OBSTACLE PROBLEMS: THEORY AND APPLICATIONS

by

Borjan Geshkovski

Contents

General notation and conventions	3
Introduction	5 7
 Part I. Obstacle problems. 1. The classical problem. 2. The parabolic problem. 3. Numerical experiments. 	14 14 29 45
 Part II. An obstacle problem with cohesion	52 52 55 59 64
Part III. Optimal control of obstacle problems8. Overview of the elliptic problem9. The parabolic problemRemarks and further topics	73 73 74 90
Appendix. Convexity, coercivity, weak lower semicontinuity. Inequalities. Monotone operator theory. 1 References. 1	93 93 98 100

 $\mathbf{2}$

General notation and conventions

- Throughout this work, $n \ge 1$ is a fixed dimension.
- All of the functions are assumed to be real-valued.
- For a point $x_0 \in \mathbb{R}^n$ and a radius r > 0, we denote by $B_r(x_0)$ the open ball $B_r(x_0) := \{x \in \mathbb{R}^n : |x x_0| < r\}.$
- We will henceforth denote by $\Omega \subset \mathbb{R}^n$ a bounded domain (open connected set) with boundary $\partial\Omega$ of class C^{∞} .

We make precise the notion of smoothness for the boundary. Let $n \geq 2$. We say that $\partial \Omega$ is of class C^k for $k \in \mathbb{N} \cup \{0\}$ if for each point $x_0 \in \partial \Omega$ there exist r > 0 and $\varphi \in C^k(\mathbb{R}^{n-1})$ such that we have

$$\Omega \cap B_r(x_0) = \{ x \in B_r(x_0) \colon x_n > \varphi(x_1, \dots, x_{n-1}) \}$$

upon reorienting the coordinates¹. The boundary is of class C^{∞} (and said to be smooth) if it is of class C^k for every $k \in \mathbb{N}$.

• For a multiindex $\alpha \in \mathbb{N}^n$ and a function $u : \Omega \to \mathbb{R}$, we define $D^{\alpha}u := \partial_{x_1}^{\alpha_1} \dots \partial_{x_n}^{\alpha_n} u$. If $k \in \mathbb{N}$, then $D^k u := \{D^{\alpha}u : \alpha \in \mathbb{N}^n, |\alpha| = k\}$, and also

$$|D^{k}u| = \left(\sum_{|\alpha|=k} |D^{\alpha}u|^{2}\right)^{\frac{1}{2}}.$$

We distinguish the special cases k = 1 where $Du = \nabla u = (u_{x_1}, \ldots, u_{x_n})$ is the gradient, and k = 2 where $D^2 u = (u_{x_i x_j})_{i,j=1}^n$ is the Hessian matrix, and $\Delta u := \operatorname{trace}(D^2 u)$.

If $m \geq 2$, for a vector field $\mathbf{u} : \Omega \to \mathbb{R}^m$, $\mathbf{u} = (u^1, \ldots, u^m)$, we set $D^{\alpha}\mathbf{u} := (D^{\alpha}u^1, \ldots, D^{\alpha}u^m)$ for any multiindex $\alpha \in \mathbb{N}^n$. The remaining definitions are identical to the scalar case. We distinguish the special case k = 1 where Du is the Jacobian matrix. If moreover m = n, then div $\mathbf{u} := \operatorname{trace}(D\mathbf{u})$. Given another vector field $\mathbf{v} : \Omega \to \mathbb{R}^m$, $\mathbf{v} = (v^1, \ldots, v^m)$, we define

$$D\mathbf{u}: D\mathbf{v} := \sum_{i=1}^m \sum_{j=1}^m u_{x_j}^i v_{x_j}^i.$$

• For any $k \in \mathbb{N}$ and $1 \leq p \leq \infty$, the Sobolev space $W^{k,p}(\Omega)$ consists of all functions $u \in L^p(\Omega)$ for which the weak derivatives $D^{\alpha}u, \alpha \in \mathbb{N}^n, |\alpha| \leq k$ exist

¹ Intuitively, one takes a ball around any point on the boundary and transforms the part of the domain in the ball to the upper half plane with a C^k function.

and are in $L^{p}(\Omega)$. We endow $W^{k,p}(\Omega)$ with the norm

$$||u||_{W^{k,p}(\Omega)} := \sum_{|\alpha| \le k} ||D^{\alpha}u||_{L^{p}(\Omega)}.$$

• For any $\alpha \in (0,1]$ and $k \in \mathbb{N} \cup \{0\}$, the Hölder space $C^{k,\alpha}(\Omega)$ consists of all functions $u \in C^k(\overline{\Omega})$ for which the norm

$$\|u\|_{C^{k,\alpha}(\Omega)} := \sum_{|\beta| \le k} \sup_{x \in \overline{\Omega}} |D^{\beta}u(x)| + \sum_{|\beta| = k} \left[D^{\beta}u\right]_{C^{0,\alpha}(\overline{\Omega})}$$

is finite. Here $[u]_{C^{0,\alpha}(\overline{\Omega})}$ denotes the α^{th} -Hölder seminorm of u:

$$[u]_{C^{0,\alpha}(\overline{\Omega})} := \sup_{x,y \in \Omega, x \neq y} \frac{|u(x) - u(y)|}{|x - y|^{\alpha}}$$

- For two normed vector spaces $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ such that $X \subseteq Y$, we write $X \hookrightarrow Y$ if X is continuously embedded in Y, meaning if the inclusion map $\iota: X \to Y$ mapping x to itself is continuous.
- We write $K \subseteq \Omega \subset \mathbb{R}^n$ if $K \subset \overline{K} \subset \Omega$ and \overline{K} is compact, and say that K is compactly contained in Ω .
- Let T > 0 and let X denote a real Banach space. For $p \in [1, \infty)$, the space $L^p(0, T; X)$ consists of all measurable functions $u : [0, T] \to X$ with

$$\|u\|_{L^{p}(0,T;X)} := \left(\int_{0}^{T} \|u(t,\cdot)\|_{X}^{p} \mathrm{d}t\right)^{\frac{1}{p}} < \infty.$$

Similarly, the space $L^{\infty}(0,T;X)$ consists of all measurable functions $u:[0,T] \to X$ with

$$\|u\|_{L^{\infty}(0,T;X)} := \underset{t \in [0,T]}{\operatorname{ess}} \sup \|u(t,\cdot)\|_{X} < \infty,$$

and $C^0([0,T];X)$ consists of all continuous functions $u:[0,T] \to X$ with

$$||u||_{C^0([0,T];X)} := \max_{t \in [0,T]} ||u(t, \cdot)||_X < \infty.$$

- For a function $u: [0,T] \to X$, the first variable will play the role of time and the latter the role of space. The time derivative will be denoted by u_t and acts on the first variable. The gradient ∇ and Laplacian Δ act only on the second variable.
- We will use the following identification of duals:

$$[L^2(0,T;H^1_0(\Omega))]' \cong L^2(0,T;H^{-1}(\Omega)).$$

• We denote by $\mathbb{1}_E$ the indicator function of the set E.

INTRODUCTION

Many phenomena in physics, biology and finance can be described by partial differential equations that display a priori unknown interfaces or boundaries. Such problems are called *free boundary problems*. One of the simplest and most important free boundary problems is the *obstacle problem*, in which, at least formally, a function u solves the Poisson or heat equation on the set where it is strictly above a certain function ψ , and equals this function elsewhere.

From a mathematical perspective, in studying obstacle problems one may ask similar questions as for classical partial differential equations. One begins by investigating the existence and uniqueness of an appropriately defined solution, and additionally, to further study the regularity of this solution. Then one seeks to conceive numerical schemes, study their convergence and conduct computer simulations of the respective problems. Understanding the regularity of the solution and of the free boundary in particular is a quite difficult question in general. For example, one may ask whether there is a regularization mechanism (as for the heat equation) that smoothness out the solution and the free boundary independently of the initial data. A key difficulty is that for solutions of elliptic or parabolic PDEs, one has an equation for a function u, and such equation forces u to be regular with respect to supplied data. In free boundary problems, such a task is more difficult as one does not have a "regularizing" equation along the free boundary, but only an equation for the solution which indirectly determines this interface.

As seen in what follows, we may rewrite obstacle problems as *variational inequalities*, from which approximation may be seen as an intuitive approach. In particular, for studying the existence, uniqueness and regularity of solutions (up to some level), we will use the *penalization method*, which may be summarized by the following scheme:

- 1. Approximate the problem by semilinear PDEs;
- 2. Show **existence** and **uniqueness** of a solution for the **approximate problems**;
- 3. Obtain estimates for the approximate solutions;
- 4. Use the estimates for a compactness argument to **pass** to the limit.

We shall in fact use a generalization of this technique for showing existence of solutions for most of the problems we will encounter, such as constrained minimization. We put an emphasis on the idea of approximation: we will approximate the "difficult" problem by a family of "simpler" problems, and then proceed as indicated.

The plan of this master's thesis is as follows.

In Part I, we revisit the well-known results on existence, uniqueness and regularity of solutions for both the elliptic (which we refer to as *classical*) and parabolic obstacle problem. We follow the books [21, 30, 41] for the classical problem, and present different methods for obtaining the aforementioned results. The parabolic problem has different variants in the literature (depending on whether the obstacle varies with time or not); we present in detail results that are briefly discussed in [2]. In fact, slightly stronger results hold. Numerical experiments are also conducted, and are based on the same technique we used for the theoretical study. The free boundary is not considered as a part of the solution, as studying its regularity is a separate topic.

In Part II, we present the paper [27] of Hintermüller, Kovtunenko and Kunisch, in which the authors investigate a steady-state obstacle problem taking into consideration molecular cohesion forces. From a mathematical perspective, this problem may be seen as a constrained minimization problem in a function space setting, and the minimizer of the objective functional in question will be shown to satisfy a particular obstacle problem.

In Part III, we present the paper [2] of Adams and Lenhart, where the authors look at the parabolic obstacle problem from the viewpoint of optimal control. Namely, the obstacle is considered as the control and one looks to drive the state (solution of the parabolic obstacle problem) to some given target profile. This is done by minimizing over all admissible obstacles the error between the state associated to the obstacle and the target, in such a way that the state is constrained to satisfy the parabolic obstacle problem. We only give an outline of the results for the elliptic case.

Acknowledgements. — I would like to thank my advisor Professor Enrique Zuazua for his guidance, advice, patience and for inviting me to stay in Bilbao for three months within his research team at DeustoTech and study this topic for my master's thesis. This research was supported by the Advanced Grant DyCon (Dynamical Control) of the European Research Council Executive Agency (ERC). I also would like to thank all of the remaining members of the DyCon team for helping me during my stay, for both mathematical and administrative issues. I also thank the authors of [27] for their suggestions and references. I finally thank Professor Marius Tucsnak for his help and advice during my studies.



FIGURE 1. The obstacle ψ , the solution u, and the free boundary $\partial \{u > \psi\}$. This figure was adapted from [44].

Motivating physical examples

We briefly describe some appearances of obstacle problems in natural phenomena. Most of these examples are discussed in the recent survey of Ros-Oton [44], as well as in the books of Rodrigues [43], Duvaut and Lions [19], Kinderlehrer and Stampacchia [30], Chipot [17] and Friedman [23].

Elasticity. — The (classical) obstacle problem may be derived from its original consideration as a problem that arises in linear elasticity theory — the mathematical study of how solids deform and become internally stressed due to prescribed loading conditions. In classical elasticity theory, a *membrane* is a thin plate which offers no resistance to bending, and acts only upon tension (stretching). We are given a homogeneous membrane occupying a domain Ω in the plane \mathbb{R}^2 ; the membrane is equally stretched in every direction by a uniform tension τ and loaded, i.e. acted upon, by a normal uniformly distributed force f.

It is natural to suppose that each point (x, y) of the membrane is displaced by an amount u(x, y) perpendicularly to the plane \mathbb{R}^2 . The boundary $\partial\Omega$ of the membrane is deformed conformly by prescribing its displacement g. In other words, we prescribe a Dirichlet boundary condition

u = g on $\partial \Omega$.

For the mathematical study we will exclusively consider homogeneous Dirichlet boundary conditions: $g \equiv 0$. Now assume that the potential energy of the deformed membrane is proportional to the increase in the area of its surface. For small deformations, we may neglect higher order derivatives and use Taylor's theorem to approximate the surface area by

$$\int_{\Omega} \sqrt{1 + u_x^2 + u_y^2} \mathrm{d}x \mathrm{d}y \approx \int_{\Omega} \left(1 + \frac{1}{2} (u_x^2 + u_y^2) \right) \mathrm{d}x \mathrm{d}y,$$

and the change in the area of the membrane is equal to

$$\frac{1}{2}\int_{\Omega}(u_x^2+u_y^2)\mathrm{d}x\mathrm{d}y = \frac{1}{2}\int_{\Omega}|\nabla u|^2\mathrm{d}x\mathrm{d}y.$$

The potential energy of deformation has the functional expression

$$\mathbf{U} = \frac{\lambda}{2} \int_{\Omega} |\nabla u|^2 \mathrm{d}x \mathrm{d}y,$$

where $\lambda > 0$ is a constant depending on the elastic properties of the membrane. For simplicity, we assume $\lambda = 1$. The work done by the external forces during the actual displacement is given by

$$\mathbf{V} = -\int_{\Omega} f u \mathrm{d}x \mathrm{d}y,$$

so that the total potential energy $\mathcal{E} := \mathbf{U} + \mathbf{V}$ by definition reads

$$\mathcal{E}[u] = \frac{1}{2} \int_{\Omega} |\nabla u|^2 \mathrm{d}x \mathrm{d}y - \int_{\Omega} f u \mathrm{d}x \mathrm{d}y.$$

To find the equilibrium position of the membrane, the minimum total potential energy principle² is applied. The problem reduces to finding, among all functions u of finite energy of deformation **U** and satisfying the Dirichlet boundary condition, the one that minimizes the potential energy \mathcal{E} . Using results from the calculus of variations, the necessary condition for the minimizer of \mathcal{E} (called the *Euler-Lagrange equation*) is given by the Poisson equation

$$-\Delta u = f$$
 in Ω .

This is known as *Dirichlet's principle*.

² Roughly stated, for conservative structural systems, of all the kinematically admissible deformations those corresponding to the equilibrium state minimize or maximize the total potential energy. Moreover, if the extremum is a minimum, the equilibrium state is stable.

For the obstacle problem, one seeks the equilibrium position with an additional constraint on the membrane. Namely, one looks for the equilibrium position of the membrane such that it lies above a body represented by

$$\{(x, y, z) \in \mathbb{R}^3 \colon z \le \psi(x, y)\}$$

with a fixed height g on the boundary. The function ψ is the *obstacle*, defined on Ω and satisfying $\psi \leq g$ on the boundary $\partial \Omega$. By the minimum total potential energy principle, this problem reduces to finding, among all functions u of finite energy of deformation \mathbf{U} , satisfying $u \geq \psi$ in Ω and the Dirichlet boundary condition, the one that minimizes the potential energy \mathcal{E} . Deducing the necessary condition for the minimizer u will require subtle arguments. We will see that it manifests as the following Euler-Lagrange equation:

$$\begin{cases} u \ge \psi & \text{in } \Omega \\ -\Delta u = f & \text{in } \{u > \psi\} \\ -\Delta u \ge f & \text{in } \Omega. \end{cases}$$

The Stefan problem. — The Stefan problem describes the temperature distribution in a homogeneous medium undergoing a phase change. An elementary example is the melting of ice submerged in a body of liquid water. In the simplest case, the temperature distribution function θ solves the homogeneous heat equation $\theta_t - \Delta \theta = 0$ in the set $\{\theta > 0\}$, and equals zero elsewhere. It was shown by Duvaut [18] that by considering the function

$$u(t,x) = \int_0^t \theta(s,x) \mathrm{d}s,$$

the Stefan problem transforms into

$$\begin{cases} u \ge 0 & \text{in } \Omega \\ u_t - \Delta u = -1 & \text{in } \{u > 0\} \\ u_t - \Delta u \ge -1 & \text{in } \Omega, \end{cases}$$

which are the Euler-Lagrange equations for the parabolic obstacle problem.

Optimal stopping and financial mathematics. — An interesting occurrence of obstacle problems is in probability and finance.

The mathematical setting for many problems in optimal control theory is the following. We are given some system whose state evolves in time according to a differential equation (deterministic or stochastic), and also certain controls which

affect the behavior of the system in some way. These controls typically either modify some parameters in the dynamics or else stop the process, or both. We are finally given a cost criterion, depending upon our choice of control and the corresponding state of the system. The goal is to discover an optimal choice of controls so that the cost criterion is minimal. We present, without much rigor, the following stochastic control model which may be found in [41, 22].

Let $\Omega \subset \mathbb{R}^n$ be a bounded domain with smooth boundary, and let $\mathbf{X} = (\mathbf{X}(t))_{t\geq 0}$ be a stochastic (diffusion) process starting at $\mathbf{X}(0) = x \in \Omega$. Let τ be a *hitting time* of $\partial\Omega$, which loosely means that τ is the first time at which $\mathbf{X}(\tau)$ "hits" (i.e. touches) the boundary $\partial\Omega$. Also, let θ be a *stopping time*, which loosely means that the process \mathbf{X} will exhibit some "behavior of interest" at this time. In particular, the hitting time τ is also a stopping time (see [**22**, p.103]).

For each θ , the expected cost of stopping **X** at time $\theta \wedge \tau := \min\{\theta, \tau\}$ is defined by

$$\mathcal{J}[\theta] := \mathbb{E}\Big[\int_0^{\theta \wedge \tau} \frac{1}{2} f(\mathbf{X}(s)) \mathrm{d}s + \psi(\mathbf{X}(\theta \wedge \tau))\Big],$$

where f and ψ are given smooth functions on $\overline{\Omega}$ and \mathbb{E} denotes the expected value. The main question is to see whether there exists a stopping time θ^* for which

$$\mathcal{J}[\theta^*] = \min_{\theta \text{ stopping time }} \mathcal{J}[\theta]$$

and if so, how can we compute it. To this end, we turn our attention to the value function

$$u(x) := \inf_{\theta} \mathcal{J}[\theta],$$

and try to figure out what u is as a function of $x \in \Omega$. Once we know u, we look to "construct" an optimal stopping time θ^* . This is known as *dynamic programming*.

In [22, p.112] it is shown that the optimality conditions for the value function u are of the form

$$\begin{cases} \max\{-Lu - f, u - \psi\} = 0 & \text{in } \Omega\\ u = \psi & \text{on } \partial\Omega \end{cases}$$

where L denotes the infinitesimal generator of the process **X** (an operator which describes the movement of the process in an infinitesimal time interval). If **X** is the Brownian motion, then $L = \frac{1}{2}\Delta$ and we see that the value function solves an obstacle problem. It can be shown that the optimal stopping time θ^* is the first hitting time of the contact region

$$\{x \in \Omega \colon u(x) = \psi(x)\}.$$



FIGURE 2. A free boundary separates the two regions: the one in which we should exercise the option, and the one in which it is better to wait. This figure was adapted from [44].

In financial mathematics, similar problems (in which we maximize rather than minimize, and often include evolution, so there is an added time derivative) appear as models for pricing American options³ (see [32]). Set $f \equiv 0$ for convenience. The function ψ represents the option's *payoff*, and the set $\{u = \psi\}$ is the *exercise region*. Notice that in this context the most important unknown to understand is the exercise region. More particularly, one looks to find and/or understand the two regions $\{u = \psi\}$ (in which one should exercise the option) and $\{u > \psi\}$ (in which one should wait and not exercise the option yet).

In finance and the theory of option evaluation however, one usually needs to consider jump processes (see [42]) instead of diffusion processes, so that discontinuous paths in the dynamics of the stock's prices are introduced. Such models would allow taking into account large price changes and in turn are a reasonable model for market fluctuations. Such models were introduced in finance in the 1970s by Nobel Prize winner R.C. Merton [38].

The operator L in these models will be singular integral operator of the form

$$Lu(x) = \lim_{\varepsilon \to 0^+} \int_{\{|x-y| > \varepsilon\}} \left[u(y) - u(x) \right] \kappa(x-y) \mathrm{d}y.$$

³ In finance, an option is a contract giving a buyer the right to sell an asset at a specified price on a specified date. There are two option styles depending on this expiration date: European (the option may only be exercised at the expiration date) and American (the option may be exercised at any time before the expiration date).

A classical example for a kernel is $\kappa(x-y) = |x-y|^{n+2s}$ for some $s \in (0,1)$. Such a choice gives $L = (-\Delta)^s$ modulo a normalization constant, where $(-\Delta)^s$ is the so-called fractional Laplacian⁴.

Potential theory, interactions in biology and materials science. — Many phenomena in biology and materials science give rise to models with interacting particles or individuals. We present without going into detail the following model formulated mathematically in [8, 15].

Let $W \in L^1_{loc}(\mathbb{R}^n)$ be a non-negative, lower semicontinuous function and consider the *interaction energy*

$$\mathcal{E}_{W}[\mu] := \frac{1}{2} \int_{\mathbb{R}^{n}} \int_{\mathbb{R}^{n}} W(x-y) \mathrm{d}\mu(y) \mathrm{d}\mu(x),$$

where μ is some regular Borel probability measure on \mathbb{R}^n . The interaction energy \mathcal{E}_W may be used to model interplay between particles via pairwise interactions. For example, m particles located at $X_1, \ldots, X_m \in \mathbb{R}^n$ have their discrete interaction energy given by

$$\mathcal{E}_W^m[X_1, \dots, X_m] := \frac{1}{2m^2} \sum_{i=1}^m \sum_{j=1, i \neq j}^m W(X_i - X_j).$$

Formally, when m is large, the discrete energy \mathcal{E}_W^m may be approximated by the continuum energy \mathcal{E}_W where $d\mu(x)$ describes is a general distribution of particles at the location $x \in \mathbb{R}^n$. In fact, for the distribution $\mu = \frac{1}{n} \sum_{i=1}^m \delta_{X_i}$ where δ_a denotes the Dirac mass at a point a, the energy $\mathcal{E}_W[\mu]$ reduces to the discrete energy \mathcal{E}_W^m .

In models arising in biology and materials science, particles, molecules or individuals in a social aggregate, like a flock of birds or a school, self-organize in order to minimize energies similar to \mathcal{E}_W^m . In these applications, the kernel W is *repulsive* in the short range, i.e. when the particles (or individuals) are very close so that they don't collide, and *attractive* in the long range, i.e. when they are far from each other so that they so that they gather to form a group or a structure. Naturally, this leads one to consider kernels of the form W(x) = w(|x|), where $w : [0, +\infty) \to (-\infty, +\infty]$ is decreasing on $[0, r_0)$ and increasing on $(r_0, +\infty)$ for some $r_0 > 0$. An example is

$$\mathcal{F}((-\Delta)^s f)(\xi) = |\xi|^{2s} \mathcal{F}(f)(\xi)$$

⁴ It is natural to define $(-\Delta)^s$ as a Fourier multiplier, namely

for $\xi \in \mathbb{R}^n$ and Schwartz functions $f \in \mathcal{S}(\mathbb{R}^n)$. It can be shown (using Fubini's theorem) that this definition coincides with the singular integral definition above. For more detail on this point, we refer to [44].

the Newtonian repulsion: $W(x) \sim \frac{1}{|x|^{n-2}}$ as $x \to 0$ for $n \ge 2$ and $W(x) \sim \log \frac{1}{|x|}$ if n = 2.

It is rigorously shown in [15] that when W is a Newtonian repulsion, the potentials $W * \mu$ associated to local minimizers μ of the interaction energy \mathcal{E}_W locally solve an obstacle problem. More particularly, let⁵

$$V(x) = \begin{cases} \frac{1}{n(n-2)\omega_n} \frac{1}{|x|^{n-2}} & \text{if } n \ge 3\\ \frac{1}{2\pi} \log \frac{1}{|x|} & \text{if } n = 2, \end{cases}$$

be the fundamental solution of the Laplace equation, thus satisfying

$$-\Delta V = \delta_0$$

and consider the function $W_a = W - V$. Thence $W = W_a + V$, and in some sense V describes the repulsive interactions whereas W_a describes the attractive interactions. For a minimizer μ of \mathcal{E}_W in some ε -ball (see [15] for definitions), consider the potential

$$u(x) := W * \mu(x) = \int_{\mathbb{R}^n} W(x - y) \mathrm{d}\mu(y).$$

Under some regularity assumptions on W_a , it is shown that for any $x_0 \in \text{supp}(\mu)$, the potential u is the unique solution to the obstacle problem

$$\begin{cases} u \ge u(x_0) & \text{in } B_{\varepsilon}(x_0) \\ -\Delta u \ge -\Delta W_a * \mu & \text{in } B_{\varepsilon}(x_0) \\ -\Delta u = -\Delta W_a * \mu & \text{in } B_{\varepsilon}(x_0) \cap \{u > u(x_0)\}. \end{cases}$$

Here $\operatorname{supp}(\mu)$ denotes the support of the Borel measure μ , defined as

$$\operatorname{supp}(\mu) := \overline{\{x \in \mathbb{R}^n \colon \mu(B_{\varepsilon}(x)) > 0 \text{ for all } \varepsilon > 0\}}.$$

5 Here $\omega_n = \frac{\pi^{n/2}}{\Gamma(\frac{n}{2}+1)}$ is the volume of the unit ball in \mathbb{R}^n .

PART I OBSTACLE PROBLEMS

1. The classical problem

The classical obstacle problem, as discussed in what precedes, is one of the most well known and motivating examples in the study of both variational inequalities and free boundary problems. Its simplest mathematical formulation is analogous to its interpretation from elasticity theory: we seek for minimizers of the *Dirichlet* energy functional

(1.1)
$$\mathcal{E}[w] := \frac{1}{2} \int_{\Omega} |\nabla w|^2 \mathrm{d}x - \int_{\Omega} f w \mathrm{d}x,$$

among all functions w belonging to the set

$$\mathcal{K}(\psi) := \{ w \in H_0^1(\Omega) \colon w \ge \psi \text{ a.e. in } \Omega \},\$$

where $\Omega \subset \mathbb{R}^n$ is a bounded domain with smooth boundary $\partial\Omega$, and the obstacle function $\psi \in H^2(\Omega) \cap C^0(\overline{\Omega}), \psi \leq 0$ on $\partial\Omega$ is given. The set $\mathcal{K}(\psi)$ is closed, convex and non-empty (see Proposition A.1 in the Appendix) and comprises those functions those functions $w \in H_0^1(\Omega)$ satisfying the unilateral (one-sided) constraint $u \geq \psi$. We also suppose that the source term $f \in L^2(\Omega)$ is given.

1.1. Existence and uniqueness of a solution. — We begin our study of the classical obstacle problem by showing existence and uniqueness of a solution. This can be done in lots of ways, the classical approach being to use tools from the calculus of variations.

Theorem 1.1. — There exists a unique function $u \in \mathcal{K}(\psi)$ satisfying

$$\mathcal{E}[u] = \inf_{w \in \mathcal{K}(\psi)} \mathcal{E}[w].$$

Proof. — The existence of a minimizer follows from the *direct method in the calculus* of variations. This method consists in exhibiting a so-called minimizing sequence, i.e. a sequence $\{u_k\}_{k=1}^{\infty} \subset \mathcal{K}(\psi)$ satisfying

$$\lim_{k \to \infty} \mathcal{E}[u_k] = \inf_{w \in \mathcal{K}(\psi)} \mathcal{E}[w].$$

Such a sequence always exists by the definition of the infimum⁶. One then uses growth properties of the functional \mathcal{E} at infinity and a compactness argument to deduce a limit for the sequence, which is then shown to be the wanted minimizer by virtue of certain continuity properties of \mathcal{E} in the weak topology.

Let $\{u_k\}_{k=1}^{\infty} \subset \mathcal{K}(\psi)$ be a minimizing sequence for \mathcal{E} . Since \mathcal{E} is coercive (see Proposition A.3 in the Appendix), meaning $\mathcal{E}[w] \to +\infty$ as $\|w\|_{H_0^1(\Omega)} \to \infty$, by contraposition we deduce that $\{u_k\}_{k=1}^{\infty}$ is bounded in $H_0^1(\Omega)$. By virtue of the Banach-Alaoglu theorem, there exist $u \in H_0^1(\Omega)$ and a subsequence $\{u_{k_j}\}_{j=1}^{\infty}$ of $\{u_k\}_{k=1}^{\infty}$ such that

$$u_{k_i} \rightharpoonup u$$
 weakly in $H_0^1(\Omega)$

as $j \to \infty$. Since $\mathcal{K}(\psi)$ is closed and convex it is weakly closed (see Theorem A.4 in the Appendix), whence $u \in \mathcal{K}(\psi)$. The weak lower semicontinuity (see Proposition A.6 in the Appendix) of \mathcal{E} reads

$$\liminf_{j\to\infty} \mathcal{E}[u_{k_j}] \ge \mathcal{E}[u],$$

and using the definitions of the minimizing sequence and the infimum, we deduce that

$$\mathcal{E}[u] \ge \inf_{w \in \mathcal{K}(\psi)} \mathcal{E}[w] \ge \liminf_{j \to \infty} \mathcal{E}[u_{k_j}] \ge \mathcal{E}[u].$$

Hence u is a minimizer of \mathcal{E} .

To demonstrate uniqueness, we argue by contradiction. Let u_1, u_2 with $u_1 \not\equiv u_2$ be two minimizers. Since $\mathcal{K}(\psi)$ is convex, $w := \frac{u_1+u_2}{2} \in \mathcal{K}(\psi)$, and by the strict convexity of \mathcal{E} (see Proposition A.2 in the Appendix), it follows that

$$\mathcal{E}[w] < \frac{1}{2}\mathcal{E}[u_1] + \frac{1}{2}\mathcal{E}[u_2].$$

This is a contradiction, since u_1 and u_2 are minimizers.

In fact, we may also extract the existence and uniqueness of a minimizer by using the properties of the functional \mathcal{E} (see the Appendix) coupled with an abstract result from the theory of convex minimization. To state this result, we will need the following definition.

Definition 1.2. — Let X be a Banach space and let $F : X \to \mathbb{R} \cup \{+\infty\}$ be some map. The set on which F is finite is called the effective domain of F

$$\operatorname{dom}(F) := \{ x \in X \colon F(x) < +\infty \}.$$

⁶ For the infimum to exist, in turn, we need \mathcal{E} to be bounded from below. This is indeed the case, as seen in the proof of Proposition A.3.

If dom $(F) \neq \emptyset$, then F is called proper.

Theorem 1.3. — Let X be a reflexive Banach space, $\mathcal{K} \subset X$ a nonempty, closed and convex subset and $F : \mathcal{K} \to \mathbb{R} \cup \{+\infty\}$ a proper, convex and lower semicontinuous map. If either \mathcal{K} is bounded or F is coercive, then there exists at least one $x^* \in \mathcal{K} \cap \operatorname{dom}(F)$ such that

$$F(x^{\star}) = \inf_{x \in \mathcal{K}} F(x).$$

If F is strictly convex, then x^* is unique.

For a proof and more detail on the topic of convex minimization, we refer to [9, Thm.2.11, p.72].

Next, we look to derive necessary conditions for the solution of the obstacle problem, namely the Euler-Lagrange equations mentioned in the introduction. We start with the following variational characterization of the solution.

Proposition 1.1. — Let $u \in \mathfrak{K}(\psi)$ be the unique minimizer of \mathcal{E} . Then

(1.2)
$$\int_{\Omega} \nabla u \cdot \nabla (v-u) \mathrm{d}x \ge \int_{\Omega} f(v-u) \mathrm{d}x$$

for all $v \in \mathfrak{K}(\psi)$. We call (1.2) an elliptic variational inequality.

Proof. — Fix $v \in \mathcal{K}(\psi)$. Then for each $\tau \in [0, 1]$,

$$u + \tau(v - u) = (1 - \tau)u + \tau v \in \mathcal{K}(\psi)$$

since $\mathcal{K}(\psi)$ is convex. Thus if we set $e(\tau) := \mathcal{E}[u + \tau(v - u)]$, we see that $e(0) = \mathcal{E}[u]$, so $e(0) \le e(\tau)$ for all $\tau \in [0, 1]$. Hence⁷

(1.3)
$$0 \le e'(0).$$

Now if $\tau \in (0, 1]$, then

$$\frac{e(\tau) - e(0)}{\tau} = \frac{1}{\tau} \int_{\Omega} \frac{|\nabla u + \tau \nabla (v - u)|^2 - |\nabla u|^2}{2} \mathrm{d}x - \int_{\Omega} f(u + \tau (v - u) - u) \mathrm{d}x$$
$$= \int_{\Omega} \left(\nabla u \cdot \nabla (v - u) + \frac{\tau |\nabla (v - u)|^2}{2} \right) \mathrm{d}x - \int_{\Omega} f(v - u) \mathrm{d}x.$$

⁷ It must be said that e' denotes the one-sided derivative at 0, namely the limit when $\tau \to 0$ with positive values. This approach does not hold for $\tau < 0$, since we cannot exploit the convexity of $\mathcal{K}(\psi)$, as $u + \tau(v - u) \notin \mathcal{K}(\psi)$ for every $v \in \mathcal{K}(\psi)$. In other words, we can only take "one-sided" variations away from the constraint. It is why this first order optimality condition manifests itself as a variational inequality, rather than a PDE in weak form.

As the integral on the right hand side is an affine function in τ , by letting $\tau \to 0$ and using (1.3) we deduce that

$$0 \le e'(0) = \int_{\Omega} \nabla u \cdot \nabla (v - u) dx - \int_{\Omega} f(v - u) dx.$$

Remark 1.4. — The previous proof also allows us to conclude that $\mathcal{E} : H_0^1(\Omega) \to \mathbb{R}$ is *Gâteaux differentiable*, meaning that for any $u \in H_0^1(\Omega)$ there exists a bounded linear operator $\delta \mathcal{E}[u; \cdot] : H_0^1(\Omega) \to \mathbb{R}$ such that

$$\delta \mathcal{E}[u;v] = \lim_{\tau \to 0} \frac{\mathcal{E}[u+\tau v] - \mathcal{E}[u]}{\tau}$$

for all $v \in H_0^1(\Omega)$. The *Gâteaux derivative* (sometimes called the *first variation*) of \mathcal{E} at $u \in H_0^1(\Omega)$ is given by

$$\delta \mathcal{E}[u; \cdot] : v \mapsto \int_{\Omega} \nabla u \cdot \nabla v \mathrm{d}x - \int_{\Omega} f v \mathrm{d}x.$$

The variational inequality (1.2), which represents the first order optimality condition, is in fact often used in the literature to state the classical obstacle problem (see [**30**]). This is due to the fact that by *Stampacchia's theorem* the converse holds⁸, so solving the variational inequality (1.2) is equivalent to minimizing \mathcal{E} . Consequently, this theorem would represent another way for showing the existence and uniqueness of a minimizer to \mathcal{E} .

Theorem 1.5 (Stampacchia). — Let H be a real Hilbert space, $\mathcal{K} \subset H$ a nonempty closed and convex subset and $\mathfrak{a} : H \times H \to \mathbb{R}$ a bilinear, continuous and H-elliptic ⁹ form. Then, given $\varphi \in H'$, there exists a unique element $u \in \mathcal{K}$ such that

$$\mathfrak{a}(u, v - u) \ge \varphi(v - u)$$

for all $v \in \mathcal{K}$. Moreover, if \mathfrak{a} is symmetric, then u is characterized by the property

$$\frac{1}{2}\mathfrak{a}(u,u) - \varphi(u) = \min_{w \in \mathcal{K}} \left\{ \frac{1}{2}\mathfrak{a}(w,w) - \varphi(w) \right\}.$$

⁸ In fact, there is no need of this theorem as one may show that the converse holds by hand. Notice that for a minimizer u and for any $v \in \mathcal{K}(\psi)$, $e(0) = \mathcal{E}[u]$ and $e(1) = \mathcal{E}[v]$ in what precedes. The idea is to show that e is nondecreasing by similar computations. Since the minimizer is unique, this would imply that e(1) > e(0), as desired.

⁹ By *H*-elliptic we mean that there exists $\gamma > 0$ such that $\mathfrak{a}(u, u) \ge \gamma ||u||_{H}^{2}$ for every $u \in H$.

We refer to [11, Thm.5.6, p.138] for a proof. Now in order to deduce the Euler-Lagrange equations for the obstacle problem, one may choose appropriate variations v and then integrate the variational inequality by parts. This would however only hold in the sense of distributions, as we need at least $H^2(\Omega)$ regularity of the solution to define the Laplacian almost everywhere in Ω . Moreover, if u is not continuous, then the set $\{u > \psi\}$ would not be open. To finish this argument, we will need estimates guaranteeing further regularity for the solution.

1.2. Further regularity. — To show that the solution of the obstacle problem is more regular, we use a *penalization method*. The main idea is to approximate the variational inequality (1.2) by a family of semilinear elliptic equations – the semilinearity originates from a *penalty function*, in which we incorporate the constraint $u \ge \psi$. This function consists of a penalty parameter multiplied by a measure of violation of the constraint, which is nonzero when the constraint is violated and is zero in the region where constraint is not violated. One then shows existence and uniqueness of a solution for each problem, and obtains adequate uniform bounds for these solutions which will assert the desired regularity. Finally, these bounds are used to "transfer" the regularity to the solution of the variational inequality by virtue of a compactness argument. The proof presented here may be found in [30, Chapter IV].

One notable advantage of this method is that we may use known results from regularity theory for elliptic equations to obtain the mentioned bounds for the approximate solutions, as the semilinear term is in general uniformly bounded. This method may also be used to show the existence of a solution for a variational inequality, and we will notably use it for the parabolic obstacle problem.

We assume henceforth that f and $\Delta \psi$ additionally satisfy

 $f \in L^p(\Omega), \quad \max\{-\Delta \psi - f, 0\} \in L^p(\Omega) \quad \text{ for a fixed } p \in [2, \infty).$

These assumptions will allow us to prove the hinted regularity result, namely that the solution u is in $W^{2,p}(\Omega)$. Note that the weakest assumption p = 2 would imply the mentioned $H^2(\Omega)$ regularity; as a matter of fact, both of the above assumptions are superfluous in this instance.

We will use the following deep result from L^p regularity theory for elliptic equations. For the complete result and proof, we refer to [26, Thm.2.5.1.1, p.128] and [24, Chap.7, Thms.7.1,7.4]. The case p = 2 is done in [21, Thm.4, p.334].

 $\mathbf{18}$

Lemma 1.6. — Let $w \in H_0^1(\Omega)$ be a weak solution¹⁰ of the Poisson equation

$$\begin{cases} -\Delta w = f & \text{ in } \Omega \\ w = 0 & \text{ on } \partial \Omega_{f} \end{cases}$$

where $f \in L^p(\Omega)$ for some $p \in [2, \infty)$. Then

$$||w||_{W^{2,p}(\Omega)} \le C ||f||_{L^p(\Omega)}$$

for some constant $C = C(p, \Omega) > 0$.

We are now in a position to state the penalized problem. For fixed $\varepsilon > 0$, we seek for a weak solution $u^{\varepsilon} \in H_0^1(\Omega)$ of

(1.4)
$$\begin{cases} -\Delta u^{\varepsilon} = \max\{-\Delta \psi - f, 0\}\beta_{\varepsilon}(u^{\varepsilon} - \psi) + f & \text{in } \Omega\\ u^{\varepsilon} = 0 & \text{on } \partial\Omega, \end{cases}$$

where the penalty function $\beta_{\varepsilon} : \mathbb{R} \to \mathbb{R}$ is assumed uniformly Lipschitz, nonincreasing and satisfies $0 \leq \beta_{\varepsilon}(\cdot) \leq 1$ on \mathbb{R} . An example is given afterwards.

Proposition 1.2. — Let $\varepsilon > 0$ be fixed and consider β_{ε} as above. Then there exists a unique weak solution $u^{\varepsilon} \in H_0^1(\Omega)$ to the penalized problem (1.4). Moreover, $u^{\varepsilon} \in W^{2,p}(\Omega)$ with the estimate

$$||u^{\varepsilon}||_{W^{2,p}(\Omega)} \le C(||f||_{L^{p}(\Omega)} + ||\max\{-\Delta\psi - f, 0\}||_{L^{p}(\Omega)}),$$

for some $C = C(p, \Omega) > 0$.

Proof. — Fix $\varepsilon > 0$. We will use tools from monotone operator theory (see Appendix B for the main results and definitions) applied to the operator $A : H_0^1(\Omega) \to H^{-1}(\Omega)$ defined by

$$A: u \mapsto Au: \varphi \mapsto \int_{\Omega} \left(\nabla u \cdot \nabla \varphi - \max\{-\Delta \psi - f, 0\} \beta_{\varepsilon} (u - \psi) \varphi - f \varphi \right) \mathrm{d}x.$$

Observe that for $u \in H_0^1(\Omega)$, the distribution Au is in $H^{-1}(\Omega)$, due to $\beta_{\varepsilon}(u-\psi)$ being in $L^{\infty}(\Omega)$, the Hölder inequality and the assumptions on f and $\max\{-\Delta \psi - f, 0\}$.

10 Recall that this means that $w \in H_0^1(\Omega)$ satisfies the weak formulation

$$\int_{\Omega} \nabla w \cdot \nabla \varphi \mathrm{d}x = \int_{\Omega} f \varphi \mathrm{d}x$$

for all $\varphi \in H_0^1(\Omega)$.

We claim that A is in fact a strictly monotone and coercive operator. Indeed, for all $u, v \in H_0^1(\Omega)$, and assume without loss of generality that $u \ge v$, we have

(1.5)
$$-\left[\beta_{\varepsilon}(u-\psi) - \beta_{\varepsilon}(v-\psi)\right](u-v) \ge 0,$$

since β_{ε} is nonincreasing. It follows that

$$\begin{aligned} \langle Au - Av, u - v \rangle \\ &= \int_{\Omega} |\nabla(u - v)|^2 \mathrm{d}x - \int_{\Omega} \max\{-\Delta \psi - f, 0\} [\beta_{\varepsilon}(u - \psi) - \beta_{\varepsilon}(v - \psi)](u - v) \mathrm{d}x \\ &\geq \int_{\Omega} |\nabla(u - v)|^2 \mathrm{d}x, \end{aligned}$$

for all $u, v \in H_0^1(\Omega)$ since $\max\{-\Delta \psi - f, 0\} \ge 0$. Here $\langle \cdot, \cdot \rangle$ denotes the duality pairing between $H^{-1}(\Omega)$ and $H_0^1(\Omega)$. This shows that A is strictly monotone and coercive. Furthermore, $u_k \to u$ in $H_0^1(\Omega)$ implies $Au_k \rightharpoonup Au$ weakly in $H^{-1}(\Omega)$ as k goes to $+\infty$, which in turn implies that A is continuous on finite dimensional subspaces of $H_0^1(\Omega)$ (as the weak and strong convergences coincide). We may apply Corollary B.5 to obtain the existence and uniqueness of $u^{\varepsilon} \in H_0^1(\Omega)$ satisfying

(1.6)
$$\langle Au^{\varepsilon}, v - u^{\varepsilon} \rangle \ge 0$$

for all $v \in H_0^1(\Omega)$. Let $w \in H_0^1(\Omega)$ and consider $v := u^{\varepsilon} + \tau w$ for $|\tau|$ small enough. We deduce that

$$\tau \langle Au^{\varepsilon}, w \rangle > 0$$

holds for both τ positive and negative and for all $w \in H^1_0(\Omega)$. Whence

$$\langle Au^{\varepsilon}, w \rangle = 0$$

for all $w \in H_0^1(\Omega)$, and since a solution to this weak formulation is also a solution to (1.6), we deduce that u^{ε} is the desired unique solution. The estimate follows from Lemma 1.6.

We now consider, for fixed $\varepsilon > 0$, the penalty function

$$\beta_{\varepsilon} : \mathbf{x} \mapsto \begin{cases} 1 & \text{if } \mathbf{x} \leq 0\\ 1 - \frac{\mathbf{x}}{\varepsilon} & \text{if } 0 \leq \mathbf{x} \leq \varepsilon\\ 0 & \text{if } \mathbf{x} \geq \varepsilon, \end{cases}$$

and turn to the main result of this subsection.

Theorem 1.7 ($W^{2,p}$ regularity). — Let $u \in \mathcal{K}(\psi)$ be the solution to the classical obstacle problem (1.2). Then $u \in W^{2,p}(\Omega)$. Moreover, if $\frac{n}{2} , then <math>u \in C^{0,\alpha}(\overline{\Omega})^{11}$ for $0 < \alpha \leq 2 - \frac{n}{p}$, and if $n then <math>u \in C^{1,\alpha}(\overline{\Omega})$ for $0 < \alpha \leq 1 - \frac{n}{p}$.

Proof. — We proceed in applying the penalty method as explained in the beginning of the subsection.

Fix $\varepsilon > 0$, and denote by u^{ε} the solution to the penalized problem (1.4). First, we claim that $u^{\varepsilon} \in \mathcal{K}(\psi)$. To this end, consider

$$\varphi := u^{\varepsilon} - \max\{u^{\varepsilon}, \psi\} \le 0.$$

Then $\varphi \in H_0^1(\Omega)$ since $u \in H_0^1(\Omega)$ and $\psi \leq 0$ on $\partial\Omega$, and observe that showing $u^{\varepsilon} \in \mathcal{K}(\psi)$ is equivalent to showing $\varphi \equiv 0$.

The weak formulation of the penalty equation (1.4) reads

$$\int_{\Omega} \left(\nabla u^{\varepsilon} \cdot \nabla \varphi - \max\{-\Delta \psi - f, 0\} \beta_{\varepsilon} (u^{\varepsilon} - \psi) \varphi - f \varphi \right) \mathrm{d}x = 0,$$

while by Green's first identity, one also has

$$\int_{\Omega} \left(\nabla \psi \cdot \nabla \varphi + \Delta \psi \varphi \right) \mathrm{d}x = 0.$$

Subtracting these two identities yields

$$\int_{\Omega} \nabla (u^{\varepsilon} - \psi) \cdot \nabla \varphi dx = \int_{\Omega} \left(\max\{-\Delta \psi - f, 0\} \beta_{\varepsilon} (u^{\varepsilon} - \psi) + \Delta \psi + f \right) \varphi dx.$$

Now using results from distribution theory [25, Lem.7.6, p.152], one has

$$\nabla \varphi = \begin{cases} \nabla (u^{\varepsilon} - \psi) & \text{ if } \varphi < 0\\ 0 & \text{ if } \varphi = 0. \end{cases}$$

Whence,

$$\int_{\Omega} |\nabla \varphi|^2 \mathrm{d}x = \int_{\{\varphi < 0\}} \left(\max\{-\Delta \psi - f, 0\} \beta_{\varepsilon} (u^{\varepsilon} - \psi) + \Delta \psi + f \right) \varphi \mathrm{d}x.$$

Observe that in the right-hand side integral, since $\varphi < 0$ we have $u^{\varepsilon} - \psi < 0$, which in turn implies that $\beta_{\varepsilon}(u^{\varepsilon} - \psi) = 1$. Plugging this in the identity above gives

$$\int_{\Omega} |\nabla \varphi|^2 \mathrm{d}x = \int_{\{\varphi < 0\}} \left(\max\{-\Delta \psi - f, 0\} + \Delta \psi + f \right) \varphi \mathrm{d}x \le 0.$$

Hence $\varphi \equiv 0$ by the Poincaré inequality, and as a consequence $u^{\varepsilon} \in \mathcal{K}(\psi)$.

¹¹ This is to be understood modulo the choice of a continuous representative.

By the estimate in Lemma 1.2, the family $\{u^{\varepsilon}\}_{\varepsilon>0}$ is bounded in $W^{2,p}(\Omega)$. So by the Banach-Alaoglu theorem, it¹² converges weakly in $W^{2,p}(\Omega)$ to some $\tilde{u} \in W^{2,p}(\Omega)$ as $\varepsilon \to 0$. Moreover, since $u^{\varepsilon} \in \mathcal{K}(\psi)$ and $\mathcal{K}(\psi)$ being weakly closed (see Theorem A.4 in the Appendix), we have $\tilde{u} \in \mathcal{K}(\psi)$. Now recall the general Morrey-Sobolev embedding (see [4, Thm.5.4, p.98] for a proof): if $k, m \in \mathbb{N} \cup \{0\}$ are integers such that $(m-1)p < n < mp < \infty$, then

$$W^{k+m,p}(\Omega) \hookrightarrow C^{k,\alpha}(\overline{\Omega}) \quad \text{ for } 0 < \alpha \le m - \frac{n}{p}$$

Consequently, for k = 0 and m = 2, we deduce that if $p > \frac{n}{2}$ then $\tilde{u} \in C^{0,\alpha}(\overline{\Omega})$ for $\alpha \leq 2 - \frac{n}{p}$. Similarly, for k = 1 and m = 1, if p > n then $\tilde{u} \in C^{1,\alpha}(\overline{\Omega})$ for $\alpha \leq 1 - \frac{n}{p}$.

To show that \tilde{u} is a solution to the obstacle problem (1.2), we use Minty's lemma (see Lemma B.2 in Appendix B) applied to the monotone operator $-\Delta - \max\{-\Delta\psi - f, 0\}\beta_{\varepsilon}(\cdot - \psi) - f$. The monotonicity of this operator follows from the positivity of the Dirichlet Laplacian $-\Delta$ and the computation done in (1.5). Let $v \in \mathcal{K}(\psi)$ and suppose $v \geq \psi + \delta$ for some $\delta > 0$. Then by Minty's lemma, the weak form for the penalized problem (1.4) is equivalent to

$$\int_{\Omega} \nabla v \cdot \nabla (v - u^{\varepsilon}) dx - \int_{\Omega} \left(\max\{-\Delta \psi - f, 0\} \beta_{\varepsilon} (v - \psi) + f \right) (v - u^{\varepsilon}) dx \ge 0.$$

Choosing $\delta > \varepsilon$ yields $\beta_{\varepsilon}(v - \psi) = 0$. Now letting $\varepsilon \to 0$, by the previously established weak convergence we obtain

$$\int_{\Omega} \nabla v \cdot \nabla (v - \tilde{u}) \mathrm{d}x \ge \int_{\Omega} f(v - \tilde{u}) \mathrm{d}x$$

for every $v \in \mathcal{K}(\psi), v \geq \psi + \delta$. We now let $\delta \to 0$ whence it follows that

$$\int_{\Omega} \nabla v \cdot \nabla (v - \tilde{u}) \mathrm{d}x \ge \int_{\Omega} f(v - \tilde{u}) \mathrm{d}x$$

for every $v \in \mathcal{K}(\psi)$. Employing Minty's lemma once more we conclude that $\tilde{u} \equiv u$.

¹² Notice that $\{u^{\varepsilon}\}_{\varepsilon>0}$ is not really a sequence, so it is not rigorous to extract subsequences and discuss convergences in the sequential sense. To avoid topological complications, by what is written we mean that we consider a sequence $\{u^{\varepsilon_k}\}_{k=1}^{\infty} \subset \{u^{\varepsilon}\}_{\varepsilon>0}$ with $\varepsilon_k \to 0$ as $k \to \infty$; one may then extract a subsequence $\{u^{\varepsilon_{k_j}}\}_{j=1}^{\infty}$ of the sequence $\{u^{\varepsilon_k}\}_{k=1}^{\infty}$ which converges to \tilde{u} in the indicated sense. As doing this repetitively is rather tedious, we will use this abuse of notation while insinuating that it is meant in this sense.

Corollary 1.3. — Let $u \in W^{2,p}(\Omega) \cap \mathcal{K}(\psi)$ be the solution to the classical obstacle problem. Then

$$||u||_{W^{2,p}(\Omega)} \le C(||f||_{L^{p}(\Omega)} + ||\max\{-\Delta\psi - f, 0\}||_{L^{p}(\Omega)}),$$

for some $C = C(p, \Omega) > 0$.

Proof. — Since $u^{\varepsilon} \rightarrow u$ weakly in $W^{2,p}(\Omega)$ as $\varepsilon \rightarrow 0$, the estimate follows from Lemma 1.2 and the weak lower semicontinuity of the norm (see Proposition A.5 in the Appendix):

$$\|u\|_{W^{2,p}(\Omega)} \le \liminf_{\varepsilon \to 0} \|u^{\varepsilon}\|_{W^{2,p}(\Omega)}.$$

1.3. Euler-Lagrange equations. — In view of the previous results, in a standard physical configuration $n \in \{1, 2, 3\}$, p = 2, the solution u of the classical obstacle problem is continuous and in $H^2(\Omega)$. Consider the set

$$\mathcal{O} := \{ x \in \Omega \colon u(x) > \psi(x) \}.$$

Notice that since u and ψ are continuous, \mathcal{O} is open. We begin by demonstrating that u is a strong solution¹³ of the Poisson equation in \mathcal{O} :

(1.7)
$$-\Delta u = f \quad \text{a.e. in } \mathcal{O}.$$

To this end, let $w \in C_c^{\infty}(\mathbb{O})$. Then for $|\tau|$ sufficiently small, $v := u + \tau w \ge \psi$, and so $v \in \mathcal{K}(\psi)$. Thus, the variational inequality (1.2) implies

$$au \int_{\mathbb{O}} (\nabla u \cdot \nabla w - fw) \mathrm{d}x \ge 0.$$

This inequality holds for all sufficiently small τ , both positive and negative, and so we deduce

$$\int_{0} (\nabla u \cdot \nabla w - fw) \mathrm{d}x = 0$$

for all $w \in C_c^{\infty}(\mathcal{O})$. Since $u \in H^2(\Omega)$, (1.7) holds by virtue of Green's first identity.

Now if $w \in C_c^{\infty}(\Omega)$ satisfies $w \ge 0$, then the variation $v := u + w \in \mathcal{K}(\psi)$ and by plugging in (1.2) we obtain

$$\int_{\Omega} (\nabla u \cdot \nabla w - fw) \mathrm{d}x \ge 0.$$

¹³ By a strong solution we mean that the u is twice weakly differentiable and satisfies the equation a.e. in Ω . For more detail, we refer to [25, Chapter 9].



FIGURE 3. The contact set and the free boundary in the classical obstacle problem. This figure was adapted from [44].

Since $u \in H^2(\Omega)$, Green's first identity yields

$$\int_{\Omega} (-\Delta u - f) w \mathrm{d}x \ge 0$$

for all nonnegative functions $w \in C_c^{\infty}(\Omega)$. Thence

(1.8) $-\Delta u \ge f$ a.e. in Ω .

To summarize, from (1.7) and (1.8) we deduce the *Euler-Lagrange equations* for the classical obstacle problem:

$$\begin{cases} u \ge \psi & \text{a.e. in } \Omega \\ -\Delta u \ge f & \text{a.e. in } \Omega \\ -\Delta u = f & \text{a.e. in } \Omega \\ u = 0 & \text{on } \partial \Omega. \end{cases}$$

Consequently, we observe that the domain Ω is split into two regions: one in which u solves the Poisson equation, and another in which the solution equals the obstacle. The latter region is called the *contact set* or the *coincidence set*. The interface $\partial \{u > \psi\}$ which separates these two regions is called the *free boundary*.

Remark 1.8 (Complementarity problem). — Combining the properties resulting from the Euler-Lagrange equations, we obtain that the solution of the obstacle problem is a function $u \in W^{2,p}(\Omega) \cap H^1_0(\Omega)$ for any $p \in [2, +\infty)$, which satisfies

$$\begin{cases} -\Delta u - f \ge 0 & \text{a.e. in } \Omega \\ u \ge \psi & \text{a.e. in } \Omega \\ (-\Delta u - f)(u - \psi) = 0 & \text{a.e. in } \Omega \\ u = 0 & \text{on } \partial \Omega. \end{cases}$$

This is known as the *complementarity problem* and uniquely characterizes the minimizers of \mathcal{E} over $\mathcal{K}(\psi)$. The complementarity problem is often written in the form

$$\begin{cases} \min\{-\Delta u - f, u - \psi\} = 0 & \text{a.e. in } \Omega\\ u = 0 & \text{on } \partial\Omega, \end{cases}$$

a non-variational formulation which is used in regularity theory for obstacle problems with more general governing operators.

Remark 1.9 (Counterexample for C^2 regularity). — An important observation is the following. One may naturally ask if the obstacle problem always has a classical solution $u \in C^2(\Omega)$. But observe that the Euler-Lagrange formulation implies that Δu skips from -f to $\Delta \psi$ across the free boundary. Hence, Δu is in general a discontinuous function. In fact, this can also be seen from the simple counter-example:

$$\Omega = B_2(0), \quad \psi(x) = 1 - |x|^2, \quad f \equiv 0.$$

1.4. Optimal regularity. — To summarize, we have shown that the solution u to the classical obstacle problem is $W^{2,p}(\Omega)$ regular when $f, \Delta \psi \in L^p(\Omega)$ for some $p \in [2, \infty)$, and if $p \in (n, \infty)$, the solution is also $C^{1,\alpha}(\overline{\Omega})$ regular with $\alpha \leq 1 - \frac{n}{p} < 1$. Under stronger assumptions on the data (f, ψ) , we may show that the case $\alpha = 1$ also holds away from the boundary $\partial \Omega$. As previously mentioned, one cannot expect a classical solution to the obstacle problem as Δu is in general a discontinuous function. The proof presented here may be found in [41, Chapter 2] and [12].

A crucial point employed for the proof is the fact that a function $w : \Omega \to \mathbb{R}$ is in $C^{0,1}(\Omega)$ (i.e. is Lipschitz continuous) if and only if $w \in W^{1,\infty}(\Omega)$, whenever $\partial\Omega$ is of class C^1 (see [21, Thm.4, p.279]). By induction, it can be shown that for any $k \in \mathbb{N}, w \in W^{k+1,\infty}(\Omega)$ if and only if $w \in C^{k,1}(\Omega)$, whenever $\partial\Omega$ is of class C^k . It would thus suffice to show that u has bounded second derivatives to obtain the desired optimal regularity.

For great simplicity, we assume that $f \equiv 0$. To illustrate our approach, observe that if the solution u to the obstacle problem is continuous, from the Euler-Lagrange

equations it follows that

$$\begin{cases} \Delta u = 0 & \text{in } \{u > \psi\} \\ \Delta u = \Delta \psi & \text{in } \{u = \psi\} \\ u \ge \psi & \text{in } \Omega. \end{cases}$$

Now due to the linearity of $-\Delta$, we reduce the obstacle problem to the case of a zero obstacle by the change of variable

$$\mathbf{u} := u - \psi.$$

We assume that $\psi \in C^{1,1}(\Omega)$ and in light of previous results we immediately deduce that $u \in C^{1,\alpha}(\overline{\Omega}) \cap H^2(\Omega)$ with $\alpha < 1$. Now by considering the open set

$$\mathcal{O} := \{ x \in \Omega \colon \mathbf{u}(x) > 0 \},\$$

it follows that the shifted solution $\mathbf{u} \in H^2(\Omega) \cap C^0(\overline{\Omega}), \mathbf{u} \geq 0$ in Ω satisfies

(1.9)
$$\Delta \mathbf{u} = -\Delta \psi \mathbb{1}_0 \quad \text{in } \Omega$$

In some sense, this configuration allows us only work inside the set \mathcal{O} , as one has $D^2 \mathbf{u} = 0$ a.e. on \mathcal{O}^c (see [25, Lem.7.7, p.152]), so the second derivative estimates follow immediately. Elsewhere, we will use the following result (see [21, Thm.7, p.29] for a proof).

Lemma 1.10. — Let
$$x_0 \in \Omega$$
 and $r > 0$ be such that $B_{2r}(x_0) \subset \Omega$. If w satisfies
 $\Delta w = -\Delta \psi$ in $B_{2r}(x_0)$,

then

(1.10)
$$\|D^2w\|_{L^{\infty}(B_r(x_0))} \le C\left(\frac{1}{r^2}\|w\|_{L^{\infty}(B_{2r}(x_0))} + \|D^2\psi\|_{L^{\infty}(B_{2r}(x_0))}\right),$$

for some C = C(n) > 0.

An important result we will use to remove the dependence on the radius is that inside a ball around a point on the free boundary, the shifted solution \mathbf{u} grows no faster than the square of the radius of this ball. To show this, we will need the following Harnack inequality for harmonic functions (see [25, Cor.9.25, p.250] for a proof).

Lemma 1.11 (Harnack inequality). — Let $w \in H^2(\Omega)$ satisfy $\Delta w = 0$ in Ω , $w \ge 0$ in Ω . Then for any $x_0 \in \Omega$ and r > 0 such that $B_{2r}(x_0) \subset \Omega$, we have

$$\sup_{x \in B_r(x_0)} w(x) \le C \inf_{B_r(x_0)} w(x)$$

 $\mathbf{26}$

for some C = C(n) > 0.

It is important to note that in this harmonic case the constant only depends on the dimension n and not on the chosen radius r (this is not necessarily true for a more general elliptic operator).

Theorem 1.12 (Quadratic growth). — Let $x_0 \in \partial \mathcal{O} \cap \Omega$ and r > 0 be such that $B_{2r}(x_0) \subset \Omega$. Then

$$\sup_{x \in B_r(x_0)} \mathbf{u}(x) \le Cr^2 \|\Delta\psi\|_{L^{\infty}(\Omega)}$$

for some C = C(n) > 0.

Proof. — Fix $x_0 \in \partial 0 \cap \Omega$ and let r > 0 be such that $B_{2r}(x_0) \subset \Omega^{14}$. We split $\mathbf{u} = u_1 + u_2$ in $B_{2r}(x_0)$, where

$$\Delta u_1 = \Delta \mathbf{u}, \quad \Delta u_2 = 0 \quad \text{in } B_{2r}(x_0)$$
$$u_1 = 0, \qquad u_2 = \mathbf{u} \quad \text{in } \partial B_{2r}(x_0).$$

We then proceed in estimating each of the functions u_1 and u_2 separately. To make the notation easier to follow, we set $m := \|\Delta \psi\|_{L^{\infty}(\Omega)}$.

We begin with u_1 : observe that

$$(1.11) \qquad \qquad |\Delta u_1(x)| \le m$$

for all $x \in B_{2r}(x_0)$. We consider the smooth function

$$\varphi: x \mapsto \frac{1}{2n}(4r^2 - |x - x_0|^2)$$

Clearly $\varphi(x) = 0$ for every $x \in \partial B_{2r}(x_0)$, and differentiating twice leads us to

$$\begin{cases} -\Delta \varphi = 1 & \text{in } B_{2r}(x_0) \\ \varphi = 0 & \text{on } \partial B_{2r}(x_0). \end{cases}$$

Combining this with (1.11) yields

$$\Delta m\varphi(x) \le \Delta u_1(x) \le -\Delta m\varphi(x)$$

for every $x \in B_{2r}(x_0)$. Since $\varphi(x) = u_1(x) = 0$ for $x \in \partial B_{2r}(x_0)$, we may apply the comparison principle for elliptic equations (see [25, Thm.3.3, p.33]) to deduce that

$$-m\varphi(x) \le u_1(x) \le m\varphi(x)$$

¹⁴ For example, one may take $r = \frac{\operatorname{dist}(x_0, \partial \Omega)}{2}$.

holds for all $x \in B_{2r}(x_0)$. Whence by using the definition of φ , we conclude that

(1.12)
$$|u_1(x)| \le m\varphi(x) \le \frac{2}{n}r^2m$$

for all $x \in B_{2r}(x_0)$.

Now notice that since u_2 is harmonic in $B_{2r}(x_0)$ and $u_2 = \mathbf{u} \ge 0$ on $\partial B_{2r}(x_0)$, we have $u_2 \ge 0$ in $B_{2r}(x_0)$ by virtue of the weak maximum principle. Also, since x_0 lies in the free boundary $\partial \mathcal{O} \cap \Omega$, one has $\mathbf{u}(x_0) = 0$. Therefore

$$u_2(x_0) = -u_1(x_0) \le \frac{2}{n}r^2m$$

holds by (1.12). We may thence apply the Harnack inequality to the above to obtain

$$\sup_{x \in B_r(x_0)} u_2(x) \le C \inf_{x \in B_r(x_0)} u_2(x) \le C u_2(x_0) \le \frac{2C}{n} r^2 m$$

for some C = C(n) > 0. Combining the estimates for u_1 and u_2 we obtain the desired result.

Theorem 1.13 ($C^{1,1}$ regularity). — Let $\mathbf{u} \in H^2(\Omega) \cap C^0(\overline{\Omega})$, $\mathbf{u} \ge 0$ in Ω be a solution to (1.9). Then $\mathbf{u} \in C^{1,1}_{loc}(\Omega)$ and

$$\|D^2\mathbf{u}\|_{L^{\infty}(K)} \leq C\left(\|\mathbf{u}\|_{L^{\infty}(\Omega)} + \|D^2\psi\|_{L^{\infty}(\Omega)}\right),$$

for any $K \subseteq \Omega$, where $C = C(n, \operatorname{dist}(K, \partial \Omega))$.

Proof. — As discussed earlier, since $\mathbf{u} \in H^2(\Omega)$ and $D^2\mathbf{u} = 0$ a.e. on \mathcal{O}^c , it suffices to prove a uniform bound for $D^2\mathbf{u}$ in $\mathcal{O} \cap K$ for any $K \Subset \Omega$.

Fix $x_0 \in \mathcal{O} \cap K$ and set $r = \text{dist}(x_0, \mathcal{O}^c)$ and $\rho = \text{dist}(K, \partial \Omega)$. Notice that such a choice for r would imply $B_r(x_0) \subset \mathcal{O}$ and thence

$$\Delta \mathbf{u} = -\Delta \psi \quad \text{in } B_r(x_0).$$

We distinguish two different cases.

If $r \ge \rho/5$, the interior derivative estimate (1.10) yields

(1.13)
$$\|D^{2}\mathbf{u}\|_{L^{\infty}(B_{r/2}(x_{0}))} \leq C(n) \left(\frac{4}{r^{2}} \|\mathbf{u}\|_{L^{\infty}(\Omega)} + \|D^{2}\psi\|_{L^{\infty}(\Omega)}\right)$$
$$\leq C(n) \left(\frac{100}{\rho^{2}} \|\mathbf{u}\|_{L^{\infty}(\Omega)} + \|D^{2}\psi\|_{L^{\infty}(\Omega)}\right).$$

If $r < \rho/5$, applying estimate (1.10) leaves us with a quadratic dependence upon r. To cancel this factor, one seeks to apply the quadratic growth estimate from Theorem 1.12. To this end, let $y_0 \in \partial B_r(x_0) \cap \partial \mathcal{O}$ be a point on the free boundary.

From the triangle inequality, it follows that $B_{4r}(y_0) \subset B_{5r}(x_0) \Subset \Omega$, so applying Theorem 1.12 gives

$$\|\mathbf{u}\|_{L^{\infty}(B_{2r}(y_0))} \leq 4C(n)r^2 \|\Delta\psi\|_{L^{\infty}(\Omega)}.$$

Now observe that by the triangle inequality, $B_r(x_0) \subset B_{2r}(y_0)$, thus

$$\|\mathbf{u}\|_{L^{\infty}(B_r(x_0))} \leq 4C(n)r^2 \|\Delta\psi\|_{L^{\infty}(\Omega)}.$$

Applying (1.10) and the above estimate respectively leads us to

$$\begin{aligned} \|D^2 \mathbf{u}\|_{L^{\infty}(B_{r/2}(x_0))} &\leq C(n) \Big(\frac{4}{r^2} \|\mathbf{u}\|_{L^{\infty}(\Omega)} + \|D^2 \psi\|_{L^{\infty}(\Omega)}\Big) \\ &\leq C(n) \Big(\|\Delta \psi\|_{L^{\infty}(\Omega)} + \|D^2 \psi\|_{L^{\infty}(\Omega)}\Big). \end{aligned}$$

By definition, one also has

$$\|\Delta\psi\|_{L^{\infty}(\Omega)} \le n \|D^{2}\psi\|_{L^{\infty}(\Omega)}$$

Therefore

(1.14)
$$||D^{2}\mathbf{u}||_{L^{\infty}(B_{r/2}(x_{0}))} \leq C(n)||D^{2}\psi||_{L^{\infty}(\Omega)}$$

when $r \ge \rho/5$. Since $x_0 \in \mathcal{O} \cap K$ was arbitrary, and one also has the estimate on $K \cap \mathcal{O}^c$, by (1.13) and (1.14) we deduce

$$\|D^{2}\mathbf{u}\|_{L^{\infty}(K)} \leq C(n) \Big(\frac{1}{\rho^{2}} \|\mathbf{u}\|_{L^{\infty}(\Omega)} + \|D^{2}\psi\|_{L^{\infty}(\Omega)}\Big).$$

Remark 1.14. — One may in fact show that the $C^{1,1}$ estimate can be extended up to the boundary of the domain $\partial\Omega$. The proof require estimates that are beyond the scope of this work, and for more detail we refer to [41, Theorem 2.17, p.45].

2. The parabolic problem

As mentioned in the introduction, it is useful to consider an evolutionary analog to the classical obstacle problem (1.2), as this would also yield a broad scope of applications. The version of the problem we present here has been studied in [2], as well as in [23]. We consider the space-time cylinder

$$Q := \Omega \times (0, T),$$

in \mathbb{R}^{n+1} , where the terminal time $T \in (0, \infty)$ is fixed and $\Omega \subset \mathbb{R}^n$ is a bounded domain with smooth boundary $\partial \Omega$. We also consider the convex set

$$\mathfrak{K}(\psi) := \{ w \in L^2(0,T; H^1_0(\Omega)) : w_t \in L^2(\Omega), w \ge \psi \text{ a.e. in } \Omega \},\$$

where we are given an obstacle $\psi \in L^2(0, T; H^2(\Omega) \cap H_0^1(\Omega))$ satisfying $\psi_t \in L^2(\Omega)$ and $\psi(0, \cdot) = 0$ in Ω . Note that if $\psi \in L^2(0, T; H_0^1(\Omega))$ and $\psi_t \in L^2(0, T; H^{-1}(\Omega))$, then $\psi \in C^0([0, T], L^2(\Omega))$, so the initial condition $\psi(0, \cdot) = 0$ in Ω makes sense. For a proof of this fact, we refer to [**21**, Thm.3, p.303]. We also assume that we are given an initial datum $u_0 \in H_0^1(\Omega)$ and a source term $f \in L^2(\Omega)$. To simplify the notations, we denote by

$$\Sigma := \partial \Omega \times (0, T)$$

the lateral boundary of the cylinder Q.

The problem we present consists in finding a solution $u \in \mathcal{K}(\psi)$ to the *parabolic* variational inequality

(2.1)
$$\begin{cases} \int_{\Omega} u_t(v-u) dx dt + \int_{\Omega} \nabla u \cdot \nabla (v-u) dx dt \ge \int_{\Omega} f(v-u) dx dt \\ u = 0 & \text{on } \Sigma \\ u(0, \cdot) = u_0 & \text{in } \Omega, \end{cases}$$

for all $v \in \mathcal{K}(\psi)$ with $v(0, \cdot) = u_0$ in Ω . The variational inequality (2.1) is called a *parabolic obstacle problem*. By arguing similarly as for ψ , for a function $u \in \mathcal{K}(\psi)$ the initial condition $u(0, \cdot) = u_0$ in Ω also makes sense.

2.1. Existence and uniqueness of a solution. — To simplify the notation for what follows, we consider the set

$$\mathcal{V} := \left\{ w \in L^2(0, T; H^2(\Omega)) \cap L^\infty(0, T; H^1_0(\Omega)) \colon w_t \in L^2(\Omega) \right\}$$

Note that a function $u \in \mathcal{V}$ would not only satisfy $u \in C^0([0,T]; L^2(\Omega))$, but even $u \in C^0([0,T]; H^1(\Omega))$ (see [21, Thm.4, p.288]). We now state the main result of this section.

Theorem 2.1. — There exists a unique solution $u \in \mathcal{K}(\psi) \cap \mathcal{V}$ to the parabolic obstacle problem (2.1).

To prove this theorem, we use a penalization method as done for the regularity study in the classical obstacle problem. Namely, we will consider a family of semilinear parabolic equations involving a penalization term, show existence and uniqueness of solutions to these problems, as well as adequate estimates which will be used to obtain a limit by a compactness argument.

It is important to mention that since the estimates only give weak convergences, due to the form of the variational inequality we may need a compact embedding theorem which will guarantee strong convergence in order to pass to the limit. However the classical result of Rellich-Kondrachov is not applicable in the parabolic setting. This issue is rectified with the following result of J.-P. Aubin and J.-L. Lions, a proof of which may be found in [7, 46, 34, 16].

Lemma 2.2 (Aubin-Lions). — Let X_1, X and X_{-1} be Banach spaces such that $X_1 \hookrightarrow X \hookrightarrow X_{-1}$, and assume that the continuous embedding $X_1 \hookrightarrow X$ is also compact. For $1 \le p, q \le \infty$, let

$$W := \{ w \in L^p(0,T;X_1) \colon w_t \in L^q(0,T;X_{-1}) \}.$$

If $p < \infty$, then the embedding $W \hookrightarrow L^p(0,T;X)$ is compact, while if $p = \infty$ and q > 1, then the embedding $W \hookrightarrow C^0([0,T];X)$ is compact.

Now for fixed $\varepsilon > 0$, the penalized problem consists in finding a weak solution u^{ε} to

(2.2)
$$\begin{cases} u_t^{\varepsilon} - \Delta u^{\varepsilon} + \beta_{\varepsilon} (u^{\varepsilon} - \psi) = f & \text{in } \Omega, \\ u^{\varepsilon} = 0 & \text{on } \Sigma \\ u^{\varepsilon} (0, \cdot) = u_0 & \text{in } \Omega, \end{cases}$$

where the penalty function $\beta_{\varepsilon}(\cdot) := \varepsilon^{-1}\beta(\cdot)$ is satisfies the assumptions $\beta \in C^1(\mathbb{R})$, $\beta(\mathbf{x}) = 0$ if and only if $\mathbf{x} \in [0, \infty)$ and $0 \leq \beta'(\cdot) \leq 1$ on \mathbb{R} . Note that the last assumption implies that β is monotone nondecreasing, and on the other hand, by the mean-value theorem β is also Lipschitz continuous with Lipschitz constant 1.

We note that by a weak solution to the semilinear parabolic equation (2.2) we mean a function $u^{\varepsilon} \in L^2(0,T; H^1_0(\Omega))$ with $u^{\varepsilon}_t \in L^2(0,T; H^{-1}(\Omega))$ such that for a.e. $0 \le t \le T$, the weak form

$$\left\langle u_t^{\varepsilon}(t,\cdot), v \right\rangle + \int_{\Omega} \nabla u^{\varepsilon}(t,x) \cdot \nabla v(x) \mathrm{d}x + \int_{\Omega} \beta_{\varepsilon} (u^{\varepsilon} - \psi)(t,x) v(x) \mathrm{d}x = \int_{\Omega} f(t,x) v(x) \mathrm{d}x$$

holds for all $v \in H_0^1(\Omega)$, and

$$u^{\varepsilon}(0,\cdot) = u_0 \quad \text{in } \Omega.$$

Here $\langle \cdot, \cdot \rangle$ denotes the duality pairing between $H^{-1}(\Omega)$ and $H^1_0(\Omega)$. For more detail on this notion, we refer to [21, Chapter 7]. To show that the above problem admits such a solution, we will use the following well-known fixed point theorem.

Theorem 2.3 (Banach's Fixed Point). — Let X be a Banach space and let S : $X \to X$ a nonlinear map. If S is a strict contraction, meaning there exists $\gamma \in [0, 1)$ such that

 $\|Sw - Sv\|_X \le \gamma \|w - v\|_X$

for all $w, v \in X$, then S has a unique fixed point.

Proposition 2.1. — For any $\varepsilon > 0$, there exists a unique weak solution $u^{\varepsilon} \in \mathcal{V}$ of (2.2).

Proof. — Fix $\varepsilon > 0$ and assume without loss of generality that $\psi \equiv 0$. Since β_{ε} is Lipschitz continuous with constant ε^{-1} and $\beta_{\varepsilon}(0) = 0$, it satisfies

(2.3)
$$|\beta_{\varepsilon}(\mathbf{x})| \leq \frac{1}{\varepsilon} |\mathbf{x}|.$$

Our strategy will be to apply Banach's fixed point theorem in the Banach space

$$X = C^0([0,T]; L^2(\Omega))$$

endowed with the norm

$$||w||_X := \max_{t \in [0,T]} ||w(t, \cdot)||_{L^2(\Omega)}.$$

Given any function $w \in X$, set $\mathbf{g}_{\varepsilon}(t, \cdot) := \beta_{\varepsilon}(w(t, \cdot))$ for $t \in [0, T]$, and in light of the growth estimate (2.3) we see that $\mathbf{g}_{\varepsilon} \in L^2(\Omega)$. Consequently, by using classical well-posedness results for linear evolution equations (see [21, Thm.3, p.356]), the problem

(2.4)
$$\begin{cases} \mathbf{y}_t - \Delta \mathbf{y} = f - \mathbf{g}_{\varepsilon} & \text{in } \Omega \\ \mathbf{y} = 0 & \text{on } \Sigma \\ \mathbf{y}(0, \cdot) = u_0 & \text{on } \Omega \end{cases}$$

has a unique weak solution

$$\mathbf{y} \in L^2(0,T; H^1_0(\Omega)), \text{ with } \mathbf{y}_t \in L^2(0,T; H^{-1}(\Omega))$$

Hence $\mathbf{y} \in X$ by the recurring argument [21, Thm.3, p.287], and since $u_0 \in H_0^1(\Omega)$, one has further regularity, namely $\mathbf{y} \in \mathcal{V}$ (see [21, Thm.5, p.360]), and in fact $\mathbf{y} \in \mathcal{V}$ also implies $\mathbf{y} \in C^0([0, T]; H^1(\Omega))$ (see [21, Thm.4, p.288]).

We may now define $S : X \to X$ by setting $Sw = \mathbf{y}$. We claim that if the terminal time T > 0 is small enough, then S is a strict contraction. To show this, let $w, \tilde{w} \in X$ and define $\mathbf{y} = Fw$ and $\tilde{\mathbf{y}} = F\tilde{w}$. Consequently, \mathbf{y} is a weak solution of (2.4) for $\mathbf{g}_{\varepsilon} = \beta_{\varepsilon}(w)$ and $\tilde{\mathbf{y}}$ satisfies an analogous weak form for $\tilde{\mathbf{g}}_{\varepsilon} = \beta_{\varepsilon}(\tilde{w})$. Now for an arbitrary $t \in [0, T]$, subtracting the weak forms with the test function $\mathbf{y}(t, \cdot) - \tilde{\mathbf{y}}(t, \cdot) \in H_0^1(\Omega)$ gives

$$\int_{\Omega} (\mathbf{y} - \tilde{\mathbf{y}})_t (\mathbf{y} - \tilde{\mathbf{y}})(t, x) dx + \int_{\Omega} |\nabla (\mathbf{y} - \tilde{\mathbf{y}})(t, x)|^2 dx = \int_{\Omega} (\mathbf{g}_{\varepsilon} - \tilde{\mathbf{g}}_{\varepsilon}) (\mathbf{y} - \tilde{\mathbf{y}})(t, x) dx.$$

 $\mathbf{32}$

The above may be equivalently be rewritten as

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_{\Omega} |(\mathbf{y} - \tilde{\mathbf{y}})(t, x)|^2 \mathrm{d}x + 2 \int_{\Omega} |\nabla(\mathbf{y} - \tilde{\mathbf{y}})(t, x)|^2 \mathrm{d}x = 2 \int_{\Omega} (\mathbf{g}_{\varepsilon} - \tilde{\mathbf{g}}_{\varepsilon})(\mathbf{y} - \tilde{\mathbf{y}})(t, x) \mathrm{d}x.$$

We estimate the right-hand side by the Young (with α) and Poincaré inequalities respectively:

$$2\int_{\Omega} (\mathbf{g}_{\varepsilon} - \tilde{\mathbf{g}}_{\varepsilon})(\mathbf{y} - \tilde{\mathbf{y}})(t, x) \mathrm{d}x \le \alpha \|\mathbf{y}(t, \cdot) - \tilde{\mathbf{y}}(t, \cdot)\|_{L^{2}(\Omega)}^{2} + \frac{1}{\alpha} \|\mathbf{g}_{\varepsilon}(t, \cdot) - \tilde{\mathbf{g}}_{\varepsilon}(t, \cdot)\|_{L^{2}(\Omega)}^{2} \\ \le \alpha C(\Omega, n) \|\mathbf{y}(t, \cdot) - \tilde{\mathbf{y}}(t, \cdot)\|_{H^{1}_{0}(\Omega)}^{2} + \frac{1}{\alpha} \|\mathbf{g}_{\varepsilon}(t, \cdot) - \tilde{\mathbf{g}}_{\varepsilon}(t, \cdot)\|_{L^{2}(\Omega)}^{2},$$

for $\alpha > 0$. Whence

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} \|\mathbf{y}(t,\cdot) - \tilde{\mathbf{y}}(t,\cdot)\|_{L^{2}(\Omega)}^{2} + \left(2 - \alpha C(\Omega,n)\right) \|\mathbf{y}(t,\cdot) - \tilde{\mathbf{y}}(t,\cdot)\|_{H^{1}_{0}(\Omega)}^{2} \\ &\leq \frac{1}{\alpha} \|\mathbf{g}_{\varepsilon}(t,\cdot) - \tilde{\mathbf{g}}_{\varepsilon}(t,\cdot)\|_{L^{2}(\Omega)}^{2}. \end{aligned}$$

Choosing $\alpha \leq \frac{2}{C(\Omega,n)}$ yields

$$\frac{\mathrm{d}}{\mathrm{d}t} \|\mathbf{y}(t,\cdot) - \tilde{\mathbf{y}}(t,\cdot)\|_{L^{2}(\Omega)}^{2} \leq \frac{1}{\alpha} \|\mathbf{g}_{\varepsilon}(t,\cdot) - \tilde{\mathbf{g}}_{\varepsilon}(t,\cdot)\|_{L^{2}(\Omega)}^{2},$$

and since $\mathbf{g}_{\varepsilon} = \beta_{\varepsilon}(u)$ and β_{ε} being Lipschitz continuous (with constant ε^{-1}), we also have

$$\frac{\mathrm{d}}{\mathrm{d}t} \|\mathbf{y}(t,\cdot) - \tilde{\mathbf{y}}(t,\cdot)\|_{L^2(\Omega)}^2 \leq \frac{1}{\alpha\varepsilon^2} \|w(t,\cdot) - \tilde{w}(t,\cdot)\|_{L^2(\Omega)}^2.$$

Consequently,

(2.5)
$$\|\mathbf{y}(\tau,\cdot) - \tilde{\mathbf{y}}(\tau,\cdot)\|_{L^{2}(\Omega)}^{2} \leq \frac{1}{\alpha\varepsilon^{2}} \int_{0}^{\tau} \|w(t,\cdot) - \tilde{w}(t,\cdot)\|_{L^{2}(\Omega)}^{2} \mathrm{d}t$$
$$\leq \frac{T}{\alpha\varepsilon^{2}} \|w - \tilde{w}\|_{X}^{2},$$

for each $\tau \in [0, T]$. Maximizing the left-hand side with respect to τ gives

$$\|Sw - S\tilde{w}\|_X^2 = \|\mathbf{y} - \tilde{\mathbf{y}}\|_X^2 \le \frac{T}{\alpha\varepsilon^2} \|w - \tilde{w}\|_X^2.$$

Choosing T sufficiently small so that $T < \alpha \varepsilon^2$ would imply that S is a strict contraction.

Now given any T > 0, in light of what precedes we select $T_1 > 0$ so that $T_1 < \alpha \varepsilon^2$ and then apply Banach's fixed point theorem to find a weak solution $u^{\varepsilon} = \mathbf{y} \in \mathcal{V}$ of the penalized problem (2.2) on the time interval $[0, T_1]$. Since $u^{\varepsilon}(t, \cdot) \in H^1_0(\Omega)$ for

a.e. $t \in [0, T_1]$, we can assume (we may redefine T_1 if needed) that $u^{\varepsilon}(T_1, \cdot) \in H_0^1(\Omega)$. Observe that the time T_1 depends only on ε . We therefore repeat the argument above to extend this solution to the time interval $[T_1, 2T_1]$. After finitely many steps we construct a final weak solution $u^{\varepsilon} \in \mathcal{V}$ existing on the full interval [0, T].

To demonstrate uniqueness, fix $\varepsilon > 0$ and let u^{ε} , \tilde{u}^{ε} be two weak solutions of the penalty problem. Then we have $\mathbf{y} = u^{\varepsilon}$, $\tilde{\mathbf{y}} = \tilde{u}^{\varepsilon}$ in (2.5), meaning

$$\|u^{\varepsilon}(\tau,\cdot) - \tilde{u}^{\varepsilon}(\tau,\cdot)\|_{L^{2}(\Omega)}^{2} \leq C(\varepsilon) \int_{0}^{\tau} \|u^{\varepsilon}(t,\cdot) - \tilde{u}^{\varepsilon}(t,\cdot)\|_{L^{2}(\Omega)}^{2} \mathrm{d}t$$

for $\tau \in [0, T]$. By virtue of the differential form of Gronwall's inequality (see Proposition A.11 in the Appendix), we conclude that $u^{\varepsilon} \equiv \tilde{u}^{\varepsilon}$.

We will need some uniform estimates in order to apply a compactness argument and pass to the limit.

Lemma 2.4. — For fixed $\varepsilon > 0$, the solution $u^{\varepsilon} \in \mathcal{V}$ to the penalized problem (2.2) satisfies

$$\begin{aligned} \underset{t \in [0,T]}{\operatorname{ess\,sup}} &\| u^{\varepsilon}(t, \cdot) \|_{H_{0}^{1}(\Omega)} + \| u^{\varepsilon} \|_{L^{2}(0,T;H^{2}(\Omega))} + \| u^{\varepsilon}_{t} \|_{L^{2}(\Omega)} \\ &\leq C \big(\| \Delta \psi \|_{L^{2}(\Omega)} + \| \psi_{t} \|_{L^{2}(\Omega)} + \| f \|_{L^{2}(\Omega)} + \| \nabla u_{0} \|_{L^{2}(\Omega)} \big), \end{aligned}$$

for some $C = C(\Omega, T) > 0$ independent of ε .

Proof. — Fix $\varepsilon > 0$. By virtue of the energy estimate [21, Thm.5, p.360]

$$\underset{t \in [0,T]}{\operatorname{ess\,sup}} \| u^{\varepsilon}(t,\cdot) \|_{H_{0}^{1}(\Omega)} + \| u^{\varepsilon} \|_{L^{2}(0,T;H^{2}(\Omega))} + \| u^{\varepsilon}_{t} \|_{L^{2}(\Omega)}$$

$$(2.6) \qquad \leq C(\Omega,T) \big(\| \beta_{\varepsilon} (u^{\varepsilon} - \psi) \|_{L^{2}(\Omega)} + \| f \|_{L^{2}(\Omega)} + \| \nabla u_{0} \|_{L^{2}(\Omega)} \big),$$

we observe that it suffices to obtain an adequate bound for the $L^2(\Omega)$ norm of β_{ε} in order to conclude. For fixed $t \in [0, T]$, using the penalty equation we obtain

$$\int_{\Omega} (\beta_{\varepsilon} (u^{\varepsilon} - \psi)(t, x))^{2} dx = \int_{\Omega} \beta_{\varepsilon} (u^{\varepsilon} - \psi) [f - u_{t}^{\varepsilon} + \Delta u^{\varepsilon}](t, x) dx$$
$$= \int_{\Omega} \beta_{\varepsilon} (u^{\varepsilon} - \psi) [\Delta (u^{\varepsilon} - \psi) + \Delta \psi](t, x) dx$$
$$- \int_{\Omega} \beta_{\varepsilon} (u^{\varepsilon} - \psi) (u^{\varepsilon} - \psi)_{t}(t, x) dx$$
$$+ \int_{\Omega} \beta_{\varepsilon} (u^{\varepsilon} - \psi) [f - \psi_{t}](t, x) dx.$$

Now, Green's first identity yields

$$\int_{\Omega} \beta_{\varepsilon} (u^{\varepsilon} - \psi) \Delta (u^{\varepsilon} - \psi)(t, x) \mathrm{d}x = -\int_{\Omega} \beta_{\varepsilon}' (u^{\varepsilon} - \psi) |\nabla (u^{\varepsilon} - \psi)|^2 (t, x) \mathrm{d}x \le 0,$$

since $\beta'_{\varepsilon} \geq 0$. By combining the previous computations we obtain

$$\int_{\Omega} (\beta_{\varepsilon} (u^{\varepsilon} - \psi)(t, x))^{2} \mathrm{d}x \leq \int_{\Omega} \beta_{\varepsilon} (u^{\varepsilon} - \psi) \Delta \psi(t, x) \mathrm{d}x - \frac{\mathrm{d}}{\mathrm{d}t} \int_{\Omega} \rho_{\varepsilon} (u^{\varepsilon} - \psi)(t, x) \mathrm{d}x + \int_{\Omega} \beta_{\varepsilon} (u^{\varepsilon} - \psi) [f - \psi_{t}](t, x) \mathrm{d}x,$$

where $\rho_{\varepsilon}' = \beta_{\varepsilon}$, $\rho_{\varepsilon} \ge 0$ on \mathbb{R} and $\rho_{\varepsilon} = 0$ on $[0, +\infty)$. Now we integrate the above inequality over [0, T]; first, observe that

$$-\int_0^T \frac{\mathrm{d}}{\mathrm{d}t} \int_\Omega \rho_\varepsilon (u^\varepsilon - \psi)(t, x) \mathrm{d}x \mathrm{d}t = -\int_\Omega \rho_\varepsilon (u^\varepsilon - \psi)(T, x) \mathrm{d}x \mathrm{d}t + \int_\Omega \rho_\varepsilon (u^\varepsilon - \psi)(0, x) \mathrm{d}x.$$

The second integral on the right hand side above is equal to 0, since $u^{\varepsilon}(0, \cdot) \geq 0$, $\psi(0, \cdot) = 0$ in Ω and $\rho_{\varepsilon} = 0$ on $[0, \infty)$. Moreover, since $\rho_{\varepsilon} \geq 0$ on \mathbb{R} and thus $\rho_{\varepsilon} \circ (u^{\varepsilon} - \psi)(T, \cdot) \geq 0$, we deduce that

$$-\int_0^T \frac{\mathrm{d}}{\mathrm{d}t} \int_{\Omega} \mathbf{\rho}_{\varepsilon} (u^{\varepsilon} - \psi)(t, x) \mathrm{d}x \mathrm{d}t \le 0.$$

Consequently,

$$\int_{\Omega} (\beta_{\varepsilon}(u^{\varepsilon} - \psi))^2 \mathrm{d}x \mathrm{d}t \leq \int_{\Omega} \beta_{\varepsilon}(u^{\varepsilon} - \psi) \Delta \psi \mathrm{d}x \mathrm{d}t + \int_{\Omega} \beta_{\varepsilon}(u^{\varepsilon} - \psi)[f - \psi_t] \mathrm{d}x \mathrm{d}t$$
$$\leq \|\beta_{\varepsilon}(u^{\varepsilon} - \psi)\|_{L^2(\Omega)} (\|\Delta \psi\|_{L^2(\Omega)} + \|f\|_{L^2(\Omega)} + \|\psi_t\|_{L^2(\Omega)}).$$

After simplification, we deduce that

(2.7)
$$\|\beta_{\varepsilon}(u^{\varepsilon} - \psi)\|_{L^{2}(\Omega)} \leq \|\Delta\psi\|_{L^{2}(\Omega)} + \|f\|_{L^{2}(\Omega)} + \|\psi_{t}\|_{L^{2}(\Omega)},$$

and by plugging this in (2.6) we obtain the desired estimate.

Remark 2.5. — The previous Lemma clearly asserts that the family of approximate solutions $\{u^{\varepsilon}\}_{\varepsilon>0}$ is bounded in $L^2(0,T;H^1_0(\Omega))$, and hence in $L^2(\Omega)$. To see this, fix $\varepsilon > 0$ and observe that

$$\|u^{\varepsilon}\|_{L^{2}(0,T;H^{1}_{0}(\Omega))}^{2} \leq T \operatorname{ess\,sup}_{t \in [0,T]} \|u^{\varepsilon}(t,\cdot)\|_{H^{1}_{0}(\Omega)}^{2}.$$

We deduce that

 $\|u^{\varepsilon}\|_{L^{2}(0,T;H^{1}_{0}(\Omega))} \leq C(\Omega, n, T) \Big(\underset{t \in [0,T]}{\operatorname{ess sup}} \|u^{\varepsilon}(t, \cdot)\|_{H^{1}_{0}(\Omega)} + \|u^{\varepsilon}\|_{L^{2}(0,T;H^{2}(\Omega))} + \|u^{\varepsilon}_{t}\|_{L^{2}(\Omega)} \Big).$

This estimate also implies the boundedness of $\{u^{\varepsilon}\}_{\varepsilon>0}$ in $L^2(\Omega)$ by the Poincaré inequality.

We now let $\varepsilon \to 0$ to prove Theorem 2.1.

Proof of Theorem 2.1. — Denote by $\{u^{\varepsilon}\}_{\varepsilon>0} \subset \mathcal{V}$ the family of solutions to the penalized problem (2.2). By virtue of the energy estimate from Lemma 2.4, we deduce that $\{u^{\varepsilon}\}_{\varepsilon>0}$ and $\{u^{\varepsilon}_{t}\}_{\varepsilon>0}$ are bounded in $L^{2}(0,T;H^{2}(\Omega))$ and $L^{2}(\Omega)$ respectively, while Remark 2.5 implies that $\{u^{\varepsilon}\}_{\varepsilon>0}$ is also bounded in $L^{2}(0,T;H^{1}_{0}(\Omega))$. Whence by the Banach-Alaoglu theorem and Proposition A.7 in the Appendix, there exists $u \in L^{2}(0,T;H^{2}(\Omega) \cap H^{1}_{0}(\Omega))$ with $u_{t} \in L^{2}(\Omega)$ such that

$$\begin{array}{ll} u_t^{\varepsilon} \rightharpoonup u_t & \text{weakly in } L^2(\mathbb{Q}) \\ u^{\varepsilon} \rightharpoonup u & \text{weakly in } L^2(0,T;H^2(\Omega)) \\ u^{\varepsilon} \rightharpoonup u & \text{weakly in } L^2(0,T;H_0^1(\Omega)) \end{array}$$

along subsequences as $\varepsilon \to 0$. Applying the Aubin-Lions lemma with p = q = 2 and the triple $X_1 = H_0^1(\Omega), X = L^2(\Omega), X_{-1} = H^{-1}(\Omega)$, we deduce that

$$u^{\varepsilon} \to u$$
 strongly in $L^2(\mathbb{Q})$

along a subsequence as $\varepsilon \to 0$. Now observe that for almost every $t \in [0, T]$, one has

$$\|\nabla (u^{\varepsilon} - u)(t, \cdot)\|_{L^{2}(\Omega)}^{2} \leq \|\Delta (u^{\varepsilon} - u)(t, \cdot)\|_{L^{2}(\Omega)}\|(u^{\varepsilon} - u)(t, \cdot)\|_{L^{2}(\Omega)}$$

by using Green's first identity and the Cauchy-Schwarz inequality. Integrating between 0 and T gives

$$\|\nabla (u^{\varepsilon} - u)\|_{L^{2}(\Omega)}^{2} \leq \|\Delta (u^{\varepsilon} - u)\|_{L^{2}(\Omega)}\|u^{\varepsilon} - u\|_{L^{2}(\Omega)},$$

again by using the Cauchy-Schwarz inequality. The $L^2(0, T; H^2(\Omega))$ boundedness of $\{u^{\varepsilon}\}_{\varepsilon>0}$ implies that $\{\Delta u^{\varepsilon}\}_{\varepsilon>0}$ is bounded in $L^2(\Omega)$ and since u^{ε} converges strongly to u in $L^2(\Omega)$, we finally deduce that $\nabla u^{\varepsilon} \to \nabla u$ strongly in $L^2(\Omega)$, i.e.

$$u^{\varepsilon} \to u$$
 strongly in $L^2(0,T; H^1_0(\Omega))$.

Now for fixed $\varepsilon > 0$, since $\beta \leq 0$ (as it is nondecreasing and equal to 0 on $[0, \infty)$), u^{ε} satisfies

$$\int_{\Omega} u_t^{\varepsilon} (v - u^{\varepsilon}) \mathrm{d}x \mathrm{d}t + \int_{\Omega} \nabla u^{\varepsilon} \cdot \nabla (v - u^{\varepsilon}) \mathrm{d}x \mathrm{d}t \ge \int_{\Omega} f(v - u^{\varepsilon}) \mathrm{d}x \mathrm{d}t,$$

36
for every $v \in \mathcal{K}(\psi)$. Observe that

$$\int_{\Omega} u_t^{\varepsilon} (v - u^{\varepsilon}) \mathrm{d}x \mathrm{d}t = \int_{\Omega} u_t^{\varepsilon} (v - u) \mathrm{d}x \mathrm{d}t - \int_{\Omega} u_t^{\varepsilon} (u^{\varepsilon} - u) \mathrm{d}x \mathrm{d}t,$$

and similarly

$$\int_{\Omega} \nabla u^{\varepsilon} \cdot \nabla (v - u^{\varepsilon}) \mathrm{d}x \mathrm{d}t = \int_{\Omega} \nabla u^{\varepsilon} \cdot \nabla (v - u) \mathrm{d}x \mathrm{d}t - \int_{\Omega} \nabla u^{\varepsilon} \cdot \nabla (u^{\varepsilon} - u) \mathrm{d}x \mathrm{d}t.$$

Hence, using the strong convergences of $\{u^{\varepsilon}\}_{\varepsilon>0}, \{\nabla u^{\varepsilon}\}_{\varepsilon>0}$, the weak convergence of $\{u^{\varepsilon}_t\}_{\varepsilon>0}$ and the boundedness of the latter two in $L^2(\Omega)$, we may let $\varepsilon \to 0$ to deduce

$$\int_{\Omega} u_t(v-u) \mathrm{d}x \mathrm{d}t + \int_{\Omega} \nabla u \cdot \nabla (v-u) \mathrm{d}x \mathrm{d}t \ge \int_{\Omega} f(v-u) \mathrm{d}x \mathrm{d}t,$$

for all $v \in \mathcal{K}(\psi)$. Using estimate (2.7) and since $\beta_{\varepsilon} = \varepsilon^{-1}\beta$, one obtains

(2.8)
$$\|\beta(u^{\varepsilon} - \psi)\|_{L^{2}(\Omega)} \leq \varepsilon C(\psi, f).$$

Because β is continuous and the strong $L^2(\Omega)$ convergence of $\{u^{\varepsilon}\}_{\varepsilon>0}$ implies a.e. convergence along some subsequence, letting $\varepsilon \to 0$ we deduce

$$\|\beta(u-\psi)\|_{L^{2}(\Omega)} = 0$$

by continuity of the norm. Recall that $\beta = 0$ only on $[0, \infty)$, thus it follows that $u \geq \psi$ a.e. in Ω and we may conclude that $u \in \mathcal{K}(\psi)$ solves the parabolic obstacle problem (2.1). Finally, since $\{u^{\varepsilon}\}_{\varepsilon>0}$ is bounded in $L^{\infty}(0, T; H_0^1(\Omega))$ and converges weakly in $L^2(0, T; H_0^1(\Omega))$ to u, we may apply Proposition A.8 from the Appendix to obtain $u \in \mathcal{V}$.

To demonstrate uniqueness, let u, \tilde{u} be two solutions. Whence, they satisfy

$$\int_{\Omega} u_t(v-u) \mathrm{d}x \mathrm{d}t + \int_{\Omega} \nabla u \cdot \nabla (v-u) \mathrm{d}x \mathrm{d}t \ge \int_{\Omega} f(v-u) \mathrm{d}x \mathrm{d}t$$

and

$$\int_{\mathbb{Q}} \tilde{u}_t(v - \tilde{u}) \mathrm{d}x \mathrm{d}t + \int_{\mathbb{Q}} \nabla \tilde{u} \cdot \nabla (v - \tilde{u}) \mathrm{d}x \mathrm{d}t \ge \int_{\mathbb{Q}} f(v - \tilde{u}) \mathrm{d}x \mathrm{d}t$$

respectively for all $v \in \mathcal{K}(\psi)$. Choosing $v = \tilde{u}$ and v = u respectively and adding up both inequalities gives

$$\int_{\Omega} (u - \tilde{u})_t (\tilde{u} - u) \mathrm{d}x \mathrm{d}t + \int_{\Omega} \nabla (u - \tilde{u}) \cdot \nabla (\tilde{u} - u) \mathrm{d}x \mathrm{d}t \ge 0.$$

By rearranging the above inequality, we deduce that

$$\frac{1}{2} \int_{\Omega} |(u - \tilde{u})_t|^2 \mathrm{d}x \mathrm{d}t + \int_{\Omega} |\nabla(u - \tilde{u})|^2 \mathrm{d}x \mathrm{d}t \le 0,$$

and by virtue of the Poincaré inequality, we conclude that $u = \tilde{u}$ a.e. in Q.

Corollary 2.2. — Let $u \in \mathcal{V} \cap \mathcal{K}(\psi)$ be the solution to the parabolic obstacle problem (2.1). Then

$$\begin{aligned} \underset{t \in [0,T]}{\operatorname{ess sup}} &\|u(t, \cdot)\|_{H_0^1(\Omega)} + \|u\|_{L^2(0,T;H^2(\Omega))} + \|u_t\|_{L^2(\Omega)} \\ &\leq C \big(\|\Delta \psi\|_{L^2(\Omega)} + \|\psi_t\|_{L^2(\Omega)} + \|f\|_{L^2(\Omega)} + \|\nabla u_0\|_{L^2(\Omega)} \big), \end{aligned}$$

holds for some $C = C(\Omega, T) > 0$. Moreover, $u \in C^0([0, T]; H^1(\Omega))$.

Proof. — Due to the weak convergences established in the previous proof, the estimate follows from Lemma 2.4 and the weak lower semicontinuity of the norms and Proposition A.8 from the Appendix. The fact that $u \in C^0([0,T]; H^1(\Omega))$ follows from [21, Thm.4, p.288].

2.2. Euler-Lagrange equations. — Assuming the existence of a continuous (space-time) solution¹⁵ u, and assuming moreover $\psi \in C^0(\Omega)$, by virtue of a variational argument we may also deduce Euler-Lagrange equations as for the classical obstacle problem. For $v \in C_c^{\infty}(\Omega)$, where

$$0 := \{ (t, x) \in Q \colon u(t, x) > \psi(t, x) \},\$$

and $|\tau|$ is sufficiently small, $v := u + \tau w \ge \psi$ and so $v \in \mathcal{K}(\psi)$. Thus, plugging v in the variational inequality gives

$$\tau \int_{\mathbb{O}} (u_t w + \nabla u \cdot \nabla w - f w) \mathrm{d}x \mathrm{d}t \ge 0.$$

Due to Green's first identity and since the above holds for τ both positive and negative, we may conclude that

(2.9)
$$u_t - \Delta u = f$$
 a.e. in \mathcal{O} .

¹⁵ We will not further discuss regularity results for the solution of the parabolic obstacle problem, since the setup for the problem is not the same in literature. In the "dynamic obstacle" case that we considered in this work, regularity is discussed for example in [40] and in [6], where the problem is referred to as the "thick" obstacle problem (compared to the "thin" problem, which we discuss on page 88). We also refer to [13, 14] for the regularity study of related problems.

Now, by setting v := u + w where $v \in C_c^{\infty}(\Omega)$, $v \ge 0$, and plugging in the variational inequality, one obtains

$$\int_{\Omega} (u_t w + \nabla u \cdot \nabla w - f w) \mathrm{d}x \mathrm{d}t \ge 0.$$

Whence, we deduce

(2.10)
$$u_t - \Delta u \ge f$$
 a.e. in Q

To summarize, from (2.9) and (2.10), we deduce the *Euler-Lagrange equations* for the parabolic obstacle problem

$$\begin{cases} u \ge \psi & \text{a.e. in } \Omega \\ u_t - \Delta u \ge f & \text{a.e. in } \Omega \\ u_t - \Delta u = f & \text{a.e. in } \Omega \\ u = 0 & \text{on } \Sigma \\ u(0, \cdot) = u_0 & \text{in } \Omega. \end{cases}$$

As for the classical obstacle problem, we may derive the *complementarity problem*

$$\begin{cases} u_t - \Delta u - f \ge 0 & \text{a.e. in } \mathcal{Q} \\ u \ge \psi & \text{a.e. in } \mathcal{Q} \\ (u_t - \Delta u - f)(u - \psi) = 0 & \text{a.e. in } \mathcal{Q} \\ u = 0 & \text{on } \Sigma \\ u(0, \cdot) = u_0 & \text{in } \Omega, \end{cases}$$

which in the literature is often written in the form

$$\begin{cases} \min\{u_t - \Delta u - f, u - \psi\} = 0 & \text{a.e. in } \Omega\\ u = 0 & \text{on } \Sigma\\ u(0, \cdot) = u_0 & \text{in } \Omega. \end{cases}$$

2.3. Euler's method for a homogeneous problem. — We now illustrate in a mostly formal way a method for studying a slightly different parabolic obstacle problem, namely one where the obstacle does not vary with time. The key idea in this method is to use the implicit Euler scheme for a finite difference approximation of the time derivative, reducing the problem to a steady-state variational inequality at each time step. The latter in turn will be shown to posses a unique solution by applying monotone operator theory. One then constructs an adequate approximation by

"gluing" the solutions at the discrete time steps, lets the mesh size go to 0 and deduces a limit by a compactness argument.

We will work in a simple case, as we assume that $f \equiv 0$ and that the obstacle does not depend on time. While such assumptions may seem restrictive, we will nonetheless observe the constructive nature of the method which renders it useful in numerical analysis and computation.

The obstacle problem we are will study consists in finding $u \in L^2(0,T; H^1_0(\Omega))$ with $u_t \in L^2(\Omega)$ such that for a.e. $t \in (0,T), u(t, \cdot) \in \mathcal{K}(\psi)$ and

(2.11)
$$\begin{cases} \int_{\Omega} u_t(v-u) dx + \int_{\Omega} \nabla u \cdot \nabla (v-u) dx \ge 0\\ u(0,\cdot) = u_0 & \text{in } \Omega, \end{cases}$$

for every $v \in \mathcal{K}(\psi)$, where

$$\mathcal{K}(\psi) := \{ v \in H^1_0(\Omega) \colon v \ge \psi \text{ a.e. in } \Omega \}$$

is the closed and convex set that appears in the classical obstacle problem. The obstacle ψ is assumed in $H^2(\Omega)$, and we will require additional regularity on the initial datum, namely we assume $u_0 \in H^2(\Omega) \cap \mathcal{K}(\psi)$. Note that if there is a solution u as above, we immediately deduce that $u \in C^0([0, T]; L^2(\Omega))$ by the recurring argument [21, Thm.3, p.303], so the initial condition $u(0, \cdot) = u_0$ in Ω makes sense.

We begin by stating the main result of this subsection.

Proposition 2.3. — There exists a unique solution to the parabolic obstacle problem (2.11).

We begin by partitioning the time interval [0, T] into m equal sub-intervals $[t_{i-1}, t_i]$ where $m \in \mathbb{N}$ is fixed, $i \in \{1, \ldots, m\}$, $t_i = ih$ and $h = \frac{T}{m}$ is the mesh size. We now consider the semidiscretized problem consisting of finding $u_i \in \mathcal{K}(\psi)$ such that

(2.12)
$$\int_{\Omega} \left(\frac{u_i - u_{i-1}}{h} \right) (v - u_i) \mathrm{d}x + \int_{\Omega} \nabla u_i \cdot \nabla (v - u_i) \mathrm{d}x \ge 0$$

for every $v \in \mathcal{K}(\psi)$ and for each $i \in \{1, \ldots, m\}$. When i = 1, we set $u_{i-1} := u_0$ where u_0 is the initial datum. We may rewrite the variational inequality (2.12) as

$$\frac{1}{h} \int_{\Omega} u_i(v - u_i) \mathrm{d}x + \int_{\Omega} \nabla u_i \cdot \nabla (v - u_i) \mathrm{d}x \ge \frac{1}{h} \int_{\Omega} u_{i-1}(v - u_i) \mathrm{d}x,$$

and since u_{i-1} is known at each time step $i \in \{1, \ldots, m\}$, we see that this corresponds to the variational inequality

$$\langle Au_i, v - u_i \rangle \ge \langle \frac{u_{i-1}}{h}, v - u_i \rangle_{L^2(\Omega)}$$

where the operator $A: \mathcal{K}(\psi) \to H^{-1}(\Omega)$ is given by

$$A: u \mapsto Au: v \mapsto \frac{1}{h} \int_{\Omega} uv dx + \int_{\Omega} \nabla u \cdot \nabla v dx.$$

Here $\langle \cdot, \cdot \rangle$ denotes the duality pairing between $H^{-1}(\Omega)$ and $H_0^1(\Omega)$. We see that A linear, thus continuous on finite dimensional subspaces, and also coercive and strictly monotone due to its inner product form. We may thence apply Theorem B.1 and deduce the existence of a unique solution $u_i \in \mathcal{K}(\psi)$ to (2.12) for each $i \in \{1, \ldots, m\}$.

We have thus obtained a family of solutions $\{u_i\}_{i=1}^m$ which may be used to form the proposed approximate solution $\mathbf{u}_m : [0,T] \to H_0^1(\Omega)$ as

$$\mathbf{u}_m(t,\cdot) := u_{i-1} + (t - t_{i-1}) \frac{u_i - u_{i-1}}{h}, \quad \text{for } t \in [t_{i-1}, t_i], \ i \in \{1, \dots, m\}.$$

Observe that $\mathbf{u}_m \in C^0([0,T]; H_0^1(\Omega))$ is affine in time over each subinterval $[t_{i-1}, t_i]$ and one also has $\mathbf{u}_m(t_{i-1}, \cdot) = u_{i-1}$ and $\mathbf{u}_m(t_i, \cdot) = u_i$, so the time variable plays the role of a homotopy parameter in some sense, connecting u_{i-1} at time t_{i-1} to u_i at time t_i . We also consider the step function $\overline{\mathbf{u}}_m : [0,T] \to H_0^1(\Omega)$ defined as

$$\overline{\mathbf{u}}_m(t,\cdot) := u_i \quad \text{for } t \in [t_{i-1}, t_i], \ i \in \{1, \dots, m\}.$$

We will use these functions to reformulate the semidiscretized variational inequality (2.12).

To show that \mathbf{u}_m converges to a solution u of the parabolic obstacle problem (2.11) as $m \to \infty$, we first establish some uniform estimates in order to pass to the limit.

Lemma 2.6. — There exist constants $C_0, C_1 > 0$ depending only on T, u_0 such that

,

(2.13)
$$\left\|\frac{u_i - u_{i-1}}{h}\right\|_{L^2(\Omega)} \le C_0$$

$$(2.14) \|\nabla u_i\|_{L^2(\Omega)} \le C_1$$

for each $m \in \mathbb{N}$ and $i \in \{1, \ldots, m\}$.

Proof. — First, by setting i = 0 and $v = u_0$ in (2.12), we obtain

$$\frac{1}{h} \int_{\Omega} (u_1 - u_0)(u_0 - u_1) \mathrm{d}x + \int_{\Omega} \nabla u_1 \cdot \nabla (u_0 - u_1) \mathrm{d}x \ge 0.$$

Slightly rearranging this inequality leads us to

$$\int_{\Omega} \nabla u_0 \cdot \nabla (u_0 - u_1) dx \ge \|\nabla (u_1 - u_0)\|_{L^2(\Omega)}^2 + \frac{1}{h} \|u_1 - u_0\|_{L^2(\Omega)}^2$$
$$\ge \frac{1}{h} \|u_1 - u_0\|_{L^2(\Omega)}^2.$$

Since $u_0 \in H^2(\Omega) \cap H^1_0(\Omega)$, we may use Green's first identity and the Cauchy-Schwarz inequality to estimate the left-hand side:

$$\int_{\Omega} \nabla u_0 \cdot \nabla (u_0 - u_1) \mathrm{d}x \le \|\Delta u_0\|_{L^2(\Omega)} \|u_1 - u_0\|_{L^2(\Omega)}$$

Combining the two estimates yields

(2.15)
$$\left\|\frac{u_1 - u_0}{h}\right\|_{L^2(\Omega)} \le \|\Delta u_0\|_{L^2(\Omega)}.$$

Now fix $j \ge 2$. The variational inequality (2.12) for i = j and $v = u_{j-1}$ reads

$$\int_{\Omega} \left(\frac{u_j - u_{j-1}}{h} \right) (u_{j-1} - u_j) \mathrm{d}x + \int_{\Omega} \nabla u_j \cdot \nabla (u_{j-1} - u_j) \mathrm{d}x \ge 0$$

rly for $i = i - 1$ and $v = u$

and similarly for i = j - 1 and $v = u_j$

$$\int_{\Omega} \left(\frac{u_{j-1} - u_{j-2}}{h} \right) (u_j - u_{j-1}) \mathrm{d}x + \int_{\Omega} \nabla u_{j-1} \cdot \nabla (u_j - u_{j-1}) \mathrm{d}x \ge 0.$$

Adding up these inequalities and switching the signs gives

$$\frac{1}{h} \int_{\Omega} (u_{j-1} - u_{j-2}) (u_j - u_{j-1}) dx \ge \|\nabla (u_j - u_{j-1})\|_{L^2(\Omega)}^2 + \frac{1}{h} \|u_j - u_{j-1}\|_{L^2(\Omega)}^2$$
$$\ge \frac{1}{h} \|u_j - u_{j-1}\|_{L^2(\Omega)}^2.$$

We may estimate the left-hand side using the Cauchy-Schwarz inequality:

$$\int_{\Omega} \left(\frac{u_{j-1} - u_{j-2}}{h} \right) (u_j - u_{j-1}) \mathrm{d}x \le \left\| \frac{u_{j-1} - u_{j-2}}{h} \right\|_{L^2(\Omega)} \|u_j - u_{j-1}\|_{L^2(\Omega)}.$$

By combining these estimates, we deduce

$$\left\|\frac{u_{j} - u_{j-1}}{h}\right\|_{L^{2}(\Omega)} \le \left\|\frac{u_{j-1} - u_{j-2}}{h}\right\|_{L^{2}(\Omega)}$$

for $j \in \{2, ..., m\}$. We may now "iterate" the above inequality and use (2.15) to obtain

$$\left\|\frac{u_j - u_{j-1}}{h}\right\|_{L^2(\Omega)} \le \|\Delta u_0\|_{L^2(\Omega)},$$

which gives (2.13).

Now note that for any $1 \leq i \leq m$,

$$u_i = (u_i - u_{i-1}) + (u_{i-1} - u_{i-2}) + \ldots + (u_1 - u_0) + u_0.$$

Coupling this with (2.13), we deduce that

$$||u_i||_{L^2(\Omega)} \le C_0 hm + ||u_0||_{L^2(\Omega)} \le C_1 T + ||u_0||_{L^2(\Omega)}$$

for every $i \in \{1, \ldots, m\}$ as $h = \frac{T}{m}$. Also, for v = 0 in (2.12) we have

$$\|\nabla u_i\|_{L^2(\Omega)}^2 \le \frac{1}{h} \int_{\Omega} (u_{i-1} - u_i) u_i \mathrm{d}x,$$

so (2.14) follows from the Cauchy-Schwarz and Poincaré inequalities.

Proof of Proposition 2.3. — First, using the bound (2.14) we observe that for a.e. $t \in (0, T)$, we have the estimates

(2.16)
$$\|\mathbf{u}_m(t,\cdot)\|_{H^1_0(\Omega)} \le 3C$$

and

(2.17)
$$\|\overline{\mathbf{u}}_m(t,\cdot)\|_{H^1_0(\Omega)} \le C$$

where $C = C(T, u_0) > 0$ is independent of m. Notice that the bound (2.13) provides a uniform estimate on the time derivative of the piecewise differentiable function \mathbf{u}_m , since the weak derivative is

$$(\mathbf{u}_m)_t = \frac{u_i - u_{i-1}}{h},$$

thus

(2.18)
$$\operatorname{ess\,sup}_{\tau\in[0,T]} \|(\mathbf{u}_m)_t(\tau,\cdot)\|_{L^2(\Omega)} \le C_0.$$

The above estimate gives

(2.19)
$$\|\mathbf{u}_m(t,\cdot) - \mathbf{u}_m(\tau,\cdot)\|_{L^2(\Omega)} \le C_0 |t-\tau| \text{ for } t, \tau \in [0,T].$$

This implies that the family $\{\mathbf{u}_m\}_{m=1}^{\infty}$ is equicontinuous, thus the Arzelà-Ascoli theorem guarantees the existence of $u \in C^0([0,T]; L^2(\Omega))$ such that

$$\mathbf{u}_m \to u$$
 strongly in $C^0([0,T]; L^2(\Omega))$

along some subsequence as $m \to \infty$. It also follows from (2.17) and the Banach-Alaoglu theorem that $\{\overline{\mathbf{u}}_m\}_{m=1}^{\infty}$ has a subsequence that converges weakly in $L^2(0,T; H_0^1(\Omega))$. Now (2.13) yields the bound

$$\|\overline{\mathbf{u}}_m(t,\cdot) - \mathbf{u}_m(t,\cdot)\|_{L^2(\Omega)} \le hC_1 + |t - t_{i-1}|C_1 \le 2hC_1 = \frac{2TC_1}{m},$$

for a.e. $t \in [0,T]$, from which it follows that $u \in L^2(0,T; H^1_0(\Omega))$ and

$$\overline{\mathbf{u}}_m \rightharpoonup u \quad \text{weakly in } L^2(0,T; H^1_0(\Omega))$$

along a subsequence as $m \to \infty$. Moreover, since $\mathcal{K}(\psi)$ is weakly closed, $u(t, \cdot) \in \mathcal{K}(\psi)$ for a.e. $t \in (0, T)$. Finally, by exploiting the bound (2.16) in a similar fashion, we deduce that $u_t \in L^2(\Omega)$ and

$$(\mathbf{u}_m)_t \rightharpoonup u_t$$
 weakly in $L^2(\mathbb{Q})$

along a subsequence as $m \to \infty$.

Now in terms of \mathbf{u}_m and $\overline{\mathbf{u}}_m$, the semidiscretized variational inequality reads

$$\int_{\Omega} (\mathbf{u}_m)_t (v - \overline{\mathbf{u}}_m) \mathrm{d}x + \int_{\Omega} \nabla \overline{\mathbf{u}}_m \cdot \nabla (v - \overline{\mathbf{u}}_m) \mathrm{d}x \ge 0$$

for all $v \in \mathcal{K}(\psi)$ and holds a.e. in [0, T]. For arbitrary points $\tau_1, \tau_2 \in [0, T]$, integrating the above from τ_1 to τ_2 gives

$$\int_{\tau_1}^{\tau_2} \int_{\Omega} (\mathbf{u}_m)_t (v - \overline{\mathbf{u}}_m) \mathrm{d}x \mathrm{d}t + \int_{\tau_1}^{\tau_2} \int_{\Omega} \nabla \overline{\mathbf{u}}_m \cdot \nabla (v - \overline{\mathbf{u}}_m) \mathrm{d}x \mathrm{d}t \ge 0$$

for all $v \in \mathcal{K}(\psi)$. Since the norm and the inner product (for fixed v) are weakly lower semicontinuous functions (see the Appendix), we have

$$\liminf_{m \to \infty} \left(\int_{\tau_1}^{\tau_2} \int_{\Omega} |\nabla \overline{\mathbf{u}}_m|^2 \mathrm{d}x \mathrm{d}t + \int_{\tau_1}^{\tau_2} \int_{\Omega} \nabla \overline{\mathbf{u}}_m \cdot \nabla(-v) \mathrm{d}x \mathrm{d}t \right)$$
$$\geq \int_{\tau_1}^{\tau_2} \int_{\Omega} |\nabla u|^2 \mathrm{d}x \mathrm{d}t - \int_{\tau_1}^{\tau_2} \int_{\Omega} \nabla u \cdot \nabla v \mathrm{d}x \mathrm{d}t,$$

by virtue of superadditivity of the limit inferior and the weak $L^2(0,T; H_0^1(\Omega))$ convergence of $\overline{\mathbf{u}}_m$. Taking the limit inferior as $m \to \infty$, we finally obtain

$$\int_{\tau_1}^{\tau_2} \int_{\Omega} u_t(v-u) \mathrm{d}t \mathrm{d}x + \int_{\tau_1}^{\tau_2} \int_{\Omega} \nabla u \cdot \nabla (v-u) \mathrm{d}x \mathrm{d}t \ge 0$$

for all $v \in \mathcal{K}(\psi)$ by using the previous limiting argument as well as the weak $L^2(\mathbb{Q})$ convergence of $(\mathbf{u}_m)_t$ and the strong $L^2(\mathbb{Q})$ convergence of $\overline{\mathbf{u}}_m$. Since the above holds for any $\tau_1, \tau_2 \in [0, T]$, we deduce that u is a solution of (2.11). Uniqueness follows by arguing in the same way as for (2.1).

3. Numerical experiments

To conduct numerical simulations of obstacle problems, we used the FEniCS software [5, 35, 37, 36]. FEniCS is a collection of open-source software components aimed at the numerical resolution of PDEs using the finite element method, and can be used from both Python and C++. As input, FEniCS takes the variational formulation of a PDE and then proceeds in discretizing and assembling the associated matrices. Once the variational formulation has been discretized, one chooses a solver (linear or nonlinear) depending on the PDE to solve the system of algebraic equations. The code for obtaining the figures is available at https://github.com/borjanG.

3.1. The finite element method. — We first give a very brief presentation of the finite element method. Assume that V is a real Hilbert space with dual V' and consider the problem

Find
$$u \in V$$
 such that $F(u) = f$,

where $F: V \to V'$ and $f \in V'$. The map F may for instance represent a differential operator such as the Dirichlet Laplacian. The weak form of the above problem reads

Find $u \in V$ such that $\langle F(u), v \rangle = \langle f, v \rangle$ for all $v \in V$,

where $\langle \cdot, \cdot \rangle$ denotes the duality pairing between V' and V. The idea behind the finite element method is to use the Galerkin technique, where one replaces the above weak form by

Find
$$u_h \in V_h$$
 such that $\langle F(u_h), v_h \rangle = \langle f, v_h \rangle$ for all $v_h \in V_h$,

where $V_h \subset V$ is a finite dimensional subspace of V (say of dimension N) with a finite dimensional dual space V'_h . In conforming finite-element spaces, i.e. when $V_h \subsetneq V$, well-posedness of the discrete problem is inherited from the infinite dimensional setting. If F is linear, then one can fix a basis $\{\phi_k\}_{k=1}^N$ of V_h , define the matrix $F_h \in \mathbb{R}^{N \times N}$ representing F on V_h (called the *stiffness matrix*) as well as the data vector $f_h = (\langle f, \phi_k \rangle)_{1 \le k \le N}$ and rewrite the discrete weak form as the linear system of equations

Find $U \in \mathbb{R}^N$ such that $F_h U = f_h$.

The vector U is the vector of *degrees of freedom* that are to be computed, which are the coefficients of the solution u_h in the basis $\{\phi_k\}_{k=1}^N$, i.e.

$$u_h(x) = \sum_{k=1}^N \mathcal{U}_k \phi_k(x).$$

The basis functions are called *shape* functions. We will use piecewise linear shape functions. The discrete spaces used in the numerical experiments in this work are based on either an unstructured triangular mesh over the ball $B_2(0) \subset \mathbb{R}^2$ consisting of 8270 points and 16136 cells, or a structured triangular mesh over the square $(-1, 1)^2$ consisting of 4225 points and 8192 cells.



FIGURE 4. Meshes for the ball $B_2(0)$ (left) and the square $(-1,1)^2$ (right).

3.2. Classical obstacle problem. — Recall that the variational formulation of the obstacle problem (1.2) is not an equation but rather an inequality. In order to use the finite element method, our approach will be to approximate the inequality via a family of penalized problems as we did for the regularity analysis. For programming purposes, we use a slightly different penalty function (proposed in [28]). Namely, for $\varepsilon > 0$ we look for a weak solution $u^{\varepsilon} \in H_0^1(\Omega)$ to the nonlinear problem

(3.1)
$$\begin{cases} -\Delta u^{\varepsilon} - \frac{1}{\varepsilon} \max\{-u^{\varepsilon} + \psi, 0\} = f & \text{in } \Omega\\ u^{\varepsilon} = 0 & \text{on } \partial\Omega. \end{cases}$$

Due to the monotonicity of the map $u \mapsto -\max\{u - \psi, 0\}$ (which is defined pointwise) it can be shown that for each $\varepsilon > 0$ there exists a unique solution to the penalized problem (3.1). One may also show that this solution converges weakly in $H_0^1(\Omega)$ to the unique solution of the classical obstacle problem. For more detail, we refer to [28].

Now for a numerical resolution of nonlinear partial differential equations of the form

$$-\Delta u + F(u) = 0,$$

there are two common approaches:

• Fixed point iteration: Pick u^0 and for $k \ge 1$, solve

$$-\Delta u^{k+1} + F(u^k) = 0.$$

In other words, one replaces u by u^{k+1} in all linear terms and by u^k in all nonlinear terms.

• Newton's method: Pick u^0 and for $k \ge 1$, solve for δu

$$-\Delta\delta u + F'(u^k)\delta u = -(\Delta u^k + F(u^k)),$$

and set $u^{k+1} := u^k + \delta u$.

Naturally the choice of which method to use depends on the differentiability properties of the nonlinearity F. In our approximation problem, the map $\mathbf{x} \mapsto \max{\{\mathbf{x}, 0\}}$ is not differentiable at the origin. We will use the Newton approach (which is automatized in FEniCS) by regularizing the max map using C^1 -approximations. Namely, for a fixed parameter $\alpha > 0$ we consider (see Figure 5)

$$\max_{\alpha} : \mathbf{x} \mapsto \begin{cases} \mathbf{x} - \frac{\alpha}{2} & \text{if } \mathbf{x} \ge \alpha \\ \frac{\mathbf{x}^2}{2\alpha} & \text{if } 0 < \mathbf{x} < \alpha \\ 0 & \text{if } \mathbf{x} \le 0. \end{cases}$$

It is readily seen that \max_{α} is monotone, convex, continuously differentiable, and the range of its derivative is included in [0, 1]. Consequently \max_{α} is Lipschitz continuous with Lipschitz constant 1 by the mean-value theorem. Moreover, the error decreases linearly with α as

$$0 \le \max(0, \mathbf{x}) - \max_{\alpha}(\mathbf{x}) \le \frac{\alpha}{2}$$

holds for any $\mathbf{x} \in \mathbb{R}$.



FIGURE 5. The penalty function and its approximation with $\alpha = 10^{-1}$.

We now present some numerical experiments. Post-processing was conducted by using the ParaView software.

Example 3.1. — As a first example we present a case where the exact solution is known. Consider the ball $\Omega = B_2(0) \subset \mathbb{R}^2$, $f \equiv 0$ and consider the obstacle

$$\psi: (x_1, x_2) \mapsto \begin{cases} \sqrt{1 - x_1^2 - x_2^2} & \text{if } x_1^2 + x_2^2 < 1 \\ 0 & \text{otherwise.} \end{cases}$$

In this case the problem has a radial solution $u(x_1, x_2) = v(r)$ with $r = \sqrt{x_1^2 + x_2^2}$. For $r \neq 0$, a simple computation gives $r_{x_i} = \frac{x_i}{r}$, and consequently by the chain rule we have

$$u_{x_i x_i} = v''(r) \frac{x_i^2}{r^2} + v'(r) \left(\frac{1}{r} - \frac{x_i^2}{r^3}\right)$$

for i = 1, 2. Therefore

$$\Delta u = v''(r) + \frac{1}{r}v'(r),$$

and if $u > \psi$ then $v''(r) + \frac{1}{r}v'(r) = 0$, whence

$$v(r) = -b\log r + c$$

where b and c are constants. If the free boundary $\partial \{u > \psi\}$ is at the position r = a, then we seek for a, b, c satisfying the nonlinear equations

$$v(a) = \psi(a), \quad v'(a) = \psi'(a), \quad v(2) = 0.$$

Clearly 0 < a < 1 (see the definition of ψ) and the equations reduce to the following nonlinear equation for a:

$$a^2(\ln 2 - \ln a) = 1 - a^2,$$

with $b = a^2(1-a^2)^{-1/2}$, and $c = b \ln 2$. One may find the root of the above equation numerically and obtain a = 0.69797, b = 0.68026 and c = 0.47152. For the penalty parameter $\varepsilon = 10^{-5}$ and approximation parameter $\alpha = 10^{-4}$ we obtain the following result.



FIGURE 6. The solution u (wireframe in color) is above the obstacle ψ (solid in white).

Example 3.2. — We now consider the square $\Omega = (-1, 1)^2 \subset \mathbb{R}^2$, the source term $f \equiv -10$ and we the "staircase" obstacle

$$\psi: (x_1, x_2) \mapsto \begin{cases} -0.2 & \text{if } x_1 \in [-1, -0.5) \\ -0.4 & \text{if } x_1 \in [-0.5, 0) \\ -0.6 & \text{if } x_1 \in [0, 0.5) \\ -0.8 & \text{if } x_1 \in [0.5, 1]. \end{cases}$$

For the penalty parameter $\varepsilon = 10^{-5}$ and approximation parameter $\alpha = 10^{-4}$ we obtain the following result.



FIGURE 7. The solution u (wireframe) is above the obstacle ψ (solid).

The negative source term would physically represent gravity (the minus sign is due to the fact that gravity points downwards). Intuitively, the membrane will sag through and touch the staircase under this load.

Example 3.3. — Finally, let $\Omega = (-1, 1)^2$, $\psi : (x_1, x_2) \mapsto -x_1^2$ and $f \equiv -4$. For the penalty parameter $\varepsilon = 10^{-5}$ and approximation parameter $\alpha = 10^{-4}$ we obtain the following result.



FIGURE 8. The solution u (wireframe) is above the obstacle ψ (solid).

3.3. Parabolic obstacle problem. — The numerical implementation of the parabolic obstacle problem is similar to the one for the heat equation. Namely, we discretize the time derivative using an implicit Euler scheme (which guarantees unconditional stability), and much like the theoretical result presented in the previous section, solve a steady-state problem at each time-step. In view of what was done for simulating the classical obstacle problem, this leads us to consider a penalized problem, namely, for fixed $\varepsilon > 0$, finding a weak solution u^{ε} to

$$\begin{cases} u_t^{\varepsilon} - \Delta u^{\varepsilon} - \frac{1}{\varepsilon} \max\{-u^{\varepsilon} + \psi, 0\} = f & \text{ in } \Omega\\ u^{\varepsilon} = 0 & \text{ on } \Sigma\\ u^{\varepsilon}(t, \cdot) = u_0 & \text{ in } \Omega. \end{cases}$$

After discretizing the time derivative using the Euler scheme and approximating the max map by the C^1 approximations proposed for the classical obstacle problem, we proceed in conducting numerical experiments.

Example 3.4. — We consider the square $\Omega = (-2,2)^2 \subset \mathbb{R}^2$, $f \equiv 0$, and we consider the static obstacle from Example 3.1. We are in a transient setting, so we also need an initial datum and a final time; take T = 16 and $u_0(x) = \psi(x)$. For $\varepsilon = 10^{-5}$ and $\alpha = 10^{-4}$, we obtain the following results.



FIGURE 9. Clockwise from top left: t = 0, 2, 8, 16. We observe that at each time the solution u (wireframe) is above the obstacle (solid), and we also observe the expected diffusion phenomenon.

PART II

AN OBSTACLE PROBLEM WITH COHESION

4. The cohesive obstacle problem

We now present a paper of M. Hintermüller, V.A. Kovtunenko and K. Kunisch [27], in which the authors investigate a variant of the classical obstacle problem, namely a steady-state obstacle problem subject to cohesion forces. As before, we consider $\Omega \subset \mathbb{R}^n$ a bounded domain with smooth boundary $\partial\Omega$. Let $\psi \in H^2(\Omega) \cap C^0(\overline{\Omega})$ with $\psi = 0$ on $\partial\Omega$ be a given obstacle. Consider a membrane which occupies the domain Ω and is clamped at the boundary $\partial\Omega$. Under a loading force $f \in L^2(\Omega)$, the membrane is in contact with the obstacle in such a way that a *cohesion* phenomenon, i.e. a mutual attraction of the molecules, occurs between these two bodies. The cohesion force is described by fixed material parameters: $\gamma > 0$ (unit of force times distance) and $\delta > 0$ (unit of distance). The problem consists in finding the normal displacement $u \in H^2(\Omega) \cap H^1_0(\Omega)$ and the normal force $\xi \in L^2(\Omega)$ of the membrane such that the pair (u, ξ) form a strong solution of

(4.1)
$$\begin{cases} -D\Delta u - f = \xi & \text{in } \Omega\\ u = 0 & \text{on } \partial\Omega, \end{cases}$$

where D > 0 is a fixed constant, and

(4.2)
$$u \ge \psi, \quad \begin{cases} \xi = 0 & \text{if } u > \psi + \delta \\ \xi = -\frac{\gamma}{\delta} & \text{if } \psi < u \le \psi + \delta \\ \xi \ge -\frac{\gamma}{\delta} & \text{if } u = \psi, \end{cases}$$

hold almost everywhere in Ω^{16} .

4.1. Physical interpretation. — In linear elasticity theory (see [43]), D denotes the *flexural rigidity*¹⁷ of the membrane and is given by

$$D = \frac{E\theta^3}{12(1-\nu^2)} > 0,$$

¹⁶ Notice that the map $u \mapsto \xi$ defined is discontinuous whenever $u(x) = \psi(x) + \delta$.

¹⁷ In layman's terms, this represents the resistance offered by a structure while undergoing bending.

where E denotes Young's modulus (a measure of the stiffness, i.e. the rigidity of a solid), ν denotes Poisson's ratio (the negative of the ratio of the signed lateral strain to the signed axial strain) and θ denotes the elastic thickness of the membrane. However, for computational convenience, we henceforth set D = 1.

The ratio $\frac{\gamma}{\delta}$ represents the *elastic limit* of the membrane, meaning the point beyond which any deformation of the membrane becomes permanent. One may thence give a natural interpretation of the conditions (4.2) which define the normal force ξ : if the membrane is sufficiently far from the obstacle, then the normal force is 0, otherwise the membrane will bend, and the deformation is permanent if the membrane is in contact with the obstacle. In fact, by virtue of the the Heaviside function $\mathcal{H} : x \mapsto \mathbb{1}_{[0,\infty)}(x)$, we may show after simple computations that the constraints in (4.2) are equivalent to the *complementarity system*

(4.3)
$$\begin{cases} u \ge \psi & \text{a.e. in } \Omega \\ \xi + \frac{\gamma}{\delta} \mathcal{H}(\delta - u + \psi) \ge 0 & \text{a.e. in } \Omega \\ \left[\xi + \frac{\gamma}{\delta} \mathcal{H}(\delta - u + \psi)\right](u - \psi) = 0 & \text{a.e. in } \Omega. \end{cases}$$

The cohesion force is described by $p := \frac{\gamma}{\delta} \mathcal{H}(\delta - u + \psi)$. The constraints formulated in the complementarity system above imply that the normal force ξ acts in the opposite direction to the cohesion force p. Models which take into consideration cohesion forces have been actively studied in recent years [**33**, **31**] in view of their applications to fracture mechanics – the field of mechanics concerned with the study of the propagation of cracks in materials.

4.2. Mathematical interpretation. — From a mathematical viewpoint, the cohesion model (4.1), (4.2) results from the minimization of an objective functional subject to contact constraints, much like the classical obstacle problem. There is however an additional unknown in the equation, and interpreting this system is not trivial at first glance. Perhaps a useful example for illustrating how such a formulation is derived is Stokes' problem from fluid dynamics. The problem consists in minimizing the functional

$$\mathcal{S}[\mathbf{w}] := \frac{1}{2} \int_{\mathcal{O}} |D\mathbf{w}|^2 \mathrm{d}x - \int_{\mathcal{O}} \mathbf{f} \cdot \mathbf{w} \mathrm{d}x$$

over all $\mathbf{w} \in \mathcal{K}$, where $\mathcal{K} := {\mathbf{w} \in H_0^1(\Omega) : \text{div } \mathbf{w} = 0 \text{ in } 0}$, $\mathcal{O} \subset \mathbb{R}^3$ is open, bounded and simply connected, and $\mathbf{f} \in L^2(\mathcal{O}; \mathbb{R}^3)$ is a given vector field. Existence and uniqueness of a minimizer $\mathbf{u} \in \mathcal{K}$ may be shown by using the direct method. We interpret \mathbf{u} as representing the *velocity field* of a steady fluid flow (continuous

motion) within the region \mathcal{O} , subject to the external force \mathbf{f} . The constraint that div $\mathbf{u} = 0$ ensures that the flow is *incompressible*¹⁸. The interesting question however is to see how the constraint manifests itself in the first order necessary condition given by the Euler-Lagrange equation. In fact, it can be shown (see [21, Thm.6, p.472]) that there exists a scalar field $p \in L^2_{loc}(\mathcal{O})$ such that

$$\int_{\mathfrak{O}} D\mathbf{u} : D\mathbf{v} \mathrm{d}x = \int_{\mathfrak{O}} \left(p \operatorname{div} \mathbf{v} + \mathbf{f} \cdot \mathbf{v} \right) \mathrm{d}x$$

for all $\mathbf{v} \in H^1(\mathcal{O}; \mathbb{R}^3)$ with compact support inside \mathcal{O} . We interpret the above variatonal formulation as saying that the pair (\mathbf{u}, p) form a weak solution of Stokes' problem

$$\begin{cases} -\Delta \mathbf{u} = \mathbf{f} - \nabla p & \text{in } \mathcal{O} \\ \operatorname{div} \mathbf{u} = 0 & \text{in } \mathcal{O} \\ \mathbf{u} = 0 & \text{on } \partial \mathcal{O}. \end{cases}$$

The function p is the *pressure* and arises as a Lagrange multiplier corresponding to the incompressibility condition div $\mathbf{u} = 0$.

The methodology for solving the cohesive obstacle problem is slightly different to Stokes' problem. While the former will also manifest itself as an optimality condition for a minimization problem, only the contact constraint will be included in the set over which we minimize. The cohesion force will be embedded in the variational formulation, and much like the classical obstacle problem, will be an inequality. Due to the form of the objective functional in the minimization problem, there will be no guarantee for uniqueness of a minimizer, and necessary and sufficient optimality conditions won't coincide.

We will first show existence of a minimizer and necessary optimality conditions, and we will derive the normal force ξ as a function of this minimizer and the cohesion force. We will also show sufficient conditions by considering the Lagrangian associated to the contact constraint, and see that the saddle-point for this Lagrangian may allow us to interpret ξ as a Lagrange multiplier.

¹⁸ This roughly means that the effects of pressure on the fluid density are zero or negligible. Thus the density and the specific volume of the fluid (i.e. the ratio of the fluid's volume to its mass) do not change during the flow.

5. Necessary conditions

The strategy we will use to derive the cohesive obstacle problem (4.1), (4.2) is analogous to the classical obstacle problem – we incorporate the constraint $u \ge \psi$ in a certain set of functions and solve a related variational problem on this set. In fact, the set in question is the same as for the classical problem:

$$\mathcal{K}(\psi) := \{ u \in H_0^1(\Omega) \colon u \ge \psi \text{ a.e. in } \Omega \}.$$

The difference for the variational problem comes from the added unknown ξ and the constraint it must satisfy, as they lack smoothness and add a nonlinearity to the problem. The associated variational inequality will remain similar nonetheless. We look for $u \in \mathcal{K}(\psi)$ such that

(5.1)
$$\int_{\Omega} \nabla u \cdot \nabla (v-u) dx - \int_{\Omega} \left[f + \frac{\gamma}{\delta} \mathcal{H}(\delta - u + \psi) \right] (v-u) dx \ge 0$$

for all $v \in \mathcal{K}(\psi)$. Inequality (5.1) is called a *hemivariational inequality*.

We begin our study by linking this problem and the equations for the cohesive problem. In fact, we may interpret them as the Euler-Lagrange equations for the obstacle problem with cohesion.

Proposition 5.1. — If there exists a solution $u \in \mathcal{K}(\psi)$ to the hemivariational inequality (5.1), then $u \in H^2(\Omega)$. Moreover, the formulation (4.1), (4.2) is equivalent to the hemivariational inequality.

Proof. — Suppose that a solution $u \in \mathcal{K}(\psi)$ of (5.1) exists. Setting

$$\mathbf{f} := f - \frac{\gamma}{\delta} \mathcal{H}(\delta - u + \psi) \in L^2(\Omega),$$

we have

$$\int_{\Omega} \nabla u \cdot \nabla (v - u) \mathrm{d}x - \int_{\Omega} \mathbf{f}(v - u) \mathrm{d}x \ge 0,$$

for all $v \in \mathcal{K}(\psi)$. This is the classical obstacle problem, and since the data (\mathbf{f}, ψ) satisfy the required assumptions, the previously established regularity results imply $u \in H^2(\Omega)$.

Suppose now that $u \in \mathcal{K}(\psi) \cap H^2(\Omega)$ satisfies (4.1), (4.2). Taking the $L^2(\Omega)$ inner product by v - u where $v \in \mathcal{K}(\psi)$ is arbitrary, and using Green's first identity, we obtain

$$\int_{\Omega} \left(\nabla u \cdot \nabla (v - u) - f(v - u) + \xi(v - u) \right) \mathrm{d}x = 0.$$

By (4.3), we have $\xi \geq -\frac{\gamma}{\delta}\mathcal{H}(\delta - u + \psi)$, and plugging this in the identity above yields (5.1). The converse can be argued with $\xi := -\Delta u - f \in L^2(\Omega)$ by choosing appropriate variations v.

It remains to be seen whether there exists a solution to the hemivariational inequality. To do so, we show in fact that it represents a necessary optimality condition for a certain minimization problem. First, we define (pointwise) the continuous, nondifferentiable¹⁹ and concave map (see Figure 10 for a visualization)

(5.2)
$$g: u \mapsto \frac{\gamma}{\delta} \min(\delta, u - \psi).$$

The following lemma will be of use in the proofs to come, and hints why we might deduce the Heaviside term from a minimization problem. For a proof, we refer the reader to the Appendix.

Lemma 5.1. — The map g satisfies

$$g(v) - g(u) \le \frac{\gamma}{\delta} \mathcal{H}(\delta - u + \psi)(v - u),$$

for all $u, v \in H_0^1(\Omega)$.

We now set up the hinted minimization problem. Consider the cost functional $\mathcal{J}: H_0^1(\Omega) \to \mathbb{R}$ defined as

$$\mathcal{J}[u] := \frac{1}{2} \int_{\Omega} |\nabla u|^2 \mathrm{d}x - \int_{\Omega} f u \mathrm{d}x + \int_{\Omega} g(u) \mathrm{d}x = \mathcal{E}[u] + \int_{\Omega} g(u) \mathrm{d}x.$$

Notice that \mathcal{J} is nonconvex and nondifferentiable solely due to the presence of g.

Proposition 5.2. — There exists at least one function $u \in \mathcal{K}(\psi)$ satisfying

(5.3)
$$\mathcal{J}[u] = \inf_{w \in \mathcal{K}(\psi)} \mathcal{J}[w]$$

Moreover, $u \in H^2(\Omega)$.

Proof. — The arguments of the proof mainly follow the direct method in the calculus of variations. We begin by observing that the map g is nonnegative over the set $\mathcal{K}(\psi)$. Hence, for an arbitrary $w \in \mathcal{K}(\psi)$ we have

$$\mathcal{J}[w] = \mathcal{E}[w] + \int_{\Omega} g(w) \mathrm{d}x \ge \mathcal{E}[w].$$

¹⁹ This follows from the definition of the min map in terms of the absolute value. In fact, this characterization also implies that g is Lipschitz continuous.

Since \mathcal{E} is coercive, the above inequality implies that \mathcal{J} is also coercive.

Now let $\{u_k\}_{k=1}^{\infty} \subset \mathcal{K}(\psi)$ be a minimizing sequence. Using the coercivity of \mathcal{J} , by contraposition we deduce that $\{u_k\}_{k=1}^{\infty}$ is bounded in $H_0^1(\Omega)$. The Banach-Alaoglu theorem asserts the existence of $u \in H_0^1(\Omega)$ such that

 $u_k \rightharpoonup u$ weakly in $H_0^1(\Omega)$

along a subsequence as $k \to \infty$. Since $\mathcal{K}(\psi)$ is weakly closed, $u \in \mathcal{K}(\psi)$, and by the Rellich-Kondrachov theorem, we also have strong $L^2(\Omega)$ convergence. The latter convergence implies that $u_k \to u$ a.e. in Ω along an additional subsequence. Whence, by using the superadditivity of the limit inferior, the weak lower semicontinuity of \mathcal{E} , and the Fatou lemma (recall that g is continuous) respectively, we have (along the subsequence)

$$\liminf_{k \to \infty} \mathcal{J}[u_k] \ge \liminf_{k \to \infty} \mathcal{E}[u_k] + \liminf_{k \to \infty} \int_{\Omega} g(u_k) \mathrm{d}x \ge \mathcal{E}[u] + \int_{\Omega} g(u) \mathrm{d}x = \mathcal{J}[u].$$

We conclude²⁰ that u minimizes \mathcal{J} over $\mathcal{K}(\psi)$. Finally, we use Propositions 5.3 and 5.1 respectively to conclude that $u \in H^2(\Omega)$.

Now we may link the hemivariational inequality to the problem of minimizing \mathcal{J} .

Proposition 5.3. — The hemivariational inequality (5.1) is the necessary optimality condition for the minimization problem (5.3).

Proof. — Let $u \in \mathcal{K}(\psi)$ be a solution of (5.3), hence

$$\mathcal{E}[u] + \int_{\Omega} g(u) \mathrm{d}x \le \mathcal{E}[v] + \int_{\Omega} g(v) \mathrm{d}x,$$

for all $v \in \mathcal{K}(\psi)$. By rearranging this inequality and using Lemma 5.1, one obtains

$$\mathcal{E}[u] - \mathcal{E}[v] \le \int_{\Omega} (g(v) - g(u)) dx \le \frac{\gamma}{\delta} \int_{\Omega} \mathcal{H}(\delta - u + \psi)(v - u) dx,$$

for all $v \in \mathcal{K}(\psi)$. Fix $\tau \in (0, 1]$ and consider $v := \tau w + (1 - \tau)u$ where $w \in \mathcal{K}(\psi)$ is arbitrary. Since $\mathcal{K}(\psi)$ is convex, $v \in \mathcal{K}(\psi)$ and plugging in the previous inequality we obtain

$$\frac{1}{\tau} \left(\mathcal{E}[u + \tau(w - u)] - \mathcal{E}[u] \right) \ge -\frac{\gamma}{\delta} \int_{\Omega} \mathcal{H}(\delta - u + \psi)(w - u) \mathrm{d}x.$$

²⁰ The above also shows that $\mathcal{J}: H^1_0(\Omega) \to \mathbb{R}$ is a weakly lower semicontinuous functional.

As $\mathcal{E}: H_0^1(\Omega) \to \mathbb{R}$ is Gâteaux differentiable with Gâteaux derivative

$$\delta \mathcal{E}[u; w - u] = \int_{\Omega} \nabla u \cdot \nabla (w - u) \mathrm{d}x - \int_{\Omega} f(w - u) \mathrm{d}x,$$

we may pass to the limit as $\tau \to 0$ to obtain (5.1).

Henceforth, we set $L^2_+(\Omega) := \{ f \in L^2(\Omega) : f \ge 0 \text{ a.e. in } \Omega \}$. We may regroup the previous results in the following theorem.

Theorem 5.2. — There exist $u \in \mathcal{K}(\psi) \cap H^2(\Omega)$ and $\lambda \in L^2_+(\Omega)$ satisfying

(5.4)
$$\int_{\Omega} \left[\nabla u \cdot \nabla v - fv + \frac{\gamma}{\delta} \mathcal{H}(\delta - u + \psi) v \right] \mathrm{d}x = \int_{\Omega} \lambda v \mathrm{d}x,$$

for all $v \in H_0^1(\Omega)$, and satisfy the complementarity condition

(5.5)
$$\int_{\Omega} \lambda(u-\psi) \mathrm{d}x = 0.$$

The function u is a solution to the hemivariational inequality (5.1), and the pair (u,ξ) where

$$\xi := \lambda - \frac{\gamma}{\delta} \mathcal{H}(\delta - u + \psi),$$

is a strong solution to the cohesive obstacle problem (4.1), (4.2).

Proof. — We take a minimizer $u \in \mathcal{K}(\psi) \cap H^2(\Omega)$ of \mathcal{J} . Its existence follows from Proposition 5.2. Now consider

$$\lambda := -\Delta u - f + \frac{\gamma}{\delta} \mathcal{H}(\delta - u + \psi)$$

Since $u \in H^2(\Omega)$, we have $\lambda \in L^2(\Omega)$, hence λ is well defined a.e. in Ω . Multiplying λ by an arbitrary test function $v \in H_0^1(\Omega)$ and integrating over Ω , we deduce

$$\int_{\Omega} \left[\nabla u \cdot \nabla v - fv + \frac{\gamma}{\delta} \mathcal{H}(\delta - u + \psi) v \right] \mathrm{d}x = \int_{\Omega} \lambda v \mathrm{d}x.$$

which is (5.4).

We may now rewrite the hemivariational inequality (5.1) equivalently as

$$\int_{\Omega} \lambda(v-u) \mathrm{d}x \ge 0,$$

for all $v \in \mathcal{K}(\psi)$, which in turn by choosing appropriate variations v implies that $\lambda \geq 0$ a.e. in Ω and

$$\int_{\Omega} \lambda(u - \psi) \mathrm{d}x = 0.$$

Hence $\lambda \in L^2_+(\Omega)$ and (5.5) also holds. We conclude that the pair (u, ξ) are a strong solution to the cohesion problem by virtue of Proposition 5.1.

6. Sufficient conditions

Since \mathcal{J} is non convex, the solution to (5.3) is not necessarily unique, therefore (5.4), (5.5) do not represent a sufficient optimality condition. To deduce sufficient optimality conditions, we consider the *Lagrangian functional* associated to the contact constraint $u \geq \psi$:

$$\mathcal{L}[u,\lambda] := \mathcal{J}[u] - \int_{\Omega} \lambda(u-\psi) \mathrm{d}x,$$

and study a related saddle-point problem.

Proposition 6.1. — If there exist $u \in H_0^1(\Omega)$ and $\lambda \in L^2_+(\Omega)$ such that the pair (u, λ) satisfies

(6.1)
$$\mathcal{L}[u,\mu] \le \mathcal{L}[u,\lambda] \le \mathcal{L}[v,\lambda]$$

for all $v \in H_0^1(\Omega)$ and $\mu \in L^2_+(\Omega)$, then $u \in \mathfrak{K}(\psi)$, u is a minimizer of \mathfrak{J} , and the pair (u, λ) satisfy (5.4) and the complementarity condition (5.5).

Proof. — Let $(u, \lambda) \in H_0^1(\Omega) \times L^2_+(\Omega)$ be a solution of saddle point problem. The left-hand side inequality in (6.1) implies

$$\int_{\Omega} (\mu - \lambda)(u - \psi) \mathrm{d}x \ge 0,$$

for every $\mu \in L^2_+(\Omega)$. By choosing appropriate variations μ , we deduce that $u \ge \psi$ a.e. in Ω , and

$$\int_{\Omega} \lambda(u - \psi) \mathrm{d}x = 0,$$

which gives (5.5). Next, by using the right-hand side inequality in (6.1) we obtain

$$\mathcal{J}[u] - \mathcal{J}[v] \le -\int_{\Omega} \lambda(v - \psi) \mathrm{d}x \le 0,$$

for all $v \in \mathcal{K}(\psi)$. This implies that $\mathcal{J}[u] \leq \mathcal{J}[v]$ for all $v \in \mathcal{K}(\psi)$, thus u is a minimizer of \mathcal{J} . Using the inequality above and Lemma 5.1 respectively leads us to

$$\begin{split} \mathcal{E}[u] - \mathcal{E}[v] - \int_{\Omega} \lambda(u - v) \mathrm{d}x &\leq \int_{\Omega} (g(v) - g(u)) \mathrm{d}x \\ &\leq \frac{\gamma}{\delta} \int_{\Omega} \mathcal{H}(\delta - u + \psi)(v - u) \mathrm{d}x, \end{split}$$

for all $v \in H_0^1(\Omega)$. Now fix $\tau \in (0,1]$ and consider $v := \tau w + (1-\tau)u$ where $w \in H_0^1(\Omega)$ is arbitrary. By plugging in the inequality above we and dividing by τ , we obtain

$$\frac{1}{\tau} \left(\mathcal{E}[u + \tau(w - u)] - \mathcal{E}[u] \right) - \int_{\Omega} \lambda(u - w) \mathrm{d}x \ge -\frac{\gamma}{\delta} \int_{\Omega} \mathcal{H}(\delta - u + \psi)(w - u) \mathrm{d}x.$$

As $\mathcal{E}: H_0^1(\Omega) \to \mathbb{R}$ is Gâteaux differentiable with Gâteaux derivative

$$\delta \mathcal{E}[u; w - u] = \int_{\Omega} \nabla u \cdot \nabla (w - u) dx - \int_{\Omega} f(w - u) dx$$

we may pass to the limit as $\tau \to 0$ in the previous inequality and then choose appropriate variations w to obtain (5.4).

In the above proof, we assumed that the saddle point problem (6.1) had a solution. Now, we look to show that this is indeed the case.

6.1. Existence of a saddle point. — To this end, we regularize the nondifferentiable functional \mathcal{J} by virtue of a family of approximate functionals which contain a regularization of the nondifferentiable term g. We show existence of a saddle point for the Lagrangian functional associated to the approximation functionals, obtain adequate estimates for the family of solutions, and conclude by taking the limit.

For $\varepsilon > 0$, let $g_{\varepsilon} \in C^1(\mathbb{R})$ be a function satisfying

$$0 \le g_{\varepsilon}(\cdot) \le c_0 < +\infty, 0 \le g'_{\varepsilon}(\cdot) \le c_1 < +\infty, g_{\varepsilon}(\cdot) = g(\cdot) + O(\varepsilon),$$

for some constants $c_0, c_1 > 0$ independent of ε . An example of such a function is

$$g_{\varepsilon}: x \mapsto \gamma \begin{cases} 1 - \frac{\varepsilon}{2} & \text{for } x \ge \psi(\mathbf{x}) + \delta \\ 1 - \frac{\varepsilon}{2} - \frac{(x - \psi(\mathbf{x}) - \delta)^2}{2\varepsilon\delta^2} & \text{for } \psi(\mathbf{x}) + \delta(1 - \varepsilon) < x < \psi(\mathbf{x}) + \delta \\ \frac{x - \psi(\mathbf{x})}{\delta} & \text{for } \psi(\mathbf{x}) \le x \le \psi(\mathbf{x}) + \delta(1 - \varepsilon), \end{cases}$$



FIGURE 10. The map g and its approximation g_{ε} for $\varepsilon = 10^{-1}$.

where $x = u(\mathbf{x})$ for $\mathbf{x} \in \Omega$. Its derivative is given by

$$g_{\varepsilon}': x \mapsto \frac{\gamma}{\delta} \begin{cases} 0 & \text{for } x \ge \psi(\mathbf{x}) + \delta \\ -\frac{x - \psi(\mathbf{x}) - \delta}{\varepsilon \delta} & \text{for } \psi(\mathbf{x}) + \delta(1 - \varepsilon) < x < \psi(\mathbf{x}) + \delta \\ 1 & \text{for } \psi(\mathbf{x}) \le x \le \psi(\mathbf{x}) + \delta(1 - \varepsilon). \end{cases}$$

We may now state the regularized minimization problem. For each $\varepsilon > 0$, we seek a function $u^{\varepsilon} \in \mathcal{K}(\psi)$ such that

(6.2)
$$\mathcal{J}_{\varepsilon}(u^{\varepsilon}) = \inf_{v \in \mathcal{K}(\psi)} \mathcal{J}_{\varepsilon}(v),$$

where

$$\mathcal{J}_{\varepsilon}[u] := \mathcal{E}[u] + \int_{\Omega} g_{\varepsilon}(u) \mathrm{d}x.$$

We begin by showing that such a function exists.

Lemma 6.1. — For every $\varepsilon > 0$ there exists a solution $u^{\varepsilon} \in \mathcal{K}(\psi) \cap H^2(\Omega)$ to the regularized minimization problem (6.2). The estimate

$$\|u^{\varepsilon}\|_{H^2(\Omega)} \le C,$$

also holds for some C > 0 independent of ε .

Proof. — Fix $\varepsilon > 0$. Since g_{ε} is nonnegative over \mathbb{R} , the arguments from Proposition 5.2 hold here as well, so there exists a solution $u^{\varepsilon} \in \mathcal{K}(\psi)$ to (6.2).

Now using the fact that u^{ε} is a minimizer, we may take the directional derivative of $\mathcal{J}_{\varepsilon}$ at u^{ε} in the direction $v - u^{\varepsilon}$ for arbitrary $v \in \mathcal{K}(\psi)$, and obtain the following first order optimality condition:

$$\int_{\Omega} \nabla u^{\varepsilon} \cdot \nabla (v - u^{\varepsilon}) \mathrm{d}x - \int_{\Omega} \left[f + g'_{\varepsilon}(u^{\varepsilon}) \right] (v - u^{\varepsilon}) \mathrm{d}x \ge 0$$

Arguing similarly as in Proposition 5.2, we deduce that $u^{\varepsilon} \in H^2(\Omega)$. The estimate follows from Corollary 1.3.

Now for fixed $\varepsilon > 0$ we define

(6.3)
$$\lambda^{\varepsilon} := -\Delta u^{\varepsilon} - f + g_{\varepsilon}'(u^{\varepsilon}).$$

In light of what precedes, $\lambda^{\varepsilon} \in L^2(\Omega)$ and it satisfies

(6.4)
$$\int_{\Omega} \left[\nabla u^{\varepsilon} \cdot \nabla v - fv + g'_{\varepsilon}(u^{\varepsilon})v \right] \mathrm{d}x = \int_{\Omega} \lambda^{\varepsilon} v \mathrm{d}x,$$

for all $v \in H_0^1(\Omega)$. From the definition of λ^{ε} and (6.4), we deduce that $\lambda^{\varepsilon} \ge 0$ a.e. in Ω and

(6.5)
$$\int_{\Omega} \lambda^{\varepsilon} (u^{\varepsilon} - \psi) \mathrm{d}x = 0$$

Now for fixed $\varepsilon > 0$, we consider the regularized Lagrangian functional

$$\mathcal{L}_{\varepsilon}[u,\lambda] := \mathcal{J}_{\varepsilon}[u] - \int_{\Omega} \lambda(u-\psi) \mathrm{d}x.$$

Analogously, we will consider a regularized saddle point problem. For fixed $\varepsilon > 0$, we seek for a pair $(u^{\varepsilon}, \lambda^{\varepsilon}) \in H_0^1(\Omega) \times L_+^2(\Omega)$ such that

(6.6)
$$\mathcal{L}_{\varepsilon}[u^{\varepsilon},\lambda] \leq \mathcal{L}_{\varepsilon}[u^{\varepsilon},\lambda^{\varepsilon}] \leq \mathcal{L}_{\varepsilon}[v,\lambda^{\varepsilon}],$$

for all $(v, \lambda) \in H_0^1(\Omega) \times L_+^2(\Omega)$.

The existence of solutions to these regularized saddle-point problems is derived by the properties of the functionals $\mathcal{L}_{\varepsilon}$ and use of so-called *mini-max theorems* (see [20, p.171]). For fixed $\varepsilon > 0$, the primal variable u^{ε} of any solution of this saddle point problem minimizes $\mathcal{J}_{\varepsilon}$, whereas the dual variable λ^{ε} satisfies (6.3) and (6.4).

We now look to pass to the limit as $\varepsilon \to 0$ to deduce existence of a solution for the saddle point problem (6.1).

Proposition 6.2. — There exists at least one pair $(u, \lambda) \in (H^2(\Omega) \cap H^1_0(\Omega)) \times L^2_+(\Omega)$ satisfying the saddle point problem (6.1).

Proof. — Let $\{u^{\varepsilon}, \lambda^{\varepsilon}\}_{\varepsilon>0} \subset (H^2(\Omega) \cap H^1_0(\Omega)) \times L^2_+(\Omega)$ be the family of solutions to the approximate problems (6.6). From (6.3), the uniform estimate from Lemma 6.1, the properties of g'_{ε} and $f \in L^2(\Omega)$, we infer that there exists C > 0 such that

$$\|\lambda^{\varepsilon}\|_{L^2(\Omega)} \le C$$

for all $\varepsilon > 0$. Since Lemma 6.1 gives a similar estimate on the family $\{u^{\varepsilon}\}_{\varepsilon>0}$, the Banach-Alaoglu theorem asserts the existence of $\lambda \in L^2(\Omega)$ and $u \in H^2(\Omega) \cap H^1_0(\Omega)$ such that

(6.7) $u^{\varepsilon} \rightharpoonup u \quad \text{weakly in } H^2(\Omega)$

(6.8)
$$u^{\varepsilon} \rightharpoonup u \quad \text{weakly in } H_0^1(\Omega)$$

(6.9)
$$u^{\varepsilon} \to u \quad \text{strongly in } L^2(\Omega)$$

(6.10) $\lambda^{\varepsilon} \rightharpoonup \lambda$ weakly in $L^2(\Omega)$

along a subsequence as $\varepsilon \to 0$ (observe that for (6.9) we have used the Rellich-Kondrachov theorem). Since $g'(\cdot) \leq c_1$ on \mathbb{R} , by using the mean value theorem we have

$$|g_{\varepsilon}(u^{\varepsilon}) - g(u)| \le c_1 |u^{\varepsilon} - u|$$

pointwise a.e., thus

(6.11)
$$g_{\varepsilon}(u^{\varepsilon}) \to g(u)$$
 strongly in $L^{2}(\Omega)$

as $\varepsilon \to 0$. Now, the right inequality in (6.6) reads

$$\int_{\Omega} \left(\frac{1}{2} |\nabla u^{\varepsilon}|^2 - f u^{\varepsilon} + g_{\varepsilon}(u^{\varepsilon}) - \lambda^{\varepsilon}(u^{\varepsilon} - \psi) \right) \mathrm{d}x \le \int_{\Omega} \left(\frac{1}{2} |\nabla v|^2 - f v + g_{\varepsilon}(v) - \lambda^{\varepsilon}(v - \psi) \right) \mathrm{d}x$$

Using the superadditivity of the limit inferior, the weak convergence (6.8) coupled with the weak lower semicontinuity of the \mathcal{E} , the strong convergences (6.9), (6.11) and the boundedness of $\{\lambda^{\varepsilon}\}_{\varepsilon>0}$, we have

$$\int_{\Omega} \left(\frac{1}{2} |\nabla u|^2 - fu + g(u) - \lambda(u - \psi) \right) \mathrm{d}x \le \liminf_{\varepsilon \to 0} \int_{\Omega} \left(\frac{1}{2} |\nabla u^\varepsilon|^2 - fu^\varepsilon + g_\varepsilon(u^\varepsilon) - \lambda^\varepsilon(u^\varepsilon - \psi) \right) \mathrm{d}x.$$

As a result of this limiting argument, taking the limit inferior in the inequality that precedes and using the weak convergence (6.10) for the right-hand side gives

$$\mathcal{L}[u,\lambda] \le \mathcal{L}[v,\lambda]$$

for all $v \in H_0^1(\Omega)$.

Since $\mathcal{K}(\psi)$ is weakly closed in $H_0^1(\Omega)$ and $L_+^2(\Omega)^{21}$ is weakly closed in $L^2(\Omega)$, we have also $u \in \mathcal{K}(\psi)$ and $\lambda \in L_+^2(\Omega)$. Thence $u \ge \psi$ and $\lambda \ge 0$ a.e. in Ω . Now recall that for every $\varepsilon > 0$,

$$\int_{\Omega} \lambda^{\varepsilon} (u^{\varepsilon} - \psi) \mathrm{d}x = 0.$$

This implies

$$\int_{\Omega} \lambda^{\varepsilon} (u^{\varepsilon} - u) \mathrm{d}x + \int_{\Omega} \lambda^{\varepsilon} (u - \psi) \mathrm{d}x = 0.$$

Similarly to previous limiting arguments, using the convergences (6.9), (6.10) and the boundedness of $\{\lambda^{\varepsilon}\}_{\varepsilon>0}$, we let $\varepsilon \to 0$ in the above identity to deduce

$$\int_{\Omega} \lambda(u - \psi) \mathrm{d}x = 0.$$

Hence

$$\mathcal{L}[u,\lambda] = \mathcal{J}[u],$$

and since $u \geq \psi$ a.e. in Ω , we also have

$$\mathcal{J}[u] \ge \mathcal{J}[u] - \int_{\Omega} \mu(u - \psi) \mathrm{d}x = \mathcal{L}[u, \mu],$$

for every $\mu \in L^2_+(\Omega)$. We thus conclude that

$$\mathcal{L}[u,\mu] \le \mathcal{L}[u,\lambda]$$

for every $\mu \in L^2_+(\Omega)$. The pair (u, λ) therefore satisfies the saddle point problem (6.1).

7. Active set method and algorithm

We will use the optimality system (5.4), (5.5) to formulate an algorithm for solving the obstacle problem with cohesion. We may rewrite (5.5) equivalently as

(7.1)
$$\lambda - \max\{0, \lambda - (u - \psi)\} = 0$$

where the max is taken pointwise. Using this formulation, we define the *active* and *inactive* sets with respect to the contact condition

$$\mathcal{A}_c = \{ x \in \Omega \colon \lambda(x) > u(x) - \psi(x) \}$$
$$\mathcal{I}_c = \{ x \in \Omega \colon \lambda(x) \le u(x) - \psi(x) \},\$$

64

²¹ Clearly $L^2_+(\Omega)$ is convex, and it is also closed by the same argument as for $\mathcal{K}(\psi)$, so we may apply Theorem A.4.

thus $\mathcal{I}_c = \Omega \setminus \mathcal{A}_c$, and with respect to the cohesion force

$$\mathcal{A}_p = \{ x \in \Omega \colon u(x) \le \psi(x) + \delta \}$$

$$\mathcal{I}_p = \{ x \in \Omega \colon u(x) > \psi(x) + \delta \},$$

thus $\mathcal{I}_p = \Omega \setminus \mathcal{A}_p$. Now using (7.1) as well as the active and inactive sets, we can rewrite the optimality system (5.4), (5.5) equivalently as

(7.2)
$$\int_{\Omega} \int_{\Omega} \left(\nabla u \cdot \nabla v - fv + pv - \lambda v \right) \mathrm{d}x = 0 \quad \text{for all } v \in H^1_0(\Omega)$$

(7.3)
$$p = \frac{\gamma}{\delta} \text{ on } \mathcal{A}_p, \quad p = 0 \text{ on } \mathcal{I}_p$$

(7.4)
$$\qquad \qquad \mathbf{l} u = \psi \text{ on } \mathcal{A}_c, \quad \lambda = 0 \text{ on } \mathcal{I}_c$$

We present the following algorithm for solving the reformulated optimality system (7.2), (7.3), (7.4).

Algorithm 1 : Primal dual active set algorithm for the cohesion problem.

1. Choose
$$\mathcal{A}_c^{-1}, \mathcal{A}_p^{-1} \subset \Omega$$
.
2. Set $\mathcal{I}_c^{-1} = \Omega \setminus \mathcal{A}_c^{-1}, \mathcal{I}_p^{-1} = \Omega \setminus \mathcal{A}_p^{-1}$.
3. Set $k = 0$.
4. While not Stop:
4.1. Solve for $u^k \in H_0^1(\Omega), \ \lambda^k \in L^2(\Omega), \ p^k \in L^2(\Omega)$:
(7.5)
$$\begin{cases} \int_{\Omega} \left(\nabla u^k \cdot \nabla v - fv + p^k v - \lambda^k v \right) dx = 0 \quad \text{for all } v \in H_0^1(\Omega) \\ p^k = \frac{\gamma}{\delta} \quad \text{on } \mathcal{A}_p^{k-1}, \quad p^k = 0 \quad \text{on } \mathcal{I}_p^{k-1} \\ u^k = \psi \quad \text{on } \mathcal{A}_c^{k-1}, \quad \lambda^k = 0 \quad \text{on } \mathcal{I}_c^{k-1}. \end{cases}$$
4.2. Update the active and inactive sets at u^k, λ^k :

$$\mathcal{A}_c^k := \{x \in \Omega : \lambda^k(x) > u^k(x) - \psi(x)\} \\ \mathcal{I}_c^k := \{x \in \Omega : \lambda^k(x) \le u^k(x) - \psi(x)\} \\ \mathcal{I}_c^k := \{x \in \Omega : u^k(x) \le \psi(x) + \delta\} \\ \mathcal{I}_p^k := \{x \in \Omega : u^k(x) > \psi(x) + \delta\}. \end{cases}$$
4.3. If $\mathcal{A}_c^k = \mathcal{A}_c^{k-1}$ and $\mathcal{A}_p^k = \mathcal{A}_p^{k-1}$ then Stop.
4.4. Else: $k := k + 1$.

We now turn to the study of the properties of this algorithm, starting by showing that the step (4.1) makes sense.

Lemma 7.1. — There exists a unique solution to the system (7.5), (7.6) (7.7).

Proof. — Fix an arbitrary iteration $k \ge 0$. We can clearly determine p^k at each step by (7.6). Now consider the closed and convex set

$$\mathcal{K}_c^{k-1}(\psi) := \{ u \in H_0^1(\Omega) \colon u = \psi \text{ a.e. on } \mathcal{A}_c^{k-1} \},\$$

as well as the minimization problem which consists of finding $u^k \in \mathcal{K}^{k-1}_c(\psi)$ such that

$$\mathcal{E}[u^k] + \int_{\mathcal{I}_c^{k-1}} p^k u^k \mathrm{d}x \le \mathcal{E}[v] + \int_{\mathcal{I}_c^{k-1}} p^k v \mathrm{d}x$$

for all $v \in \mathcal{K}_c^{k-1}(\psi)$. By virtue of the properties of the functional \mathcal{E} (see the Appendix) we apply Theorem 1.3 to obtain the existence of a unique solution $u^k \in \mathcal{K}_c^{k-1}(\psi)$. Now using the fact that u^k is a minimizer, by differentiating the cost functional at u^k in the direction $v - u^k$ for arbitrary $v \in \mathcal{K}_c^{k-1}(\psi)$, we obtain the first order necessary and sufficient optimality condition

$$\int_{\Omega} \left(\nabla u^k \cdot \nabla (v - u^k) - f(v - u^k) + p^k (v - u^k) \right) \mathrm{d}x \ge 0.$$

Choosing the test functions $v := u^k \pm \varphi$ for arbitrary $\varphi \in C_c^{\infty}(\Omega)$ with $\operatorname{supp}(\varphi) \subset \mathfrak{I}_c^{k-1}$ yields

(7.8)
$$-\Delta u^k - f + p^k = 0 \quad \text{in } \mathcal{I}_c^{k-1}$$

in the sense of distributions, whence $\Delta u^k \in L^2(\mathfrak{I}_c^{k-1})$. Since $u^k \in \mathcal{K}_c^{k-1}(\psi)$, we also have $\Delta u^k = \Delta \psi$ in \mathcal{A}_c^{k-1} in the sense of distributions, so $\Delta u^k \in L^2(\mathcal{A}_c^{k-1})$ as well. Consequently, $\Delta u^k \in L^2(\Omega)$. We may now define

(7.9)
$$\lambda^k := -\Delta u^k - f + p^k \in L^2(\Omega).$$

Notice that having determined λ^k in such a way, the identity (7.8) implies that $\lambda^k = 0$ in \mathcal{I}_c^{k-1} , which corresponds to (7.7). Taking the $L^2(\Omega)$ scalar product by $v \in H_0^1(\Omega)$ in (7.9) and applying Green's first identity gives (7.5).

We now show that most of the iterates of Algorithm 1 are monotonic. We shall see at the end of this part that the strategy of the proof is in fact more important than the actual result.

66

Lemma 7.2. — If $\mathfrak{I}_p^{-1} = \emptyset$, then the iterates $(u^k, \mathcal{A}_c^k, p^k, \mathcal{A}_p^k)$ of Algorithm 1 are monotonic in the following sense:

$$\psi \leq u^{1} \leq \ldots \leq u^{k-1} \leq u^{k} \leq \ldots$$
$$\Omega \supseteq \mathcal{A}_{c}^{0} \supseteq \ldots \supseteq \mathcal{A}_{c}^{k-1} \supseteq \mathcal{A}_{c}^{k} \supseteq \ldots$$
$$\frac{\gamma}{\delta} = p^{0} \geq p^{1} \geq \ldots \geq p^{k-1} \geq p^{k} \geq \ldots$$
$$\Omega = \mathcal{A}_{p}^{-1} \supseteq \mathcal{A}_{p}^{0} \supseteq \ldots \supseteq \mathcal{A}_{p}^{k-1} \supseteq \mathcal{A}_{p}^{k} \supseteq \ldots$$

Proof. — The proof will be done by induction. For $k \ge 1$ we define the differences $\delta_u^{k-1} := u^k - u^{k-1}, \quad \delta_\lambda^{k-1} := \lambda^k - \lambda^{k-1}, \quad \delta_p^{k-1} := p^k - p^{k-1}.$

We proceed in three steps.

Step 1: The assertion (7.6) which determines p^k can be rewritten equivalently as

$$p^k = \frac{\gamma}{\delta} \mathbb{1}_{A_p^{k-1}}.$$

So if $\mathcal{A}_p^{k-1} \subseteq \mathcal{A}_p^{k-2}$, then $\delta_p^{k-1} \leq 0$ a.e. in Ω . In particular, this property is satisfied for k = 1 since the initialization $\mathcal{I}_p^{-1} = \emptyset$ implies $\mathcal{A}_p^0 \subseteq \mathcal{A}_p^{-1} = \Omega$. Thus $\delta_p^0 \leq 0$ a.e. in Ω .

Step 2: Observe that for $k \ge 0$, showing $\mathcal{A}_c^k \subseteq \mathcal{A}_c^{k-1}$ is equivalent to showing $\mathcal{I}_c^{k-1} \subseteq \mathcal{I}_c^k$. Now if $x \in \mathcal{I}_c^{k-1}$, from (7.7) we have that $\lambda^k(x) = 0$. Thus to show this set inclusion, it would suffice to show that $u^k(x) - \psi(x) \ge 0$ for a.e. $x \in \mathcal{I}_c^{k-1}$.

Now fix $k \ge 1$. Using the property (7.7) once again, we see that for $x \in \Omega$, we either have $\lambda^{k-1}(x) = 0$ or $u^{k-1}(x) = \psi(x)$. Using the definition of the active and passive sets with respect to the contact condition, we obtain the following mutually exclusive possibilities:

- If $\lambda^{k-1} = 0$, then $u^{k-1} < \psi$ a.e. in \mathcal{A}_c^{k-1} and $u^{k-1} > \psi$ a.e. in \mathcal{I}_c^{k-1} ;
- If $u^{k-1} = 0$, then $\lambda^{k-1} > 0$ a.e. in \mathcal{A}_c^{k-1} and $\lambda^{k-1} \leq 0$ a.e. in \mathcal{I}_c^{k-1} .

Whence it follows that

$$u^{k-1} \le \psi$$
 and $\lambda^{k-1} \ge 0$ a.e. in \mathcal{A}_c^{k-1} ,

as well as

(7.10) $u^{k-1} \ge \psi$ and $\lambda^{k-1} \le 0$ a.e. in \mathcal{I}_c^{k-1} .

Using (7.7) as well as the above, we deduce that

$$\delta_u^{k-1} = u^k - u^{k-1} \ge u^k - \psi \ge 0 \quad \text{ a.e. in } \mathcal{A}_c^{k-1}$$

and

(7.11)
$$\delta_{\lambda}^{k-1} = \lambda^k - \lambda^{k-1} \ge \lambda^k \ge 0 \quad \text{a.e. in } \mathcal{I}_c^{k-1}.$$

Now taking the difference of the iterates at steps k and k-1 in (7.5) gives

(7.12)
$$\Delta(\delta_u^{k-1}) = \delta_p^{k-1} - \delta_\lambda^{k-1} \quad \text{a.e. in } \Omega.$$

If $\delta_p^{k-1} \leq 0$ a.e. in Ω (note that by Step 1, this would be true if $\mathcal{A}_p^{k-1} \subseteq \mathcal{A}_p^{k-2}$), then using (7.11) and (7.12) we deduce that

$$\Delta(\delta_u^{k-1}) \le 0 \quad \text{ a.e. in } \mathcal{I}_c^{k-1}.$$

By virtue of the weak maximum principle, the minimum of δ_u^{k-1} is attained on the boundary $\partial \mathcal{I}_c^{k-1}$. Since $\delta_u^{k-1} \in H_0^1(\Omega)$, one also has $\delta_u^{k-1} = 0$ on $\partial \mathcal{I}_c^{k-1} \cap \partial \Omega$, and $\delta_u^{k-1} \ge 0$ on $\partial \mathcal{A}_c^{k-1} \cap \partial \mathcal{I}_c^{k-1}$. Therefore $\delta_u^{k-1} \ge 0$ a.e. on $\partial \mathcal{I}_u^{k-1}$. Consequently

$$\delta_u^{k-1} \ge 0$$
 a.e. in Ω .

Now for a.e. $x \in \mathcal{I}_c^{k-1}$, by (7.10), we have $u^{k-1}(x) \ge \psi(x)$, so by what precedes we also have $u^k(x) \ge \psi(x)$. The remark made at the beginning of this step implies that

$$\mathcal{A}_{c}^{k} \subseteq \mathcal{A}_{c}^{k-1}$$

From $\delta_u^{k-1} \ge 0$ it also follows that $u^{k-1}(x) \le u^k(x) < \psi(x) + \delta$ for a.e. $x \in \mathcal{A}_p^k$. Hence

$$\mathcal{A}_p^k \subseteq \mathcal{A}_p^{k-1}$$

Step 3: Notice that in Step 2, we only worked on the assumption $\mathcal{A}_p^{k-1} \subseteq \mathcal{A}_p^{k-2}$ for an arbitrary $k \geq 1$. Under this assumption, Step 1 and 2 give the desired monotonicity properties. Since $\mathcal{A}_p^{k-1} \subseteq \mathcal{A}_p^{k-2}$ holds for k = 1 as remarked in Step 1, we look to conclude by induction.

Let k > 1 and assume that $\mathcal{A}_p^{k-1} \subseteq \mathcal{A}_p^{k-2}$. Then Step 1 yields $\delta_p^{k-1} \leq 0$ a.e. in Ω , while Step 2 yields $\delta_u^{k-1} \geq 0$ a.e. in Ω . But the latter also gives $\mathcal{A}_c^k \subseteq \mathcal{A}_c^{k-1}$ and $\mathcal{A}_p^k \subseteq \mathcal{A}_p^{k-1}$, which finishes the proof.

The following result is the reason behind the stopping rule set in the Algorithm.

Lemma 7.3. — If $\mathcal{A}_c^{k^*} = \mathcal{A}_c^{k^*-1}$ and $\mathcal{A}_p^{k^*} = \mathcal{A}^{k^*-1}$ at some step $k^* \in \mathbb{N} \cup \{0\}$, then the pair (u^{k^*}, λ^{k^*}) satisfies the optimality system (5.4), (5.5).

68

Proof. — If $\mathcal{A}_c^{k^\star} = \mathcal{A}_c^{k^{\star}-1}$, then $u^{k^\star} = \psi$ a.e. in \mathcal{A}^{k^\star} and $\lambda^{k^\star} = 0$ a.e. in \mathcal{I}^{k^\star} . Using also the definition of the active set $\mathcal{A}_c^{k^\star}$ and the passive set $\mathcal{I}_c^{k^\star}$, we deduce $\lambda^{k^\star} \ge 0$ and $u^{k^\star} \ge \psi$ a.e. in Ω respectively. The first remark also gives

$$\int_{\Omega} \lambda^{k^{\star}} (u^{k^{\star}} - \psi) \mathrm{d}x = \int_{\mathcal{A}^{k^{\star}}} \lambda^{k^{\star}} (u^{k^{\star}} - \psi) \mathrm{d}x + \int_{\mathcal{I}^{k^{\star}}} \lambda^{k^{\star}} (u^{k^{\star}} - \psi) \mathrm{d}x = 0$$

Hence u^{k^*} and λ^{k^*} satisfy the condition (5.5). Now if $\mathcal{A}_p^{k^*} = \mathcal{A}^{k^*-1}$, then $p^{k^*} = \frac{\gamma}{\delta} \mathcal{H}(\delta - u^{k^*} + \psi)$, which coupled with (7.5) implies (5.4). Thus the pair $(u^{k^{\star}}, \lambda^{k^{\star}})$ satisfies (5.4), (5.5).

It is important to note however that these results do not imply the convergence of the iterates to the solutions (u, λ, p) , since we do not know the rate of increase of $\{u^k\}_{k=0}^{\infty}$ and $\{\lambda^k\}_{k=0}^{\infty}$ is not necessarily monotone. To show the convergence of Algorithm 1, we discretize the system using the finite element method and show that convergence holds in the finite dimensional setting by using finite dimensional analogs of Lemma 7.2 and Lemma 7.3.

7.1. Finite element discretization. — For simplicity, let us assume that n = 2and the boundary $\partial \Omega$ is polyhedral. For a fixed discretization parameter $N \in \mathbb{N}$, we consider a mesh of triangles $\mathcal{T} = \{T\}$ of Ω and vertices $\{x_i\}_{i=1}^N \subset \Omega$ such that

$$\overline{\Omega} = \bigcup_{T \in \mathcal{T}} T$$
 and $\mu(\Omega) = \sum_{T \in \mathcal{T}} \mu(T).$

Here μ denotes the Lebesgue measure in \mathbb{R}^2 . The partition \mathcal{T} is assumed to be conforming or compatible, i.e. the intersection of any two triangles T_1 and T_2 in \mathcal{T} is either empty or an edge. We will discretize Algorithm 1 in such a way that the active and inactive sets for the discretized problem can be entirely determined by the vertices of the mesh $\{x_i\}_{i=1}^N$ and the values of the discretized functions u^N, λ^N at these points. We assume that the stiffness matrix $R = (R_{ij}) \in \mathbb{R}^{N \times N}$, which corresponds to the finite element discretization of the Dirichlet Laplacian $-\Delta$ is non-singular (i.e. det $R \neq 0$) and that for every partitioning of R into blocks

$$R = \begin{bmatrix} R_{AA} & R_{AI} \\ R_{IA} & R_{II} \end{bmatrix},$$

 $R_{II}^{-1} \geq 0$ and $R_{IA} \leq 0$ hold component-wise. For example, if $R_{ij} \leq 0$ for $i \neq j$ and $\Re(\lambda) \geq 0$ for any eigenvalue λ , then this holds. In some sense, this will serve as a "discrete" weak maximum principle in the proof of the convergence.

We approximate a function $u \in H_0^1(\Omega)$ by

$$u(x) \approx \sum_{i=1}^{N} u_i^N \phi_i(x)$$

where $\{\phi_i\}_{i=1}^N \in (H_0^1(\Omega))^N$ is some finite element basis. We discretize the load and normal forces by using the projection operator $\operatorname{proj}_N : L^2(\Omega) \to \mathbb{R}^N$, given by

$$(\operatorname{proj}_N f)_i := \int_{\Omega} f(x)\phi_i(x) \mathrm{d}x$$

for $i \in \{1, \ldots, N\}$. In particular, for

$$f(x) \approx \sum_{j=1}^{N} f_j^N \phi_j(x),$$

we have $\operatorname{proj}_N f = M f^N$, where M is the mass matrix:

$$(M_{ij})_{1 \le i,j \le N} = (\langle \phi_i, \phi_j \rangle_{L^2(\Omega)})_{1 \le i,j \le N}.$$

The representation of $\operatorname{proj}_N p$ where $p = \frac{\gamma}{\delta} \mathcal{H}(\delta - u + \psi)$ is more delicate since it involves a nonlinearity with the Heaviside function. Given u, we define the active set

$$\mathcal{A} = \{ x \in \Omega \colon \mathcal{H}(\delta - u + \psi)(x) = 1 \},\$$

whence $p = \frac{\gamma}{\delta} \mathbb{1}_{\mathcal{A}}$. We approximate \mathcal{A} by

$$\widetilde{\mathcal{A}} \approx \bigcup_{T \in \mathcal{T}, T \subset \mathcal{A}} T,$$

and then we approximate $\operatorname{proj}_N p = \frac{\gamma}{\delta} \operatorname{proj}_N \mathbb{1}_A$ by

$$(\operatorname{proj}_{N}\mathbb{1}_{\mathcal{A}})_{i} = \int_{\Omega} \mathbb{1}_{\mathcal{A}}(x)\phi_{i}(x)\mathrm{d}x \approx \int_{\Omega} \mathbb{1}_{\widetilde{\mathcal{A}}}(x)\phi_{i}(x)\mathrm{d}x = (\operatorname{proj}_{N}\mathbb{1}_{\widetilde{\mathcal{A}}})_{i}$$

for $i \in \{1, \ldots, N\}$. We will consider the discrete active set $\mathcal{A}^N = \{x_i\}_{i=1}^N \cap \mathcal{A}$, and we may then determine the discrete cohesion force $p^N = \frac{\gamma}{\delta} \mathbb{1}_{\mathcal{A}^N}$ at the vertices of the mesh. Let us note that by unisolvence, knowledge of the discrete active nodal points uniquely determines the active finite element cells T and the approximate set $\widetilde{\mathcal{A}}$. Therefore

$$(\pi(\mathbb{1}_{\mathcal{A}^N}))_i := \int_{\Omega} \mathbb{1}_{\widetilde{\mathcal{A}}}(x)\phi_i(x)\mathrm{d}x = (\mathrm{proj}_N \mathbb{1}_{\widetilde{\mathcal{A}}})_i$$

 $\mathbf{70}$

is well defined. Hence for given \mathcal{A}^N we have $\pi(p^N) = \frac{\gamma}{\delta}\pi(\mathbb{1}_{\mathcal{A}^N})$. For the convergence analysis, we assume that $\pi(\mathbb{1}_{\mathcal{A}^N})$ is nonnegative (component-wise) for every partition \mathcal{A}^N , and

$$\pi(\mathbb{1}_{\mathcal{A}^N}) \geq \pi(\mathbb{1}_{\mathcal{B}^N})$$
 if and only if $\mathcal{A}^N \supseteq \mathcal{B}^N$.

This is satisfied for the continuous and piecewise-linear finite elements on a regular grid. In the following, we remove the superscript N for notation simplicity. In the finite dimensional subspace, the reference problem (7.2), (7.3), (7.4) takes the matrix form

(7.13)
$$(Ru - Mf + \pi(p) - \lambda = 0)$$

(7.14)
$$\begin{cases} p = \frac{\gamma}{\delta} \text{ on } \mathcal{A}_p, \quad p = 0 \text{ on } \mathcal{I}_p \end{cases}$$

(7.15)
$$\mathbf{l} u = \psi \text{ on } \mathcal{A}_c, \quad \lambda = 0 \text{ on } \mathcal{I}_c$$

Analogously, the system (7.5), (7.6), (7.7) in Algorithm 1 can be expressed as

(7.16)
$$\begin{cases} Ru^k - Mf + \pi(p^k) - \lambda^k = 0 \end{cases}$$

(7.17)
$$\begin{cases} p^k = \frac{\gamma}{\delta} \text{ on } \mathcal{A}_p^{k-1}, \quad p^k = 0 \text{ on } \mathcal{I}_p^{k-1} \\ k = 0 \text{ on } \mathcal{I}_p^{k-1} \end{cases}$$

(7.18)
$$(u^{\kappa} = \psi \text{ on } \mathcal{A}_{c}^{\kappa-1}, \quad \lambda^{\kappa} = 0 \text{ on } \mathcal{I}_{c}^{\kappa-1}.$$

Taking into account all of these assumptions, we have the following

Taking into account all of these assumptions, we have the following convergence result.

Theorem 7.4. — If $\mathfrak{I}_p^{-1} = \emptyset$, then the iterates (u^k, λ^k, p^k) of Algorithm 1 written in the form (7.16), (7.17), (7.18) converge to a solution $(u^{k^*}, \lambda^{k^*}, p^{k^*})$ of (7.13), (7.14), (7.15) in a finite number of steps $k^* \in \mathbb{N}$, and satisfy

$$\psi \leq u^{1} \leq \ldots \leq u^{k-1} \leq u^{k} \leq \ldots$$
$$\{x_{k}\}_{k=1}^{N} \supseteq \mathcal{A}_{c}^{0} \supseteq \ldots \supseteq \mathcal{A}_{c}^{k-1} \supseteq \mathcal{A}_{c}^{k} \supseteq \ldots$$
$$\frac{\gamma}{\delta} = p^{0} \geq p^{1} \geq \ldots \geq p^{k-1} \geq p^{k} \geq \ldots$$
$$\{x_{k}\}_{k=1}^{N} = \mathcal{A}_{p}^{-1} \supseteq \mathcal{A}_{p}^{0} \supseteq \ldots \supseteq \mathcal{A}_{p}^{k-1} \supseteq \mathcal{A}_{p}^{k} \supseteq \ldots$$

Proof. — The idea is to repeat the arguments of the proof for Lemma 7.2, essentially by replacing the "a.e. inequalities" with "component-wise inequalities". Notably, one would use the assumptions on the stiffness matrix R instead of the weak maximum principle in the argument. Having deduced the monotonicity properties, we apply Lemma 7.3 to deduce the convergence.

For $k \geq 1$, we define the following difference vectors in \mathbb{R}^N :

$$\delta_u^{k-1} := u^k - u^{k-1}, \quad \delta_\lambda^{k-1} := \lambda^k - \lambda^{k-1} \quad \delta_p^{k-1} = \pi(p^k) - \pi(p^{k-1}).$$

The discrete analog of Step 1 in the proof of Lemma 7.2 remains true when replacing the a.e. arguments by component-wise ones for the involved vectors.

From (7.16) we derive the identity

$$R\delta_u^{k-1} = \delta_\lambda^{k-1} - \delta_p^{k-1},$$

which is the matrix version of (7.12) in Step 2 in the proof of Lemma 7.2, replacing $-\Delta$ by R. Distinguishing the values of δ_u and $\delta_{\lambda} - \delta_p$ on \mathcal{A}_c^{k-1} and \mathcal{I}_c^{k-1} , we split the system in the following way:

$$\begin{bmatrix} R_{\mathcal{A}_{c}^{k-1}\mathcal{A}_{c}^{k-1}} & R_{\mathcal{A}_{c}^{k-1}\mathcal{J}_{c}^{k-1}} \\ R_{\mathcal{J}_{c}^{k-1}\mathcal{A}_{c}^{k-1}} & R_{\mathcal{J}_{c}^{k-1}\mathcal{J}_{c}^{k-1}} \end{bmatrix} \begin{bmatrix} (\delta_{u}^{k-1})_{\mathcal{A}_{c}^{k-1}} \\ (\delta_{u}^{k-1})_{\mathcal{J}_{c}^{k-1}} \end{bmatrix} = \begin{bmatrix} (\delta_{\lambda}^{k-1} - \delta_{p}^{k-1})_{\mathcal{A}_{c}^{k-1}} \\ (\delta_{\lambda}^{k-1} - \delta_{p}^{k-1})_{\mathcal{J}_{c}^{k-1}} \end{bmatrix},$$

and deduce the equality

(7.19)
$$R_{\mathcal{J}_{c}^{k-1}\mathcal{J}_{c}^{k-1}}(\delta_{u}^{k-1})_{\mathcal{J}_{c}^{k-1}} = -R_{\mathcal{J}_{c}^{k-1}\mathcal{A}_{c}^{k-1}}(\delta_{u}^{k-1})_{\mathcal{A}_{c}^{k-1}} + (\delta_{\lambda}^{k-1} - \delta_{p}^{k-1})_{\mathcal{J}_{c}^{k-1}}.$$

By inverting the left-hand side, we obtain

$$(7.20) \quad (\delta_u^{k-1})_{\mathcal{I}_c^{k-1}} = -R_{\mathcal{I}_c^{k-1}\mathcal{I}_c^{k-1}}^{-1}R_{\mathcal{I}_c^{k-1}\mathcal{A}_c^{k-1}}(\delta_u^{k-1})_{\mathcal{A}_c^{k-1}} + R_{\mathcal{I}_c^{k-1}\mathcal{I}_c^{k-1}}^{-1}(\delta_\lambda^{k-1} - \delta_p^{k-1})_{\mathcal{I}_c^{k-1}}.$$

From (7.18), we have either $\lambda^{k-1} = 0$ or $u^{k-1} = \psi$, thus $\delta_u^{k-1} \ge 0$ on \mathcal{A}_c^{k-1} , and $\delta_\lambda^{k-1} \ge 0$ on \mathcal{I}_c^{k-1} . If $\delta_p^{k-1} \le 0$, then $\delta_\lambda^{k-1} - \delta_p^{k-1} \ge 0$ on \mathcal{I}_c^{k-1} . The assumptions on the blocks of R yield $\delta_u^{k-1} \ge 0$ on \mathcal{I}_c^{k-1} . Consequently $\delta_u^{k-1} \ge 0$ for every vertex of the mesh. But if $\delta_u^{k-1} \ge 0$ then $\mathcal{A}_p^k \subseteq \mathcal{A}_p^{k-1}$ and $\delta_p^{k-1} \le 0$ due to the assumption made on π . Repeating the induction argument from Step 3, we infer the monotonicity properties of the iteration process.

The monotonic y of the active set iterates in the finite dimensional subspace guarantees that the stopping rule is satisfied after a finite number of steps. A finite dimensional analog of Lemma 5 yields the desired convergence. \Box
PART III

OPTIMAL CONTROL OF OBSTACLE PROBLEMS

As discussed in the introduction, obstacle problems may be used to model many natural or social phenomena. We have however only considered such problems with a given obstacle satisfying certain assumptions. A question that one may naturally ask is whether it is possible to obtain an optimal solution of an obstacle problem by means of choosing the obstacle²². For example, we may be interested in finding an obstacle that drives this solution to some given target u_d and satisfies some reasonable assumptions (for example it is of "minimal energy"). Hence, we find ourselves in the context of *optimal control*: the *control* (which we seek to minimize) is the obstacle ψ , and the associated *state u* is the solution to an obstacle problem.

We will only discuss optimal control of the parabolic obstacle problem presented in Section 4, following [2]. For the sake of completeness, we very briefly review the setting and main results of the elliptic case.

8. Overview of the elliptic problem

Optimal control of the obstacle in the elliptic case has been studied in [1, 3]. The problem of interest is minimizing the objective functional

$$\mathcal{J}[\psi] := \frac{1}{2} \int_{\Omega} (\sigma(\psi) - u_d)^2 \mathrm{d}x + \int_{\Omega} |\nabla \psi|^2 \mathrm{d}x,$$

over all obstacles $\psi \in H_0^1(\Omega)$, where $\sigma : H_0^1(\Omega) \to \mathcal{K}(\psi)$ maps an obstacle ψ to the solution of the variational inequality for the obstacle problem (1.2), and $u_d \in L^2(\Omega)$ is a given target profile. The authors first considered the case where the source term $f \equiv 0$. Existence and uniqueness of an optimal control ψ^* are shown, and the main result states that the optimal state $\sigma(\psi^*)$ coincides with the optimal control, i.e. $\sigma(\psi^*) = \psi^*$. The optimal control is then shown to be characterized by adjoint variables and approximation; one considers the the penalized problem for the variational inequality as the state equation, then derives some stationary conditions for the approximate controls and states, and obtains uniform estimates in order to pass to the limit. This methodology is followed in the parabolic problem we present next. When the source term $f \leq 0$, it is shown that the above result

²² Optimal control of the source term on the right-hand side for a fixed obstacle has also been considered in the literature, see [10]. We will only be interested in controlling the obstacle.

persists, otherwise, different cases based on the sign of the target profile u_d and the shape of the domain Ω are deduced. An active set algorithm in the nature of the cohesion problem from Part II is presented in [29].

9. The parabolic problem

We assume, if otherwise not stated, that the framework is the same as the one set up in Section 4. We consider the control set

$$\mathcal{U} := \left\{ \psi \in L^2(0,T; H^2(\Omega) \cap H^1_0(\Omega)) \colon \psi_t \in L^2(\Omega), \psi(0,\cdot) = 0 \text{ in } \Omega \right\}.$$

By virtue of the repeated argument [21, Thm.5, p.382], the initial condition installed in \mathcal{U} makes sense. Given a control $\psi \in \mathcal{U}$ and a source term $f \in L^2(\mathbb{Q})$, the corresponding state u is defined as the solution of the parabolic obstacle problem (2.1). We seek for an obstacle $\psi^* \in \mathcal{U}$ which minimizes the error in L^2 norm between the corresponding state u^* and a given target profile $u_d \in L^2(\mathbb{Q})$, and one which does so with the least energy. Namely, the problem in question is the (quadratic) minimization problem under variational inequality constraints, which consists of finding $\psi^* \in \mathcal{U}$ such that

(9.1)
$$\mathcal{J}[\psi^{\star}] = \inf_{\psi \in \mathcal{U}} \mathcal{J}[\psi],$$

where

$$\mathcal{J}[\psi] := \int_{\Omega} (\sigma(\psi) - u_d)^2 \mathrm{d}x \mathrm{d}t + \int_{\Omega} \left(|\Delta \psi|^2 + |\psi_t|^2 \right) \mathrm{d}x \mathrm{d}t,$$

and where $\sigma : \mathcal{U} \to \mathcal{K}(\psi)$ denotes the solution map which maps an obstacle $\psi \in \mathcal{U}$ to the solution $u \in \mathcal{K}(\psi)$ of the parabolic variational inequality (2.1).

9.1. Existence of an optimal control. — We begin our study by showing the existence of an optimal control, i.e. a minimizer of the functional \mathcal{J} defined above.

Theorem 9.1. — There exists a solution $\psi^* \in \mathcal{U}$ to the minimization problem (9.1).

Proof. — The arguments of the proof follow the direct method. Since $\mathcal{J}[\psi] \geq 0$ for all $\psi \in \mathcal{U}$, we may invoke a minimizing sequence $\{\psi_k\}_{k=1}^{\infty} \subset \mathcal{U}$ of \mathcal{J} . Observe that

$$\lim_{k\to\infty}\mathcal{J}[\psi_k] = \inf_{\psi\in \mathcal{U}}\mathcal{J}[\psi] \leq \mathcal{J}[\psi_1]$$

and so the sequence $\{\mathcal{J}[\psi_k]\}_{k=1}^{\infty}$ is bounded by some constant independent of k. Due to the form of \mathcal{J} , we immediately deduce that the sequences $\{(\psi_k)_t\}_{k=1}^{\infty}$ and $\{\Delta\psi_k\}_{k=1}^{\infty}$ are bounded in $L^2(\Omega)$, and by a Green's first identity and Cauchy-Schwarz argument,

we deduce that $\{\psi_k\}_{k=1}^{\infty}$ is bounded in $L^2(0, T; H_0^1(\Omega))$ (hence in $L^2(\Omega)$ as well). The Banach-Alaoglu theorem and Proposition A.7 in the Appendix assert the existence of $\psi^* \in \mathcal{U}$ such that

$$\begin{split} \psi_k &\to \psi^* \quad \text{strongly in } L^2(\mathbb{Q}) \\ \psi_k &\rightharpoonup \psi^* \quad \text{weakly in } L^2(0, T; H^1_0(\Omega)) \\ (\psi_k)_t &\to \psi^*_t \quad \text{weakly in } L^2(\mathbb{Q}) \\ \Delta \psi_k &\rightharpoonup \Delta \psi^* \quad \text{weakly in } L^2(\Omega) \end{split}$$

along subsequences as $k \to \infty$. It remains to be seen whether $\sigma(\psi_k)$ converges to the solution $\sigma(\psi^*)$ of the parabolic variational inequality (2.1) in order to conclude the direct method argument. For $k \in \mathbb{N}$, consider $u_k = \sigma(\psi_k)$. Using the energy estimates in Proposition 2.4, Remark 2.5 with the established bounds on $\{\psi_k\}_{k=1}^{\infty}$ and its derivatives, and arguing as in the proof of Theorem 2.1, we deduce a limit $u^* \in \mathcal{V}$ such that

$$u_k \to u^{\star}$$
 strongly in $L^2(\mathbb{Q})$
 $u_k \to u^{\star}$ strongly in $L^2(0, T; H_0^1(\Omega))$
 $(u_k)_t \to u_t^{\star}$ weakly in $L^2(\mathbb{Q})$
 $\Delta u_k \to \Delta u^{\star}$ weakly in $L^2(\mathbb{Q})$

along subsequences as $k \to \infty$. Now consider a test function $v \in \mathcal{K}(\psi^*)$. Note that a priori $v \notin \mathcal{K}(\psi_k)$ for $k \in \mathbb{N}$, so u_k would not immediately satisfy a parabolic variational inequality for such a test function. To rectify this problem, for $k \in \mathbb{N}$, we consider $v_k := \max\{v, \psi_k\}$. Then $v_k \in \mathcal{K}(\psi_k)$ and satisfies

(9.2)
$$\int_{Q} (u_k)_t (v_k - u_k) \mathrm{d}x \mathrm{d}t + \int_{Q} \nabla u_k \cdot \nabla (v_k - u_k) \mathrm{d}x \mathrm{d}t \ge \int_{Q} f(v_k - u_k) \mathrm{d}x \mathrm{d}t$$

for all $k \in \mathbb{N}$. To pass to the limit, we need to establish some convergence of $\{v_k\}_{k=1}^{\infty}$. Recall that the distributional derivative of the piecewise smooth function $\mathbf{x} \mapsto |\mathbf{x}|$ is the function sgn : $\mathbf{x} \mapsto \frac{|\mathbf{x}|}{\mathbf{x}}$, $\mathbf{x} \neq 0$. Now using the chain rule [25, Thm.7.8, p.153] and the fact that

$$v_k = \frac{|v - \psi_k| + (v + \psi_k)}{2},$$

we deduce that

$$\|\nabla v_k\|_{L^2(\Omega)} \le \|\nabla (v - \psi_k)\|_{L^2(\Omega)}$$

for all $k \in \mathbb{N}$. Since $\{\psi_k\}_{k=1}^{\infty}$ is bounded in $L^2(0,T; H_0^1(\Omega))$, so is $\{v_k\}_{k=1}^{\infty}$, and by the Banach-Alaoglu theorem we deduce

$$v_k \rightarrow v$$
 weakly in $L^2(0,T; H_0^1(\Omega))$

along a subsequence as $k \to \infty$. The same argument may be used to show that $\{(v_k)_t\}_{k=1}^{\infty}$ is bounded in $L^2(\Omega)$, whence by the Aubin-Lions lemma we obtain

$$v_k \to v$$
 strongly in $L^2(\Omega)$

as $k \to \infty$. Now notice that

$$\int_{\Omega} (u_k)_t (v_k - u_k) \mathrm{d}x \mathrm{d}t = \int_{\Omega} (u_k)_t (v_k - v) \mathrm{d}x \mathrm{d}t + \int_{\Omega} (u_k)_t (v - u^*) \mathrm{d}x \mathrm{d}t + \int_{\Omega} (u_k)_t (u^* - u_k) \mathrm{d}x \mathrm{d}t$$

as well as

$$\int_{\Omega} \nabla u_k \cdot \nabla (v_k - u_k) \mathrm{d}x \mathrm{d}t = \int_{\Omega} \nabla (u_k - u^\star) \cdot \nabla v_k \mathrm{d}x \mathrm{d}t + \int_{\Omega} \nabla u^\star \cdot \nabla v_k \mathrm{d}x \mathrm{d}t - \int_{\Omega} |\nabla u_k|^2 \mathrm{d}x \mathrm{d}t,$$
and

$$\int_{\Omega} f(v_k - u_k) \mathrm{d}x \mathrm{d}t = \int_{\Omega} f(v_k - v) \mathrm{d}x \mathrm{d}t + \int_{\Omega} f(v - u^*) \mathrm{d}x \mathrm{d}t + \int_{\Omega} f(u^* - u_k) \mathrm{d}x \mathrm{d}t.$$

Using the strong $L^2(\Omega)$ convergence of $\{u_k\}_{k=1}^{\infty}$, $\{\nabla u_k\}_{k=1}^{\infty}$ and $\{v_k\}_{k=1}^{\infty}$, as well as the weak convergence and boundedness of $\{(u_k)_t\}_{k=1}^{\infty}$ in $L^2(\Omega)$ and of $\{v_k\}_{k=1}^{\infty}$ in $L^2(0,T; H_0^1(\Omega))$, we may pass to the limit in (9.2) to obtain

$$\int_{\Omega} u_t^{\star}(v - u^{\star}) \mathrm{d}x \mathrm{d}t + \int_{\Omega} \nabla u^{\star} \cdot \nabla (v - u^{\star}) \mathrm{d}x \mathrm{d}t \ge \int_{\Omega} f(v - u^{\star}) \mathrm{d}x \mathrm{d}t$$

for any $v \in \mathcal{K}(\psi^*)$. Since $u_k \geq \psi_k$ a.e. in Ω , the strong $L^2(\Omega)$ convergence implies a.e. convergence along a subsequence, thence $u^* \geq \psi^*$ a.e. in Ω . Hence $u^* \in \mathcal{K}(\psi^*)$ solves (2.1) and $u^* = \sigma(\psi^*)$. Now using the weak lower semicontinuity of the $L^2(\Omega)$ norm coupled with the weak convergences of $\{\Delta \psi_k\}_{k=1}^{\infty}$ and $\{(\psi_k)_t\}_{k=1}^{\infty}$ and since $\sigma(\psi_k) \to \sigma(\psi^*)$ as $k \to \infty$, we have

$$\mathcal{J}[\psi^{\star}] \leq \liminf_{k \to \infty} \mathcal{J}[\psi_k],$$

from which it follows that

$$\mathcal{J}[\psi^{\star}] \leq \inf_{\psi \in \mathcal{U}} \mathcal{J}[\psi].$$

 \square

Thus ψ^{\star} is a minimizer of \mathcal{J} .

Remark 9.2. — It is not obvious whether the functional \mathcal{J} satisfies convexity properties. We therefore cannot immediately conclude on the possible uniqueness of an optimal control.

9.2. Characterizing an optimal control. — We now look to characterize an optimal control ψ^* as well as its associated state $\sigma(\psi^*)$. We seek to differentiate the functional \mathcal{J} in some sense at the minimizer ψ^* and deduce stationary conditions in terms of Euler-Lagrange equations. However, showing that the map σ is differentiable may prove to be a difficult task, and we thence cannot immediately differentiate in the sense of Gâteaux. Rather, we will use the penalized problem (2.2) set up in Section 4. We will study the properties of the approximate solution map, introduce adjoint variables and a sequence of approximate controls. We will then use some uniform estimates and deduce conditions on $(\psi^*, \sigma(\psi^*))$ by taking the limit.

The following theorem gives us some differentiability for the approximate solution map $\sigma_{\varepsilon} : \psi \mapsto u^{\varepsilon}$, mapping an obstacle $\psi \in \mathcal{U}$ to the solution $u^{\varepsilon} \in \mathcal{V}$ of the penalized problem (2.2).

Theorem 9.3. — Let $\varepsilon > 0$ be fixed. The map σ_{ε} has a weak directional derivative in $L^2(0,T; H_0^1(\Omega))$, in the sense that given an obstacle $\psi \in \mathcal{U}$ and a direction $\mathbf{v} \in L^2(\Omega)$ such that $\psi + h\mathbf{v} \in \mathcal{U}$, there exists $\xi^{\varepsilon} \in L^2(0,T; H_0^1(\Omega))$ such that

$$\frac{\sigma_{\varepsilon}(\psi + h\mathbf{v}) - \sigma_{\varepsilon}(\psi)}{h} \rightharpoonup \xi^{\varepsilon} \quad weakly \ in \ L^{2}(0, T; H^{1}_{0}(\Omega))$$

along a subsequence as $h \to 0$. Moreover, ξ^{ε} satisfies $\xi^{\varepsilon}_t \in L^2(0,T; H^{-1}(\Omega))$ and is a weak solution of

(9.3)
$$\begin{cases} \xi_t^{\varepsilon} - \Delta \xi^{\varepsilon} + \beta_{\varepsilon}' (u^{\varepsilon} - \psi) (\xi^{\varepsilon} - \mathbf{v}) = 0 & \text{in } \Omega \\ \xi^{\varepsilon} = 0 & \text{on } \Sigma \\ \xi^{\varepsilon} (0, \cdot) = 0 & \text{in } \Omega. \end{cases}$$

Remark 9.4. — Notice that the initial condition $\xi^{\varepsilon}(0, \cdot) = 0$ in Ω would make sense in such a setup by the recurring argument [**21**, Thm.5, p.382].

Proof. — Fix $\varepsilon > 0$. Set $u^{\varepsilon,h} := \sigma_{\varepsilon}(\psi + h\mathbf{v})$, $u^{\varepsilon} := \sigma_{\varepsilon}(\psi)$, and let $t \in (0,T)$ be fixed. We consider the cylinder $\Omega_t = \Omega \times (0,t)$, on which the following identity holds a.e.

(9.4)
$$u_t^{\varepsilon,h} - u_t^{\varepsilon} - \Delta(u^{\varepsilon,h} - u^{\varepsilon}) + \beta_{\varepsilon}(u^{\varepsilon,h} - (\psi + h\mathbf{v})) - \beta_{\varepsilon}(u^{\varepsilon} - \psi) = 0.$$

The choice of this domain will allow us to delete the time derivative when integrating, as seen in the computations. Multiplying (9.4) by $(u^{\varepsilon,h} - u^{\varepsilon})$, integrating over Q_t

and using Green's first identity gives

(9.5)
$$\int_{\Omega_{t}} (u^{\varepsilon,h} - u^{\varepsilon})_{t} (u^{\varepsilon,h} - u^{\varepsilon}) \mathrm{d}x \mathrm{d}s + \int_{\Omega} |\nabla (u^{\varepsilon,h} - u^{\varepsilon})|^{2} \mathrm{d}x \mathrm{d}s$$
$$= -\frac{1}{\varepsilon} \int_{\Omega_{t}} \left[\beta (u^{\varepsilon,h} - (\psi + h\mathbf{v})) - \beta (u^{\varepsilon} - \psi) \right] (u^{\varepsilon,h} - u^{\varepsilon}) \mathrm{d}x \mathrm{d}s.$$

We invoke the following trick

(9.6)
$$\frac{\mathrm{d}}{\mathrm{d}\theta}\beta(\theta(u^{\varepsilon,h}-\psi-h\mathbf{v})+(1-\theta)(u^{\varepsilon}-\psi)) \\ = \beta'(\theta(u^{\varepsilon,h}-\psi-h\mathbf{v})+(1-\theta)(u^{\varepsilon}-\psi))(u^{\varepsilon,h}-u^{\varepsilon}-h\mathbf{v})$$

and plug it in the right-hand side integral above to obtain (9.7)

$$\begin{aligned} &-\frac{1}{\varepsilon} \int_{\Omega_t} \left(\beta(u^{\varepsilon,h} - (\psi + h\mathbf{v})) - \beta(u^{\varepsilon} - \psi) \right) (u^{\varepsilon,h} - u^{\varepsilon}) \mathrm{d}x \mathrm{d}s \\ &= -\frac{1}{\varepsilon} \int_{\Omega_t} \int_0^1 \frac{\mathrm{d}}{\mathrm{d}\theta} \beta(\theta(u^{\varepsilon,h} - \psi - h\mathbf{v}) + (1 - \theta)(u^{\varepsilon} - \psi)) (u^{\varepsilon,h} - u^{\varepsilon}) \mathrm{d}\theta \mathrm{d}x \mathrm{d}s \\ &= -\frac{1}{\varepsilon} \int_{\Omega_t} \int_0^1 \beta'(\theta(u^{\varepsilon,h} - \psi - h\mathbf{v}) + (1 - \theta)(u^{\varepsilon} - \psi)) \left((u^{\varepsilon,h} - u^{\varepsilon})^2 - h\mathbf{v}(u^{\varepsilon,h} - u^{\varepsilon}) \right) \mathrm{d}\theta \mathrm{d}x \mathrm{d}s. \end{aligned}$$

Observe that since $u^{\varepsilon,h}(0,\cdot) - u^{\varepsilon}(0,\cdot) = 0$, we have

$$\int_{\Omega_t} (u^{\varepsilon,h} - u^{\varepsilon})_t (u^{\varepsilon,h} - u^{\varepsilon})(s,x) \mathrm{d}x \mathrm{d}s = \frac{1}{2} \int_{\Omega} \left[(u^{\varepsilon,h} - u^{\varepsilon})^2 (s,x) \right]_0^t \mathrm{d}x$$
$$= \frac{1}{2} \int_{\Omega} (u^{\varepsilon,h} - u^{\varepsilon})^2 (t,x) \mathrm{d}x.$$

Combining this with (9.5) and (9.7) gives

$$\frac{1}{2} \int_{\Omega} (u^{\varepsilon,h} - u^{\varepsilon})^{2}(t,x) dx + \int_{\Omega_{t}} |\nabla(u^{\varepsilon,h} - u^{\varepsilon})|^{2} dx ds$$

$$= \frac{1}{\varepsilon} \int_{\Omega_{t}} \int_{0}^{1} \beta'(\theta(u^{\varepsilon,h} - \psi - h\mathbf{v}) + (1 - \theta)(u^{\varepsilon} - \psi)) (h\mathbf{v}(u^{\varepsilon,h} - u^{\varepsilon})) d\theta dx ds$$

$$- \frac{1}{\varepsilon} \int_{\Omega_{t}} \int_{0}^{1} \beta'(\theta(u^{\varepsilon,h} - \psi - h\mathbf{v}) + (1 - \theta)(u^{\varepsilon} - \psi))(u^{\varepsilon,h} - u^{\varepsilon})^{2} d\theta dx ds$$

$$(9.8) \leq \frac{h}{\varepsilon} \int_{\Omega_{t}} \mathbf{v}(u^{\varepsilon,h} - u^{\varepsilon}) dx ds,$$

since $0 \leq \beta'(\cdot) \leq 1$ on \mathbb{R} . We may estimate the right-hand side using the Cauchy-Schwarz and the Young inequalities respectively:

(9.9)
$$\frac{h}{\varepsilon} \int_{\mathfrak{Q}_{t}} \mathbf{v}(u^{\varepsilon,h} - u^{\varepsilon}) \mathrm{d}x \mathrm{d}s \leq \frac{h}{\varepsilon} \|\mathbf{v}\|_{L^{2}(\mathfrak{Q}_{t})} \|u^{\varepsilon,h} - u^{\varepsilon}\|_{L^{2}(\mathfrak{Q}_{t})} \leq \frac{h^{2}}{2\varepsilon^{2}} \|\mathbf{v}\|_{L^{2}(\mathfrak{Q}_{t})}^{2} + \frac{1}{2} \|u^{\varepsilon,h} - u^{\varepsilon}\|_{L^{2}(\mathfrak{Q}_{t})}^{2}.$$

Now observe that since $|\nabla(u^{\varepsilon,h} - u^{\varepsilon})|^2 \ge 0$, (9.8) and (9.9) imply

$$\int_{\Omega} (u^{\varepsilon,h} - u^{\varepsilon})(t,x) \mathrm{d}x \le \frac{h^2}{\varepsilon^2} \|\mathbf{v}\|_{L^2(\Omega)}^2 + \|u^{\varepsilon,h} - u^{\varepsilon}\|_{L^2(\Omega_t)}^2.$$

The integral form of Gronwall's inequality (see Proposition A.12 in the Appendix) then gives

$$\int_{\Omega} (u^{\varepsilon,h} - u^{\varepsilon})(t,x) \mathrm{d}x \le \frac{h^2}{\varepsilon^2} \|\mathbf{v}\|_{L^2(\Omega)}^2 (1 + Te^T),$$

and by integrating between 0 and T one finally obtains

$$||u^{\varepsilon,h} - u^{\varepsilon}||^{2}_{L^{2}(\Omega)} \le \frac{h^{2}}{\varepsilon^{2}} ||\mathbf{v}||^{2}_{L^{2}(\Omega)} (1 + Te^{T})T.$$

Plugging this in (9.9) yields

(9.10)
$$\frac{h}{\varepsilon} \int_{\mathcal{Q}_t} \mathbf{v}(u^{\varepsilon,h} - u^{\varepsilon}) \mathrm{d}x \mathrm{d}s \le \frac{h^2}{2\varepsilon^2} \|\mathbf{v}\|_{L^2(\mathcal{Q})}^2 (1 + T(1 + Te^T)).$$

Since $|\nabla(u^{\varepsilon,h} - u^{\varepsilon})|^2 \ge 0$, coming back to (9.8) we obtain by what precedes

$$\frac{1}{2} \Big(\int_{\Omega} (u^{\varepsilon,h} - u^{\varepsilon})^2 (t,x) \mathrm{d}x + \int_{\mathcal{Q}_t} |\nabla (u^{\varepsilon,h} - u^{\varepsilon})|^2 \mathrm{d}x \mathrm{d}s \Big) \le \frac{h^2}{2\varepsilon^2} \|\mathbf{v}\|_{L^2(\mathcal{Q})}^2 (1 + T(1 + Te^T)).$$

In particular, this inequality implies

$$\int_0^T \int_\Omega |\nabla (u^{\varepsilon,h} - u^{\varepsilon})|^2 \mathrm{d}x \mathrm{d}t \le \frac{h^2}{\varepsilon^2} \|\mathbf{v}\|_{L^2(\Omega)}^2 (1 + T(1 + Te^T)).$$

By rearranging the above estimate, we deduce

(9.11)
$$\left\|\frac{u^{\varepsilon,h}-u^{\varepsilon}}{h}\right\|_{L^{2}(0,T;H^{1}_{0}(\Omega))} \leq \frac{C(T)}{\varepsilon} \|\mathbf{v}\|_{L^{2}(\Omega)}.$$

Thus $\left\{\frac{u^{\varepsilon,h}-u^{\varepsilon}}{h}\right\}_{h>0}$ is bounded in $L^2(0,T;H_0^1(\Omega))$. By the Banach-Alaoglu theorem, there exists $\xi^{\varepsilon} \in L^2(0,T;H_0^1(\Omega))$ such that

$$\frac{u^{\varepsilon,h}-u^{\varepsilon}}{h} \rightharpoonup \xi^{\varepsilon} \quad \text{ weakly in } L^2(0,T;H^1_0(\Omega))$$

along a subsequence as $h \to 0$. By the weak lower semicontinuity of the norm (see the Appendix), we also obtain

$$\|\xi^{\varepsilon}\|_{L^{2}(0,T;H^{1}_{0}(\Omega))} \leq \frac{C(T)}{\varepsilon} \|\mathbf{v}\|_{L^{2}(\Omega)},$$

which can be rewritten as

(9.12)
$$\|\nabla\xi^{\varepsilon}\|_{L^{2}(\Omega)} \leq \frac{C(T)}{\varepsilon} \|\mathbf{v}\|_{L^{2}(\Omega)}.$$

We now look to show that $\xi_t^{\varepsilon} \in L^2(0,T; H^{-1}(\Omega))$. To this end, we will estimate the quantity

$$\left\|\left(\frac{u^{\varepsilon,h}-u^{\varepsilon}}{h}\right)_t\right\|_{L^2(0,T;H^{-1}(\Omega))},$$

and conclude by letting $h \to 0$. Using the identity (9.4), we observe that

$$(9.13) \qquad \left\| \left(\frac{u^{\varepsilon,h} - u^{\varepsilon}}{h} \right)_t \right\|_{L^2(0,T;H^{-1}(\Omega))} \leq \left\| \Delta \left(\frac{u^{\varepsilon,h} - u^{\varepsilon}}{h} \right) \right\|_{L^2(0,T;H^{-1}(\Omega))} \\ + \left\| \frac{\beta_{\varepsilon} (u^{\varepsilon,h} - \psi - h\mathbf{v}) - \beta_{\varepsilon} (u^{\varepsilon} - \psi)}{h} \right\|_{L^2(\Omega)}.$$

We now estimate the two norms on the right-hand side. For the first norm, we conclude by (9.11):

$$\left\|\Delta\left(\frac{u^{\varepsilon,h}-u^{\varepsilon}}{h}\right)\right\|_{L^{2}(0,T;H^{-1}(\Omega))} \leq \left\|\nabla\left(\frac{u^{\varepsilon,h}-u^{\varepsilon}}{h}\right)\right\|_{L^{2}(\Omega)} \leq \frac{C(T)}{\varepsilon} \|\mathbf{v}\|_{L^{2}(\Omega)}.$$

On the other hand, by using the trick (9.6) and the fact that $0 \leq \beta'(\cdot) \leq 1$ on \mathbb{R} , we have

$$\begin{split} &\int_{\Omega} \Big(\frac{\beta_{\varepsilon}(u^{\varepsilon,h} - \psi - h\mathbf{v}) - \beta_{\varepsilon}(u^{\varepsilon} - \psi)}{h} \Big)^{2} \mathrm{d}x \mathrm{d}t \\ &= \frac{1}{h^{2}} \int_{\Omega} \Big(\int_{0}^{1} \beta_{\varepsilon}'(\theta(u^{\varepsilon,h} - \psi - h\mathbf{v}) + (1 - \theta)(u^{\varepsilon} - \psi))(u^{\varepsilon,h} - u^{\varepsilon} - h\mathbf{v}) \mathrm{d}\theta \Big)^{2} \mathrm{d}x \mathrm{d}t \\ &\leq \frac{1}{\varepsilon^{2}} \int_{\Omega} \Big(\frac{u^{\varepsilon,h} - u^{\varepsilon} - h\mathbf{v}}{h} \Big)^{2} \mathrm{d}x \mathrm{d}t \\ &\leq \frac{2}{\varepsilon^{2}} \int_{\Omega} \Big[\Big(\frac{u^{\varepsilon,h} - u^{\varepsilon}}{h} \Big)^{2} + \mathbf{v}^{2} \Big] \mathrm{d}x \mathrm{d}t. \end{split}$$

From the above and (9.11), it follows that

$$\int_{\Omega} \left(\frac{\beta_{\varepsilon}(u^{\varepsilon,h} - \psi - h\mathbf{v}) - \beta_{\varepsilon}(u^{\varepsilon} - \psi)}{h} \right)^2 \mathrm{d}x \mathrm{d}t \le \frac{C}{\varepsilon^2} \|\mathbf{v}\|_{L^2(\Omega)}^2$$

holds for some C = C(T) > 0. Therefore

$$\left\| \left(\frac{u^{\varepsilon,h} - u^{\varepsilon}}{h} \right)_t \right\|_{L^2(0,T;H^{-1}(\Omega))} \le \frac{C}{\varepsilon} \| \mathbf{v} \|_{L^2(\Omega)}$$

for some C = C(T) > 0. Arguing by compactness just as with the previous convergence, we deduce that $\xi_t^{\varepsilon} \in L^2(0,T; H^{-1}(\Omega))$ and

$$\left(\frac{u^{\varepsilon,h}-u^{\varepsilon}}{h}\right)_t \rightharpoonup \xi_t^{\varepsilon} \quad \text{weakly in } L^2(0,T;H^{-1}(\Omega))$$

along a subsequence as $h \to 0$. Using both of the established weak convergences and the Aubin-Lions lemma, we finally deduce that

(9.14)
$$\frac{u^{\varepsilon,h} - u^{\varepsilon}}{h} \to \xi^{\varepsilon} \quad \text{strongly in } L^2(\mathfrak{Q})$$

along a subsequence as $h \to 0$.

To deduce the equation, we multiply the identity (9.4) by an arbitrary test function $\varphi \in L^2(0,T; H^1_0(\Omega))$ and integrate over Ω to obtain

(9.15)
$$\int_{\Omega} \left(\frac{u^{\varepsilon,h} - u^{\varepsilon}}{h}\right)_{t} \varphi dx dt + \int_{\Omega} \nabla \left(\frac{u^{\varepsilon,h} - u^{\varepsilon}}{h}\right) \cdot \nabla \varphi dx dt + \int_{\Omega} \left(\frac{\beta_{\varepsilon}(u^{\varepsilon,h} - (\psi + h\mathbf{v})) - \beta_{\varepsilon}(u^{\varepsilon} - \psi)}{h}\right) \varphi dx dt = 0.$$

Observe that in order to pass to the limit in (9.15), we only need to investigate the third integral term, as the established weak convergences would account for the first two. It is now easier argue with the original notation $u^{\varepsilon,h} = \sigma_{\varepsilon}(\psi + h\mathbf{v}), u^{\varepsilon} = \sigma_{\varepsilon}(\psi)$. The strong L^2 convergence (9.14) implies that

(9.16)
$$\frac{\sigma_{\varepsilon}(\psi + h\mathbf{v}) - \sigma_{\varepsilon}(\psi)}{h} \to \xi^{\varepsilon} \quad \text{a.e. in } \mathcal{Q}$$

along a further subsequence as $h \to 0$. If we set $\zeta(x) := \sigma_{\varepsilon}(x) - x$, we deduce the following identity:

$$\frac{\beta_{\varepsilon}(u^{\varepsilon,h} - (\psi + h\mathbf{v})) - \beta_{\varepsilon}(u^{\varepsilon} - \psi)}{h} = \frac{\beta_{\varepsilon}(\zeta(\psi + h\mathbf{v})) - \beta_{\varepsilon}(\zeta(\psi))}{h}.$$

Now by virtue of (9.16), notice that

$$\frac{\zeta(\psi + h\mathbf{v}) - \zeta(\psi)}{h} \to \xi^{\varepsilon} - \mathbf{v} \quad \text{a.e. in } \mathcal{Q}$$

as $h \to 0$. Consequently, since $\beta_{\varepsilon} \in C^1(\mathbb{R})$, using the previous convergence and the chain rule, we obtain

$$\frac{\beta_{\varepsilon}(\zeta(\psi + h\mathbf{v})) - \beta_{\varepsilon}(\zeta(\psi))}{h} \to \beta_{\varepsilon}'(\zeta(\psi))(\xi^{\varepsilon} - \mathbf{v}) \quad \text{ a.e. in } \Omega$$

as $h \to 0$, or equivalently,

$$\frac{\beta_{\varepsilon}(u^{\varepsilon,h} - (\psi + h\mathbf{v})) - \beta_{\varepsilon}(u^{\varepsilon} - \psi)}{h} \to \beta_{\varepsilon}'(u^{\varepsilon} - \psi)(\xi^{\varepsilon} - \mathbf{v}) \quad \text{ a.e. in } \mathcal{Q}.$$

We may now let $h \to 0$ in (9.15) by using both of the established weak convergences for the first two integrals and the Lebesgue dominated convergence theorem for the third to deduce that

$$\int_0^T \langle \xi_t^{\varepsilon}(t,\cdot), \varphi(t,\cdot) \rangle dt + \int_{\Omega} \nabla \xi^{\varepsilon} \cdot \nabla \varphi dx dt + \int_{\Omega} \beta_{\varepsilon}'(u^{\varepsilon} - \psi)(\xi^{\varepsilon} - \mathbf{v})\varphi dx dt = 0$$

holds for all $\varphi \in L^2(0,T; H^1_0(\Omega))$. Whence ξ^{ε} is a weak solution of (9.3).

We now consider the set

$$\mathcal{W} := \{ w \in H^1(\Omega) \colon w = 0 \text{ on } \Sigma, w(0, \cdot) = 0 \text{ in } \Omega \}$$

endowed with the $\|\cdot\|_{H^1(\Omega)}$ norm, and let \mathcal{W}' be its dual. Recall that the Sobolev space $H^1(\Omega)$ consists of all functions $u \in L^2(0,T; H^1(\Omega))$ such that $u_t \in L^2(0,T; H^1(\Omega))$, and is endowed with the norm

$$||u||_{H^1(\Omega)}^2 := ||u||_{L^2(\Omega)}^2 + ||u_t||_{L^2(\Omega)}^2$$

It is also worth noting that the initial value installed in \mathcal{W} makes sense since $u \in H^1(\Omega)$ implies $u \in C^0([0,T]; L^2(\Omega))$ (see [21, Thm.1, p.303]).

As discussed in what precedes, we will look to differentiate the functional associated to the approximate problem. Fix $\varepsilon > 0$ and consider:

$$\mathcal{J}_{\varepsilon}[\psi] := \int_{\Omega} (\sigma_{\varepsilon}(\psi) - u_d)^2 \mathrm{d}x \mathrm{d}t + \int_{\Omega} \left(|\Delta \psi|^2 + |\psi_t|^2 \right) \mathrm{d}x \mathrm{d}t$$

for $\psi \in \mathcal{U}$. The existence of a minimizer to this functional follows from arguments similar to the proof of Theorem 9.1. As it shall be seen in the computations that follow, it will be useful to introduce an adjoint variable. We have the following Lemma.

Lemma 9.5. — Let $\varepsilon > 0$ be fixed. Given a minimizer $\psi^{\varepsilon} \in \mathcal{U}$ of $\mathcal{J}_{\varepsilon}$, there exists a unique adjoint state $p^{\varepsilon} \in L^2(0,T; H_0^1(\Omega))$ with $p_t^{\varepsilon} \in L^2(0,T; H^{-1}(\Omega))$ such that p^{ε} is a weak solution of

(9.17)
$$\begin{cases} -p_t^{\varepsilon} - \Delta p^{\varepsilon} + \beta_{\varepsilon}' (u^{\varepsilon} - \psi^{\varepsilon}) p^{\varepsilon} = u^{\varepsilon} - u_d & \text{in } \Omega\\ p^{\varepsilon} = 0 & \text{on } \Sigma\\ p^{\varepsilon}(T, \cdot) = 0 & \text{in } \Omega, \end{cases}$$

and the estimate

$$(9.18) \quad \sup_{t \in [0,T]} \|p^{\varepsilon}(t,\cdot)\|_{L^{2}(\Omega)} + \|p^{\varepsilon}\|_{L^{2}(0,T;H^{1}_{0}(\Omega))} + \|\beta_{\varepsilon}'(u^{\varepsilon}-\psi^{\varepsilon})p^{\varepsilon}\|_{\mathcal{W}'} + \|p^{\varepsilon}_{t}\|_{\mathcal{W}'} \le C$$

holds for some constant $C = C(T, u_d) > 0$ independent of ε .

Proof. — Fix $\varepsilon > 0$. As the adjoint equation (9.17) is linear, $\beta'_{\varepsilon}(u^{\varepsilon} - \psi^{\varepsilon}) \in L^{\infty}(\Omega)$ and $u^{\varepsilon} - u_d \in L^2(\Omega)$, by virtue of the change of variable $t \leftarrow T - t$ and [21, Thm.3, p.378] we deduce the existence of a unique weak solution $p^{\varepsilon} \in L^2(0, T; H^1_0(\Omega))$ with $p_t^{\varepsilon} \in L^2(0, T; H^{-1}(\Omega))$.

We now look to show the uniform estimate. Recall that in Remark 2.5, we showed that the family of approximate solutions $\{u^{\varepsilon}\}_{\varepsilon>0}$ is bounded in $L^2(\Omega)$. Since $u_d \in L^2(\Omega)$, there exists some constant C > 0 such that the following estimate of the right-hand side of (9.17)

$$(9.19) ||u^{\varepsilon} - u_d||_{L^2(\mathfrak{Q})} \le C$$

holds for all $\varepsilon > 0$. Consider for fixed $t \in [0, T)$ the cylinder $\Omega^t := \Omega \times (t, T)$. The weak form of the adjoint equation for the test function p^{ε} reads

$$\int_{\mathcal{Q}^t} \left(-p_t^{\varepsilon} p^{\varepsilon} + |\nabla p^{\varepsilon}|^2 + \beta_{\varepsilon}' (u^{\varepsilon} - \psi) (\psi^{\varepsilon})^2 \right) \mathrm{d}x \mathrm{d}s = \int_{\mathcal{Q}^t} p^{\varepsilon} (u^{\varepsilon} - u_d) \mathrm{d}x \mathrm{d}s$$

Since $\beta'(\cdot) \geq 0$ on \mathbb{R} , we also have

$$\int_{\mathbb{Q}^t} \beta_{\varepsilon}'(u^{\varepsilon} - \psi)(p^{\varepsilon})^2 \mathrm{d}x \mathrm{d}s \ge 0,$$

thus

$$\int_{\Omega^t} (-p_t^{\varepsilon} p^{\varepsilon} + |\nabla p^{\varepsilon}|^2) \mathrm{d}x \mathrm{d}s \le \int_{\Omega^t} p^{\varepsilon} (u^{\varepsilon} - u_d) \mathrm{d}x \mathrm{d}s.$$

Similarly as in the previous proof, observe that

(9.20)
$$-\int_{\Omega^t} (p_t^{\varepsilon} p^{\varepsilon}) \mathrm{d}x \mathrm{d}s = -\frac{1}{2} \int_{\Omega} \left[(p^{\varepsilon})^2 \right]_t^T \mathrm{d}x = \frac{1}{2} \int_{\Omega} p^{\varepsilon} (t, x)^2 \mathrm{d}x$$

since $p^{\varepsilon}(T, \cdot) = 0$ in Ω . Whence,

(9.21)
$$\int_{\Omega} p^{\varepsilon}(t,x)^{2} \mathrm{d}x + \int_{\Omega^{t}} |\nabla p^{\varepsilon}|^{2} \mathrm{d}x \mathrm{d}s \leq 2 \int_{\Omega^{t}} p^{\varepsilon}(u^{\varepsilon} - u_{d}) \mathrm{d}x \mathrm{d}s.$$

We may estimate the right-hand side by using the Cauchy-Schwarz and Young inequalities respectively:

$$2\int_{\mathbb{Q}^t} p^{\varepsilon} (u^{\varepsilon} - u_d) \mathrm{d}x \mathrm{d}s \leq 2 \|p^{\varepsilon}\|_{L^2(\mathbb{Q}^t)} \|u^{\varepsilon} - u_d\|_{L^2(\mathbb{Q}^t)}$$
$$\leq \|p^{\varepsilon}\|_{L^2(\mathbb{Q}^t)}^2 + \|u^{\varepsilon} - u_d\|_{L^2(\mathbb{Q}^t)}^2.$$

As $|\nabla p^{\varepsilon}|^2 \ge 0$, the above and (9.21) yield

$$\int_{\Omega} p^{\varepsilon}(t,x)^2 \mathrm{d}x \le \|p^{\varepsilon}\|_{L^2(\Omega^t)}^2 + \|u^{\varepsilon} - u_d\|_{L^2(\Omega^t)}^2.$$

Now using the integral form of Gronwall's inequality (see Proposition A.12 in the Appendix) and integrating between 0 and T, we obtain

$$\int_{\Omega} p^{\varepsilon} (u^{\varepsilon} - u_d) \mathrm{d}x \mathrm{d}s \le C(T) \| u^{\varepsilon} - u_d \|_{L^2(\Omega)}^2.$$

Plugging this in (9.21) gives

$$\int_{\Omega} p^{\varepsilon}(t,x)^2 \mathrm{d}x + \int_{\Omega} |\nabla p^{\varepsilon}|^2 \mathrm{d}x \mathrm{d}t \le C(T) \|u^{\varepsilon} - u_d\|_{L^2(\Omega)}^2,$$

which allows us to conclude that

(9.22)
$$\sup_{t \in [0,T]} \|p^{\varepsilon}(t, \cdot)\|_{L^{2}(\Omega)} + \|\nabla p^{\varepsilon}\|_{L^{2}(\Omega)} \le C \|u^{\varepsilon} - u_{d}\|_{L^{2}(\Omega)}$$

holds for some C = C(T) > 0.

Now let $\varphi \in \mathcal{W}$ be arbitrary. From the weak form of the adjoint equation, we see that

$$\begin{split} \left| \int_{\Omega} \beta_{\varepsilon}'(u^{\varepsilon} - \psi^{\varepsilon}) p^{\varepsilon} \varphi \mathrm{d}x \mathrm{d}t \right| &= \left| \int_{\Omega} \left((u^{\varepsilon} - u_d) \varphi - p^{\varepsilon} \varphi_t + \nabla p^{\varepsilon} \cdot \nabla \varphi \right) \mathrm{d}x \mathrm{d}t \right| \\ &\leq \|u^{\varepsilon} - u_d\|_{L^2(\Omega)} \|\varphi\|_{L^2(\Omega)} + \|p^{\varepsilon}\|_{L^2(\Omega)} \|\varphi_t\|_{L^2(\Omega)} \\ &+ \|\nabla p^{\varepsilon}\|_{L^2(\Omega)} \|\nabla \varphi\|_{L^2(\Omega)}. \end{split}$$

Using the Poincaré inequality and (9.22), we obtain

$$\int_{\Omega} \beta_{\varepsilon}'(u^{\varepsilon} - \psi^{\varepsilon}) p^{\varepsilon} \varphi \mathrm{d}x \mathrm{d}t \Big| \le C(T) \|u^{\varepsilon} - u_d\|_{L^2(\Omega)} \|\varphi\|_{W^{1/2}}$$

Consequently, one has

$$\|\beta_{\varepsilon}'(u^{\varepsilon}-\psi^{\varepsilon})p^{\varepsilon}\|_{\mathcal{W}'} \le C(T)\|u^{\varepsilon}-u_d\|_{L^2(\Omega)}.$$

Similarly, from the weak form of the adjoint equation for the test function $\varphi \in \mathcal{W}$, the estimate on the gradient (9.22) and the \mathcal{W}' estimate above, we deduce

$$\|p_t^{\varepsilon}\|_{\mathcal{W}'} \le C(T) \|u^{\varepsilon} - u_d\|_{L^2(\Omega)}.$$

Combining all of these estimates, we conclude by virtue of (9.19) that

$$\sup_{t\in[0,T]} \|p^{\varepsilon}(t,\cdot)\|_{L^{2}(\Omega)} + \|\nabla p^{\varepsilon}\|_{L^{2}(\Omega)} + \|\beta_{\varepsilon}'(u^{\varepsilon}-\psi^{\varepsilon})p^{\varepsilon}\|_{\mathcal{W}'} + \|p_{t}^{\varepsilon}\|_{\mathcal{W}'} \le C$$

for some C = C(T) > 0 independent of ε .

We now define the notion of solution of an equation that a minimizer ψ^{ε} of $\mathcal{J}_{\varepsilon}$ will be shown to satisfy.

Definition 9.6. — Fix $\varepsilon > 0$. Given $p^{\varepsilon} \in L^2(0,T; H_0^1(\Omega))$, the function $\psi^{\varepsilon} \in U$ is a weak solution of

(9.23)
$$\begin{cases} -\psi_{tt}^{\varepsilon} + \Delta^2 \psi^{\varepsilon} + \beta_{\varepsilon}' (u^{\varepsilon} - \psi^{\varepsilon}) p^{\varepsilon} = 0 & \text{in } \Omega \\ \psi^{\varepsilon} = \Delta \psi^{\varepsilon} = 0 & \text{on } \Sigma \\ \psi^{\varepsilon}(0, \cdot) = \psi_t^{\varepsilon}(T, \cdot) = 0 & \text{in } \Omega, \end{cases}$$

 $\textit{if } \Delta\psi^{\varepsilon} \in L^2(0,T;H^1_0(\Omega)), \ \psi^{\varepsilon}_t \in L^2(\mathbb{Q}), \ \psi^{\varepsilon}_{tt} \in \mathcal{W}' \ \textit{and}$

$$\int_{\Omega} \left(\psi_t^{\varepsilon} \varphi_t - \nabla \Delta \psi^{\varepsilon} \cdot \nabla \varphi + \beta_{\varepsilon}' (u^{\varepsilon} - \psi^{\varepsilon}) p^{\varepsilon} \varphi \right) \mathrm{d}x \mathrm{d}t = 0$$

holds for all $\varphi \in \mathcal{W}$.

Proposition 9.1. — Let $\varepsilon > 0$ be fixed and let $\psi^{\varepsilon} \in \mathcal{U}$ be a minimizer of $\mathcal{J}_{\varepsilon}$. Then ψ^{ε} is a weak solution of equation (9.23) in the stated sense, with associated state $u^{\varepsilon} = \sigma_{\varepsilon}(\psi^{\varepsilon})$, and the estimate and

(9.24)
$$\|\psi_t^{\varepsilon}\|_{L^2(0,T;H_0^1(\Omega))} + \|\Delta\psi^{\varepsilon}\|_{L^2(0,T;H_0^1(\Omega))} + \|\psi_{tt}^{\varepsilon}\|_{W'} \le C$$

holds for some constant $C = C(T, u_d) > 0$ independent of ε .

Remark 9.7. — By the Poincaré inequality, the estimates from the previous lemma and this proposition would also imply that $\{p^{\varepsilon}\}_{\varepsilon>0}$, $\{\Delta\psi^{\varepsilon}\}_{\varepsilon>0}$ and $\{\psi^{\varepsilon}_{t}\}_{\varepsilon>0}$ are bounded in $L^{2}(\Omega)$.

Proof. — Fix $\varepsilon > 0$ and set $u^{\varepsilon} := \sigma_{\varepsilon}(\psi^{\varepsilon})$ and $u^{\varepsilon,h} := \sigma_{\varepsilon}(\psi^{\varepsilon} + h\mathbf{v})$ for an arbitrary direction $\mathbf{v} \in C_c^{\infty}(\Omega)$ with $\mathbf{v}(0, \cdot) = 0$ and h > 0. Since ψ^{ε} is a minimizer of $\mathcal{J}_{\varepsilon}$, one has

$$\mathcal{J}_{\varepsilon}[\psi^{\varepsilon}] \leq \mathcal{J}_{\varepsilon}[\psi^{\varepsilon} + h\mathbf{v}],$$

which in turn implies

$$0 \le \liminf_{h \to 0^+} \frac{\mathcal{J}_{\varepsilon}[\psi^{\varepsilon} + h\mathbf{v}] - \mathcal{J}_{\varepsilon}[\psi^{\varepsilon}]}{h}$$

After some computations, using the weak and strong convergences of the difference quotients established in Theorem 9.3 as well as Green's first identity, we deduce

$$\liminf_{h \to 0^+} \frac{\mathcal{J}_{\varepsilon}[\psi^{\varepsilon} + h\mathbf{v}] - \mathcal{J}_{\varepsilon}[\psi^{\varepsilon}]}{h} = 2 \int_{\Omega} \left(\xi^{\varepsilon} (u^{\varepsilon} - u_d) + \Delta \psi^{\varepsilon} \Delta \mathbf{v} + \psi_t^{\varepsilon} \mathbf{v}_t \right) \mathrm{d}x \mathrm{d}t.$$

Using the adjoint equation (9.17) and Green's first identity once more, we also have

$$\begin{split} \int_{\Omega} \left(\xi^{\varepsilon} (u^{\varepsilon} - u_d) + \Delta \psi^{\varepsilon} \Delta \mathbf{v} + \psi^{\varepsilon}_t \mathbf{v}_t \right) \mathrm{d}x \mathrm{d}t \\ &= \int_{\Omega} \left(-\xi^{\varepsilon} p_t^{\varepsilon} + \nabla \xi^{\varepsilon} \cdot \nabla p^{\varepsilon} + \xi^{\varepsilon} \beta_{\varepsilon}' (u^{\varepsilon} - \psi^{\varepsilon}) p^{\varepsilon} + \Delta \psi^{\varepsilon} \Delta \mathbf{v} + \psi^{\varepsilon}_t \mathbf{v}_t \right) \mathrm{d}x \mathrm{d}t \\ &= \int_{\Omega} \left(\xi^{\varepsilon}_t p^{\varepsilon} + \nabla \xi^{\varepsilon} \cdot \nabla p^{\varepsilon} + \xi^{\varepsilon} \beta_{\varepsilon}' (u^{\varepsilon} - \psi^{\varepsilon}) p^{\varepsilon} + \Delta \psi^{\varepsilon} \Delta \mathbf{v} + \psi^{\varepsilon}_t \mathbf{v}_t \right) \mathrm{d}x \mathrm{d}t, \end{split}$$

and note that to obtain the last integral we used the initial conditions $\xi^{\varepsilon}(0, \cdot) = 0$ and $p^{\varepsilon}(T, \cdot) = 0$. Finally, using equation (9.3), we deduce

$$\int_{\Omega} \left(\xi_t^{\varepsilon} p^{\varepsilon} + \nabla \xi^{\varepsilon} \cdot \nabla p^{\varepsilon} + \xi^{\varepsilon} \beta_{\varepsilon}' (u^{\varepsilon} - \psi^{\varepsilon}) p^{\varepsilon} + \Delta \psi^{\varepsilon} \Delta \mathbf{v} + \psi_t^{\varepsilon} \mathbf{v}_t \right) \mathrm{d}x \mathrm{d}t$$
$$= \int_{\Omega} \left(\beta_{\varepsilon}' (u^{\varepsilon} - \psi^{\varepsilon}) \mathbf{v} p^{\varepsilon} + \Delta \psi^{\varepsilon} \Delta \mathbf{v} + \psi_t^{\varepsilon} \mathbf{v}_t \right) \mathrm{d}x \mathrm{d}t.$$

Whence,

$$0 \leq \int_{\Omega} \left(\beta_{\varepsilon}' (u^{\varepsilon} - \psi^{\varepsilon}) \mathbf{v} p^{\varepsilon} + \Delta \psi^{\varepsilon} \Delta \mathbf{v} + \psi_t^{\varepsilon} \mathbf{v}_t \right) \mathrm{d}x \mathrm{d}t.$$

Since **v** was taken arbitrary we deduce the weak form of (9.23). We conclude that ψ^{ε} satisfies (9.23) with the stated boundary conditions (see [2] for more detail). Now, multiplying (9.23) by $\Delta \psi^{\varepsilon}$, integrating over Ω and using Green's first identity, we have

$$\int_{\mathcal{Q}} \left(\psi_{tt}^{\varepsilon} \Delta \psi^{\varepsilon} + |\nabla \Delta \psi^{\varepsilon}|^2 \right) \mathrm{d}x \mathrm{d}t = \int_{\mathcal{Q}} \beta_{\varepsilon}' (u^{\varepsilon} - \psi^{\varepsilon}) p^{\varepsilon} \Delta \psi^{\varepsilon} \mathrm{d}x \mathrm{d}t.$$

Using Green's first identity once more as well as the adjoint equation (9.17), the above may be rewritten as

$$\int_{\mathbb{Q}} \left(|\nabla \psi_t^{\varepsilon}|^2 + |\nabla \Delta \psi^{\varepsilon}|^2 \right) \mathrm{d}x \mathrm{d}t = \int_{\mathbb{Q}} (u^{\varepsilon} - u_d + p_t^{\varepsilon} + \Delta p^{\varepsilon}) \Delta \psi^{\varepsilon} \mathrm{d}x \mathrm{d}t.$$

From the \mathcal{W}' estimates above and the Cauchy-Schwarz inequality, we deduce the $L^2(0,T; H^1_0(\Omega))$ estimates for ψ^{ε}_t and $\Delta \psi^{\varepsilon}$. Using (9.23) we deduce the \mathcal{W}' estimate for ψ^{ε}_{tt} (see [2] for more detail on this point).

Now observe that the approximate adjoint equation and the equation for the minimizer of the approximate objective functional have a term in common. By rewriting, we deduce that $(\psi^{\varepsilon}, p^{\varepsilon})$ satisfy

$$\begin{cases} -p_t^{\varepsilon} - \Delta p^{\varepsilon} + \psi_{tt}^{\varepsilon} - \Delta^2 \psi^{\varepsilon} = u^{\varepsilon} - u_d & \text{in } \Omega\\ p^{\varepsilon} = \psi^{\varepsilon} = \Delta \psi^{\varepsilon} = 0 & \text{on } \Sigma\\ p^{\varepsilon}(T, \cdot) = \psi^{\varepsilon}(0, \cdot) = \psi_t^{\varepsilon}(T, \cdot) = 0 & \text{in } \Omega, \end{cases}$$

in the weak \mathcal{W}' sense defined above. We now look to use the established estimates for a compactness argument and let $\varepsilon \to 0$ in the above equation. First, we define the notion of solution for the equation that the limit functions ψ^* and p will be shown to satisfy.

Definition 9.8. — The pair (ψ^*, p) is a weak solution of

(9.25)
$$\begin{cases} -p_t - \Delta p + \psi_{tt}^{\star} - \Delta^2 \psi^{\star} = u^{\star} - u_d & \text{in } \Omega \\ p = \psi = \Delta \psi^{\star} = 0 & \text{on } \Sigma \\ p(T, \cdot) = \psi^{\star}(0, \cdot) = \psi_t^{\star}(T, \cdot) = 0 & \text{in } \Omega, \end{cases}$$

where $u^{\star} = \sigma(\psi^{\star})$, if $\psi^{\star} \in \mathcal{U}$, $\Delta \psi^{\star} \in L^{2}(0,T; H^{1}_{0}(\Omega))$, $\psi^{\star}_{t} \in L^{2}(0,T; H^{1}_{0}(\Omega))$, $\psi^{\star}_{tt} \in \mathcal{W}'$, $p \in L^{2}(0,T; H^{1}_{0}(\Omega))$, $p_{t} \in \mathcal{W}'$ and $\int_{\Omega} \left(p\varphi_{t} + \nabla p \cdot \nabla \varphi + \nabla \Delta \psi^{\star} \cdot \nabla \varphi - \psi^{\star}_{t} \varphi_{t} \right) \mathrm{d}x \mathrm{d}t = \int_{\Omega} (u^{\star} - u_{d}) \varphi \mathrm{d}x \mathrm{d}t$

holds for all $\varphi \in \mathcal{W}$.

Theorem 9.9. — There exist an optimal control ψ^* (i.e. a minimizer of \mathfrak{J}) and an adjoint state p such that the pair (ψ^*, p) is a weak solution of (9.25). Moreover, the corresponding state $u^* = \sigma(\psi^*)$ satisfies the parabolic obstacle problem (2.1).

Proof. — Consider the family of minimizers $\{\psi^{\varepsilon}\}_{\varepsilon>0} \subset \mathcal{U}$ of the approximate objective functionals, as well as the family of solutions $\{p^{\varepsilon}\}_{\varepsilon>0} \subset L^2(0,T; H^1_0(\Omega))$ of the family of adjoint problems (9.17). In light of what precedes, for fixed $\varepsilon > 0$ these approximations satisfy the weak formulation

(9.26)
$$\int_{\Omega} \left(p^{\varepsilon} \varphi_t + \nabla p^{\varepsilon} \cdot \nabla \varphi + \nabla \Delta \psi^{\varepsilon} \cdot \nabla \varphi - \psi_t^{\varepsilon} \varphi_t \right) \mathrm{d}x \mathrm{d}t = \int_{\Omega} (u^{\varepsilon} - u_d) \varphi \mathrm{d}x \mathrm{d}t$$

for all $\varphi \in \mathcal{W}$. Arguing as in previous proofs, from the estimates in Proposition 9.1 and the lemma before, we deduce the existence of $\psi^* \in \mathcal{U}$ with $\Delta \psi^* \in L^2(0,T; H^1_0(\Omega)), \ \psi^*_t \in L^2(0,T; H^1_0(\Omega)), \ \psi^*_{tt} \in \mathcal{W}'$, and $p \in L^2(0,T; H^1_0(\Omega))$ with $p_t \in \mathcal{W}'$, and with the following convergences in particular

$$\begin{array}{ll} p^{\varepsilon} \rightharpoonup p & \text{weakly in } L^{2}(\Omega) \\ p^{\varepsilon} \rightharpoonup p & \text{weakly in } L^{2}(0,T;H^{1}_{0}(\Omega)) \\ \psi^{\varepsilon}_{t} \rightharpoonup \psi^{\star}_{t} & \text{weakly in } L^{2}(\Omega) \\ \Delta \psi^{\varepsilon} \rightharpoonup \Delta \psi^{\star} & \text{weakly in } L^{2}(0,T;H^{1}_{0}(\Omega)) \end{array}$$

along subsequences as $\varepsilon \to 0$. Using also previously established convergences of the approximate solutions $\{u^{\varepsilon}\}_{\varepsilon>0}$ (see the proof of Theorem 2.1), we may pass to the limit in (9.26) to deduce that (ψ^{\star}, p) satisfy (9.25) in the defined weak sense.

To show that u^* solves the parabolic obstacle problem, we argue as in Theorem 9.1. Let $v \in \mathcal{K}(\psi^*)$ be an arbitrary test function. Since a priori $v \notin \mathcal{K}(\psi^{\varepsilon})$, we cannot say that u^{ε} satisfies a parabolic variational inequality for such v and for each $\varepsilon > 0$. Rather, fix $\varepsilon > 0$ and set $v^{\varepsilon} = \max\{v, \psi^{\varepsilon}\} \in \mathcal{K}(\psi^{\varepsilon})$. Then since u^{ε} solves the penalization problem (2.2), one has (as in the proof of Theorem 9.1)

$$\int_{\Omega} u_t^{\varepsilon} (v^{\varepsilon} - u^{\varepsilon}) \mathrm{d}x \mathrm{d}t + \int_{\Omega} \nabla u^{\varepsilon} \cdot \nabla (v^{\varepsilon} - u^{\varepsilon}) \mathrm{d}x \mathrm{d}t \ge \int_{\Omega} f(v^{\varepsilon} - u^{\varepsilon}) \mathrm{d}x \mathrm{d}t,$$

and in Theorem 9.1, we also showed that

 $\begin{array}{ll} v^{\varepsilon} \rightharpoonup v & \mbox{ weakly in } L^2(0,T;H^1_0(\Omega)) \\ v^{\varepsilon} \rightarrow v & \mbox{ strongly in } L^2(\Omega) \end{array}$

along subsequences as $\varepsilon \to 0$. We may use these convergences as well as the weak $L^2(\mathfrak{Q})$ convergence of $\{u_t^{\varepsilon}\}_{\varepsilon>0}$ and strong $L^2(\mathfrak{Q})$ convergence of $\{\nabla u^{\varepsilon}\}_{\varepsilon>0}$ (see the proof of Theorem 2.1) to let $\varepsilon \to 0$ in the above inequality and deduce

$$\int_{\Omega} u_t^{\star}(v - u^{\star}) \mathrm{d}x \mathrm{d}t + \int_{\Omega} \nabla u^{\star} \cdot \nabla (v - u^{\star}) \mathrm{d}x \mathrm{d}t \ge \int_{\Omega} f(v - u^{\star}) \mathrm{d}x \mathrm{d}t$$

Recall the estimate (2.7):

$$\|\beta(u^{\varepsilon}-\psi^{\star})\|_{L^{2}(\Omega)} \leq \varepsilon C(\psi^{\star},f).$$

Using the properties of β and the strong $L^2(\mathbb{Q})$ convergence of $\{u^{\varepsilon}\}_{\varepsilon>0}$ as in the proof of Theorem 9.1, we may let $\varepsilon \to 0$ and deduce that $u^* \in \mathcal{K}(\psi^*)$, and we may therefore conclude that $u^* = \sigma(\psi^*)$, i.e. u^* solves the parabolic obstacle problem (2.1).

Finally, we show that ψ^* is a minimizer of \mathcal{J} . Since for each $\varepsilon > 0$, ψ^{ε} is a minimizer for $\mathcal{J}_{\varepsilon}$, we have

$$\mathcal{J}_{\varepsilon}[\psi^{\star}] \geq \mathcal{J}_{\varepsilon}[\psi^{\varepsilon}].$$

Since $\sigma_{\varepsilon}(\psi^{\star}) \to u^{\star}$ strongly in $L^2(\mathbb{Q})$ (see the proof of Theorem 2.1), we also have

$$\mathcal{J}[\psi^{\star}] = \limsup_{\varepsilon \to 0} \mathcal{J}_{\varepsilon}[\psi^{\star}] \geq \limsup_{\varepsilon \to 0} \mathcal{J}_{\varepsilon}[\psi^{\varepsilon}]$$

By weak lower semicontinuity of the norms, we deduce

$$\liminf_{\varepsilon \to 0} \mathcal{J}_{\varepsilon}[\psi^{\varepsilon}] \ge \mathcal{J}[\psi^{\star}].$$

Therefore ψ^{\star} is a minimizer of \mathcal{J} .

Remarks and further topics

We discuss some interesting topics concerning obstacle problems that have not been covered.

• An important problem which is beyond the scope of this work is the mathematical analysis of the free boundary

$$\Gamma(u) := \partial \{u > \psi\}.$$

This topic is an interplay of analysis of PDEs and geometric measure theory, and has (and still is, for different variants of the problem) been an important challenge in the mathematical study of obstacle problems. The main question is to understand the geometry and the regularity of the free boundary, as a priori this could be a very irregular object. The main results state that the free boundary is of finite perimeter, all the points on the free boundary are classified in regular and singular points. The set of regular points is an open subset of the free boundary, and is C^{∞} , while the singular points are in some sense very "rare" (the contact set has Lebesgue density 0 at these points). The regularity study and results are presented in [12] and [41]

• Given its variational formulation, it is natural to solve the obstacle problem numerically by using the finite element method. As for all PDEs, an important question is to know the error (in both energy and max-norms) for the approximations of the solution (and even for the free boundary). In the setup of Section I, for an "appropriate" mesh \mathcal{T} of Ω and a finite dimensional approximation $\mathcal{K}_{\mathcal{T}} \subset \mathcal{K}$, the discrete problem consists of finding $u_{\mathcal{T}} \in \mathcal{K}_{\mathcal{T}}$ such that

$$\langle \nabla u_{\mathcal{T}}, \nabla (v - u_{\mathcal{T}}) \rangle_{L^2(\Omega)} \ge \langle f, v - u_{\mathcal{T}} \rangle_{L^2(\Omega)}$$

for all $v \in \mathcal{K}_{\mathcal{T}}$. It can be shown that this problem also has a unique solution, with the error estimates

$$\|\nabla(u - u_{\mathcal{T}})\|_{L^{2}(\Omega)} \lesssim h_{\mathcal{T}}(\|f\|_{L^{2}(\Omega)} + \|\psi\|_{H^{2}(\Omega)})$$

and

$$\|u - u_{\mathcal{T}}\|_{L^{\infty}(\Omega)} < Ch_{\mathcal{T}}^2 |\log h_{\mathcal{T}}| \|u\|_{W^{2,\infty}(\Omega)},$$

where $h_{\mathcal{T}}$ denotes the mesh size (defined as the maximum of the diameters of every cell in the mesh). Pointwise error estimates are usually derived by analyzing the error in mesh cells near the contact set of the continuous solution $\{u = \psi\}$, and by subsequently applying a discrete maximum principle. For more detail, we refer to [**39**].

• An important variant (for both physics and mathematics) of the obstacle problem is the so-called *thin obstacle problem*. Consider the hyperplane $\mathcal{M} := \{x \in \mathbb{R}^n : x_n = 0\}$ and let $\Omega^+ \subset B(0, 1) \cap \{x_n \ge 0\}$ be a domain in \mathbb{R}^n with smooth boundary. The problem arises when minimizing

$$\mathcal{D}[u] := \int_{\Omega^+} |\nabla u|^2 \mathrm{d}x$$

among all functions $u \in \{w \in H_0^1(\Omega^+) : w \ge \varphi \text{ on } \mathcal{M} \cap \Omega^+\}$. Here the function $\varphi : \mathcal{M} \to \mathbb{R}$ satisfying $\varphi < 0$ on $\mathcal{M} \cap \partial \Omega^+$ is called the *thin obstacle*, because u is constrained to stay above the obstacle φ only on the (n-1)-dimensional hyperplane \mathcal{M} and not on the entire *n*-dimensional domain Ω^+ . Considering a slight change of the boundary conditions, namely u = 0 on $\partial \Omega^+ \setminus \mathcal{M}$ and $u \ge \varphi$ on \mathcal{M} yields the so-called *Signorini problem*. This problem arises in linear elasticity theory, namely in the study of the equilibrium of an elastic membrane that rests above a very thin object. Existence and uniqueness of solutions is guaranteed by similar arguments as for the classical problem. It is shown in [41, Chapter 9] that the optimal regularity for the solution is $C_{\text{loc}}^{1,\alpha}(\Omega^+ \cup \mathcal{M})$. The proof uses a penalization method, but requires results that are beyond the scope of this work. In the literature, the problem is often stated in its complementarity form

$$\begin{cases} \Delta u = 0 & \text{in } \Omega^+ \cap \{x_n > 0\} \\ \min\{-\partial_{x_n} u, u - \varphi\} = 0 & \text{on } \Omega^+ \cap \mathcal{M}. \end{cases}$$

- As mentioned in the introduction, some phenomena give rise to obstacle problems with integro-differential operators such as the fractional Laplacian $(-\Delta)^s$, $s \in (0, 1)$. In recent years, many of the known results for the classical and parabolic obstacle problems have found their analog in the case where the Laplacian is changed with its fractional counterpart. The regularity theory for the elliptic and parabolic fractional obstacle problems has been studied in [45] and [13] respectively, and convergence analysis for the finite element discretization may be found in [39].
- We did not manage to present numerical experiments for computing the optimal control (obstacle) for the elliptic or parabolic obstacle problem. A problem that has more often been considered in the mathematical and numerical literature is minimizing the L^2 error between the state and the target profile

$$\frac{1}{2} \|y - y_d\|_{L^2}^2 + \frac{\alpha}{2} \|u\|_{L^2}^2$$



FIGURE 11. The free boundary and the contact set for the thin obstacle problem. This figure was adapted from [44].

(here α is small) subject to a variational inequality constraint on the state

$$\langle Ly, v - y \rangle \ge \langle f + u, v - y \rangle, \quad v \ge \psi,$$

where the obstacle ψ is given, and L is an elliptic or parabolic operator. Such problems fall in the realm of *mathematical programs with equilibrium constraints*. An optimal control (should it exist) may be computed, for example, by considering a penalization of the variational inequality and reduce the problem to PDE-constrained minimization, and then solve this problem using the *adjoint method*.

APPENDIX

Convexity, coercivity, weak lower semicontinuity

Proposition A.1. — The set

$$\mathcal{K}(\psi) := \{ u \in H^1_0(\Omega) \colon u \ge \psi \ a.e. \ in \ \Omega \}$$

is convex, closed and is non-empty.

Proof. — Let $u, v \in \mathcal{K}(\psi)$. Then for any $\theta \in [0,1]$, we have $\theta u + (1-\theta)v \ge \theta \psi + (1-\theta)\psi = \psi$. Thus $\mathcal{K}(\psi)$ is convex.

To show that $\mathcal{K}(\psi)$ is closed, let $\{u_k\}_{k=0}^{\infty} \subset \mathcal{K}(\psi)$ be a sequence converging strongly in $H_0^1(\Omega)$ to some $u \in H_0^1(\Omega)$. By the Poincaré inequality, $u_k \to u$ strongly in $L^2(\Omega)$ as $k \to \infty$. Moreover, there exists a subsequence $\{u_{k_j}\}_{j=0}^{\infty}$ of $\{u_k\}_{k=0}^{\infty}$ such that $u_{k_j} \to u$ almost everywhere in Ω as $j \to \infty$. But for every $j \in \mathbb{N}$, one has $u_{k_j} \ge \psi$ a.e. in Ω , hence by letting $j \to \infty$ we deduce that $u \in \mathcal{K}(\psi)$.

To show that $\mathcal{K}(\psi)$ is non-empty, observe that since $\psi \in H^2(\Omega) \cap C^0(\overline{\Omega}), \psi^+ \in H^1(\Omega)$ (see [25, Lem.7.6, p.152]) where $\psi^+ = \max\{\psi, 0\}$ denotes the positive part of ψ . Since $\psi \leq 0$ on $\partial\Omega$, we have $\psi^+ = 0$ on $\partial\Omega$ in the trace sense. This implies $\psi^+ \in H^1_0(\Omega)$, and since $\psi^+ \geq \psi$, we deduce that $\psi^+ \in \mathcal{K}(\psi)$.

We now present some properties of the Dirichlet energy functional

$$\mathcal{E}[u] := \frac{1}{2} \int_{\Omega} |\nabla u|^2 \mathrm{d}x - \int_{\Omega} f u \mathrm{d}x,$$

that we used in various proofs.

Proposition A.2 (Convexity of \mathcal{E}). — The functional $\mathcal{E} : H_0^1(\Omega) \to \mathbb{R}$ is convex and strictly convex.

Proof. — Let $u, v \in H_0^1(\Omega)$ and $\theta \in [0, 1]$ be arbitrary. Then by the triangle inequality,

$$\mathcal{E}[\theta u + (1-\theta)v] \le \frac{\theta^2}{2} \int_{\Omega} |\nabla u|^2 \mathrm{d}x + \frac{(1-\theta)^2}{2} \int_{\Omega} |\nabla v|^2 \mathrm{d}x - \theta \int_{\Omega} f u \mathrm{d}x - (1-\theta) \int_{\Omega} f v \mathrm{d}x.$$

Since $t^2 \leq t$ on [0, 1], the convexity of \mathcal{E} follows.

Recall that \mathcal{E} is strictly convex if

$$\mathcal{E}[\theta u + (1-\theta)v] < \theta \mathcal{E}[u] + (1-\theta)\mathcal{E}[v],$$

for $u, v \in H_0^1(\Omega)$ with $u \not\equiv v$ and $\theta \in (0, 1)$. Observe that if $u \not\equiv v$, then at least either $u \not\equiv 0$ or $v \not\equiv 0$. Thus the strict convexity also follows from the inequality above, as $t^2 < t$ on (0, 1).

Proposition A.3 (Coercivity of \mathcal{E}). — The functional $\mathcal{E} : H_0^1(\Omega) \to \mathbb{R}$ is coercive, meaning

$$\mathcal{E}[u] \to +\infty \quad as \quad \|u\|_{H^1_0(\Omega)} \to +\infty.$$

Proof. — Let u be an arbitrary element of $H_0^1(\Omega)$. One obtains

$$\begin{split} \mathcal{E}[u] &\geq \frac{1}{2} \|u\|_{H_0^1(\Omega)}^2 - \Big| \int_{\Omega} f u \mathrm{d}x \Big| \\ &\geq \frac{1}{2} \|u\|_{H_0^1(\Omega)}^2 - \|f\|_{L^2(\Omega)} \|u\|_{L^2(\Omega)} \\ &\geq \Big(\frac{1}{2} \|u\|_{H_0^1(\Omega)} - C(\Omega, n) \|f\|_{L^2(\Omega)} \Big) \|u\|_{H_0^1(\Omega)} \end{split}$$

by virtue of the Cauchy-Schwarz and Poincaré inequalities respectively. Now observe that as $||u||_{H_0^1(\Omega)} \to +\infty$, at some point one clearly has $\frac{1}{2}||u||_{H_0^1(\Omega)} > c(\Omega, n)||f||_{L^2(\Omega)}$. Thus, \mathcal{E} is coercive.

9.3. Weak topology results. — We now present some helpful results for showing weak lower semicontinuity of various functionals. Henceforth X denotes a real Banach space.

Definition A.1. — A map $f : X \to (-\infty, +\infty]$ is said to be strongly (resp. weakly) lower semicontinuous if for every $\alpha \in \mathbb{R}$ the set

$$\{x \in X \colon f(x) \le \alpha\}$$

is closed (resp. weakly closed).

If f is strongly (resp. weakly) lower semicontinuous, then for every sequence $\{x_k\}_{k=0}^{\infty} \subset X$ such that $x_k \to x$ strongly in X (resp. $x_k \to x$ weakly in X) for some $x \in X$ as $k \to \infty$, we have

$$f(x) \le \liminf_{k \to \infty} f(x_k),$$

and conversely.

Theorem A.4. — Let \mathcal{K} be a convex subset of X. Then \mathcal{K} is weakly closed if and only if it is strongly closed.

Proof. — Convexity is actually not needed to show that if \mathcal{K} is weakly closed then it is strongly closed. In effect, this follows merely from the fact that strong (i.e. norm) convergence implies weak convergence.

Now suppose that \mathcal{K} is strongly closed. We will show that $X \setminus \mathcal{K}$ is weakly open. Let $x_0 \in X \setminus \mathcal{K}$. By the Hahn-Banach theorem (second geometric form, see [11, Thm.1.7, p.7]), there exists a closed hyperplane strictly separating $\{x_0\}$ and \mathcal{K} , meaning there exist some $f \in X'$ and $\alpha \in \mathbb{R}$ such that

(9.27)
$$f(x_0) < \alpha < f(y)$$
 for all $y \in \mathcal{K}$.

Now set

$$\mathcal{O} := \{ x \in X \colon f(x) < \alpha \}.$$

Then $\mathcal{O} = f^{-1}((\alpha, +\infty))$ is weakly open, and from (9.27), it follows that $x_0 \in \mathcal{O}$ and $\mathcal{O} \cap \mathcal{K} = \emptyset$. Hence $\mathcal{O} \subset X \setminus \mathcal{K}$ and so $X \setminus \mathcal{K}$ is weakly open.

Corollary A.5. — Assume that $f: X \to (-\infty, +\infty]$ is convex and strongly lower semicontinuous. Then f is weakly lower semicontinuous.

Proof. — For every $\alpha \in \mathbb{R}$ the set

$$\mathcal{K} = \{ x \in X \colon f(x) \le \alpha \}$$

is strongly closed and convex. By the previous Theorem, it is also weakly closed and thus f is weakly lower semicontinuous.

It may be difficult to show that a function is weakly lower semicontinuous. In practice, we use the above Corollary to continuous and convex functions f and deduce that f is weakly lower semicontinuous. For example, the function f(x) = ||x|| is convex and strongly continuous, thus it weakly lower semicontinuous.

Proposition A.6 (Weak lower semicontinuty of \mathcal{E}). — The functional \mathcal{E} : $H_0^1(\Omega) \to \mathbb{R}$ is weakly lower semicontinuous.

Proof. — We proceed by analyzing separately both terms. In light of the previous Corollary and the subsequent remarks, we only need to investigate the second integral term as the first is the $\|\cdot\|_{H^1_0(\Omega)}$ norm. Observe that the map $F: H^1_0(\Omega) \to \mathbb{R}$ defined as

$$F: u\mapsto \int_\Omega -fu\mathrm{d} x$$

is linear, and also bounded by the Cauchy-Schwarz and Poincaré inequalities (recall that $f \in L^2(\Omega)$). It is thus convex and strongly continuous, therefore weakly lower semicontinuous.

Now assume $u_k \rightharpoonup u$ weakly in $H_0^1(\Omega)$. Since the limit inferior is superadditive, it follows that

$$\begin{split} \mathcal{E}[u] &= \frac{1}{2} \|u\|_{H_0^1(\Omega)} + F(u) \leq \liminf_{k \to \infty} \|u_k\|_{H_0^1(\Omega)} + \liminf_{k \to \infty} F(u_k) \\ &\leq \liminf_{k \to \infty} \left(\frac{1}{2} \|u_k\|_{H_0^1(\Omega)} + F(u_k)\right) \\ &= \liminf_{k \to \infty} \mathcal{E}[u_k]. \end{split}$$

Proposition A.7. — Assume that

$$u_k \rightharpoonup u \quad weakly \ in \ L^2(0,T; H^1_0(\Omega))$$
$$(u_k)_t \rightharpoonup v \quad weakly \ in \ L^2(0,T; H^{-1}(\Omega))$$

as $k \to \infty$. Then $v = u_t$.

This is Problem 4 from [21, p.425].

Proof. — Let $\varphi \in C_c^{\infty}(\Omega)$ and $w \in H_0^1(\Omega)$. Then the map $t \mapsto \varphi(t, \cdot)w$ is in $L^2(0, T; H_0^1(\Omega))$. Now denoting by $\langle \cdot, \cdot \rangle$ the duality pairing between $H^{-1}(\Omega)$ and $H_0^1(\Omega)$, we have

$$\left\langle \int_0^T \varphi_t(t,\cdot)u(t,\cdot)\mathrm{d}t, w \right\rangle = \int_0^T \langle \varphi_t(t,\cdot)u(t,\cdot), w \rangle \mathrm{d}t$$
$$= \int_0^T \langle u(t,\cdot), \varphi_t(t,\cdot)w \rangle \mathrm{d}t.$$

The weak $L^2(0,T; H^1_0(\Omega))$ convergence gives

$$\int_0^T \langle u(t,\cdot), \varphi_t(t,\cdot)w \rangle dt = \lim_{k \to \infty} \int_0^T \langle u_k(t,\cdot), \varphi_t(t,\cdot)w \rangle dt$$
$$= \lim_{k \to \infty} \Big\langle \int_0^T \varphi_t(t,\cdot)u_k(t,\cdot)dt, w \Big\rangle.$$

By definition of the weak derivative, we have

$$\lim_{k \to \infty} \left\langle \int_0^T \varphi_t(t, \cdot) u_k(t, \cdot) dt, w \right\rangle = \lim_{k \to \infty} \left\langle -\int_0^T \varphi(t, \cdot) (u_k)_t(t, \cdot) dt, w \right\rangle$$
$$= \lim_{k \to \infty} \int_0^T -\langle (u_k)_t(t, \cdot), \varphi(t, \cdot) w \rangle dt.$$

Using the weak $L^2(0,T; H^{-1}(\Omega))$ convergence of the derivatives, we obtain

$$\lim_{k \to \infty} \int_0^T -\langle (u_k)_t(t, \cdot), \varphi(t, \cdot)w \rangle dt = \int_0^T -\langle v(t, \cdot), \varphi(t, \cdot)w \rangle dt$$
$$= \left\langle -\int_0^T \varphi(t, \cdot)v(t, \cdot)dt, w \right\rangle$$

Combining the previous chain of equalities, we deduce

$$\left\langle \int_0^T \varphi_t(t,\cdot) u(t,\cdot) \mathrm{d}t + \int_0^T \varphi(t,\cdot) v(t,\cdot) \mathrm{d}t, w \right\rangle = 0.$$

Since the above holds for all $w \in H_0^1(\Omega)$ and for every test function $\varphi \in C_c^{\infty}(\Omega)$, we obtain the desired result. \Box

Proposition A.8. — Suppose H is a real Hilbert space and

 $u_k \rightharpoonup u$ weakly in $L^2(0,T;H)$.

Suppose further that there exists C > 0 such that

$$\operatorname{ess\,sup}_{t\in[0,T]} \|u_k(t,\cdot)\|_H \le C,$$

for all $k \in \mathbb{N}$. Then

$$\operatorname{ess\,sup}_{t\in[0,T]} \|u(t,\cdot)\|_H \le C.$$

This is Problem 5 from [21, p.425].

Proof. — Let $v \in H$ and $0 \le a \le b \le T$ be given. From the assumed bound, we clearly have

$$\int_{a}^{b} \langle v, u_{k}(t, \cdot) \rangle_{H} \mathrm{d}t \leq C ||v||_{H} |b-a|,$$

and from the assumed weak convergence, we also deduce

$$\int_{a}^{b} \langle v, u(t, \cdot) \rangle_{H} \mathrm{d}t \le C \|v\|_{H} |b-a|$$

Now notice that

$$\int_0^T |\langle v, u(t, \cdot) \rangle_H|^2 \mathrm{d}t \le T ||v||_H ||u||_{L^2(0,T;H)}$$

by applying Cauchy-Schwarz twice, hence $t \mapsto \langle v, u(t, \cdot) \rangle_H$ is in $L^2(0, T; H)$. Now the bound before last may be rewritten as

$$\frac{1}{|b-a|} \int_{a}^{b} \langle v, u(t, \cdot) \rangle_{H} \mathrm{d}t \le C \|v\|_{H},$$

whence the Lebesgue differentiation theorem yields

$$\langle v, u(t, \cdot) \rangle_H \le C \|v\|_H$$

for a.e. $t \in (0,T)$ and for all $v \in H$. Choosing $v = u(t, \cdot)$ shows the claim.

Inequalities

Proposition A.9 (Young). — Let $1 < p, q < \infty$, $\frac{1}{p} + \frac{1}{q} = 1$. Then $ab \leq \frac{a^p}{p} + \frac{b^q}{q}$

for a, b > 0.

Proof. — Since $x \mapsto e^x$ is convex, it follows that

$$ab = e^{\ln a + \ln b} = e^{\frac{1}{p}\ln a^p + \frac{1}{q}\ln b^q} \le \frac{1}{p}e^{\ln a^p} + \frac{1}{q}e^{\ln b^q} = \frac{a^p}{p} + \frac{b^q}{q}.$$

We will often use Young's inequality and its following variant in the case p = q = 2.

Proposition A.10 (Young with ε). — Let $1 < p, q < \infty$, $\frac{1}{p} + \frac{1}{q} = 1$. Then

$$ab \le \varepsilon a^p + \frac{(\varepsilon p)^{-q/p}b}{q}$$

for a, b > 0 and $\varepsilon > 0$.

Proof. — Observe that

$$ab = \left((\varepsilon p)^{1/p} a \right) \left((\varepsilon p)^{-1/p} b \right),$$

so applying Young's inequality yields the desired result.

Proposition A.11 (Gronwall, differential form). — Let f be a nonnegative, absolutely continuous function on [0, T], which satisfies the differential inequality

$$f'(t) \le \varphi(t)f(t) + g(t)$$

for a.e. $t \in [0,T]$, where φ, g are nonnegative, integrable functions on [0,T]. Then

$$f(t) \le e^{\int_0^t \varphi(s) \mathrm{d}s} \left(f(0) + \int_0^t g(s) \mathrm{d}s \right)$$

for all $t \in [0,T]$. In particular, if $g \equiv 0$ on [0,T] and f(0) = 0, then $f \equiv 0$ on [0,T].

Proof. — Using the assumed differential inequality, one has

$$\frac{\mathrm{d}}{\mathrm{d}s}\Big(f(s)e^{\int_0^s\varphi(\tau)\mathrm{d}\tau}\Big) = e^{-\int_0^s\varphi(\tau)\mathrm{d}\tau}\big(f'(s) - \varphi(s)f(s)\big) \le e^{-\int_0^s\varphi(\tau)\mathrm{d}\tau}g(s)$$

for a.e. $s \in [0, T]$. Now for each $t \in [0, T]$, we integrate the above between 0 and t to obtain

$$f(t)e^{-\int_0^t \varphi(\tau)\mathrm{d}\tau} - f(0) \le \int_0^t e^{-\int_0^s \varphi(\tau)\mathrm{d}\tau} g(s)\mathrm{d}s.$$

Since φ is nonnegative, rearranging the above we obtain the desired result.

Proposition A.12 (Gronwall, integral form). — Let ζ be a nonnegative, integrable function on [0,T] which satisfies the integral inequality

$$\zeta(t) \le C_1 \int_0^t \zeta(s) \mathrm{d}s + C_2$$

for a.e. $t \in [0,T]$ and for some constants $C_1, C_2 \ge 0$. Then

$$\zeta(t) \le C_2 (1 + C_1 t e^{C_1 t})$$

for a.e. $t \in [0,T]$. In particular, if $C_2 = 0$, then $\zeta = 0$ a.e. on [0,T].

Proof. — Consider

$$f(t) := \int_0^t \zeta(s) \mathrm{d}s.$$

Then $f' \leq C_1 f + C_2$ a.e. in [0, T] by assumption. Due to the differential form of Gronwall's inequality, one has

$$f(t) \le e^{C_1 t} (f(0) + C_2 t) = C_2 t e^{C_1 t}.$$

Then by the assumed integral inequality,

$$\zeta(t) \le C_1 f(t) + C_2 \le C_2 (1 + C_1 t e^{C_1 t}).$$

We now prove Lemma 5.1 giving an inequality for the map

$$g: u \mapsto \frac{\gamma}{\delta} \min\{\delta, u - \psi\}.$$

Proof of Lemma 5.1. — Let $u, v \in H_0^1(\Omega)$ be arbitrary. One has

$$g(v) - g(u) = \frac{\gamma}{\delta} \begin{cases} 0 & \text{if } v \ge \psi + \delta \text{ and } u \ge \psi + \delta \\ \delta - u + \psi & \text{if } v \ge \psi + \delta \text{ and } u < \psi + \delta \\ v - \psi - \delta & \text{if } v < \psi + \delta \text{ and } u \ge \psi + \delta \\ v - u & \text{if } v < \psi + \delta \text{ and } u < \psi + \delta \end{cases}$$
$$= \frac{\gamma}{\delta} \begin{cases} \mathcal{H}(\delta - u + \psi)(v - u) & \text{if } v \ge \psi + \delta \text{ and } u \ge \psi + \delta \\ \mathcal{H}(\delta - u + \psi)(\delta - u + \psi) & \text{if } v \ge \psi + \delta \text{ and } u < \psi + \delta \\ v - \psi - \delta & \text{if } v < \psi + \delta \text{ and } u \ge \psi + \delta \\ \mathcal{H}(\delta - u + \psi)(v - u) & \text{if } v < \psi + \delta \text{ and } u \ge \psi + \delta \end{cases}$$

Now observe that $\delta - u + \psi \leq v - u$ when $v \geq \psi + \delta$, and $v - \psi - \delta < \mathcal{H}(\delta - u + \psi)(v - u)$ when $v < \psi + \delta$ and $u \geq \psi + \delta$. Whence from the above computation it follows that

$$g(v) - g(u) \le \frac{\gamma}{\delta} \mathcal{H}(\delta - u + \psi)(v - u).$$

Monotone operator theory

We briefly present some results from monotone operator theory related to the study of variational inequalities in a Hilbert space. The property of monotonicity is often useful for showing the existence of a solution to a variational inequality. We will see that a variational inequality associated to a monotone operator over a convex set enjoying certain continuity properties can be solved. In general, if the convex set is unbounded, it will be necessary to add hypotheses of coercivity to achieve the existence of a solution.

Let H be a real Hilbert space, a priori not identified with its dual H', and denote by $\langle \cdot, \cdot \rangle$ their duality pairing. Let $\mathcal{K} \subset H$ be a closed convex and non-empty set; further assumptions will be made subsequently.

Definition B.1. — A map $A : \mathcal{K} \to H'$ is called monotone if

$$\langle Au - Av, u - v \rangle \ge 0$$

for all $u, v \in \mathcal{K}$. A is called strictly monotone if

 $\langle Au - Av, u - v \rangle = 0$ implies u = v.

Definition B.2. — A map $A : \mathcal{K} \to H'$ is continuous on finite dimensional subspaces if for any finite dimensional subspace $\mathcal{M} \subset H$ the map $A : \mathcal{K} \cap \mathcal{M} \to H'$ is weakly continuous.

Definition B.3. — A map $A: \mathcal{K} \to H'$ is coercive if there exists $\varphi \in \mathcal{K}$ such that

$$\frac{\langle Au - A\varphi, u - \varphi \rangle}{\|u - \varphi\|_H} \to +\infty \quad as \quad \|u\|_H \to +\infty$$

for any $u \in \mathcal{K}$.

We prove the following simple but powerful lemma due to Minty.

Lemma B.2 (Minty). — Let $\mathcal{K} \subset H$ be a closed and convex set, and let $A : \mathcal{K} \to H'$ be monotone and continuous on finite dimensional subspaces. Then $u \in \mathcal{K}$ satisfies

(9.28)
$$\langle Au, v - u \rangle \ge 0 \quad \text{for all } v \in \mathcal{K}$$

if and only if it satisfies

(9.29)
$$\langle Av, v - u \rangle \ge 0 \quad \text{for all } v \in \mathcal{K}.$$

Proof. — Assume that $u \in \mathcal{K}$ solves (9.28). By monotonicity of A,

$$0 \le \langle Av - Au, v - u \rangle = \langle Av, v - u \rangle - \langle Au, v - u \rangle$$

for all $v \in \mathcal{K}$. Thus,

$$0 \le \langle Au, v - u \rangle \le \langle Av, v - u \rangle$$

for all $v \in \mathcal{K}$.

Conversely, assume that u solves (9.29). Let $w \in \mathcal{K}$ and set $v := u + \tau(w - u) \in \mathcal{K}$ for $\tau \in (0, 1]$ since \mathcal{K} is convex. Hence by (9.29)

$$\langle A(u+\tau(w-u)), \tau(w-u) \rangle \ge 0,$$

whence it follows that

$$\langle A(u+\tau(w-u)), w-u \rangle \ge 0$$

for all $w \in \mathcal{K}$. Since A is weakly continuous on the intersection of \mathcal{K} and the finite dimensional subspace spanned by u and w, we let $\tau \to 0$ to obtain (9.28).

We now give without proof the main existence results (see [30]).

Theorem B.1. — Let $\mathcal{K} \subset H$ be a closed, bounded and convex, and let $A : \mathcal{K} \to H'$ be monotone and continuous on finite dimensional subspaces. Then there exists $u \in \mathcal{K}$ solving the variational inequality

$$\langle Au, v - u \rangle \ge 0$$

for all $v \in \mathcal{K}$. Moreover, if A is strictly monotone, then u is unique.

Corollary B.4. — Let $\mathcal{K} \subset H$ be a closed, bounded, convex and non-empty set, and let $F : \mathcal{K} \to \mathcal{K}$ be a proper map. Then the set of fixed points of F is closed and non-empty.

Corollary B.5. — Let $\mathfrak{K} \subset H$ be a closed, convex and non-empty set and let $A : \mathfrak{K} \to H'$ be monotone, coercive and continuous on finite dimensional subspaces. Then there exists $u \in \mathfrak{K}$ such that

$$\langle Au, v-u \rangle \ge 0$$

for all $v \in \mathcal{K}$.

References

- ADAMS, D., AND LENHART, S. An obstacle control problem with a source term. Appl Math Optim 47 (2002), 79–95.
- [2] ADAMS, D., AND LENHART, S. Optimal control of the obstacle for a parabolic variational inequality. J Math Anal Appl 268 (2002), 602–614.
- [3] ADAMS, D., LENHART, S., AND YONG, J. Optimal control of the obstacle for an elliptic variational inequality. *Appl Math Optim 38* (1998), 121–140.
- [4] ADAMS, R. Sobolev Spaces. Academic Press, 1975.
- [5] ALNÆS, M. S., BLECHTA, J., HAKE, J., JOHANSSON, A., KEHLET, B., LOGG, A., RICHARDSON, C., RING, J., ROGNES, M. E., AND WELLS, G. N. The fenics project version 1.5. Archive of Numerical Software 3, 100 (2015).
- [6] ATHANASOPOULOS, I., CAFFARELLI, L., AND MILAKIS, E. Parabolic obstacle problems: Quasi-convexity and regularity. *ArXiv e-prints* (2016).
- [7] AUBIN, J.-P. Un théorème de compacité. C.R. Acad. Sci. Paris 256 (1963), 5042– 5044.
- [8] BALAGUÉ, D., CARRILLO, J., LAURENT, T., AND RAOUL, G. Dimensionality of local minimizers of the interaction energy. Arch. Ration. Mech. Anal. 209, 3 (2013), 1055–1088.
- [9] BARBU, V., AND PRECUPANU, T. Convexity and Optimization in Banach Spaces. Springer Monographs in Mathematics. Springer Netherlands, 2012.
- [10] BERGOUNIOUX, M. Optimal control of an obstacle problem. Appl Math Optim. 36 (1997), 147–172.
- [11] BREZIS, H. Functional analysis, Sobolev spaces and partial differential equations. Universitext. Springer, 2011.
- [12] CAFFARELLI, L. The obstacle problem revisited. J. Fourier Anal. Appl. 4, 4-5 (1998), 383–402.
- [13] CAFFARELLI, L., AND FIGALLI, A. Regularity of solutions to the parabolic fractional obstacle problem. J. Reine Angew. Math 680 (2011), 191–233.
- [14] CAFFARELLI, L., AND FRIEDMAN, A. Continuity of the temperature in the Stefan problem. *Indiana Univ. Math. J.* 28, 1 (1979), 53–70.
- [15] CARRILLO, J., DELGADINO, M., AND MELLET, A. Regularity of local minimizers of the interaction energy via obstacle problems. *Comm. Math. Phys.* 343, 3 (2016), 747–781.
- [16] CHEN, X., JÜNGEL, A., AND LIU, J.-G. A note on Aubin-Lions-Dubinskii lemmas. Acta Appl. Math. 133 (2014), 33–43.
- [17] CHIPOT, M. Variational Inequalities and Flow in Porous media. Appl. Math. Sci. 52, Springer-Verlag, New York, 1984.
- [18] DUVAUT, G. Résolution d'un problème de Stefan (fusion d'un bloc de glace a zero degrées). C.R. Acad. Sci. Paris 276 (1973), 1461–1463.

- [19] DUVAUT, G., AND LIONS, J. Inequalities in Mechanics and Physics, vol. 219 of Grundlehren der Matematischen Wissenschaften. Springer-Verlag, Berlin-New York, 1976.
- [20] EKELAND, I., AND TEMAM, R. Convex analysis and Variational Problems, vol. 28 of Classics in Applied Mathematics. Society for Industrial and Applied Mathematics, Philadelphia, PA, 1999.
- [21] EVANS, L. Partial Differential Equations, vol. 19 of Graduate Studies in Mathematics. American Mathematical Society, Providence, RI, 1998.
- [22] EVANS, L. An Introduction to Stochastic Differential Equations. American Mathematical Society, Providence, RI., 2013.
- [23] FRIEDMAN, A. Variational Principles and Free Boundary Problems, second ed. Robert E. Krieger Publishing Co., Inc., Malabar, FL, 1982.
- [24] GIAQUINTA, M., AND MARTINAZZI, L. An Introduction to the Regularity theory for Elliptic Systems, Harmonic Maps and Minimal Graphs, second ed., vol. 11. Edizioni della Normale, Pisa, 2012.
- [25] GILBARG, D., AND TRUDINGER, N. Elliptic Partial Differential Equations of Second Order. Classics in Mathematics. Springer-Verlag, Berlin, 2001.
- [26] GRISVARD, P. Elliptic Problems in Nonsmooth Domains, vol. 24 of Monographs and Studies in Mathematics. Advanced Publishing, Boston, MA, 1985.
- [27] HINTERMÜLLER, M., KOVTUNENKO, V., AND KUNSCH, K. Obstacle problems with cohesion: a hemivariational inequality approach and its efficient numerical solution. *SIAM J.Optim 21* (2011), 491–516.
- [28] HINTERMÜLLER, M., AND KOPACKA, I. A smooth penalty approach and a nonlinear multigrid algorithm for elliptic mpecs. *Comput. Optim. Appl.* 50, 1 (2011), 111–145.
- [29] ITO, K., AND KUNISCH, K. Optimal control of obstacle problems by H^1 obstacles. Appl Math Optim. 56 (2007), 1–17.
- [30] KINDERLEHRER, D., AND STAMPACCHIA, G. An Introduction to Variational Inequalities and their Applications, vol. 31 of Classics in Applied Mathematics. Society for Industrial and Applied Mathematics, Philadelphia, PA, 1980.
- [31] KOVTUNENKO, V. A hemivariational inequality in crack problems. Optimization 60, 8-9 (2011), 1071–1089.
- [32] LAURENCE, P., AND SALSA, S. Regularity of the free boundary of an American option on several assets. Comm. Pure Appl. Math. 62 (2009), 969–994.
- [33] LEUGERING, G., PRECHTEL, M., STEINMAN, P., AND STINGL, M. A cohesive crack propagation model: mathematical theory and numerical solution. *Commun. Pure App. Anal.* 12, 4 (2013), 1705–1729.
- [34] LIONS, J.-L. Quelques methodes de résolution des problèmes aux limites non linéaires. Dunod; Gauthier-Villars, Paris, 1969.

- [35] LOGG, A., MARDAL, K.-A., WELLS, G. N., ET AL. Automated Solution of Differential Equations by the Finite Element Method. Springer, 2012.
- [36] LOGG, A., AND WELLS, G. N. Dolfin: Automated finite element computing. ACM Transactions on Mathematical Software 37, 2 (2010).
- [37] LOGG, A., WELLS, G. N., AND HAKE, J. DOLFIN: a C++/Python Finite Element Library. Springer, 2012, ch. 10.
- [38] MERTON, R. Option pricing when the underlying stock returns are discontinuous. J. Finan. Econ. 5 (1976), 125–144.
- [39] NOCHETTO, R., OTÁROLA, E., AND SALGADO, A. Convergence rates for the classical, thin and fractional elliptic obstacle problems. *Philos. Transact. A Math. Phys. Eng. Sci.* 373 (2015).
- [40] PETROSYAN, A., AND SHAHGHOLIAN, H. Parabolic obstacle problems applied to finance. Contemp. Math. 439 (2007), 117–133.
- [41] PETROSYAN, A., SHAHGHOLIAN, H., AND URALTSEVA, N. Regularity of Free Boundaries in Obstacle-type Problems, vol. 136 of Graduate Studies in Mathematics. American Mathematical Society, Providence, RI, 2012.
- [42] PHAM, H. Optimal stopping, free boundary, and American option in a jump-diffusion model. Appl. Math. Optim. 35 (1997), 145–164.
- [43] RODRIGUES, J. Obstacle Problems in Mathematical Physics, vol. 114 of Notas de Matemática. North-Holland Publishing Co., Amsterdam, 1987.
- [44] ROS-OTON, X. Obstacle problems and free boundaries: an overview. SeMa Journal (2017), 1–21.
- [45] SILVESTRE, L. Regularity of the obstacle problem for a fractional power of the Laplace operator. *Comm. Pure Appl. Math.* 60 (2006), 67–112.
- [46] SIMON, J. Compact sets in the space $L^p(0,T;B)$. Ann. Mat. Pura Appl. 146 (1987), 65–96.

Borjan Geshkovski

Master 2 Mathématiques et Applications Analyse, Équations aux dérivées partielles, Probabilités Université de Bordeaux

E-MAIL: borjan.geshkovski@etu.u-bordeaux.fr

BORJAN GESHKOVSKI