L^∞-ERROR ESTIMATES FOR THE OBSTACLE PROBLEM REVISITED

C. CHRISTOF[§]

Abstract. In this paper, we present an alternative approach to a priori L^∞-error estimates for the piecewise linear finite element approximation of the classical obstacle problem. Our approach is based on stability results for discretized obstacle problems and on error estimates for the finite element approximation of functions under pointwise inequality constraints. As an outcome, we obtain the same order of convergence proven in several works before. In contrast to prior results, our estimates can, for example, also be used to study the situation where the function space is discretized but the obstacle is not modified at all.

Key words. A priori error analysis, Linear finite elements, Obstacle problem

1. Introduction. This paper is concerned with a priori L^∞-error estimates for the piecewise linear finite element approximation of the classical obstacle problem

$$(P)\begin{cases} \min \dfrac{1}{2}\displaystyle\int_\Omega \nabla v \cdot \nabla v \,\mathrm{d}x - \langle f, v\rangle \\ \text{s.t. } v \in K := \{z \in H_0^1(\Omega) : z \geq \psi \text{ a.e. in } \Omega\}. \end{cases}$$

Pointwise error estimates for the problem (P) have been studied by various authors before. They are typically derived by analyzing the error in mesh cells near the contact set of the continuous solution (i.e., the set where the solution and the obstacle coincide) and by subsequently applying the discrete maximum principle of Raviart-Ciarlet (cf. [5]). We only mention [2,8,13,15] as references. In this paper, we take a more global perspective and demonstrate that a priori L^∞-error estimates for the problem (P) can also be obtained as corollaries of a more general stability result for discretized obstacle problems. Our method of proof has the advantage that the resulting error estimates are more flexible than their classical counterparts. They can, for example, also handle curved obstacles in the discrete setting. Moreover, our approach illustrates that the problem of estimating the approximation error for (P) is, in fact, a problem of sensitivity analysis. This interpretation turns out to be very advantageous when it comes to analyzing the behavior of the approximation error in lower L^p-norms and the limitations of the piecewise linear finite element method. The alternative viewpoint provided by our analysis and the flexibility of our estimates were, for example, of major importance for the construction of two counterexamples found in a companion paper [4] which demonstrate that the convergence rates obtained for the L^∞-error are – at least in the one-dimensional setting – also optimal if the L^p-error, $p > 1$, is considered. The latter implies in particular that the Aubin-Nitsche trick does not work for the obstacle problem. We refer to [4] for details on this topic.

The outline of this paper is the following: In Section 2, we clarify the notation, address the used discretization scheme, and recall basic results about the solvability of the obstacle problem and the regularity of its solution. In the subsequent section, we introduce the concept of discrete supersolutions and use it to study the stability of the approximate problems obtained from the finite element discretization. We will see

[§]Faculty of Mathematics, Technische Universität Dortmund, Dortmund-Germany.

here that, in contrast to the continuous setting, the solution operator of a discretized obstacle problem is not Lipschitz as a function of the obstacle if the latter is allowed to be curved. Section 4 is devoted to the a priori error analysis in the L^∞-norm. Here, it is demonstrated that L^∞-error estimates for the obstacle problem follow straightforwardly from the stability results of Section 3 if the Ritz projection of the continuous solution is identified with the solution of an appropriately defined discrete problem. The order of convergence that we ultimately obtain in this section is the same as in the classical works of Nitsche [15] and Baiocchi [2]. Lastly, in Section 5 we conclude our investigation with some remarks and a discussion of open problems. The appendix of this paper contains results about one-sided finite element approximations that are needed for our argumentation. The theorems found there may also be of independent interest.

2. Preliminaries. In what follows, Ω will always denote a bounded Lipschitz domain in $\mathbb{R}^d$, where $d \in \mathbb{N}$ is arbitrary but fixed. Furthermore, we will use the standard abbreviations $H_0^1(\Omega)$, $W^{m,p}(\Omega)$, $C^{m,\gamma}(\overline{\Omega})$ and $H^{-1}(\Omega)$ for the Sobolev spaces on Ω, the Hölder spaces on the closure $\overline{\Omega}$ and the dual of $H_0^1(\Omega)$ w.r.t. the L^2-inner product. The pairing between elements of $H_0^1(\Omega)$ and $H^{-1}(\Omega)$ will be denoted with $\langle .,. \rangle$. We refer to [1] and [7] for details.

As already mentioned, the objective of this paper is to study the classical unilateral obstacle problem with zero boundary conditions: Given an $f \in H^{-1}(\Omega)$ (the force) and a measurable function $\psi : \Omega \to \mathbb{R}$ (the obstacle) find the solution to

$$(P) \begin{cases} \min \ \dfrac{1}{2} a(v,v) - \langle f, v \rangle \\ \text{s.t. } \ v \in K := \{ z \in H_0^1(\Omega) : z \geq \psi \text{ a.e. in } \Omega \}. \end{cases}$$

The bilinear form a appearing here is defined to be

$$a : H_0^1(\Omega) \times H_0^1(\Omega) \to \mathbb{R}, \quad (v,w) \mapsto \int_\Omega \nabla v \cdot \nabla w \, \mathrm{d}x.$$

Using that a is coercive (due to the inequality of Poincaré-Friedrichs), it is easy to prove that (P) admits a unique solution provided the admissible set K is not empty. A detailed analysis shows the following:

THEOREM 2.1. *If $\psi : \Omega \to \mathbb{R}$ is a measurable function such that K is not empty, then for all $f \in H^{-1}(\Omega)$ there is a unique solution $u \in K$ to the problem (P) and this solution is also uniquely determined by the variational inequality*

$$u \in K : \quad a(u, u - v) \leq \langle f, u - v \rangle \quad \forall v \in K. \tag{2.1}$$

Moreover, the solution map $S : H^{-1}(\Omega) \ni f \mapsto u \in H_0^1(\Omega)$ is Lipschitz continuous. If, further, there exists a $2 \leq q < \infty$ such that ψ, f and Ω satisfy

- *$f \in L^q(\Omega)$, $\psi \in W^{2,q}(\Omega)$ and $\operatorname{tr} \psi \leq 0$ a.e. on $\partial\Omega$,*
- *there exists a constant $C = C(\Omega, q) > 0$ such that for all functions $v \in H_0^1(\Omega)$ with $\Delta v \in L^q(\Omega)$ it holds*

$$\|v\|_{W^{2,q}} \leq C \|\Delta v\|_{L^q}, \tag{2.2}$$

then the solution u is in $W^{2,q}(\Omega)$ and there exists a constant $C' = C'(\Omega, q)$ such that

$$\|u\|_{W^{2,q}} \leq C' \left(\|f\|_{L^q} + \| \max(-\Delta\psi - f, 0)\|_{L^q} \right).$$

Proof. The unique solvability of (P), the characterization of the solution u by (2.1) and the Lipschitz continuity of the solution operator follow from standard results like the well-known theorem of Lions-Stampacchia (see, e.g., [12, Chapter II]). The $W^{2,q}$-regularity of the solution can be obtained with an approximation argument. We refer to [12, Chapter IV] for details. □

A constant $C(\Omega, q)$ with property (2.2) exists, for example, if the domain Ω has a $C^{1,1}$-boundary and $2 \leq q < \infty$ (see [9, Theorem 9.15, Lemma 9.17]) or if Ω is a polygon with largest interior angle α and $2 \leq q < (1 - \pi/(2\alpha))^{-1}$ (see [10, Theorem 4.4.4.13] and [11, Theorem 2.2.3, Theorem 2.4.3.]). It should be noted that the solution u to the problem (P) will in general not possess higher derivatives than stated in the last theorem even if the obstacle ψ and the force f are smooth. If we consider, for example, the situation $\Omega = (-2, 2)$, $f(x) = 0$ and $\psi(x) = 1 - x^2$, then the solution u is a spline whose second derivatives are discontinuous at the boundary of the set $\{u = \psi\}$ where the solution and the obstacle coincide. This illustrates that higher order finite elements provide little practical advantages in the case of problem (P) (at least as far as non-adaptive methods are concerned) and explains, why it makes sense to restrict the analysis to piecewise linear functions.

Having dealt with the existence, the uniqueness and the regularity of the exact solution, we now turn our attention to the discretization. First, let us recall some basic concepts (cf. [3]):

DEFINITION 2.2. *If $\Omega \subset \mathbb{R}^d$ is a bounded domain with a Lipschitz boundary, then a collection $\mathcal{T} = \{T_i\}$ of finitely many closed d-dimensional simplices T_i is called a triangulation of Ω if the following holds:*

- $\bigcup T_i = \overline{\Omega}$,
- *If $\mathcal{C}_i$ denotes the set of all vertices of a simplex $T_i \in \mathcal{T}$ and* conv(...) *denotes the convex hull of a set, then for all $T_i, T_j \in \mathcal{T}$ it is true that*

$$T_i \cap T_j = \operatorname{conv}(\mathcal{C}_i \cap \mathcal{C}_j).$$

If a Lipschitz domain admits a triangulation, we call it a d-dimensional polyhedron. A familiy $\mathcal{F} = \{\mathcal{T}_h\}_{0<h\leq h_0}$ of triangulations is called quasi-uniform if there are positive constants ρ_1 and ρ_2 such that for all $0 < h \leq h_0$ it holds

$$\max\{\operatorname{diam} T : T \in \mathcal{T}_h\} \leq \rho_1 h \quad \text{and} \quad \min\{\operatorname{diam} B_T : T \in \mathcal{T}_h\} \geq \rho_2 h. \tag{2.3}$$

Here, B_T denotes the largest ball contained in a simplex T.

To approximate (P), we will consider finite-dimensional minimization problems of the form

$$(P_h) \begin{cases} \min \dfrac{1}{2} a(v_h, v_h) - \langle f_h, v_h \rangle \\ \text{s.t. } v_h \in K_h := \{z_h \in V_h^0 : z_h \geq \psi_h \text{ a.e. in } \Omega_h\} \end{cases}.$$

Our standing assumptions are as follows:

Assumption 2.3.

- $\{\Omega_h\}_{0<h\le h_0}$ *is a family of* d*-dimensional polyhedra with* $\Omega_h \subseteq \Omega$ *for all* h,
- $\{\mathcal{T}_h\}_{0<h\le h_0}$ *is a quasi-uniform family of triangulations such that each* $\mathcal{T}_h$ *is a triangulation of* Ω_h *for all* h,
- $V_h^0 := \{v \in C(\overline{\Omega}) : v|_T$ *is affine for all* $T \in \mathcal{T}_h$ *and* $v|_{\overline{\Omega}\setminus\Omega_h} = 0\}$,
- $\psi_h : \Omega_h \to \mathbb{R}$ *is measurable for all* h,
- $f_h \in H^{-1}(\Omega)$ *for all* h.

Note that the condition $\Omega_h \subseteq \Omega$ ensures that V_h^0 is a subspace of $H_0^1(\Omega)$. This implies in particular that the quantities $a(v_h, v_h)$ and $\langle f_h, v_h\rangle$ are well-defined. For brevity's sake, in what follows we will often suppress the range of the mesh width h, i.e., we will write $\{\Omega_h\}$ instead of $\{\Omega_h\}_{0<h\le h_0}$, $h > 0$ instead of $h_0 \ge h > 0$ etc. Using again the theorem of Lions-Stampacchia, it is straightforward to prove:

Theorem 2.4. *If the admissible set* K_h *is not empty, then for all* $f_h \in H^{-1}(\Omega)$ *there exists one and only one solution* $u_h \in V_h^0$ *to the problem* (P_h) *and this solution is also uniquely determined by the variational inequality*

$$u_h \in K_h : \quad a(u_h, u_h - v_h) \le \langle f_h, u_h - v_h\rangle \quad \forall v_h \in K_h. \tag{2.4}$$

Moreover, the solution operator $S_h : H^{-1}(\Omega) \ni f_h \mapsto u_h \in H_0^1(\Omega)$ *is Lipschitz continuous with a Lipschitz constant independent of* h.

It should be noted that we do not assume ψ_h to be an element of our finite element space. This will be of major importance in Section 4.

3. Discrete Supersolutions and Stability Results. To estimate the error between the continuous solution u and the finite element approximation u_h, we will study the sensitivity of the solution map $(f_h, \psi_h) \mapsto u_h$ associated with the discrete problem (P_h). The main tool of our stability analysis will be a variant of the discrete maximum principle of Raviart-Ciarlet that is tailored to the study of the variational inequality (2.4). More precisely, we will make use of the following concept:

Definition 3.1. *A function* g_h *is called a discrete supersolution of the problem* (P_h) *if it holds:*

- $g_h \in V_h := \{v \in C(\overline{\Omega_h}) : v|_T$ *is affine for all* $T \in \mathcal{T}_h\}$,
- $a(g_h, v_h) \le \langle f_h, v_h\rangle$ *for all* $v_h \in V_h^0$ *with* $v_h \le 0$ *in* Ω_h,
- $g_h \ge \psi_h$ *a.e. in* Ω_h,
- $g_h \ge 0$ *on* $\partial\Omega_h$.

The expression $a(g_h, v_h)$ appearing in the second point of the above definition is, of course, to be understood as

$$a(g_h, v_h) := \int_{\Omega_h} \nabla g_h \cdot \nabla v_h \, \mathrm{d}x.$$

In what follows, we will make frequent use of this slight abuse of notation.

Note that Definition 3.1 extends the concept of supersolutions employed in [12] straightforwardly to the discrete setting. The main idea in the following is to prove that discrete supersolutions exhibit broadly the same behavior as their continuous counterparts, i.e., to show that a discrete supersolution g_h majorizes (at least in some sense) the solution u_h of the problem (P_h). To obtain such a result, we have to restrict our analysis to triangulations of a special type:

DEFINITION 3.2. *A triangulation $\mathcal{T}_h$ of Ω_h is said to satisfy the condition (Z) if*

$$a(\varphi_h^i, \varphi_h^j) = \int_{\Omega_h} \nabla\varphi_h^i \cdot \nabla\varphi_h^j \mathrm{d}x \leq 0 \quad \forall i \neq j \text{ with } x_j \notin \partial\Omega_h. \tag{3.1}$$

Here, $\{x_i\}$ denotes the set of all vertices of the triangulation $\mathcal{T}_h$ (including those on the boundary $\partial\Omega_h$) and $\{\varphi_h^i\}$ denotes the nodal basis of the space V_h (i.e., the basis with $\varphi_h^i(x_l) = \delta_{il}$ for all nodes x_l).

The condition (Z) expresses that the system matrix arising from the finite element discretization has to be a Z-matrix. (It is easy to see that it is even an M-matrix in this case). It should be noted that assumptions of the type (Z) are well-known in the context of discrete maximum principles (see, e.g., [5]). In our approach, the non-negativity condition (3.1) will come into play very naturally in the proof of Theorem 3.4. As the following lemma shows, the triangulations satisfying the condition (Z) can be characterized precisely in terms of certain geometric features:

LEMMA 3.3 ([19]).

- *If $d = 1$, then every triangulation satisfies (Z).*
- *If $d = 2$, then (Z) is satisfied if and only if for each edge E of $\mathcal{T}_h$ with $E \not\subseteq \partial\Omega_h$ it holds*

$$\theta_E^{T_1} + \theta_E^{T_2} \leq \pi.$$

  Here, $\theta_E^{T_1}, \theta_E^{T_2} \in (0, \pi)$ denote the angles that oppose E in the adjacent mesh cells T_1 and T_2 (see Figure 3.1).
- *If $d > 2$, then (Z) is satisfied if and only if for all edges E of $\mathcal{T}_h$ with $E \not\subseteq \partial\Omega_h$ it holds*

$$\sum_{T \supset E} \mathcal{H}^{d-2}(\kappa_E^T) \cot\theta_E^T \geq 0.$$

  Here, for every $T = \mathrm{conv}(p_1, ..., p_{d+1}) \in \mathcal{T}_h$ and every $E = \mathrm{conv}(p_i, p_j) \subset T$ the quantities κ_E^T and θ_E^T are defined by

$$\kappa_E^T := S_i \cap S_j \quad \text{and} \quad \theta_E^T := \measuredangle(S_i, S_j),$$

  where S_i and S_j denote the $(d-1)$-dimensional simplices

$$S_i := \mathrm{conv}(p_1, ..., p_{i-1}, p_{i+1}, ..., p_{d+1})$$

  and

$$S_j := \mathrm{conv}(p_1, ..., p_{j-1}, p_{j+1}, ..., p_{d+1})$$

  and $\measuredangle(S_i, S_j) \in (0, \pi)$ denotes the angle enclosed by S_i and S_j (or the normal vectors of S_i and S_j, to be more precise). With $\mathcal{H}^{d-2}(.)$, we mean the $(d-2)$-dimensional Hausdorff measure.

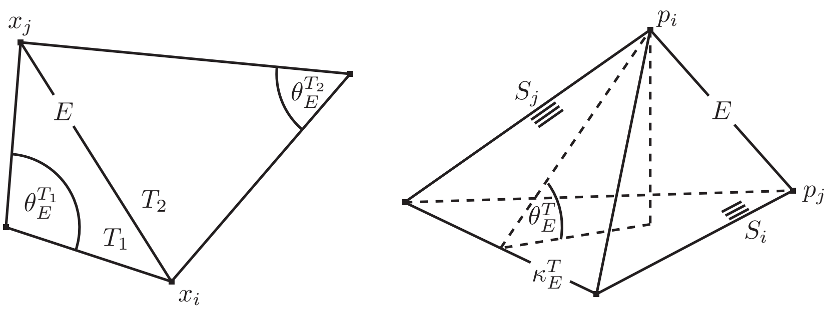

FIG. 3.1. *The geometric situation in Lemma 3.3 for $d = 2$ and $d = 3$ (cf. [19]).*

Using the results about one-sided finite element approximations found in the appendix of this paper, we can prove:

THEOREM 3.4. *Assume that the admissible set K_h of the problem (P_h) is not empty and that $\mathcal{T}_h$ satisfies (Z). Assume further that the obstacle ψ_h in (P_h) satisfies*

$$\psi_h \in C(\overline{\Omega_h}) \quad \textit{and} \quad \psi_h|_T \in C^{1,\gamma}(T) \quad \forall T \in \mathcal{T}_h$$

for some $\gamma \in (0, 1]$ and let ρ_1 be the constant in (2.3). Then (P_h) admits a unique solution u_h and for every supersolution g_h of (P_h) it is true that

$$u_h \leq g_h + \frac{\sqrt{d}}{1+\gamma}\, \rho_1^{1+\gamma} h^{1+\gamma} \max_{T \in \mathcal{T}_h} |\psi_h|_{C^{1,\gamma}(T)} \quad \textit{in } \Omega_h, \tag{3.2}$$

where

$$|\psi_h|_{C^{1,\gamma}(T)} := \max_{k=1,\ldots,d} \sup_{x \neq y \in T} \frac{|\partial_k \psi_h(x) - \partial_k \psi_h(y)|}{\|x - y\|^\gamma}.$$

The above theorem shows that discrete supersolutions at least approximately behave as expected: They are larger than the solution u_h modulo an error that depends on the mesh width h and the curvature of the obstacle ψ_h. The inequality $u_h \leq g_h$, i.e., the behavior observed in the continuous setting (cf. [12, Theorem II6.4]), is only obtained if ψ_h is an element of the space V_h.

Proof of Theorem 3.4. The existence of the solution u_h follows straightforwardly from Theorem 2.4. To prove inequality (3.2), we will use an argument similar to that employed in the continuous setting (cf. [12]). In a first step, we define $g_h' := g_h + C$, where

$$C := \frac{\sqrt{d}}{1+\gamma}\, \rho_1^{1+\gamma} h^{1+\gamma} \max_{T \in \mathcal{T}_h} |\psi_h|_{C^{1,\gamma}(T)}.$$

Note that g_h' is again a supersolution since the addition of a positive constant to the function g_h has no effect on the properties in Definition 3.1. We now consider the unique element v_h in V_h with

$$v_h(x_i) = \min(u_h(x_i), g_h'(x_i)) \quad \forall x_i, \tag{3.3}$$

where $\{x_i\}$ again denotes the set off all vertices of $\mathcal{T}_h$ (including those on the boundary $\partial\Omega_h$). Since $u_h \in V_h^0$, the function v_h can be identified with an element of V_h^0 and from (3.3) we readily obtain that $v_h \le u_h$ holds everywhere in Ω_h. Furthermore, it follows from our construction that $v_h \ge \psi_h$. To see this, we use that according to Theorem A.4 from the appendix (applied to $z := u_h - \psi_h$), for every mesh cell $T \in \mathcal{T}_h$ with vertices $p_1, ..., p_{d+1}$ we can find an affine linear function ψ_h^T on T such that $\psi_h \le \psi_h^T \le u_h$ holds on T and such that

$$0 \le \psi_h^T(p_k) - \psi_h(p_k) \le \frac{\sqrt{d}}{1+\gamma}\rho_1^{1+\gamma} h^{1+\gamma} |\psi_h|_{C^{1,\gamma}(T)} \quad \forall k = 1, .., d+1.$$

This yields that v_h satisfies

$$v_h(p_k) = \min(u_h(p_k), g_h(p_k) + C) \ge \min(\psi_h^T(p_k), \psi_h(p_k) + C) \ge \psi_h^T(p_k)$$

for all $k = 1, ..., d+1$. From the affine linearity of v_h and ψ_h^T on T, it now follows $v_h \ge \psi_h^T \ge \psi_h$ which implies $v_h \ge \psi_h$ on Ω_h as claimed. From the second property in Definition 3.1 and the variational inequality (2.4), we may now deduce:

$$a(g_h', v_h - u_h) \le \langle f_h, v_h - u_h \rangle \quad \text{and} \quad a(u_h, u_h - v_h) \le \langle f_h, u_h - v_h \rangle .$$

If we add these inequalities and define $y_i := u_h(x_i) - g_h'(x_i)$ for all nodes x_i, we obtain (using the properties of v_h)

$$\begin{aligned} 0 &\ge a(u_h - g_h', u_h - v_h) \\ &= \sum_{x_i} y_i \max(0, y_i) a(\varphi_h^i, \varphi_h^i) + \sum_{x_i \ne x_j} y_i \max(0, y_j) a(\varphi_h^i, \varphi_h^j) \\ &= \sum_{x_i} \max(0, y_i)^2 a(\varphi_h^i, \varphi_h^i) + \sum_{x_i \ne x_j \text{ and } x_j \notin \partial\Omega_h} y_i \max(0, y_j) a(\varphi_h^i, \varphi_h^j), \end{aligned} \tag{3.4}$$

where $\{\varphi_h^i\}$ again denotes the nodal basis of V_h. Because of the condition (Z), however, we also know that for all i, j with $x_i \ne x_j$ and $x_j \notin \partial\Omega_h$ it holds

$$y_i \max(0, y_j) a(\varphi_h^i, \varphi_h^j) \ge \max(0, y_i) \max(0, y_j) a(\varphi_h^i, \varphi_h^j).$$

Thus, (3.4) implies

$$0 \ge \sum_{x_i} \sum_{x_j} \max(0, y_i) \max(0, y_j) a(\varphi_h^i, \varphi_h^j) = a(u_h - v_h, u_h - v_h) \ge 0$$

and consequently

$$u_h(x_i) - v_h(x_i) = \max(0, u_h(x_i) - g_h'(x_i)) = 0 \quad \forall x_i.$$

Using again the piecewise linearity of the involved functions, we may deduce

$$u_h \le g_h' = g_h + C = g_h + \frac{\sqrt{d}}{1+\gamma}\rho_1^{1+\gamma} h^{1+\gamma} \max_{T \in \mathcal{T}_h} |\psi_h|_{C^{1,\gamma}(T)} \quad \text{in } \Omega_h.$$

This completes the proof. ∎

Theorem 3.4 allows to analyze the sensitivity of the solution u_h with respect to perturbations of the obstacle ψ_h and the force f_h:

Theorem 3.5. *Consider two discrete obstacle problems of the form*

$$(P_{h,i}) \begin{cases} \min \dfrac{1}{2}\, a(v_h, v_h) - \langle f_{h,i}, v_h\rangle \\ \text{s.t. } v_h \in K_{h,i} := \{z_h \in V_h^0 : z_h \geq \psi_{h,i} \text{ a.e. in } \Omega_h\} \end{cases}, \qquad i = 1, 2,$$

and assume that:

- $f_{h,1}, f_{h,2} \in H^{-1}(\Omega)$,
- *the underlying triangulation* $\mathcal{T}_h$ *satisfies* (Z),
- $\psi_{h,1}, \psi_{h,2} \in C(\overline{\Omega_h})$ *and* $K_{h,1} \neq \emptyset$, $K_{h,2} \neq \emptyset$,
- *there exist* $\gamma_1, \gamma_2 \in (0, 1]$ *such that* $\psi_{h,i}|_T \in C^{1,\gamma_i}(T)$ *for all* $T \in \mathcal{T}_h$, $i = 1, 2$.

Let ρ_1 *be the constant in* (2.3). *Then* $(P_{h,1})$ *and* $(P_{h,2})$ *admit unique solutions* $u_{h,1}$ *and* $u_{h,2}$ *and there exists a constant* $C > 0$ *independent of* h *such that*

$$\begin{aligned} &\|(u_{h,1} - u_{h,2})^+\|_{L^\infty} \\ &\quad \leq \|(\psi_{h,1} - \psi_{h,2})^+\|_{L^\infty} + C\, r(h) \|f_{h,1} - f_{h,2}\|_{H^{-1}} \\ &\qquad + \frac{\sqrt{d}}{1+\gamma_1} (\rho_1 h)^{1+\gamma_1} \max_{T \in \mathcal{T}_h} |\psi_{h,1}|_{C^{1,\gamma_1}(T)} \end{aligned} \tag{3.5}$$

and

$$\begin{aligned} &\|(u_{h,1} - u_{h,2})^-\|_{L^\infty} \\ &\quad \leq \|(\psi_{h,1} - \psi_{h,2})^-\|_{L^\infty} + C r(h) \|f_{h,1} - f_{h,2}\|_{H^{-1}} \\ &\qquad + \frac{\sqrt{d}}{1+\gamma_2} (\rho_1 h)^{1+\gamma_2} \max_{T \in \mathcal{T}_h} |\psi_{h,2}|_{C^{1,\gamma_2}(T)}. \end{aligned} \tag{3.6}$$

Here, $v^+ := \max(0, v)$ *and* $v^- := \min(0, v)$ *denote the positive and the negative part of a function, respectively, and* $r(h)$ *is defined by:*

$$r(h) := \begin{cases} 1 & \text{if } d = 1 \\ (1 + |\log h|)^{1/2} & \text{if } d = 2 \\ h^{1-d/2} & \text{if } d \geq 3 \end{cases}. \tag{3.7}$$

Proof. The unique solvability of the problems $(P_{h,1})$ and $(P_{h,2})$ is a straightforward consequence of Theorem 2.4. It remains to prove the estimates (3.5) and (3.6). If we assume first that $f_{h,1} = f_{h,2}$ and define $g_{h,1} := u_{h,2} + \|(\psi_{h,1} - \psi_{h,2})^+\|_{L^\infty}$, then it certainly holds $g_{h,1} \in V_h$, $g_{h,1} \geq 0$ on $\partial\Omega_h$ and

$$g_{h,1} \geq u_{h,2} + \psi_{h,1} - \psi_{h,2} \geq \psi_{h,1} \text{ in } \Omega_h.$$

From the variational inequality associated with $(P_{h,2})$ and the definition of $a(.,.)$, it follows further that $g_{h,1}$ satisfies

$$a(g_{h,1}, v_h) = a(u_{h,2}, v_h) = a(u_{h,2}, u_{h,2} - (u_{h,2} - v_h)) \leq \langle f_{h,2}, v_h\rangle = \langle f_{h,1}, v_h\rangle$$

for every $v_h \in V_h^0$ with $v_h \leq 0$ in Ω_h. Thus, $g_{h,1}$ is a supersolution for $(P_{h,1})$ and we may deduce from our last theorem that

$$u_{h,1} \leq g_{h,1} + \frac{\sqrt{d}}{1+\gamma_1} (\rho_1 h)^{1+\gamma_1} \max_{T \in \mathcal{T}_h} |\psi_{h,1}|_{C^{1,\gamma_1}(T)} \quad \text{in } \Omega_h,$$

which implies

$$\|(u_{h,1}-u_{h,2})^+\|_{L^\infty} \le \|(\psi_{h,1}-\psi_{h,2})^+\|_{L^\infty} + \frac{\sqrt{d}}{1+\gamma_1}(\rho_1 h)^{1+\gamma_1} \max_{T\in\mathcal{T}_h} |\psi_{h,1}|_{C^{1,\gamma_1}(T)}.$$

This is exactly (3.5) for $f_{h,1} = f_{h,2}$. The inequality (3.6) is obtained analogously if we interchange the roles of $u_{h,1}$ and $u_{h,2}$. This proves the claim for discrete obstacle problems with identical forces. Assume now that $f_{h,1} \neq f_{h,2}$ and denote with $u_{h,i,j}$ the solution of the discrete obstacle problem with obstacle $\psi_{h,i}$ and force $f_{h,j}$, then it follows from the triangle inequality, the Lipschitz property in Theorem 2.4 and well-known inverse estimates (see, e.g., [3, Section 4.5, 4.9]) that

$$\begin{aligned}
&\|(u_{h,1}-u_{h,2})^+\|_{L^\infty}\\
&\quad\le \|(u_{h,1,1}-u_{h,2,1})^+\|_{L^\infty} + \|u_{h,2,1}-u_{h,2,2}\|_{L^\infty}\\
&\quad\le \|(u_{h,1,1}-u_{h,2,1})^+\|_{L^\infty} + C_1 r(h)\|u_{h,2,1}-u_{h,2,2}\|_{H^1}\\
&\quad\le \|(\psi_{h,1}-\psi_{h,2})^+\|_{L^\infty} + C_2 r(h)\|f_{h,1}-f_{h,2}\|_{H^{-1}}\\
&\qquad\qquad + \frac{\sqrt{d}}{1+\gamma_1}\rho_1^{1+\gamma_1} h^{1+\gamma_1} \max_{T\in\mathcal{T}_h} |\psi_{h,1}|_{C^{1,\gamma_1}(T)},
\end{aligned}$$

where C_1, C_2 are constants independent of h. This proves (3.5) in the general case. The estimate (3.6) is again obtained by interchanging roles. □

As the above result shows, for fixed f_h the solution operator $\psi_h \mapsto u_h$ of the problem (P_h) is not Lipschitz continuous as a function from (a subset of) $L^\infty(\Omega)$ to $L^\infty(\Omega)$. We only obtain a Lipschitz-like estimate with an error that again depends on the mesh width h and the curvature of the involved obstacles. This is a major difference to the continuous setting where it can be shown easily that the solutions u_1 and u_2 of two obstacle problems with L^∞-obstacles ψ_1 and ψ_2 and identical forces satisfy $\|u_1 - u_2\|_{L^\infty} \le C\|\psi_1 - \psi_2\|_{L^\infty}$ (cf. [12, Theorem IV8.5]). It should be noted that neither the continuous solution u nor the obstacle ψ or the domain Ω have been relevant for the derivation of (3.5) and (3.6). Up to now, we have solely worked with the discrete problems.

4. $\mathbf{L^\infty}$-Error Estimates. A priori estimates for the error $\|u-u_h\|_{L^\infty}$ can be derived straightforwardly from Theorem 3.5. We just have to observe the following:

LEMMA 4.1. *If $u \in H_0^1(\Omega)$ is the solution to the obstacle problem (P) and $R_h u$ the Ritz projection of u, i.e., the unique element of V_h^0 satisfying*

$$a(R_h u, v_h) = a(u, v_h) \quad \forall v_h \in V_h^0, \tag{4.1}$$

then $R_h u$ is also the unique solution to the discrete obstacle problem

$$(Q_h)\begin{cases} \min\ \dfrac{1}{2}a(v_h,v_h) - \langle f, v_h\rangle \\ \text{s.t. } v_h \in V_h^0 \textit{ and } v_h \ge \psi + R_h u - u \textit{ a.e. in } \Omega_h \end{cases}.$$

Recall that we have assumed $\Omega_h \subseteq \Omega$ (see Assumption 2.3). This ensures that u and ψ are defined everywhere in Ω_h and that the constraint in (Q_h) makes sense.

Proof of Lemma 4.1. The Ritz projection $R_h u$ is obviously admissible for (Q_h) and because of (4.1) and the variational inequality (2.1), it holds

$$a(R_h u, R_h u - v_h) = a(u, R_h u - v_h) = a(u, u - (u - R_h u + v_h)) \le \langle f, R_h u - v_h\rangle$$

for all $v_h \in V_h^0$ with $v_h \geq \psi + R_h u - u$, i.e., $u - R_h u + v_h \geq \psi$. This shows that $R_h u$ is indeed the solution to (Q_h) and completes the proof (cf. Theorem 2.4). ☐

Note that Lemma 4.1 holds without any further assumptions on the regularity of the functions u and ψ and that the obstacle $\psi + R_h u - u$ appearing in (Q_h) is typically not piecewise linear. By applying Theorem 3.5 to (Q_h) and the problem (P_h) used for the finite element approximation, we obtain:

THEOREM 4.2. *Assume that (P) admits a solution u and denote with $R_h u$ the Ritz projection of u as defined in (4.1). Suppose further that the following is satisfied:*

- *$u, \psi \in C(\overline{\Omega})$, $\psi_h \in C(\overline{\Omega_h})$ and $K_h \neq \emptyset$,*
- *$\exists \gamma_1, \gamma_2 \in (0,1]$ with $\psi_h|_T \in C^{1,\gamma_1}(T)$ and $u|_T, \psi|_T \in C^{1,\gamma_2}(T)$ for all $T \in \mathcal{T}_h$,*
- *the triangulation $\mathcal{T}_h$ satisfies (Z).*

Then (P_h) admits a unique solution u_h and there exists a constant $C > 0$ independent of h such that

$$\begin{aligned} &\|(u - u_h)^-\|_{L^\infty(\Omega_h)} \\ &\quad \leq \|(u - R_h u)^-\|_{L^\infty(\Omega_h)} + \|(\psi_h - \psi + u - R_h u)^+\|_{L^\infty(\Omega_h)} \\ &\quad\quad + Cr(h)\|f - f_h\|_{H^{-1}(\Omega)} + \frac{\sqrt{d}}{1+\gamma_1}(\rho_1 h)^{1+\gamma_1} \max_{T \in \mathcal{T}_h} |\psi_h|_{C^{1,\gamma_1}(T)} \end{aligned} \tag{4.2}$$

and

$$\begin{aligned} &\|(u - u_h)^+\|_{L^\infty(\Omega_h)} \\ &\quad \leq \|(u - R_h u)^+\|_{L^\infty(\Omega_h)} + \|(\psi_h - \psi + u - R_h u)^-\|_{L^\infty(\Omega_h)} \\ &\quad\quad + Cr(h)\|f - f_h\|_{H^{-1}(\Omega)} + \frac{\sqrt{d}}{1+\gamma_2}(\rho_1 h)^{1+\gamma_2} \max_{T \in \mathcal{T}_h} |\psi - u|_{C^{1,\gamma_2}(T)}. \end{aligned} \tag{4.3}$$

Here, $r(h)$ is again defined by (3.7).

With the above theorem we have reduced the problem of finding an a priori estimate for the error $\|u - u_h\|_{L^\infty}$ to that of estimating the L^∞-error between u and the Ritz projection $R_h u$. The approximation properties of $R_h u$, however, have been studied by numerous authors and estimates for the quantity $\|u - R_h u\|_{L^\infty(\Omega_h)}$ are well-known. The following result can be found, for example, in [17]:

LEMMA 4.3. *Assume that $\partial\Omega$ is smooth, that $u \in H_0^1(\Omega) \cap C(\overline{\Omega})$ and that there exists a constant $\delta > 0$ independent of h such that*

$$\max_{x \in \partial\Omega_h} \operatorname{dist}(x, \partial\Omega) \leq \delta h^2.$$

Then there exists a constant $C > 0$ independent of h such that

$$\|u - R_h u\|_{L^\infty(\Omega_h)} \leq C|\log h|^\alpha \inf_{v_h \in V_h^0} \|u - v_h\|_{L^\infty(\Omega_h)}$$

with $\alpha = 0$ for $d = 1$ and $\alpha = 1$ for $d > 1$.

Combining Theorem 4.2 and Lemma 4.3 yields:

COROLLARY 4.4. *Let the assumptions of Theorem 4.2 and Lemma 4.3 hold. Then there exists a constant $C > 0$ independent of h such that*

$$\begin{aligned} &\|(u-u_h)^-\|_{L^\infty(\Omega_h)} \\ &\quad \le \|(\psi-\psi_h)^-\|_{L^\infty(\Omega_h)} + \frac{\sqrt{d}}{1+\gamma_1}(\rho_1 h)^{1+\gamma_1} \max_{T\in\mathcal{T}_h} |\psi_h|_{C^{1,\gamma_1}(T)} \\ &\qquad + Cr(h)\|f-f_h\|_{H^{-1}(\Omega)} + C|\log h|^\alpha \inf_{v_h\in V_h^0} \|u-v_h\|_{L^\infty(\Omega_h)} \end{aligned} \tag{4.4}$$

and

$$\begin{aligned} &\|(u-u_h)^+\|_{L^\infty(\Omega_h)} \\ &\quad \le \|(\psi-\psi_h)^+\|_{L^\infty(\Omega_h)} + \frac{\sqrt{d}}{1+\gamma_2}(\rho_1 h)^{1+\gamma_2} \max_{T\in\mathcal{T}_h} |\psi-u|_{C^{1,\gamma_2}(T)} \\ &\qquad + Cr(h)\|f-f_h\|_{H^{-1}(\Omega)} + C|\log h|^\alpha \inf_{v_h\in V_h^0} \|u-v_h\|_{L^\infty(\Omega_h)}, \end{aligned} \tag{4.5}$$

where α and $r(h)$ are defined as before.

As a consequence of Corollary 4.4, we obtain in particular:

COROLLARY 4.5. *Assume that:*

- *$\partial\Omega$ is smooth,*
- *$f \in L^q(\Omega)$ and $\psi \in W^{2,q}(\Omega)$ for some $\max(d,2) < q < \infty$,*
- *$\operatorname{tr}\psi \le 0$ a.e. on $\partial\Omega$,*
- *there exists a constant $\delta > 0$ independent of h such that*

$$\max_{x\in\partial\Omega_h} \operatorname{dist}(x,\partial\Omega) \le \delta h^2,$$

- *the triangulation $\mathcal{T}_h$ satisfies (Z).*

Suppose further that one of the following holds:

a) ψ_h is equal to the Lagrange interpolant $I_h\psi \in V_h$ of ψ and $K_h \neq \emptyset$.
b) ψ_h is equal to the restriction $\psi|_{\Omega_h}$ and $K_h \neq \emptyset$.

Then (P) and (P_h) admit unique solutions u and u_h, it holds $u \in H_0^1(\Omega)\cap W^{2,q}(\Omega)$ and there exists a constant $C > 0$ independent of h such that

$$\begin{aligned} &\|u-u_h\|_{L^\infty(\Omega_h)} \\ &\quad \le C|\log h|^\alpha h^{2-d/q}(\|f\|_{L^q(\Omega)} + \|\psi\|_{W^{2,q}(\Omega)}) + Cr(h)\|f-f_h\|_{H^{-1}(\Omega)}, \end{aligned} \tag{4.6}$$

where α and $r(h)$ are defined as before.

Proof. The unique solvability of the problems (P) and (P_h) and the $W^{2,q}$-regularity of the solution u are direct consequences of Theorem 2.1 and Theorem 2.4. The error estimate (4.6) follows straightforwardly from (4.4) and (4.5). We just have to employ standard results about the accuracy of the Lagrange interpolant (as found, e.g., in [3, Theorem 4.4.20]) and the embedding $W^{2,q}(\Omega) \hookrightarrow C^{1,1-d/q}(\overline{\Omega})$. ∎

Note that in case a), (4.6) is the 'standard' L^∞-error estimate for the obstacle problem that is usually found in the literature (cf. [2,13,15]).

5. Concluding Remarks. The method that we have employed in the last two sections to derive a priori error estimates for the obstacle problem (P) has some advantages that we would like to point out here:

First of all, our approach is more flexible than the traditional one since we do not require ψ_h to be the Lagrange interpolant $I_h\psi$ of the continuous obstacle ψ (or an element of the finite element space at all, cf. Theorem 4.2). Moreover, we can treat the case $\Omega_h \subset \Omega$ with ease since the relation between the domains Ω and Ω_h is completely irrelevant for the stability analysis that our proofs are based on (cf. Theorem 3.5).

Second, our results provide slightly more information about the behavior of the approximation error than those found in the literature. We obtain, for example, in a natural way separate estimates for the quantities $(u-u_h)^+$ and $(u-u_h)^-$ that allow to study in greater detail how the accuracy of the finite element method is affected by the choice of ψ_h (cf. (4.2) and (4.3)).

Lastly, our approach demonstrates that the problem of estimating the error between the continuous solution u and the finite element approximation u_h can be identified with a problem of sensitivity analysis: If we know how the solution of the discrete obstacle problem (Q_h) changes when the obstacle $\psi + R_h u - u$ is replaced with ψ_h, then we also know how the quantities $u - u_h$ and $u - R_h u$, i.e., the errors associated with the constraint and the unconstraint setting, are related to each other and vice versa. Note that this interpretation is only possible when the obstacles in the discrete problems are allowed to be arbitrary measurable functions (cf. the definition of (Q_h)).

The above perspective on the a priori error analysis turns out to be very advantageous when error estimates in lower L^p-norms are considered. In [4], for example, it was used to construct two counterexamples which demonstrate (among other things) that the estimate (4.6) is optimal in the one-dimensional case in the sense that there exist situations where the assumptions of Corollary 4.5 are satisfied and where it holds $\|u-u_h\|_{L^p(\Omega_h)} = \operatorname{ord}(h^{2-1/q})$ for all $1 \le p \le \infty$. Interestingly, the latter is true regardless of whether the Lagrange interpolant $I_h\psi$ or the restriction $\psi|_{\Omega_h}$ is chosen as ψ_h in (P_h). We refer to [4] for a detailed discussion of this topic.

It should be noted that the situation is much less clear in higher dimensions and that it is (at least to the author's best knowledge) presently unknown if an L^p-error estimate of the form $\|u-u_h\|_{L^p(\Omega_h)} = \mathcal{O}(h^\gamma)$ with $\gamma > 2 - d/q$ and $\gamma > 1$ can be obtained for an obstacle problem with $u, \psi \in W^{2,q}(\Omega)$ if the dimension is greater than one. A further open question is whether the condition (Z) can be weakened. The results found in [6] indicate that the latter might be the case and that it might be sufficient to assume that (3.1) holds in an appropriately chosen subset of Ω_h to derive Theorem 3.4 (cf. also the results in [16]). A proof of this conjecture, however, is still pending.

Appendix A. Finite Element Approximation under Inequality Constraints.

In this section, we prove the approximation results that we have used in the proof of Theorem 3.4. We will be mainly concerned with the following task:

> *Assume that T is a closed d-dimensional simplex with vertices $p_1, ..., p_{d+1}$ and let z be a function satisfying $0 \le z \in C^{1,\gamma}(T)$ for some $0 < \gamma \le 1$. Find an affine linear function z_T such that $0 \le z_T \le z$ holds in T and such that z_T approximates z as accurately as possible.*

The above approximation problem has already been studied by Mosco in [14] and Strang in [18] for one- and two-dimensional H^2-functions. The method of proof that we will employ in this section is closely related to the approach of these two authors. Our analysis, however, also covers the higher-dimensional case.

To construct an approximation z_T with the desired properties, we introduce a partial order on the set of affine linear functions on T (analogously to [14]) and define:

Definition A.1. *Let $T \subset \mathbb{R}^d$ and $0 \le z \in C^{1,\gamma}(T)$, $0 < \gamma \le 1$, be as above and let*

$$V := \{v : T \to \mathbb{R} : v \text{ affine linear with } 0 \le v \le z \text{ in } T\}.$$

Then a function $v \in V$ is called a maximal element of V if for every affine linear function w with $w \not\equiv 0$ and $w \ge 0$ in T it holds $v + w \notin V$.

Using standard arguments, it is easy to prove:

Lemma A.2. *The set V always admits at least one maximal element.*

Proof. If we denote with $p_1, ..., p_{d+1}$ the vertices of T and define

$$U = \left\{ v \in \mathbb{R}^{d+1} : 0 \le \sum_{i=1}^{d+1} \lambda_i v_i \le z\left(\sum_{i=1}^{d+1} \lambda_i p_i\right) \text{ for all } \lambda_i \ge 0 \text{ with } \sum_{i=1}^{d+1} \lambda_i = 1 \right\},$$

then U is closed, non-empty and bounded. Thus, the function $f(v) := \sum_{i=1}^{d+1} v_i$ attains its supremum in U in some v. Because of its maximality, this v has to satisfy $v+w \notin U$ for all $w \in \mathbb{R}^{d+1} \setminus \{0\}$ with $w \ge 0$ (componentwise). This, however, implies that the affine linear map

$$x = \sum_{i=1}^{d+1} \lambda_i p_i \in T \mapsto \sum_{i=1}^{d+1} \lambda_i v_i$$

defined in the barycentric coordinates w.r.t. the vertices p_i of the simplex T is a maximal element of V. □

To estimate the difference between z and a maximal element of V, we observe the following:

Lemma A.3. *Let $T \subset \mathbb{R}^d$ and $0 \le z \in C^{1,\gamma}(T)$, $0 < \gamma \le 1$, be as above and denote with $p_1, ..., p_{d+1}$ the vertices of T. Let v be a maximal element of V and define*

$$E(v) := \{\zeta \in T : z(\zeta) = v(\zeta)\}.$$

Then $E(v)$ is not empty and if it holds $E(v) \subseteq \operatorname{conv}(p_1, ..., p_{k-1}, p_{k+1}, ..., p_{d+1})$, where $\operatorname{conv}(...)$ denotes the convex hull, then there exists a $\zeta \in E(v)$ such that

$$\nabla(z - v)(\zeta) \cdot (p_k - \zeta) = 0. \tag{A.1}$$

Proof. The non-emptiness of $E(v)$ follows trivially from the maximality of v. To prove the second part of the lemma, we assume w.l.o.g. that $k = d+1$, that $\operatorname{conv}(p_1, ..., p_d) \subset \mathbb{R}^{d-1} \times \{0\}$, and that $p_{d+1} \in \mathbb{R}^{d-1} \times (0, \infty)$. Since all $\zeta \in E(v)$ are global minima of the function $z - v \in C^1(T)$, it necessarily holds

$$D(\zeta) := \nabla(z - v)(\zeta) \cdot (p_{d+1} - \zeta) \ge 0 \quad \forall \zeta \in E(v). \tag{A.2}$$

From the compactness of $E(v)$ and the continuity of the function $D : T \to \mathbb{R}$, we obtain that there exists a $\zeta' \in E(v)$ with

$$D(\zeta') = \min_{\zeta \in E(v)} D(\zeta) =: m \geq 0. \tag{A.3}$$

If the minimum m is zero, then the claim is obviously true. If $m > 0$, then there exists an open ball $B(\zeta)$ around each $\zeta \in E(v)$ such that $D(x) > m/2$ holds for all $x \in T \cap B(\zeta)$ and we may define

$$\begin{aligned} &B'(\zeta) \\ &\quad := B(\zeta) \cap \left\{\lambda x + (1-\lambda)p_{d+1} : x \in B(\zeta) \cap T \cap \left(\mathbb{R}^{d-1} \times \{0\}\right) \text{ and } \lambda \in (0,1]\right\} \end{aligned}$$

and

$$E'(v) := \bigcup_{\zeta \in E(v)} B'(\zeta).$$

Note that it follows from our construction that $E(v) \subset E'(v) \subset T$ (cf. Figure A.1). Moreover, $E'(v)$ is relatively open in T and it holds

$$\nabla(z - v)(x) \cdot (p_{d+1} - x) > \frac{1}{2}m \quad \forall x \in E'(v).$$

Suppose now that c is a constant satisfying $0 < c < c' := m/(4\, x_d(p_{d+1}))$, where $x_d(p_{d+1}) > 0$ denotes the d-th coordinate of the point p_{d+1}, and consider the function $v_c(x) := v(x) + c\, x_d$. Then v_c is obviously affine and it holds

$$\nabla(z - v_c)(x) \cdot (p_{d+1} - x) \geq \frac{1}{2}m - c\, x_d(p_{d+1}) \geq \frac{1}{4}m \quad \forall x \in E'(v).$$

From $z - v_c = z - v \geq 0$ on $T \cap (\mathbb{R}^{d-1} \times \{0\})$, the mean value theorem, and the fact that for every $x \in E'(v)$ the line between x and the unique $x' = x'(x)$ with $x' \in T \cap (\mathbb{R}^{d-1} \times \{0\})$ and $x \in \operatorname{conv}(x', p_{d+1})$ is contained in $E'(v)$ (cf. Figure A.1), it follows further

$$\begin{aligned} (z - v_c)(x) &\geq (z - v_c)(x) - (z - v_c)(x') \\ &= \int_0^1 \nabla(z - v_c)(x' + t(x - x')) \cdot (x - x')\mathrm{d}t \\ &= \int_0^1 \nabla(z - v_c)(x' + t(x - x')) \cdot \frac{(p_{d+1} - (x' + t(x - x'))}{\|p_{d+1} - (x' + t(x - x')\|} \|x - x'\| \,\mathrm{d}t \\ &\geq \frac{1}{4}m \int_0^1 \frac{\|x - x'\|}{\|p_{d+1} - (x' + t(x - x')\|}\mathrm{d}t \\ &\geq 0 \qquad \forall x \in E'(v). \end{aligned}$$

To avoid a contradiction with the maximality of v, it now has to hold that for every $0 < c < c'$ there exists at least one $x_c \in T \setminus E'(v)$ with $(z - v_c)(x_c) < 0$. The set $T \setminus E'(v)$, however, is compact. This implies that we can find a sequence $c_n \to 0$ such that x_{c_n} converges to an $x_0 \in T \setminus E'(v)$ and such a limit x_0 has to satisfy

$$0 \geq \lim_{c_n \to 0}(z - v_{c_n})(x_{c_n}) = \lim_{c_n \to 0}(z - v)(x_{c_n}) = (z - v)(x_0) \geq 0,$$

i.e., $x_0 \in E(v) \subset E'(v)$. This is a contradiction to $x_0 \in T \setminus E'(v)$ and shows that the minimum m in (A.3) cannot be positive. ∎

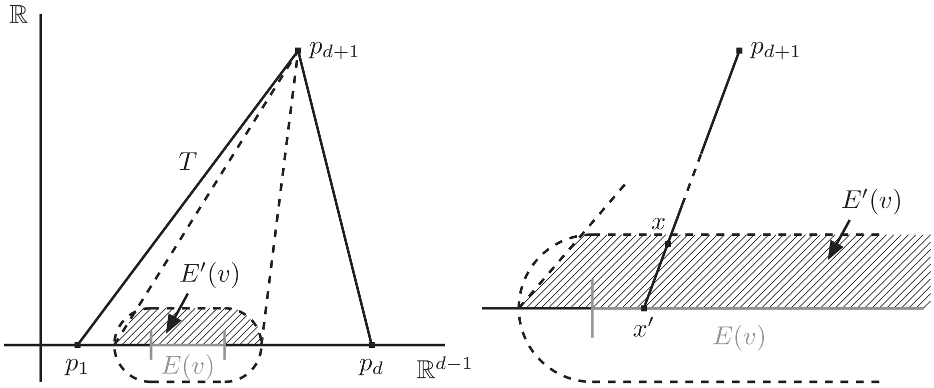

FIG. A.1. *The geometric situation in the proof of Lemma A.3.*

The intuition behind the proof of Lemma A.3 is clear: If $v \in V$ is a function with $E(v) \subseteq \operatorname{conv}(p_1, ..., p_{k-1}, p_{k+1}, ..., p_{d+1})$ such that there is no $\zeta \in E(v)$ with (A.1), then we can increase the value $v(p_k)$ without violating the constraint $0 \le v \le z$ and v cannot be a maximal element. Using Lemma A.3, we obtain:

THEOREM A.4. *Let $T \subset \mathbb{R}^d$, $p_1, ..., p_{d+1}$ and $0 \le z \in C^{1,\gamma}(T)$ be as before. Then there exists an affine linear function $v : T \to \mathbb{R}$ such that $0 \le v \le z$ and*

$$z(p_k) - v(p_k) \le \frac{\sqrt{d}}{1+\gamma} \operatorname{diam}(T)^{1+\gamma} |z|_{C^{1,\gamma}(T)} \quad \forall k = 1, .., d+1.$$

Proof. If v is an arbitrary maximal element of V and p_k a vertex of T, then there are three possibilities: If $p_k \in E(v)$, then it holds $v(p_k) = z(p_k)$ and the claim is certainly true. If, on the other hand, $p_k \notin E(v)$ and there exist $\zeta \in E(v)$ and $\varepsilon > 0$ such that $\zeta + \epsilon(p_k - \zeta), \zeta - \epsilon(p_k - \zeta) \in T$, then (A.2) implies

$$\nabla(z - v)(\zeta) \cdot (p_k - \zeta) = 0$$

and we may compute

$$\begin{aligned}
(z - v)(p_k) &= (z - v)(p_k) - (z - v)(\zeta) \\
&= \int_0^1 \Big[\nabla(z - v)(\zeta + t(p_k - \zeta)) - \nabla(z - v)(\zeta)\Big] \cdot (p_k - \zeta) \mathrm{d}t \\
&= \int_0^1 \Big[\nabla z(\zeta + t(p_k - \zeta)) - \nabla z(\zeta)\Big] \cdot (p_k - \zeta) \mathrm{d}t \\
&\le \int_0^1 \frac{\|\nabla z(\zeta + t(p_k - \zeta)) - \nabla z(\zeta)\|}{\|t(p_k - \zeta)\|^\gamma} \|p_k - \zeta\|^{\gamma+1} t^\gamma \, \mathrm{d}t \\
&\le |z|_{C^{1,\gamma}(T)} \sqrt{d} \int_0^1 \|p_k - \zeta\|^{\gamma+1} t^\gamma \, \mathrm{d}t \\
&\le \frac{\sqrt{d}}{1+\gamma} \operatorname{diam}(T)^{1+\gamma} |z|_{C^{1,\gamma}(T)}. \qquad \text{(A.4)}
\end{aligned}$$

This proves the claim in the second case. If, lastly, $p_k \notin E(v)$ and there are no $\zeta \in E(v)$ and $\varepsilon > 0$ such that $\zeta + \epsilon(p_k - \zeta), \zeta - \epsilon(p_k - \zeta) \in T$, then it necessarily holds $E(v) \subseteq \operatorname{conv}(p_1, ..., p_{k-1}, p_{k+1} ..., p_{n+1})$ and we may employ Lemma A.3 to obtain that there is some $\zeta \in E(v)$ with

$$\nabla(z - v)(\zeta) \cdot (p_k - \zeta) = 0.$$

A calculation analogous to (A.4) now yields the claim. This completes the proof. ☐

It should be noted that there is no straightforward way to construct the approximation v appearing in the last theorem from the Lagrange interpolant of the function z since it is in general unclear how the interpolant has to be modified such that both the constraints $v \leq z$ and $v \geq 0$ are satisfied. We conclude our investigation with the following result about global approximations:

Corollary A.5. *Let $\Omega \subset \mathbb{R}^d$ be a bounded polyhedric domain with a quasi-uniform family of triangulations $\{\mathcal{T}_h\}$ and let ρ_1 and ρ_2 be defined as in Definition 2.2. Assume that $z \in W^{2,q}(\Omega) \cap H_0^1(\Omega)$ for some $d < q < \infty$ and suppose that a family of functions $\{w_h\}$ is given such that*

$$z \leq w_h \text{ a.e. in } \Omega \quad \text{and} \quad w_h \in V_h^0 \quad \forall h > 0,$$

where $V_h^0 := \{v \in C(\overline{\Omega}) : v|_T \text{ is affine for all } T \in \mathcal{T}_h \text{ and } v|_{\partial\Omega} = 0\}$. Then there exists a family of approximations $\{z_h\}$ satisfying

$$z \leq z_h \leq w_h \text{ a.e. in } \Omega \quad \text{and} \quad z_h \in V_h^0$$

for all h such that

$$\|z - z_h\|_{L^\infty} \leq Ch^{2-d/q}\|z\|_{W^{2,q}}$$

holds with a constant C independent of h.

Proof. Let h be arbitrary but fixed, then it follows from $W^{2,q}(\Omega) \hookrightarrow C^{1,1-d/q}(\overline{\Omega})$ and Theorem A.4 that for every $T \in \mathcal{T}_h$ there exists an affine linear $v_T : T \to \mathbb{R}$ such that $0 \leq v_T \leq w_h - z$ and

$$0 \leq (w_h - z)(p_k) - v_T(p_k) \leq \frac{\sqrt{d}}{2 - d/q} \operatorname{diam}(T)^{2-d/q} |z|_{C^{1,1-d/q}(T)}$$

holds for all vertices p_k of T. We now define v_h to be the unique element of V_h^0 with

$$v_h(x_i) = \min_{T \in \mathcal{T}_h : x_i \in T} v_T(x_i)$$

for all mesh nodes x_i. This v_h certainly satisfies $0 \leq v_h \leq v_T \leq w_h - z$ on every mesh cell T and

$$\begin{aligned} 0 \leq (w_h - z)(x_i) - v_h(x_i) &\leq \frac{\sqrt{d}}{2 - d/q} \left(\max_{T \in \mathcal{T}_h : x_i \in T} \operatorname{diam}(T)^{2-d/q} |z|_{C^{1,1-d/q}(T)} \right) \\ &\leq \frac{\sqrt{d}}{2 - d/q} \rho_1^{2-d/q} h^{2-d/q} \max_{T \in \mathcal{T}_h} |z|_{C^{1,1-d/q}(T)} \end{aligned}$$

for all nodes x_i of the mesh. Defining $z_h := w_h - v_h$, we now obtain a function with $z_h \in V_h^0$, $z \leq z_h \leq w_h$ and

$$\|I_h z - z_h\|_{L^\infty} \leq \max_{x_i} |z(x_i) - z_h(x_i)| \leq Ch^{2-d/q}|z|_{C^{1,1-d/q}(\overline{\Omega})}$$

for some $C = C(d, q, \rho_1) > 0$. The claim now follows from the triangle inequality, the Sobolev embedding $W^{2,q}(\Omega) \hookrightarrow C^{1,1-d/q}(\overline{\Omega})$ and well-known error estimates for the L^∞-error of the Lagrange interpolant ([3, Theorem 4.4.20]). ∎

REFERENCES

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

[2] Baiocchi, C.: *Estimation d'erreur dans L^∞ pour les inéquations à obstacle*, Lecture Notes in Mathematics, 606, 27-34 (1977)

[3] Brenner, S.C., and L.R. Scott: *The Mathematical Theory of Finite Element Methods*, 3. Edition, Springer-Verlag, New York (2008)

[4] Christof, C., and C. Meyer: *A Note on A Priori L^p-Error Estimates for the Obstacle Problem*, Ergebnisbericht des Instituts für Angewandte Mathematik Nr. 543, TU Dortmund, Faculty of Mathematics (2016)

[5] Ciarlet, P.G.: *Discrete Maximum Principle for Finite-Difference Operators*, Aequationes Mathematicae, 4, 338-352 (1970)

[6] Draganescu, A., T.F. Dupont and L.R. Scott: *Failure of the Discrete Maximum Principle for an Elliptic Finite Element Problem*, Mathematics of Computation, 74, 1-23 (2005)

[7] Evans, L.C.: *Partial Differential Equations*, 2. Edition, AMS, Providence, RI (2010)

[8] Finzi-Vita, S.: *L^∞-Error Estimates for Variational Inequalities with Hölder-continuous Obstacle*, RAIRO Analyse Numérique, 16, 27-37 (1982)

[9] Gilbart, D., and N.S. Trudinger: *Elliptic Partial Differential Equations of Second Order*, Reprint of the 1998 Edition, Springer-Verlag, Berlin/Heidelberg/New York (2001)

[10] Grisvard, P.: *Elliptic Problems in Nonsmooth Domains*, Pitman, London (1985)

[11] Grisvard, P.: *Singularities in Boundary Value Problems*, Springer-Verlag, Berlin/Heidelberg/New York (1992)

[12] Kinderlehrer, D., and G. Stampacchia: *An Introduction to Variational Inequalities and Their Applications*, SIAM, Philadelphia (2000)

[13] Meyer, C., and O. Thoma: *A Priori Finite Element Error Analysis for Optimal Control of the Obstacle Problem*, SIAM Journal on Numerical Analysis, 51, 605-628 (2013)

[14] Mosco, U., and G. Strang: *One-Sided Approximation and Variational Inequalities*, Bulletin of the American Mathematical Society, 80, 308-312 (1974)

[15] Nitsche, J.: *L^∞-Convergence of Finite Element Approximations*, Lecture Notes in Mathematics, 606, 261-274 (1977)

[16] Nochetto, R.H.: *Pointwise accuracy of a finite element method for nonlinear variational inequalities*, Numerische Mathematik, 54, 601 - 618 (1989)

[17] Schatz, A.H., and L.B. Wahlbin: *On the Quasi-Optimality in L^∞ of the H_0^1-Projection into Finite Element Spaces*, Mathematics of Computation, 38, 1-22 (1982)

[18] Strang, G.: *One-Sided Approximation and Plate Bending*, Computing Methods in Applied Sciences and Engineering Part 1, Lecture Notes in Computer Science, 10, 140-155 (1974)

[19] Xu, J., and L. Zikatanov: *A Monotone Finite Element Scheme for Convection Diffusion Equations*, Mathematics of Computation, 68, 1429-1446 (1999)