varphi Pu\|_{L_2(K)}, \end{aligned} \tag{11.10}$$

and by (11.5)

$$II_K \leq \underline{\varphi}_K^{-1} C_i h_K^2 \|D^2((\varphi^2)_i Pu)\|_{L_2(K)}.$$

But $(\varphi^2)_i$ and Pu are linear on K, so that

$$\begin{aligned}\|D^2((\varphi^2)_i Pu)\|_{L_2(K)} &\leq 2\|\nabla(\varphi^2)_i\|_{L_\infty(K)}\|\nabla Pu\|_{L_2(K)}\\ &\leq 2C_{inv,1}h_K^{-1}\|\nabla\varphi^2\|_{L_\infty(K)}\|Pu\|_{L_2(K)},\end{aligned}$$

where we used that $\|\nabla\psi_i\|_{L_\infty(K)} \leq \|\nabla\psi\|_{L_\infty(K)}$ for all $\psi \in \mathcal{W}$ and the inverse estimate (11.5) in the last step. By the same argument as in (11.10), we get

$$II_K \leq 4C_i C_{inv,1}\nu(1-\nu)^{-2}\|\varphi Pu\|_{L_2(K)}.$$

Hence we conclude that

$$\begin{aligned}\|\varphi^{-1}(\varphi^2 Pu - (\varphi^2 Pu)_i)\|_{L_2(\Omega)} &\leq (1 + 2C_iC_{inv,1})2\nu(1-\nu)^{-2}\|\varphi Pu\|_{L_2(\Omega)}\\ &\leq C\nu\|\varphi Pu\|_{L_2(\Omega)}.\end{aligned} \tag{11.11}$$

By (11.11) it also follows that

$$\begin{aligned}\|\varphi^{-1}(\varphi^2 Pu)_i\|_{L_2(\Omega)} &\leq \|\varphi Pu\|_{L_2(\Omega)} + \|\varphi^{-1}(\varphi^2 Pu - (\varphi^2 Pu)_i)\|_{L_2(\Omega)}\\ &\leq (1 + C\nu)\|\varphi Pu\|_{L_2(\Omega)}.\end{aligned} \tag{11.12}$$

By combining (11.7), (11.11), and (11.12) we get

$$\|\varphi Pu\|_{L_2(\Omega)} \leq C\|\varphi u\|_{L_2(\Omega)}.$$

for sufficiently small ν. □

The next lemma is an extension of Corollary 7.4 in [3], where it is proved that $\|Pu\|_{L_p(\Omega)} \leq C\|u\|_{L_p}$ for sufficiently small $\|\nabla\tilde{h}\|_{L_\infty(\Omega)}$, and the proof is similar. We recall that, under assumption (3.1), there is a constant $C_{inv,2}$ such that

$$\|\chi\|_{L_\infty(K)} \leq C_{inv,2}h_K^{-d/2}\|\chi\|_{L_2(K)}, \quad \forall\chi \in S. \tag{11.13}$$

Lemma 11.3. *Assume that the family $\mathcal{F}$ satisfies (3.1) for some constant c_0. Then there exist a constant C such that for any triangulation $\mathcal{T} \in \mathcal{F}$ with sufficiently small δ and for any $\varphi \in F_{\mathcal{T}}^{\nu_{\mathcal{F}}/3}$, where $\nu_{\mathcal{F}}$ is the constant in Lemma 11.2, we have*

$$\|\varphi Pu\|_{L_p(\Omega)} \leq C\|\varphi u\|_{L_p(\Omega)}, \quad \forall u \in L_p(\Omega), \quad 1 \leq p \leq \infty.$$

Proof. Let $\mu = \nu_{\mathcal{F}}/6$ and let $c_1 = \sqrt{1-\delta}$ be the constant in Lemma 11.1. We begin by proving the case $p = \infty$. Let $x_0 \in \Omega$ and let $K_0 \in \mathcal{T}$ be a simplex such that $x_0 \in \overline{K_0}$. Let $\sigma(x)$ be defined by

$$\sigma(x) = \sqrt{|x - x_0|^2 + \theta^2}, \quad \theta = h_{\mathcal{T}}(x_0)/\mu. \tag{11.14}$$

Then, by (11.13),

$$\begin{aligned}|\varphi(x_0)Pu(x_0)| &\leq \varphi(x_0)\|Pu\|_{L_\infty(K_0)}\\ &\leq C_{inv,2}h_{K_0}^{-d/2}\overline{\varphi}_{K_0}\|Pu\|_{L_2(K_0)}\\ &\leq C_{inv,2}\frac{\overline{\sigma}_{K_0}^2}{h_{K_0}^{d/2}}\frac{\overline{\varphi}_{K_0}}{\underline{\varphi}_{K_0}}\|\varphi\sigma^{-2}Pu\|_{L_2(K_0)}.\end{aligned} \tag{11.15}$$

We will now check that the weight function $\varphi\sigma^{-2}$ satisfies the assumptions of Lemma 11.2. Let $K \in \mathcal{T}$. We note that for $\delta \leq \nu_{\mathcal{F}}/(12c_0)$ we have $2c_0\delta \leq \nu_{\mathcal{F}}/6 = \mu$ so that, by Lemma 11.1 and (11.14),

$$c_1^2 h_K^2 \leq h_{\mathcal{T}}(x)^2 \leq 2\big(|x-x_0|^2(2c_0\delta)^2 + h_{\mathcal{T}}(x_0)^2\big) \leq 2\mu^2\sigma^2(x), \quad \forall x \in K, \quad \forall K \in \mathcal{T}.$$

Further, by direct calculation,

$$|\nabla\sigma^2| \leq 2|\sigma|,$$

and therefore

$$|\nabla\sigma(x)^{-2}| = |\sigma^{-4}\nabla\sigma^2| \leq 2\sigma^{-3} \leq 2\sqrt{2}\mu c_1^{-1} h_K^{-1}\sigma(x)^{-2}, \quad \forall x \in K, \quad \forall K \in \mathcal{T}.$$

Using also 11.2 we get

$$\begin{aligned}|\nabla(\varphi\sigma^{-2})(x)| &\leq |\sigma^{-2}\nabla\varphi| + |\varphi\nabla\sigma^{-2}| \leq \frac{\nu_{\mathcal{F}}}{3} h_K^{-1}\varphi\sigma^{-2} + 2\sqrt{2}c_1^{-1}\mu h_K^{-1}\varphi\sigma^{-2}\\ &\leq (\nu_{\mathcal{F}}/3 + 2\sqrt{2}c_1^{-1}\mu)h_K^{-1}\varphi\sigma^{-2}\\ &\leq \nu_{\mathcal{F}} h_K^{-1}(\varphi\sigma^{-2})(x), \quad \forall x \in K, \quad \forall K \in \mathcal{T},\end{aligned}$$

where we in the last inequality assumed that $\delta \leq 1/2$ so that $c_1 \geq 1/\sqrt{2}$. Hence, for sufficiently small δ we have that $\varphi\sigma^{-2} \in F_{\mathcal{T}}^{\nu_{\mathcal{F}}}$ and we may use Lemma 11.2 with φ replaced by $\varphi\sigma^{-2}$ to get

$$\|\varphi\sigma^{-2}Pu\|_{L_2(\Omega)} \leq C\|\varphi\sigma^{-2}u\|_{L_2(\Omega)} \leq C\|\varphi u\|_{L_\infty(\Omega)}\|\sigma^{-2}\|_{L_2(\Omega)}. \tag{11.16}$$

Now by direct calculation

$$\|\sigma^{-2}\|_{L_2(\Omega)} \leq C(h_{\mathcal{T}}(x_0)/\mu)^{(d-4)/2}. \tag{11.17}$$

Further

$$\overline{\sigma}_{K_0}^2 \leq h_{K_0}^2 + \mu^{-2}h_{\mathcal{T}}(x_0)^2, \tag{11.18}$$

and as in the proof of Lemma 11.2, see (11.9),

$$\overline{\varphi}_{K_0}/\underline{\varphi}_{K_0} \leq (1-\nu)^{-1}. \tag{11.19}$$

Since x_0 is arbitrary we get by combining (11.15), (11.16), (11.17), (11.18), and (11.19)

$$\begin{aligned}\|\varphi Pu\|_{L_\infty(\Omega)} &\leq CC_{inv,2}\mu^{(4-d)/2}\frac{h_{K_0}^2 + (\mu)^{-2}h_{\mathcal{T}}(x_0)^2}{h_{K_0}^{d/2}h_{\mathcal{T}}(x_0)^{(4-d)/2}}(1-\nu)^{-1}\|\varphi u\|_{L_\infty(\Omega)}\\ &\leq C\|\varphi u\|_{L_\infty(\Omega)}.\end{aligned}$$

This is the case $p = \infty$. By Lemma 11.2 we also have

$$\|\varphi Pu\|_{L_2(\Omega)} \leq C\|\varphi u\|_{L_2(\Omega)}. \tag{11.20}$$

By application of the Riesz-Thorin interpolation theorem to the linear operator $Tf = \varphi P(\varphi^{-1}f)$ we conclude

$$\|\varphi Pu\|_{L_p(\Omega)} \leq C\|\varphi u\|_{L_p(\Omega)}, \quad 2 \leq p \leq \infty.$$

For $1 \leq p \leq 2$ the estimate follows by a duality argument. Let q be such that $1/p + 1/q = 1$, $1 \leq p \leq 2$. We then have

$$\|\varphi Pu\|_{L_p(\Omega)} = \sup_{\|v\|_{L_q}=1}(\varphi Pu, v) \leq \|\varphi u\|_{L_p(\Omega)} \sup_{\|v\|_{L_q}=1}\|\varphi^{-1}P(\varphi v)\|_{L_q(\Omega)}.$$

But

$$|\nabla\varphi^{-1}(x)| = |\varphi^{-2}\nabla\varphi| \le \varphi^{-2}\frac{\nu_{\mathcal{F}}}{3}h_K^{-1}\varphi = \frac{\nu_{\mathcal{F}}}{3}h_K^{-1}\varphi^{-1}(x), \quad \forall x \in K, \quad \forall K \in \mathcal{T},$$

so that $\varphi^{-1} \in F_{\mathcal{T}}^{\nu_{\mathcal{F}}/3}$. Since $q \ge 2$ we may use (11), with φ replaced by φ^{-1}, to get

$$\|\varphi^{-1}P(\varphi v)\|_{L_q(\Omega)} \le C\|\varphi^{-1}\varphi v\|_{L_q(\Omega)} = C\|v\|_{L_q(\Omega)},$$

which completes the proof. □

We also have a result for the derivative.

Lemma 11.4. *Assume that the family $\mathcal{F}$ satisfies* (3.1) *for some constant c_0. Then there exist a constant C such that for any triangulation $\mathcal{T} \in \mathcal{F}$ with sufficiently small δ and any $\varphi \in F_{\mathcal{T}}^{\nu_{\mathcal{F}}/6}$ we have*

$$\|\varphi\nabla Pu\|_{L_p(\Omega)} \le C\|h^{-1}\varphi u\|_{L_p(\Omega)}, \quad \forall u \in L_p(\Omega), \quad 1 \le p \le \infty. \tag{11.21}$$

Proof. By (11.5), Lemma 11.1, and (11.9) we get

$$\begin{aligned}\|\varphi\nabla Pu\|^p_{L_p(\Omega)} &= \sum_{K\in\mathcal{T}} \|\varphi\nabla Pu\|^p_{L_p(K)} \le \sum_{K\in\mathcal{T}} C^p_{inv,1}\overline{\varphi}^p_K/h^p_K\|Pu\|^p_{L_p(K)} \\ &\le \sum_{K\in\mathcal{T}} C^p_{inv,1}\frac{\overline{\varphi}^p_K}{\underline{\varphi}^p_K}\frac{\overline{h_{\mathcal{T}}}^p_K}{h^p_K}\|h_{\mathcal{T}}^{-1}\varphi Pu\|^p_{L_p(K)} \\ &\le C\|h_{\mathcal{T}}^{-1}\varphi Pu\|^p_{L_p(\Omega)} \le C\|h_{\mathcal{T}}^{-1}\varphi u\|^p_{L_p(\Omega)},\end{aligned}$$

where we used Lemma 11.3, with φ replaced by $h_{\mathcal{T}}^{-1}\varphi$, in the last equality. This is possible for $\delta \le \min(1/2, \nu_{\mathcal{F}}/(12\sqrt{2}c_0))$, since by direct calculation

$$|\nabla(h_{\mathcal{T}}^{-1}\varphi)|_K| \le (\nu_{\mathcal{F}}/6 + 2c_0\delta/\sqrt{1-\delta})h_K^{-1}(h_{\mathcal{T}}^{-1}\varphi) \le \frac{\nu_{\mathcal{F}}}{3}h_K^{-1}h_{\mathcal{T}}^{-1}\varphi, \quad \forall K \in \mathcal{T},$$

so that $h_{\mathcal{T}}^{-1}\varphi \in F_{\mathcal{T}}^{\nu_{\mathcal{F}}/3}$. Finally by Lemma 11.1, we have

$$\|h_{\mathcal{T}}^{-1}\varphi u\|_{L_p(\Omega)} \le C\|h^{-1}\varphi u\|_{L_p(\Omega)}.$$

□

We now state the main result of this section, a lemma concerning the approximation property of the L_2-projection in a weighted L_p-norm, $1 \le p \le \infty$. We use this lemma in the proof of Lemma 7.1. We note that the result depends on two conditions on the triangulation through the parameters c_0, see (3.1), and δ, see (3.2). Recall that $h(x) = h_K$ for $x \in K$.

Lemma 11.5. *Assume that the family $\mathcal{F}$ satisfies the condition* (3.1) *for some constant c_0. Then there exist a constant C such that for any triangulation $\mathcal{T} \in \mathcal{F}$ with sufficiently small δ we have*

$$\|h^{-2}(I-P)u\|_{L_p(\Omega)} \le C\|D^2u\|_{L_p(\Omega)}, \quad \forall u \in H^1_0 \cap W^2_p, \quad 1 \le p \le \infty,$$

$$\|h^{-1}\nabla(I-P)u\|_{L_p(\Omega)} \le C\|D^2u\|_{L_p(\Omega)}, \quad \forall u \in H^1_0 \cap W^2_p, \quad 1 \le p \le \infty.$$

Proof. Let $\pi u = u_i \in S$ denote the Lagrange interpolant of the function u. We first note that

$$\|h^{-2}(I-P)u\|_{L_p(\Omega)} \le \|h^{-2}(I-\pi)u\|_{L_p(\Omega)} + \|h^{-2}P(\pi - I)u\|_{L_p(\Omega)},$$

where, by (11.5),

$$\|h^{-2}(I-\pi)u\|_{L_p(\Omega)} \le C\|D^2u\|_{L_p(\Omega)}.$$

Further, by Lemma 11.1,

$$\begin{aligned}\|h^{-2}P(\pi - I)u\|_{L_p(\Omega)} &\le C\|h_{\mathcal{T}}^{-2}P(\pi - I)u\|_{L_p(\Omega)} \le C\|h_{\mathcal{T}}^{-2}(\pi - I)u\|_{L_p(\Omega)}\\ &\le C\|h^{-2}(\pi - I)u\|_{L_p(\Omega)} \le C\|D^2u\|_{L_p(\Omega)},\end{aligned}$$

where we used Lemma 11.3, with φ replaced by $h_{\mathcal{T}}^{-2}$, in the second step. This is possible for $\delta \le \min(1/2, \nu_{\mathcal{F}}/(6\sqrt{2}c_0))$, since for such δ we have $|\nabla h_{\mathcal{T}}^{-2}|_K| \le 2c_0\delta c_1^{-1} h_K^{-1} h_{\mathcal{T}}^{-2} \le \frac{\nu_{\mathcal{F}}}{3} h_K^{-1} h_{\mathcal{T}}^{-2}$ for all $K \in \mathcal{T}$, so that $h_{\mathcal{T}}^{-2} \in F_{\mathcal{T}}^{\nu_{\mathcal{F}}/3}$. This proves the first statement of the lemma. Similarly,

$$\|h^{-1}\nabla(I - P)u\|_{L_p(\Omega)} \le \|h^{-1}\nabla(I - \pi)u\|_{L_p(\Omega)} + \|h^{-1}\nabla P(\pi - I)u\|_{L_p(\Omega)},$$

where

$$\|h^{-1}\nabla(I - \pi)u\|_{L_p(\Omega)} \le C\|D^2u\|_{L_p(\Omega)},$$

and

$$\begin{aligned}\|h^{-1}\nabla P(\pi - I)u\|_{L_p(\Omega)} &\le C\|h_{\mathcal{T}}^{-1}\nabla P(\pi - I)\|_{L_p(\Omega)} \le C\|h^{-1}h_{\mathcal{T}}^{-1}(\pi - I)\|_{L_p(\Omega)}\\ &\le C\|h^{-2}(\pi - I)\|_{L_p(\Omega)} \le C\|D^2u\|_{L_p(\Omega)}.\end{aligned}$$

In the second step we used Lemma 11.4, with φ replaced by $h_{\mathcal{T}}^{-1}$, which is allowed for $\delta \le \min(1/2, \nu_{\mathcal{F}}/(12\sqrt{2}c_0))$, since, for such δ, $|\nabla h_{\mathcal{T}}^{-1}|_K| \le 2c_0\delta c_1^{-1} h_K^{-1} h_{\mathcal{T}}^{-1} \le \frac{\nu_{\mathcal{F}}}{6} h_K^{-1} h_{\mathcal{T}}^{-1}$ for all $K \in \mathcal{T}$, so that $h_{\mathcal{T}}^{-1} \in F_{\mathcal{T}}^{\nu_{\mathcal{F}}/6}$. □

References

1. M. Crouzeix and V. Thomée, *The stability in L_p and W_p^1 of the L_2-projection onto finite element function spaces*, Math. Comp. **48** (1987), 521–532.
2. K. Eriksson and C. Johnson, *Adaptive finite element methods for parabolic problems I: A linear model problem*, SIAM J. Numer. Anal. **28** (1991), 43–77.
3. ———, *Adaptive finite element methods for parabolic problems II: Optimal error estimates in $L_\infty L_2$ and $L_\infty L_\infty$*, SIAM J. Numer. Anal. **32** (1995), 706–740.
4. M. Fassihi, *L^p integrability of the second order derivatives of Green potentials in convex domains*, Preprint 1998-22, Department of Mathematics, Chalmers University of Technology and Göteborg University.
5. D. A. French, S. Larsson, and R. H. Nochetto, *Pointwise a posteriori error analysis for an adaptive penalty finite element method for the obstacle problem*, Preprint 2000:58, Department of Mathematics, Chalmers University of Technology and Göteborg University.
6. A. Friedman, *Stochastic Differential Equations and Applications*, Academic Press, New York, 1976.
7. ———, *Variational Principles and Free-boundary Problems*, Wiley, New York, 1982.
8. S. J. Fromm, *Potential space estimates for Green potentials in convex domains*, Proc. Amer. Math. Soc. **119** (1993), 225–233.
9. P. Grisvard, *Elliptic Problems in Nonsmooth Domains*, Pitman, Boston, 1985.
10. M. Hieber and J. Prüss, *Heat kernels and maximal L^p-L^q estimates for parabolic evolution equations*, Comm. Partial Differential Equations **22** (1997), 1647–1669.
11. R. H. Nochetto, *Pointwise a posteriori error estimates for elliptic problems on highly graded meshes*, Math. Comp. **64** (1995), 1–22.
12. A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Springer-Verlag, New York, 1983.
13. R. Scholz, *Numerical solution of the obstacle problem by the penalty method*, Computing **32** (1984), 297–306.
14. ———, *Numerical solution of the obstacle problem by the penalty method. II. Time-dependent problems*, Numer. Math. **49** (1986), 255–268.

Department of Mathematics, Chalmers University of Technology and Göteborg University, SE–412 96 Göteborg, Sweden

E-mail address: boman@math.chalmers.se